



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES

DEPARTAMENTO DE EDUCACIÓN

TESIS

EVIDENCIAS DE VALIDEZ DEL EXAMEN INSTITUCIONAL *PLACEMENT ENGLISH TEST* DE LA UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES: RELACIÓN CON OPINIONES DE DOCENTES Y ESTUDIANTES

PRESENTA

José Armando Collazo García

PARA OBTENER EL GRADO DE MAESTRÍA EN INVESTIGACIÓN EDUCATIVA

TUTORA

Dra. Guadalupe Ruíz Cuellar

CO-TUTOR:

Dr. Luis Horacio Pedroza Zúñiga

COMITÉ TUTORIAL

Lic. Felipe Martínez Rizo

Dr. Adán Moisés García Medina

Aguascalientes, Ags, marzo de 2018



UNIVERSIDAD AUTONOMA
DE AGUASCALIENTES
CENTRO DE CIENCIAS SOCIALES
Y HUMANIDADES

DEC. CCS Y H OF. N° 0266
Asunto: Conclusión de Tesis

**DRA. EN ADMÓN. MARÍA DEL CARMEN MARTÍNEZ SERNA
DIRECTORA GENERAL DE INVESTIGACIÓN Y POSGRADO
P R E S E N T E.**

Por este conducto le informo que el documento final de Tesis/Trabajo Práctico Titulado: "EVIDENCIAS DE VALIDEZ DEL EXAMEN INSTITUCIONAL PLACEMENT ENGLISH TEST DE LA UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES: RELACIÓN CON OPINIONES DE DOCENTES Y ESTUDIANTES", presentado por el sustentante JOSÉ ARMANDO COLLAZO GARCÍA con ID. 120476, egresado de la MAESTRÍA EN INVESTIGACIÓN EDUCATIVA, cumple las normas y lineamientos establecidos institucionalmente para presentar el examen de grado.

Sin más por el momento, aprovecho la oportunidad para enviarle un cordial saludo.

**ATENTAMENTE
"SE LUMEN PROFERRE"
Aguascalientes, Ags. A 5 de Marzo del 2018**

**DRA. GRISELDA ALICIA MACÍAS IBARRA
DECANA**

c.c.p. Dr. Francisco Javier Pedroza Cabrera. Secretario de Investigación y Posgrado del CCS y H.
c.c.p. Dra. Victoria Eugenia Gutiérrez Marfileño. Secretaria Técnica de la Maestría en Inv. Educativa
c.c.p. Mtra. Imelda Jiménez García. Jefa del Depto. De Control Escolar
c.c.p. Lic. José Armando Collazo García. Egresado de la Maestría en Investigación Educativa
c.c.p. Archivo



UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES

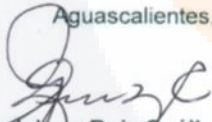
DRA. GRISELDA ALICIA MACÍAS IBARRA
DECANA DEL CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES
P R E S E N T E


Por medio de la presente, como comité tutorial designado del estudiante **JOSÉ ARMANDO COLLAZO GARCÍA** con ID 120476 quien realizó la tesis titulada: **EVIDENCIAS DE VALIDEZ DEL EXAMEN INSTITUCIONAL PLACEMENT ENGLISH TEST DE LA UNIVERSIDAD AUTÓNOMA DE AGUASCALIENTES: RELACIÓN CON OPINIONES DE DOCENTES Y ESTUDIANTES**, y con fundamento en el Artículo 175 Apartado II del Reglamento General de Docencia, nos permitimos emitir el **VOTO APROBATORIO** para que él pueda proceder a su impresión. De igual manera, el estudiante podrá continuar con el procedimiento administrativo para la obtención del grado en el programa de Maestría en Investigación Educativa.


Ponemos lo anterior a su digna consideración y sin otro particular por el momento, le enviamos un cordial saludo.

ATENTAMENTE
"SE LUMEN PROFERRE"

Aguascalientes, Ags., a 01 de marzo de 2018


Dra. Guadalupe Ruiz Cuéllar
Tutora de tesis


Dr. Luis Horacio Pedroza Zúñiga
Co-tutor de tesis


Lic. Felipe Martínez Rizo
Integrante Comité Tutorial


Dr. Adán Moisés García Medina
Integrante Comité Tutorial

c.c.p. Interesado
c.c.p. Secretaría Técnica de la Maestría en Investigación Educativa



2 MAR 2018

AZUCENA LIRA

AGRADECIMIENTOS

Que las siguientes líneas sirvan como un agradecimiento a todas aquellas personas que creyeron en mí y de quienes recibí apoyo durante este proceso.

En primer lugar, quiero agradecer a las instituciones y programas que me dieron la oportunidad de llevar a cabo este proyecto, pues esto no sería posible sin el apoyo económico y la confianza brindada por parte de las mismas. Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT, México) por la convocatoria que me ha permitido cumplir con la meta de estudiar una maestría, apoyándome con una beca que me favoreció de tal forma que pude dedicarme con constancia y motivación a la adquisición de conocimientos sobre la investigación educativa. Agradezco a la Universidad Autónoma de Aguascalientes (UAA) por el subsidio con que me apoyó además de haberme otorgado la confianza para realizar un posgrado dentro de sus programas de investigación.

También agradezco al Departamento de Educación de la UAA, a su coordinación, siempre al pendiente y al servicio de sus estudiantes; a sus docentes, expertos en temas educativos que de manera profesional han compartido sus conocimientos con todos aquellos interesados en formarse como investigadores educativos; y a la línea de evaluación, en especial al Dr. Luis Horacio Pedroza Zúñiga y a la Dra. María Guadalupe Ruíz Cuellar, de quienes recibí un apoyo constante y supieron cómo guiarme en la presente investigación. Gracias a mi comité, el Lic. Felipe Martínez Rizo y el Dr. Adán Moisés García Medina, de quienes tuve la fortuna de ser testigo de su dedicación, brillantez, y experiencia en la investigación educativa al recibir retroalimentación y comentarios que me ayudaron a realizar este proyecto.

No quiero dejar pasar la oportunidad de agradecer a mis compañeros, en especial a Luis Ernesto Alemán Macías, Aldo Raúl Díaz Medina, Eduardo Gómez Gómez, Carlos Francisco Lozano Corrales, Indira Viridiana Medina Mendoza, Iris Alicia Meza Muñoz, y José Fernando Ramos Leal, con quienes intercambié charlas interesantes y divertidos momentos, además de recibir ayuda y asesoría en más de una ocasión de su parte. Más que compañeros, personas sin cuya amistad esta experiencia no habría sido la misma.

Por último, agradezco infinitamente a mi madre y a mi hermana, personas a quienes amo más allá de lo que las palabras pueden expresar. A ellas dedico éste y todos los logros en mi vida.

Dedicatoria

Dedico mi tesis a mi madre, Rebeca García Ramírez, y a mi hermana, Olivia Collazo García, pues ellas me han dado apoyo incondicional e irremplazable, no sólo en éste, sino en todos los proyectos de mi vida. .

Son mi familia, mi motivación, la representación más sublime de amor y cariño con que la vida me ha premiado.

Para ustedes es mi esfuerzo, pues son mi inspiración para ser alguien mejor, y ningún logro mío es suficiente para poder expresarles lo profundamente agradecido que estoy con todo lo que han hecho por mí.

Porque el único éxito que me importa es el que les otorga felicidad y bienestar, el que las haga sentirse orgullosas de la persona que soy.

Madre, ha hecho bien las cosas, y hemos llegado a donde estamos gracias a usted.

Nada nos debe, y en cambio, nosotros se lo debemos todo.

Las amo inmensamente, y nada de esto tendría sentido para mí si ustedes no estuvieran conmigo.

Me hacen falta tantas palabras para expresarles la gratitud y el amor que siento por ustedes, así que seguiré en busca del éxito que les regrese un poco de toda la felicidad que me han dado en mi vida.

Como lo dijo André Maurois, “Sin una familia, el hombre, solo en el mundo, tiembla de frío”.

INDICE GENERAL

INDICE GENERAL..... 1

INDICE DE TABLAS..... 2

RESUMEN..... 6

ABSTRACT..... 7

INTRODUCCIÓN..... 8

Capítulo 1. PLANTEAMIENTO DEL PROBLEMA..... 10

1.1 Antecedentes..... 10

1.2 Problematización..... 16

1.3 Objetivos y preguntas de investigación 18

1.4 Justificación..... 19

Capítulo 2. REFERENTES TEÓRICOS..... 21

2.1 Validez y Evidencias de Validez 21

2.2 El Marco Común Europeo de Referencia para las Lenguas..... 39

Capítulo 3. METODOLOGÍA DE LA INVESTIGACIÓN..... 57

3.1 Participantes..... 58

3.2 Instrumentos..... 60

3.3 Procedimiento..... 68

Capítulo 4. RESULTADOS 79

4.1 Resultados de la Encuesta para la Autopercepción de Habilidades en Inglés ... 79

4.2 Resultados de la Encuesta para el Juicio Docente 90

4.3 Resultados de la Rúbrica de Evaluación de Desempeño..... 95

4.4 Análisis de Correlación..... 102

Capítulo 5. DISCUSIÓN DE RESULTADOS 117

5.1 Autopercepción de Habilidades en Inglés..... 117

5.2 Juicio Docente..... 125

5.3 Desempeño de los Estudiantes en su Examen Intermedio..... 129

Capítulo 6. Conclusiones..... 132

Referencias..... 138

ANEXOS..... 146

INDICE DE TABLAS

Tabla 1. Facetas de validez.....25

Tabla 2. Tabla de niveles del MCER 42

Tabla 3. Tabla de comprensión auditiva por niveles 47

Tabla 4 Tabla de comprensión lectora por niveles 51

Tabla 5. Tabla de contenidos gramaticales por nivel del MCER 56

Tabla 6. Cantidad de participantes por instrumento 59

Tabla 7. Instrumentos de obtención de información e informantes 67

Tabla 8. Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de comprensión lectora. 83

Tabla 9. Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de comprensión auditiva 86

Tabla 10. Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de gramática. 89

Tabla 11. Resultados de fiabilidad en Encuestas para la Autopercepción de Habilidades en Inglés 90

Tabla 12. Porcentajes de respuesta para Comprensión Lectora en las Encuestas para el Juicio Docente 91

Tabla 13. Porcentajes de respuesta para Comprensión Auditiva en las Encuestas para el Juicio Docente 92

Tabla 14. Porcentajes de respuesta para Gramática en las Encuestas para el Juicio Docente..... 93

Tabla 15. Porcentajes de respuesta para la Valoración Global en las Encuestas para el Juicio Docente 94

Tabla 16. Distribución de exámenes intermedios analizados con la rúbrica95

Tabla 17. Porcentajes de alumnos que pueden realizar actividades de Comprensión Lectora de su nivel 97

Tabla 18. Porcentajes de alumnos que pueden realizar actividades de Comprensión Auditiva de su nivel 99

Tabla 19. Porcentajes de alumnos que pueden realizar actividades de Gramática de su nivel..... 101

Tabla 20. Porcentajes de alumnos que pueden realizar actividades de las tres habilidades de su nivel, de acuerdo con lo visto en cada habilidad 101

Tabla 21. Correlación de Pearson entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Comprensión Lectora 104

Tabla 22. Correlación de Spearman entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Comprensión Auditiva 105

Tabla 23. Correlación de Pearson entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Gramática 106

Tabla 24. Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Comprensión Lectora..... 107

Tabla 25. Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Comprensión Auditiva..... 108

Tabla 26. Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Gramática..109

Tabla 27. Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la Valoración General del Desempeño del Estudiante 110

Tabla 28. Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Comprensión Lectora.....112

Tabla 29. Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Comprensión Auditiva.....113

Tabla 30. Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Gramática 114

Tabla 31. Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Valoración General de Capacidad para Actividades del Nivel Asignado 115

Tabla 32. Valores de Correlación entre Autoevaluaciones de Habilidades en Inglés y Exámenes de Colocación encontrados en otros Estudios 121

Resumen

El presente es un proyecto en el que se recuperaron evidencias de validez basada en relación con otras variables para el examen institucional de colocación y acreditación *Placement English Test* (PLET), de la Universidad Autónoma de Aguascalientes (UAA). El objetivo del estudio es aportar evidencias de validez a los argumentos, usos, e interpretaciones de los puntajes que los sustentantes del PLET obtienen. La obtención de evidencias que contribuyen a la validez de ello resultó en tres instrumentos: 1) la Encuesta para la Autopercepción de Habilidades en Inglés, 2) la Encuesta para el Juicio Docente, y 3) la Rúbrica de Evaluación de Desempeño. La información recuperada en estos instrumentos fue llevada a un análisis de correlación de variables, donde se revisó la asociación entre lo que dice el puntaje del PLET y lo que dicen los datos en cada instrumento. Se pudo apreciar que los valores de correlación son positivos y oscilan entre asociaciones moderadas y muy altas. La información obtenida por medio del análisis de resultados en el examen intermedio de los cursos en que son colocados los estudiantes es la que apoya de manera más notoria al argumento de que el PLET puede ser utilizado para colocar a sus sustentantes de forma adecuada con respecto de sus habilidades, mientras que la autopercepción que el estudiante tiene de sus habilidades y el juicio que el docente desarrolla sobre el desempeño de sus estudiantes, ambos son de asociación moderada con el mismo argumento del PLET.

Palabras clave: INSTRUMENTOS DE EVALUACIÓN, IDIOMA EXTRANJERO, VALIDEZ DE LAS PRUEBAS.

Abstract

In the present project, validity evidences based on relations to other variables were gathered to explore the validity of the institutional exam Placement English Test (PLET) for placement and accreditation of the English language in the Universidad Autonoma de Aguascalientes (UAA). The objective of this study is to contribute with validity evidences to the arguments, uses, and interpretations from the scores obtained by PLET test-takers. The obtainment of evidences that support the validity of the previously mentioned elements from PLET scores resulted in three instruments: 1) the Survey for the Self-perception of the English Language Skills (Encuesta para la Autopercepción de Habilidades en Inglés), 2) the Survey for Teacher's Judgment (Encuesta para el Juicio Docente), and 3) the Rubric of Performance Evaluation (Rúbrica de Evaluación de Desempeño). The information gathered with these instruments was submitted to a variables correlation analysis, where the association between what PLET scores and what the data in each instrument inform were explored. It was possible to see that the correlation coefficients were all positive and go from moderate to very high relations. The information obtained through the data analysis in the midterm exam with the Rubric of Performance Evaluation is what supports in a more relevant way the statement that the PLET can be used to place its test-takers in an appropriate way regarding their proficiency, while the students' self-perception on their language skills and the teacher judgment built from his students' performance, both are of moderate association to the same statement about PLET.

Keywords: EVALUATION INSTRUMENTS, FOREIGN LANGUAGE, TEST VALIDITY.

INTRODUCCIÓN

El documento que aquí se presenta contiene el resultado de una investigación de dos años para obtener el grado de Maestría en Investigación Educativa.

En esta investigación se planteó como objetivo explorar la correlación entre diferentes variables y la colocación de estudiantes a partir de sus resultados en el examen institucional de la Universidad Autónoma de Aguascalientes (UAA), el *Placement English Test* (PLET). Dicho examen evalúa el desempeño de sus sustentantes en tres habilidades del idioma inglés, a saber, comprensión auditiva, comprensión lectora, y gramática. Lo anterior se menciona debido a que la correlación de las variables en el estudio se da por medio de puntajes obtenidos a partir de la información proporcionada por los participantes en relación con estas tres habilidades.

La investigación parte de la necesidad por conocer la validez del PLET en relación con otras variables. Se recopiló información sobre distintas investigaciones que han explorado evidencias de este tipo de validez, misma actividad que ayudó a tener una mejor comprensión sobre los procesos metodológicos para la obtención de información, que a su vez pudiera ser puesta a un análisis de correlación con los puntajes de un examen estandarizado. Esta búsqueda de información se sintetiza en el segundo capítulo de la tesis, sobre referentes teóricos, donde, a grandes rasgos, se abordan el tema de validez, la validez basada en la relación con otras variables, la validez basada en argumentos, y la información sobre las habilidades del idioma inglés que se estudiaron en este proyecto, desde la perspectiva del Marco Común Europeo de Referencia para las Lenguas (MCER).

Con un diseño correlacional, esta es una investigación de tipo metodológico. Los instrumentos utilizados para la recolección de información de este proyecto consistieron en 1)

TESIS TESIS TESIS TESIS TESIS

una encuesta sobre la percepción que los estudiantes tenían sobre sus habilidades en el idioma inglés, quienes, en el momento de la aplicación de dicho instrumento, estudiaban en un curso de inglés al que fueron ubicados de acuerdo con sus resultados en el PLET; 2) una encuesta para los docentes de estos cursos, en la que expresaron su juicio sobre lo adecuado del desempeño de algunos de sus estudiantes con relación a lo esperado para el nivel al que fueron colocados; y 3) una rúbrica con el fin de valorar el desempeño demostrado por los estudiantes en un examen intermedio de su curso de inglés, específicamente en la sección de gramática, comprensión lectora, y comprensión auditiva.

Los resultados apuntaron a una correlación moderada entre algunas variables y alta entre otras. Bajo la perspectiva de la validez basada en argumentos, estos resultados apoyan argumentos tales como que el PLET es un examen de inglés cuyas interpretaciones son válidas de acuerdo con el grado de correlación que presentaron con otras variables que exploran las mismas habilidades que el instrumento evalúa, así como que éste puede ser utilizado para la ubicación de estudiantes dentro de alguno de los primeros cuatro niveles del Marco Común Europeo de Referencia para las Lenguas (MCER), ya sea como usuario básico o usuario independiente del lenguaje.

El proceso metodológico que se siguió fue adecuado para los fines del proyecto de investigación, aunque hubo aspectos relacionados con el tamaño de la muestra que pudieron ser mejorados para obtener información de más casos y a su vez hacer más robusta la correlación de datos.

Capítulo 1. Planteamiento del Problema

1.1 Antecedentes

La importancia del inglés en ésta, la era de la información, es tal que se ha llegado a la consideración de que el no saber inglés es una forma de analfabetización. La Facultad de Química de la UNAM publicó un documento en el 2013 titulado *¿Cuál es la importancia del aprendizaje del idioma inglés?*, en el cual mencionan que un 80% de los datos de almacenamiento electrónico en el mundo se encuentran disponibles solamente en inglés, lo que evidencia el impacto que no saber inglés trae consigo. Estamos hablando de que un poco más de tres cuartas partes de las fuentes de información y conocimiento en el mundo digital se encuentran en inglés; tres cuartas partes cuyo acceso es sólo posible para aquellos que dominan el inglés en cierto nivel.

A partir de esta panorámica, se puede entender que la importancia por aprender inglés se traduce en la necesidad por dominarlo de una forma inteligible debido a las oportunidades de mejoramiento en distintos aspectos personales y sociales que éste proporciona. Sin embargo, una serie de factores pueden dificultar o facilitar la tarea de aprender inglés. Estos factores van desde cuestiones intrínsecas del individuo hasta cuestiones meramente externas. De estos factores parte uno que interesa a esta investigación: la colocación en niveles de inglés a partir de una prueba estandarizada.

Se puede considerar que si un estudiante se encuentra cursando un nivel de inglés al que no pertenece, ya sea porque éste es muy alto o muy bajo para el nivel de dominio del idioma con que el estudiante cuenta, puede perjudicar su desempeño y aprendizaje en el curso. Muchas instituciones en todo el mundo, principalmente de educación superior (Kokhan, 2012; 2013),

utilizan pruebas estandarizadas con la finalidad de colocar estudiantes en cursos de inglés basándose en los resultados que cada estudiante obtuvo en dicha prueba. Una de estas instituciones es la UAA.

La UAA atiende anualmente a una cantidad considerable de estudiantes que desean estudiar el nivel superior en el estado de Aguascalientes. Como parte de los requisitos que la UAA establece para que sus estudiantes puedan obtener el grado de licenciatura, se encuentra el aprender una segunda lengua (UAA, s.f.). Con la finalidad de apoyar al cumplimiento de este requisito, la UAA cuenta con el Programa Institucional de Lenguas Extranjeras (PILE), cuyo objetivo es ofrecer cursos de inglés a los estudiantes de la institución, y de esta forma puedan conseguir los créditos que su carrera requiera para cubrir este requisito de egreso.

En la mayoría de las carreras en la UAA se busca que los alumnos terminen su formación con al menos un nivel A2 de dominio del inglés, mientras que en otras, en su mayoría ingenierías, se les pide cubrir un nivel B1, y para el caso de carreras como Gestión Turística o Comercio Internacional, se pide cubrir el nivel B2. Para ofrecer los cursos mencionados en el párrafo anterior, el PILE necesita conocer el nivel de inglés que tienen los estudiantes de nuevo ingreso, y con base en ello poder darles una colocación acorde a sus conocimientos y dominio del idioma (PILE, 2015). Para ello, la institución cuenta con dos opciones; una de ellas es que los estudiantes que hayan presentado un examen *Test of English as a Foreign Language* (TOEFL), un *Test of English for International Communication* (TOEIC), o un *Cambridge English: First* (FCE), en un periodo no mayor a dos años, presenten su papeleta con el puntaje obtenido en cualquiera de los dos exámenes y dependiendo de su puntuación se les pueda hacer la colocación apropiada, o si es que exceden el puntaje requerido por su carrera, se les pueda hacer la

acreditación del requisito. La otra opción, que es a la que recurren la mayoría de los estudiantes, es presentar el examen de colocación y/o acreditación ofrecido por la institución.

Para colocar a los estudiantes en un nivel de Inglés, se habían utilizado una serie de exámenes estandarizados, principalmente el TOEIC. Durante varios años, los docentes encargados de impartir estos cursos de Inglés ofrecidos por el PILE, reportaban que no todos los estudiantes eran colocados en un nivel adecuado, lo que llevó a los encargados del programa a elaborar un examen interno que colocara con mayor precisión a los alumnos en un nivel determinado; a este examen se le llamó *Placement English Test* (PLET).

1.1.1 *Placement English Test*. Como se plantea en párrafos anteriores de este capítulo, el PILE del Departamento de Apoyo a la Formación Integral (DAFI) en la UAA, tomó la decisión de crear un nuevo instrumento de evaluación para la colocación de los estudiantes de nuevo ingreso, ya que, basados en los comentarios hechos por algunos de sus docentes de Inglés, el PILE reporta que la prueba que se usaba con anterioridad para dicho fin (TOEIC) no ubicaba a una gran cantidad de estudiantes en el nivel al que realmente pertenecían debido a sus competencias y conocimientos del idioma inglés. Es de esta forma que el PLET fue diseñado.

El PLET es una prueba diseñada en el año 2015 en la UAA por parte del PILE, con la intención de contar con un instrumento interno de evaluación del idioma inglés. El fin de este instrumento es:

certificar el dominio del idioma inglés a los estudiantes de educación superior, de acuerdo al Marco Común Europeo, (...) [y] diagnosticar el nivel de competencia comunicativa del idioma inglés para ubicar a los sustentantes en uno de los 8 programas ofrecidos por PILE. (García-Medina & Pedroza-Zúñiga, 2014, p. 3).

TESIS TESIS TESIS TESIS TESIS

A partir de los propósitos que tiene esta prueba, este instrumento se denomina como una prueba de colocación y/o acreditación. Profundizando más en la descripción de este instrumento, García-Medina y Pedroza-Zúñiga (2015) declaran que el PLET es una prueba de alto impacto, de lápiz y papel, con cuadernillo y hoja de respuesta de lector óptico por separado; mide competencias de la lengua inglesa; [es] estandarizado, con reactivos de opción múltiple de respuesta cerrada. Las tareas evaluativas incluyen reactivos a partir de textos escritos y material auditivo. (p. 3).

Los resultados del examen cuentan con ciertas evidencias de validez a las que se sumó la de criterio o relación con otras variables. El examen cuenta ya con evidencias de validez de contenido y constructo. La evidencia de validez basada en relación con otras variables se estudió por medio de esta investigación.

De acuerdo con la *American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME)* (2014), la validez es la consideración fundamental para el diseño y evaluación de un examen, ya que indica si un instrumento de evaluación está evaluando aquello que pretende, o, por el contrario, está examinando constructos de algo para lo que el instrumento no fue diseñado.

La cantidad de investigaciones hechas sobre validación en instrumentos de evaluación es sin duda inmensa, pero por cuestiones de interés y relación con el tema de investigación de este estudio se señalarán solamente investigaciones que se han realizado sobre evidencias de validez basadas en relación con otras variables en cuanto a exámenes de colocación.

Papageorgiou y Cho (2014) mencionan que la mayoría de los estudios realizados con respecto a exámenes de colocación se ha centrado en educación superior. Esto se debe a que es

TESIS TESIS TESIS TESIS TESIS

en la universidad y en los colegios universitarios donde se da el mayor uso de este tipo de instrumentos para conocer las competencias y conocimientos que los estudiantes de nuevo ingreso tienen en distintas áreas. Estudios recientes (Kokhan, 2012; 2013) señalan el incremento que universidades en los Estados Unidos han visto con respecto al ingreso de estudiantes de otros países, lo que obliga a estas instituciones a usar pruebas estandarizadas para medir el conocimiento que los estudiantes tienen acerca del uso del idioma inglés, esto como una forma de predecir su éxito académico, ya que, al ser universidades norteamericanas, las clases son impartidas en inglés.

Por citar un ejemplo de estudios realizados para conocer los diferentes tipos de evidencia de validez que un examen de colocación de inglés puede tener, podemos mencionar el de Fulcher (1997), en el que se investiga sobre la validez de contenido, la validez concurrente (una de las formas de evidencia basada con relación a otras variables) y la validez aparente del examen de colocación que el *English Language Institute* utilizaba para dar colocación en cursos de inglés a estudiantes internacionales que deseaban cursar estudios superiores en la Universidad de Surrey. En el mismo, Fulcher da una serie de evidencias sobre el diseño de la prueba y cómo en el proceso se iban cumpliendo ciertos criterios de validez desde una perspectiva estadística y de contenido. El autor de este estudio concluye con que es necesario seguir estudiando la validez sobre este tipo de pruebas o exámenes, ya que colocar estudiantes que usan el inglés como una segunda lengua en lugares que no corresponden a su nivel de dominio del idioma, puede afectar su desempeño en sus estudios universitarios.

Con respecto a la validez predictiva de los exámenes de colocación de idiomas, otra de las formas de evidencia de validez basada en relación con otras variables, se han realizado algunos estudios (Lee & Greene, 2007; y Papageorgiou & Cho, 2014). En estos estudios se da un

TESIS TESIS TESIS TESIS TESIS

proceso similar para cumplir con la evidencia de validez predictiva de los instrumentos de colocación en que cada estudio se enfocó. Dicho proceso se basa en recurrir a la opinión de docentes sobre el desempeño que los estudiantes tenían en el curso o nivel de inglés en el que habían sido colocados, así como la comparación con evaluaciones que los estudiantes tomaban durante el curso. En el caso del estudio realizado por Lee y Greene (2007), se optó por la utilización de las puntuaciones del examen TOEFL. Para el estudio hecho por Pageorgiou y Cho (2007), no sólo se toma en cuenta el juicio realizado por los maestros de acuerdo a si el estudiante estaba en el nivel adecuado, además, los maestros comentaban qué nivel deberían tomar sus estudiantes, mismo que se comparaba con el nivel asignado por medio de la prueba TOEFL® Junior™ Standard, aplicada a nivel de educación secundaria.

En el estudio que ocupó al presente proyecto de investigación se atendió la validez en relación con otras variables, la cual se divide en concurrente y predictiva. De acuerdo con la literatura consultada, la validez concurrente es utilizada en exámenes que, como el PLET, son relativamente nuevos. Con respecto de la validez en relación con otras variables, se han generado distintos estudios (Armstrong, 1994; Fulcher, 1997; Le, 2011; Wall, Clapham & Alderson, 1994) en los que para contribuir con evidencia de validez a las interpretaciones de resultados obtenidos en exámenes de colocación de inglés, se ha recurrido a distintos métodos. Se pueden comparar los resultados del examen de colocación nuevo con los de otros exámenes estandarizados. Por ejemplo, comúnmente se compara con exámenes como TOEIC o TOEFL, pues son exámenes cuyos resultados cuentan con diferentes tipos de evidencia de validez, incluida la validez concurrente y predictiva. Otra forma de hacerlo es comparando los resultados obtenidos en el examen de colocación de inglés y otros exámenes hechos específicamente para cada medida, como lo menciona Moritoshi (2001) en un estudio sobre el TOEIC, esto con la intención de

TESIS TESIS TESIS TESIS TESIS

analizar la consistencia de los resultados en el examen de colocación. Una forma más de hacerlo, como Fulcher (1997) menciona, es al comparar resultados con los obtenidos en autoevaluaciones de desempeño por parte de los estudiantes, así como actividades o exámenes diseñados por el docente con el fin de evaluar las habilidades y competencias comunicativas con que el estudiante cuenta, y ver el grado de correlación que guardan correlación con la colocación designada por el examen.

Sin embargo, a pesar de que la investigación sobre validez predictiva y concurrente vaya incrementando, la mayoría de los estudios han sido realizados en Estados Unidos y algunos países de Europa y Asia, mientras que en México, de acuerdo con la revisión de literatura, no existe, o no se ha publicado, evidencia sobre este tipo de estudios con respecto a exámenes de colocación de inglés.

1.2 Problematización

La enseñanza del idioma inglés es un proceso que requiere de una correlación entre el contenido de las clases y el desempeño en habilidades del idioma que un estudiante tiene, así como con la metodología que el docente emplea para enseñar el idioma. Si un estudiante se encuentra tomando clase en un nivel al que no pertenece, esto puede afectar su desempeño en el aprendizaje del idioma inglés.

Suponiendo que un estudiante sea colocado en un nivel superior al que en verdad domina (siendo este el problema que, según personal administrativo del PILE, varios docentes reportaban), esto presentará para el estudiante y para el maestro una dificultad importante, pues implica que el maestro, en el más optimista de los casos, explique de nueva cuenta contenidos que se supone todo el grupo domina, para intentar regularizar a aquellos estudiantes que hayan

TESIS TESIS TESIS TESIS TESIS

sido colocados erróneamente, lo cual representa un problema para los estudiantes que sí fueron colocados en el nivel adecuado, pues no se hace uso del tiempo necesario para que ellos puedan seguir desarrollando y mejorando sus habilidades. También puede suceder que el maestro decida enfocarse en aquellos estudiantes que fueron colocados apropiadamente, dando como resultado que aquellos que fueron asignados a un nivel superior al que en realidad poseen, se vean en desventaja y exista una alta probabilidad de que tengan que repetir el curso o desertar y tratar de buscar otras alternativas para regularizarse y cumplir con el requisito de aprender una segunda lengua. En el caso de que el estudiante sea colocado en un nivel inferior al que en realidad tiene, puede tener repercusiones relacionadas con pasar un semestre trabajando con habilidades que ya tiene y no empleándolo para desarrollar nuevas competencias pertenecientes a niveles más altos de inglés.

Otra de las posibles situaciones que se pueden presentar si el instrumento de colocación no lograra su propósito con propiedad, es que los estudiantes alcancen la acreditación cuando en realidad no cuentan con las habilidades y el nivel de competencia lingüística como para haber obtenido ese resultado. Lo que una situación tal implica es que los estudiantes contarán con una certificación del dominio del idioma que no comprueba su desempeño verdadero, y se enfrentarán a las demandas del mundo laboral sin contar con las habilidades necesarias para hacer uso del idioma inglés como su formación se los exige.

Es por estas razones que la existencia de evidencias de validez en un instrumento que supone asignar una colocación en niveles de inglés, representa un elemento sumamente importante para la formación de los estudiantes en la UAA. Los estudiantes deben de contar con una educación de calidad que les prepare para las exigencias del mundo laboral globalizado,

siendo el aprendizaje del idioma inglés una de las herramientas más útiles que les apoyará en su desempeño profesional.

1.3 Objetivos y Preguntas de Investigación

El objetivo principal de esta investigación es aportar evidencia de validez basada en relación con otras variables al examen institucional de colocación y acreditación PLET de la UAA. Para cumplir con este objetivo, las siguientes preguntas de investigación fueron guías fundamentales para orientar este estudio a la meta que se fijó:

1. ¿Cuál es el grado de asociación que existe entre el resultado del examen PLET y la percepción de los estudiantes sobre su nivel de dominio en las habilidades de comprensión lectora y auditiva, además de la sub-habilidad de gramática, del idioma inglés?
2. ¿Cuál es el grado de asociación que existe entre el resultado del examen PLET y el juicio de los maestros sobre la colocación de estudiantes en un determinado nivel de dominio del idioma inglés, basado en el desempeño visto de sus estudiantes en el curso?
3. ¿Cuál es la asociación que existe entre el resultado del examen PLET y el desempeño demostrado por parte de los estudiantes en las secciones de comprensión lectora, comprensión auditiva y gramática en el examen intermedio del curso?

1.4 Justificación

Como ya se ha precisado en la problematización, es necesario contribuir con evidencias de validez que comprueben que un instrumento de evaluación cumple con el propósito para el que fue hecho. En el caso de un examen como el PLET, que se propone colocar y acreditar a estudiantes de la UAA dependiendo de los resultados que estos hayan obtenido en la prueba, la justificación para llevar a cabo este estudio se basa en la importancia de ampliar las evidencias de validez que den cuenta de que el uso que se le da a los resultados obtenidos del PLET están siendo apropiados, acertados, válidos, y confiables. En la problematización ya se hacía mención de las consecuencias que trae el no colocar a los estudiantes en un nivel acorde a sus habilidades en el idioma inglés, así como acreditar a estudiantes que no cuentan con las habilidades que una acreditación en el idioma inglés exige dependiendo de la carrera que hayan decidido cursar. De acuerdo con Bachman y Douglas (1990, 2004; 2003 en Le, 2011), los exámenes de colocación son considerados de mediano impacto teniendo en cuenta el tipo de decisión que se hace sobre el desempeño realizado por los sustentantes. De esta forma, la validez es fundamental para este tipo de pruebas, ya que, considerando el impacto que tienen en la vida de los sustentantes, si el estudiante no es colocado en un nivel apropiado, puede perder la oportunidad de desarrollar sus habilidades comunicativas en el idioma de una manera satisfactoria, lo que consecuentemente le llevará a perderse de oportunidades de desarrollo personal y profesional, como ya se mencionaba en los primeros párrafos de esta introducción. Le (2011) menciona otros problemas que surgen al colocar estudiantes en niveles a los que el nivel de desarrollo en sus competencias comunicativas en el idioma inglés no pertenecen. Dichos problemas se ven representados en la cantidad de dinero, tiempo y esfuerzo que el estudiante tendrá que invertir en aprender el lenguaje, llegando incluso a tener repercusiones emocionales como la frustración o desmotivación de estar en un

TESIS TESIS TESIS TESIS TESIS

nivel al que no se corresponde. Además, los maestros que imparten los cursos de inglés también se ven afectados en cuanto a la instrucción de los contenidos en un programa de cierto nivel de inglés, pues dificulta el satisfacer las necesidades de los alumnos que sí cuentan con las habilidades lingüísticas para ese nivel si hay compañeros con un nivel más bajo o incluso más alto.

La falta de evidencias de validez y confiabilidad en un instrumento como el PLET puede generar problemas que varían en nivel de impacto entre los miles de estudiantes que ingresan a la UAA y toman el examen institucional de colocación y/o acreditación. Colocar a estudiantes en un nivel inapropiado, así como dar acreditación a estudiantes que no cuentan con las habilidades comunicativas suficientes como para desenvolverse en el mundo laboral globalizado, son problemas que pretenden evitarse con la implementación de este instrumento relativamente nuevo, además de, como ya se había referenciado en el párrafo anterior, facilitar la instrucción de contenidos tanto para maestros como para estudiantes. Para evitar lo anteriormente mencionado, es importante verificar en qué medida el instrumento logra este propósito. Por ello, este estudio se justificó bajo la importancia que tiene conocer la relación entre la colocación que ofrece el PLET y la opinión que tienen los docentes con respecto al nivel en que sus alumnos fueron asignados, juicio que se basó en el juicio del profesor sobre el desempeño de los estudiantes en clase, así como la percepción que los estudiantes tienen sobre su competencia en la lengua y el desempeño mostrado en el examen intermedio de sus cursos.

La UAA se beneficia de este estudio en términos de mejorar el instrumento PLET, pues se pretende que al ser un examen institucional pueda ser usado con la seguridad de que está al nivel de otros exámenes estandarizados, y aún más importante, de que evalúe y coloque con exactitud a sus estudiantes. Este trabajo de investigación además sirve para incrementar los

estudios de validez realizados con exámenes que cuentan con la naturaleza del PLET (de colocación de inglés), los cuales, con lo revisado en la literatura, son muy reducidos, y tal vez inexistentes en el contexto mexicano. También aporta a cuestiones metodológicas, pues si bien hay evidencia de estudios realizados sobre evidencias de validez en relación con otras variables, en la revisión de literatura no se encontró algún estudio que maneje tanto los juicios de maestros como la autopercepción por parte de los estudiantes en cuanto a su competencia en el lenguaje, tal vez se ha hecho de manera separada, pero no en conjunto como se tiene pensado realizar este estudio.

Capítulo 2. Referentes Teóricos

2.1 Validez y Evidencias de Validez

La validez es un concepto clave en la evaluación educativa. El concepto de validez se refiere a “el grado en que la evidencia y la teoría apoyan las interpretaciones de los resultados de una prueba vinculados por los usos propuestos de las pruebas.” (AERA, APA, & NCME, 2014, p. 11). Linn (2010) simpatiza con la definición anterior, y agrega que “la validez es la consideración más importante en la evaluación de las interpretaciones de los resultados de evaluaciones educacionales, y los usos que son hechos de esos resultados.” (p. 181). Por otro lado, Mendoza-Ramos (2015) menciona que desde un enfoque lógico-argumentativo, la validez “consiste en todo un proceso de inferencias que deben encontrarse sustentadas en la evidencia y en el respaldo de las mismas.” (p. 1). A partir de estas definiciones, la validez puede ser referida como aquello que da cuenta de que un instrumento evalúe lo que pretende evaluar; esto implica que no se evalúe ni menos ni más allá de lo que ha sido definido en la realidad que pretende medirse.

2.1.1 Desarrollo del concepto de validez. El concepto de validez ha sufrido ciertos cambios en cuanto a la manera en que es entendido y aplicado a cuestiones de educación.

De acuerdo con Newton y Shaw (2014), la evolución del concepto de validez puede dividirse en cuatro momentos principales, que son: 1) la génesis de validez (mediados de 1800 a 1951); 2) la fragmentación de la validez (1952 a 1974); 3) la (re) unificación de la validez (1975 a 1999); y 4) la deconstrucción de validez (2000 a 2012).

Un periodo de gestación (antes de 1921). Todo se remonta a la segunda mitad del siglo XIX, cuando los exámenes de ingreso a universidades cobraron mayor presencia. La necesidad por contar con exámenes imparciales, se optó por la introducción de exámenes escritos, y a esto se sumó el uso de ciertas técnicas estadísticas que en un principio Galton, con su trabajo *Hereditary Genius* (1869), proponía en el campo de estudio de la herencia, en el que sugería estudiar diferencias individuales como una disciplina por sí misma. Esto se convertiría en una aproximación empírica a la validación. A partir de esto, varios pioneros en el terreno de la psicometría desarrollaron esta aproximación estadística de Galton; entre ellos destaca Charles Spearman, pues fue quien más influyó en cuanto al método de correlación de información psicológica, que a su vez podía ser trasladado al terreno educativo.

Para inicios del siglo XX, la tendencia por el desarrollo e implementación de pruebas de logro y selección de individuos creció en diferentes áreas y países desarrollados, desde lo educativo hasta cuestiones de trabajo. Para el terreno en educación, Edward Thorndike ayudó a establecer una sub-disciplina de la psicología que sería denominada psicología educacional.

El descontento que varios investigadores tenían con respecto a la tradición de los exámenes de logro escolar, denominándoles como poco confiables, pues “los resultados [de los

exámenes de logro] no podían ser interpretados de forma justa por las mentes humanas.” (p. 29), fue un momento importante para que el concepto de validez naciera.

De esta forma, se buscaron opciones de pruebas más objetivas, que denotaran una mayor imparcialidad en cuanto a las interpretaciones de sus resultados. “El periodo entre 1910 a 1920 fue una década de incremento de atención sobre la validez de las mediciones mentales, aunque no era necesariamente descrita como tal.” (p. 29).

Para este momento, ya existía una división clara de sub-dominios en cuanto a la medición educacional y psicológica, los cuales eran ‘las comunidades profesionales’, cuyo uso de evaluaciones tenían propósitos tales como diagnóstico clínico, evaluación de logro académico, y selección de individuos, y ‘las comunidades científicas’, donde las pruebas eran utilizadas y diseñadas con el fin de explorar características varias de la personalidad.

Aunado a esta división de dominios, el reconocimiento del coeficiente de correlación empezó a establecerse ampliamente, pues se consideraba una herramienta para determinar la calidad de las evaluaciones.

Un periodo de cristalización (después de 1921). En una búsqueda por establecer estándares en cuanto a los procedimientos y conceptos de medición educativa, en 1921 la Asociación de Directores de Investigación Educativa, inician la promoción del establecimiento de una consistencia en la aplicación de estos procesos y conceptos pertenecientes a la medición en educación. Compartiendo este interés por mejorar la estandarización de las cuestiones anteriormente mencionadas, la Asociación Estadounidense de Psicología (APA, por sus siglas en inglés), también había realizado ya un par de intentos por conseguir este objetivo, en 1895 y 1906.

TESIS TESIS TESIS TESIS TESIS

Después de un proceso de discusión entre un comité de expertos, la Asociación Nacional de Directores de Investigación educativa, llegaron a lo que podría ser considerado la ‘definición clásica’ de validez: “el grado en que un examen mide lo que se propone medir.” (Ruch, 1924 en Newton y Shaw, 2014).

De cualquier forma, a partir de la definición clásica, la validez empezó a entenderse en términos de correlación. La validez empezó a ser definida en términos de un acuerdo observado entre puntuaciones de exámenes y puntuaciones de una medida de criterio.

La fragmentación de la validez: 1952 – 1974. En este periodo se divide el término de validez tanto en validez de criterio como validez de contenido. La validez de criterio se enfocaba en determinar el grado en que la puntuación de una prueba podía predecir un comportamiento fuera del examen, mientras que la validez de contenido era entendida como el grado en que había un empate entre el contenido del examen y el contenido de otra variable.

En 1955, un artículo de Cronbach y Meehl, ayudó a la introducción de lo que sería la noción de validez de constructo, la cual parecía ser comprendida como una evidencia de validez que debería ser aplicada siempre que la validez de contenido y la validez de criterio no pudieran ser empleadas en determinadas situaciones. Para este momento ya eran cuatro los ‘tipos’ de validez (de contenido, predictiva, concurrente, y de constructo) que caracterizaban este periodo en que se entendían diferentes tipos de validez, alejándose de la definición núcleo que sugeriría en todo caso que estas no son más que evidencias de validez, no tipos de validez separadas.

Es de esta forma que en este periodo se da una visión trinitaria de validez. Sin embargo, debido a que, como lo menciona Borsboom (2015), se creía que todas las puntuaciones de exámenes tenían la intención de medir constructos teóricos, la validez de constructo “tomó su

lugar como la nueva pieza central de la teoría de validez.” (p. 2), resultando en una inestabilidad conceptual de la perspectiva trinitaria del concepto de validez.

La (re)unificación de la validez: (1975 a 1999). Como su nombre lo sugiere, este periodo se vio caracterizado por la intención de reunir los diferentes tipos de validez bajo un concepto que las caracterizara como evidencias de un mismo objeto: la validez misma.

Durante este periodo, como Panahi (2014) y Newton y Shaw (2014) lo mencionan, Messick sugería la unión de los tipos de validez bajo el nombre de validez de constructo, pues se creía que este tipo de evidencia era una base para todos los elementos relacionados con el proceso de validación. Es de esta forma que para Messick la validez tenía que ver con el grado en que la evidencia empírica y teórica apoya lo adecuadas que son las inferencias que se hacen de resultados de exámenes.

La siguiente tabla apoya lo que se dice anteriormente.

Tabla 1

Facetas de validez (Messick, 1989^a;1989^b, en Panahi, 2014, p. 333).

Uso del Test	Interpretación del Test	
Base Evidencial	Validez de Constructo	Validez de Constructo + Relevancia/Utilidad
Base Consecuente	Implicaciones Importantes	Consecuencias Sociales

Al apreciar este modelo, es posible deducir que es la validez de constructo la que, en el pensamiento de Messick, une el concepto de validez.

La deconstrucción de la validez: 2000 – 2012. Parte de la justificación para proceder a una deconstrucción en torno al concepto de validez, como lo expresan Newton y Shaw (2014), se debe a que “no sólo [se buscaba] descomponer el concepto en componentes más manipulables, sino también para remover potencialmente (o al menos para reposicionar) el concepto de constructo.” (p. 136).

Así, se llega a una aclaración que hoy en día es utilizada en cuestiones de validez, y es la que Linn (2010) hace sobre el uso del término de validez, y esta es que un instrumento no puede ser válido o inválido, lo son los usos y las interpretaciones que se hacen de aquellos resultados arrojados por la medición que se ha realizado. Así que “los juicios deben ser hechos para determinar si el grado de validez es adecuado para un fin particular.” (Linn, 2010, p. 181).

Es evidente entonces que la validez, como término, ha sufrido una serie de cambios en cuanto a su conceptualización a lo largo de la historia, pues se ha pretendido darle un sentido unificado al concepto, y no segmentado como se hizo al principio. En la década de 1950, señala Linn (2010), había dos nociones de validez que prevalecían, las cuales eran la validez predictiva y la validez de contenido. Esto no quiere decir que la validez tenga distintos tipos dentro de sí; existen evidencias de validez, pero no son categorías que dividen a la validez en distintos tipos.

Sin embargo, todavía para 1954, se seguía refiriendo a “tipos de validez”. Linn (2010) señala que la APA, amplió los tipos de evidencia de validez a cuatro, los cuales fueron validez de contenido, validez concurrente, validez predictiva, y validez de constructo. El haber ampliado los tipos de evidencia de validez se debe a que son varios los criterios que tienen que cumplir con

una forma de validez para que los resultados que se obtengan del instrumento que se desee emplear resulten válidos en más de un sentido. Messick (1989, en Linn, 2010) da una definición de validez parecidas a las anteriores, sólo que agrega palabras como “adecuación” y “propiedad” para referirse a que la validez apoya con evidencia empírica a estos conceptos referidos a las inferencias y acciones que se hacen a partir de las puntuaciones de pruebas o de otras formas de evaluación. Así, podemos ver que ya desde 1989 se iba formando la concepción que se tiene hoy en día sobre lo que es la validez, misma que puede encontrarse en la más reciente de las ediciones del *Standards for psychological and educational testing*, usada al principio de este apartado.

2.1.2 La validez basada en argumentos. Como ya se ha visto, la validez de los usos e interpretaciones que se hacen sobre los resultados de una prueba puede ser abordada desde distintas evidencias. Un enfoque que, según Mendoza-Ramos (2015), ha cobrado presencia en estudios de validez de exámenes de alto impacto, en especial en la evaluación de segundas lenguas, es el de validez desde la lógica argumentativa. Mendoza-Ramos analiza tres diferentes modelos de argumentación; el de Toulmin, Kane, y Mislevy. Respecto al modelo de Toulmin (1958, en Mendoza-Ramos 2015), nos comenta que el proceso de argumentación parte de dos elementos: “una aserción (claim) y los hechos o datos (data) que fungen como base o fundamento para validar dicha aserción.” (p. 173). Respecto a la importancia de los datos (data), se refiere a que el uso primordial que de ellos debemos hacer es generar proposiciones, entendidas como reglas, principios e inferencias, a las que se les llama garantías (warrants).

En el 2005, Bachman (en Mendoza-Ramos, 2015), reformula un par de preguntas formuladas por Spoky (1981) para atender a consideraciones en la evaluación de lenguaje. Las preguntas que Bachman propuso para atender la validez en exámenes de lenguajes fueron “¿Qué

tan convincente es el argumento para usar la evaluación de esta forma?; y la segunda “¿Qué tan creíble es la evidencia que apoya ese argumento?” (Mendoza-Ramos, 2015, p. 175), atendiendo así a la importancia que tiene el fundamentar las aserciones que hacemos sobre los usos y resultados de un examen, tomando en cuenta los elementos propuestos por Toulmin. Sin embargo, el proceso de validación de exámenes basado en la lógica argumentativa, no fue considerado principalmente por ninguno de estos autores, sino hasta Kane, en 1992.

Kane (1992, p.2, en Mendoza-Ramos, 2015) concibe al enfoque basado en argumentación como “un marco sustentado en la interpretación para recolectar y presentar evidencia de varias suposiciones e inferencias involucradas en la interpretación.” (p. 176). De esta forma, es posible concluir que toda interpretación hecha a partir del resultado de un examen debe de estar sustentada por un argumento.

Kane (2013) menciona cuatro tipos de inferencia, los cuales son a) la inferencia de los resultados observados, que se basa en analizar que la evaluación fue apropiada y aplicada de manera correcta, de manera que el resultado que presente el sustentante sustente esta afirmación; b) la inferencia de generalización, que refiere a la confiabilidad de los resultados del examen, reconociendo que éste no puede evaluar todo el desempeño del sustentante en un área de conocimiento, sino más bien una muestra pequeña del ,mismo; c) la inferencia de extrapolación, que refiere a la afirmación que se hace sobre que los resultados del examen reflejan lo que el sustentante puede hacer en situaciones reales del día a día, teniendo la evidencia empírica y analítica como sustentos de esta afirmación; y d) la inferencia de toma de decisiones, siendo la única inferencia de carácter político, pues se ve relacionado a los estatutos que las instituciones que aplican el examen establecen para determinar las habilidades que un sustentante debe de tener para desempeñar cierta función.

El último modelo que Mendoza-Ramos (2015) analiza es el de Mislevy, Stenberg y Almond (2003), el cual expresa que los exámenes deben de ser diseñados centrándolos en evidencia, para la cual hay cuatro componentes necesarios. El primero de ellos se refiere al análisis del dominio en el que se desea llevar a cabo la evaluación; el segundo es la descripción del dominio, expresando qué, cómo y con qué se pretende evaluar; el tercero refiere al marco conceptual y los procesos que le dan confiabilidad a una prueba; y el cuarto que es la consideración del resultado del examen y todos los agentes involucrados en el mismo.

La importancia del enfoque basado en argumentos recae en la demanda de evidencias que sustenten tanto el diseño como el uso que se le dé a los resultados de un examen. Para esta investigación es necesario mencionar este enfoque, pues se pretende dar sustento a una serie de afirmaciones en torno a la colocación que se designa a los sustentantes conforme a su puntuación en el examen PLET. Dichas afirmaciones construyen una cadena de justificaciones, que comienzan desde la colocación misma del examen, justificada con el juicio que los maestros construyen a partir de su percepción en el desempeño de los estudiantes, el cual se justifica con las producciones realizadas por los estudiantes, que a su vez justifican la opinión que los estudiantes tiene sobre su nivel de competencia en el idioma.

2.1.3 Fuentes de evidencia de validez. Al hablar de validez, es comprensible que no todos los fenómenos y los instrumentos que los miden cuenten con un proceso de validación de interpretaciones genérico. Por ello, al hablar de validez, es importante hablar sobre los tipos de evidencia que existen. De acuerdo con la *AERA et al.*(2014) las evidencias de validez son cinco, y son las siguientes:

- TESIS TESIS TESIS TESIS TESIS
1. Evidencia basada en el contenido de la prueba: En esta evidencia se debe definir el constructo y crear especificaciones sobre el o los contenidos a ser evaluados; de aquí, debe de existir una relación entre esas especificaciones y el contenido de la prueba.
 2. Evidencia basada en procesos de respuesta: Esta evidencia refiere a los procesos mentales usados por los sustentantes de una prueba.
 3. Evidencia basada en estructura interna: Esta evidencia, como su nombre lo refiere, se centra en un análisis de la estructura interna que dé fe de la relación de dicha estructura con las interpretaciones que se hagan de los resultados de la prueba.
 4. Evidencia basada en relación con otras variables: La finalidad de este tipo de evidencia es aportar información de relaciones del instrumento con otras medidas del constructo que se está evaluando.
 5. Evidencia basada en las consecuencias de evaluar: Este tipo de evidencia hace referencia al impacto de la prueba.

Linn (2010) También menciona estas categorías de evidencia de validez, y concluye que todas y cada una de ellas son necesarias para poder decir que existe validez en las interpretaciones que se están haciendo de los resultados de una prueba, a pesar de que, según él, existan objeciones en contra de considerar al último tipo de validez aquí mencionado como “relevante” para la validez.

2.1.4 Evidencia de validez basada en relación con otras variables. La validez basada en relación con otras variables, en ocasiones referida como evidencia de criterio, se refiere a la validez basada en la relación entre los resultados de una evaluación y otras variables. En la

revisión de literatura puede apreciarse que este tipo de evidencia de validez puede ser identificada como predictiva y concurrente. Kane (2001) menciona que la validez siempre ha requerido de alguna medida de criterios que aporten un valor real a la variable de interés, o que al menos se acerque de mejor forma al valor real. Se debe de entender que estos criterios con los cuales se pretende establecer una relación a los resultados de un examen, se basan en el desempeño del sustentante en la realidad evaluada, ya sea antes, durante, o después de la aplicación del examen. Estos criterios, al estar basados en la competencia o desempeño del estudiante en cierta área de conocimiento, pueden ser los resultados de otras evaluaciones que estén destinadas a medir esa misma área. Entonces, hablamos de una variedad de elementos, como lo son, en el caso de la escuela, exámenes durante y al final de un curso, trabajos académicos, e incluso los juicios tanto por parte de docentes como estudiantes en cuanto a cómo estos últimos se desempeñan en el área de conocimiento evaluada en el examen cuyos resultados, y los usos de los mismo, pretenden ser validados por este tipo de evidencias.

Rescatando la idea principal de la definición que el consorcio AERA et. al (2014) ofrece sobre este tipo de evidencia de validez de relación con otras variables, se menciona que el constructo debe estar relacionado con otras variables, además de que esta relación deberá de ser analizada, siendo estas variables externas al examen mismo. “Las variables externas pueden incluir medidas de algunos criterios que se espera el examen pueda predecir, así como relaciones con otros exámenes que miden los mismos constructos, y exámenes que midan constructos similares o diferentes.” (AERA, APA & NCME, 2014, P. 16). Numerosos estudios (Bayliss, 2006; Scott-Calyton, 2012; Wang, Niemi, & Wang, 2007; Armstrong, 2001; Karakaya & Tavsancil, 2008; Muyskens, Betts, Lau, & Martson, 2009; Cotton & Conrow, 1998; y Alnahdi, 2015) han abordado evidencias de validez desde una perspectiva predictiva en cuanto a los

TESIS TESIS TESIS TESIS TESIS

resultados de algunos exámenes que son usados para colocación en niveles de inglés. El procedimiento que han seguido es básicamente comparar los resultados de un examen de colocación con el promedio parcial o final que los estudiantes obtienen tanto en cursos de inglés como en sus carreras, de manera que la relación entre colocación y desempeño pueda ser valorada. Otros estudios (Armstrong, 1994; Lee & Greene, 2007; Geranpayeh, 2003; Riazi, 2014; y Kokhan, 2012) han abordado la validez en relación con otras variables comparando los resultados de un examen de colocación con otros exámenes o evaluaciones que miden los mismos constructos, dándole así una perspectiva de carácter concurrente, mientras que otros estudios (Lee & Greene, 2007; y Powers, Kim, & Weng, 2008), siguiendo esta misma perspectiva, han utilizado autoevaluaciones hechas por los estudiantes como criterio para establecer relación con los resultados de exámenes de colocación. Un estudio en particular (Papageorgiou & Cho, 2014), utiliza el juicio emitido por los maestros con respecto a la colocación de estudiantes después de un periodo de cuatro semanas conociendo y evaluando el desempeño de los estudiantes.

De acuerdo con la literatura revisada, la evidencia de validez en relación con otras variables tiene una tendencia a abordarse desde la comparación de los resultados de un examen con variables como los resultados de otros exámenes, los promedios de calificación que obtienen los estudiantes al terminar un curso, y también la percepción que estos mismos tienen sobre su desempeño, ya sea antes, durante o después de que el examen ha sido aplicado.

2.1.4.1 Juicio docente. Como se ha especificado en el planteamiento del problema, el objetivo, y las preguntas de este trabajo de investigación, así como se aborda en la metodología, el juicio por parte de docentes con respecto al desempeño de los estudiantes fue un punto clave

TESIS TESIS TESIS TESIS TESIS

para el desarrollo de este proyecto. Por esta razón, es necesario definir qué se entiende por juicio docente como concepto.

De acuerdo con Brehmer (1987, en Denicolo & Kompf, 2005) y su Teoría de Juicio Social, un juicio puede definirse como un proceso en el que aquél que desempeña el papel de juez hace uso de sus principios e ideales para realizar inferencias sobre casos individuales. Denicolo y Kompf (2005) conectan esta definición de juicio con el juicio meramente docente, y determinan que los docentes son capaces de desempeñar este proceso de inferencias a partir de la información que estos obtienen por medio de la observación que hacen tanto del desempeño y prácticas de un estudiante como de su comportamiento en el aula. Respecto a este último punto, otros autores (Gerbel & Semmel, 1985 en Hamilton & Shinn, 2003) han discutido la relevancia que el comportamiento que los estudiantes presentan en clase constituye como factor para la construcción de un juicio por parte del docente. Así mismo, Hoge y Coladarci (1989) consideran que la certeza de estos juicios depende de tres factores: a) el maestro mismo, b) la tarea o competencia juzgada, y c) el estudiante juzgado, siendo este último factor con el que están de acuerdo con los otros autores que han discutido la definición del concepto de juicio docente y su elaboración como proceso cognitivo.

La importancia que poseen los juicios emitidos por maestros es tal que “muchas decisiones instructivas son determinadas por los juicios subjetivos de los maestros sobre el logro alcanzado por sus estudiantes.” (Südkamp, Kaiser, & Möller, 2014, p. 5). Naturalmente, si el impacto que los juicios que los docentes emiten sobre sus alumnos es tan grande como para demandar que el maestro modifique su instrucción, o como para predecir el éxito que sus estudiantes tendrán (Fuller, 2000), e incluso como para afectar la forma en que sus estudiantes se

desempeñan (Südkamp, Kaiser, & Möller, 2014), es necesario preguntarse qué tan confiables son estos juicios y, de serlo, porque resultan confiables.

DuPaul y Rapport (1991) reportan que los juicios emitidos por docentes pueden tener una mayor comprensión del desempeño de los estudiantes respecto a un contenido académico de lo que podrían demostrar exámenes estandarizados de logro académico. Al respecto, Daugherty (1995) menciona que parte de la confiabilidad de estos juicios se debe a que “los maestros platican con sus estudiantes, guían su trabajo, preguntan y responden sus preguntas, les observan, les ayudan, motivan y retan.” (p. 72). Con base en estos comentarios, puede concluirse que los maestros son una fuente confiable cuando se trata de obtener información sobre el nivel de desempeño de los estudiantes, debido a que ellos pueden percibir elementos del desempeño que una prueba estandarizada no. A pesar de lo anterior dicho, algunos estudios (Fuller, 2000; Martínez, Stecher, & Borko, 2009) han señalado que los juicios de docentes suelen estar fuertemente relacionados con el desempeño que los estudiantes han demostrado en exámenes estandarizados. Esto es que los maestros pueden, en cierta forma, predecir cuáles estudiantes tendrán puntutaciones sobresalientes, promedio, y bajas, en una prueba estandarizada.

En un estudio realizado en Australia para conocer las prácticas docentes, Meiers, Ozolins y McKenzie (2007) hacen notar una serie de elementos que deben estar presentes para considerar que los juicios hechos por docentes están fundamentados de alguna forma. Entre estos elementos se mencionan a) el conocimiento y entendimiento de los docentes con respecto a los estándares locales de evaluación; b) la planeación curricular; oportunidades para que los estudiantes alcancen el logro establecido en los estándares; c) evidencia del logro del estudiante; y d) la validación de los juicios docentes. Estos elementos ayudan a comprender que el concepto de juicio docente no es simplemente algo subjetivo que depende más sobre lo que el docente podría

TESIS TESIS TESIS TESIS TESIS

llegar a percibir erróneamente del desempeño de un estudiante, sino algo fundamentado, por ejemplo, en producciones por parte de los estudiantes que apoyen el juicio del maestro, o por ejemplo, que el docente entienda qué elementos del desempeño del estudiante son los que debe considerar al momento de elaborar un juicio sobre éste.

Con lo anterior, puede entenderse que el juicio docente sobre el nivel de logro de sus estudiantes en algún área de conocimiento no se limita al comportamiento que los alumnos desarrollan dentro del aula, ni tampoco está limitado a lo que el maestro considere apropiado en sus producciones, aunque parte del juicio puede verse influenciado por factores como los que se mencionan con anterioridad. Sin embargo, la subjetividad de estos juicios, de acuerdo con Martínez, Stecher y Borko (2009), tiene que ver con los antecedentes del docente y el contexto en que se desempeña el juicio por parte de éste, pues “no todos los maestros pueden juzgar el desempeño de sus estudiantes con la misma certeza y validez.” (p. 80). De esta forma, se puede entender de una manera más general por qué los juicios que los maestros emiten pueden variar, e incluso ser inexactos, razón por la cual, es necesario tomar en cuenta los elementos mencionados en el párrafo anterior, de manera que los juicios sean más confiables de lo que podría ser si sólo se le diera la oportunidad al docente de emitir un juicio sin necesidad de tomar en cuenta ningún tipo de fundamento.

Para concluir con este apartado, las ideas a rescatar sobre el juicio docente son las siguientes. Primero, hay que comprender al juicio docente como un proceso cognitivo que se fundamenta en la interacción que tiene día a día con sus estudiantes. Segundo, que el juicio docente puede ser una fuente confiable al momento de predecir el nivel de logro académico que podrán alcanzar sus estudiantes. Tercero, que estos juicios deben fundamentarse no sólo en el comportamiento que el estudiante desempeña en el aula, sino en cosas más concretas, como

TESIS TESIS TESIS TESIS TESIS

producciones por parte de los estudiantes donde puedan demostrar su logro académico en una o varias áreas de conocimiento; esto hace que el juicio del docente sea más preciso, y las producciones forman parte de aquellos elementos en que el docente se fundamenta para emitir dicho juicio. Y cuarto, los juicios emitidos por docentes pueden ser influenciados por la formación académica y personal del profesor.

2.1.4.2. Autoevaluación estudiantil. La autoevaluación por parte del estudiante es otra de las formas en que nos aproximamos en este estudio para analizar la certeza con que un examen de colocación ubica a sus sustentantes. Por esta razón, es necesario discutir qué se entiende por autoevaluación además de algunas características de la misma.

Márquez, Medina y Perez (2013), definen a la autoevaluación como “el proceso evaluativo que realiza el mismo estudiante con lo cual conoce el desarrollo y alcance de su propio proceso en relación con unos objetivos trazados previamente.” (p.58). A partir de esta definición, en la que se considera a la autoevaluación como un proceso, podemos rescatar la idea que se apunta sobre la evaluación que el estudiante hace sobre su desarrollo en cierta habilidad o conocimiento.

Al respecto, Martínez, Tellado y Rasposo (2013) argumentan que el estudiante, como “agente activo de su propia evaluación al mismo tiempo que aprendiz del contenido que se evalúa (...) toma conciencia de sí mismo y sus metas” (p.374), de tal forma que el alumno tiene la responsabilidad de conocer cuál es o ha sido su alcance tomando como referencia objetivos que pueden ser planteados por él mismo o por alguien más, como el maestro o un programa.

Oscarson (2013), menciona que “mientras es verdad que el estimado que una persona hace sobre su propia habilidad es subjetivo, eso no necesariamente significa que es no confiable,

o que no es importante ni carente de valor” (p. 1), pues los estudiantes son testigos conscientes de aquello que saben hacer con el lenguaje, así como de aquellos límites que tienen en el mismo.

En este punto es necesario hacer la aclaración que Le-Blanc y Painchaud (1985) hacen al decir que la autoevaluación no es autocalificarse. Para empezar, como Taras (2015) comenta, una calificación no representa de forma directa lo que el estudiante conoce sobre algún área de conocimiento. Taras (2015) y Le-Blanc y Painchaud (1985) parecen estar de acuerdo al expresar que, el hecho de que los estudiantes sepan que pueden asignarse una calificación, hace que la autoevaluación pierda confiabilidad, ya que los estudiantes se preocupan más por una buena calificación que por ser honestos al momento de manifestar su percepción con respecto de su conocimiento.

De cualquier forma, siendo un problema el cuán confiable puede resultar la autoevaluación, Gardner (1999) sugiere que la autoevaluación sea acompañada por una revisión o monitoreo ocasional por parte de maestros, esto para motivar al estudiante a ser honesto con lo que expresa.

En distintos estudios (Blue, 1994; Gardner 1999; Oscarson, 2013; Taras, 2015) se evidencia una revisión de distintas investigaciones donde se ha correlacionado la autoevaluación que hacen los estudiantes sobre su conocimiento en algún área con dos tipos de evaluación, como los resultados de un examen o una evaluación por parte de sus maestros, mostrando una asociación que, si bien los autores simpatizan en decir que no es perfecta, sí resulta significativa y suficiente como para determinar que la autoevaluación es una herramienta confiable al momento de indagar en el conocimiento que un estudiante tiene sobre su aprendizaje o desempeño en cierta área de conocimiento.

Una de las discusiones presentes en la revisión de literatura es el por qué unos estudiantes tienden a sobreestimar su desempeño mientras que otros tienden a infravalorarlo.

Blue (1988, en Blue 1994) habla sobre un estudio en el que se sugería que “algunas nacionalidades podían estar más inclinadas a sobreestimar o subestimar su nivel que otras.” (p. 7). A pesar de esta sugerencia, no se profundiza en tratar de averiguar qué hace que los individuos de cierta nacionalidad tiendan a subestimar o sobrestimar su nivel de dominio en la lengua inglesa por medio de las autoevaluaciones. Boud (1986, en Taras, 2015) menciona que los estudiantes novatos tendían a subestimar ligeramente su nivel de dominio en el lenguaje, mientras que los estudiantes avanzados sobrestimaban ligeramente. Además, Taras (2015) agrega, tras haber analizado otros estudios (Brown y Harris, 2013) que parece haber una tendencia por parte de los estudiantes que tienen un desempeño pobre en autoevaluarse con mayor certeza que los que tienen un desempeño adecuado.

A pesar de que lo anterior puede hacer perder la confiabilidad en la autoevaluación, Oscarson (2013) dice que

el hecho más significativo [de la autoevaluación] es que la propia perspectiva del usuario del lenguaje sobre su habilidad para usarlo es única y, en principio, ilimitada. Nunca podrá ser completamente reemplazada por la visión de nadie más (...) [pues] sólo el usuario mismo ha tenido experiencia de primera mano del grado de efectividad con que el lenguaje ha funcionado en sus encuentros con él. (p. 11).

Se podría llegar a la conclusión de que el problema de la confiabilidad en la autoevaluación no está en relación con lo que el alumno reconoce saber, sino en la honestidad al expresarlo. Como ya se ha mencionado en párrafos anteriores (Taras, 2015; Le-Blanc y

Painchaud, 1985; Gardner, 1999), la honestidad del alumno va en función del peso que tiene la autoevaluación, y la recomendación que queda es no asignar ningún valor que represente un impacto numérico en la calificación sobre el trayecto académico del estudiante.

2.2 El Marco Común Europeo de Referencia para las Lenguas

Otro de los ejes de este proyecto de investigación es el Marco Común Europeo de Referencia para las Lenguas (MCER), por lo que se abordará una descripción del mismo en el que se incluya información tal como el contexto histórico, los objetivos y usos, así como información con respecto a algunas de las competencias comunicativas descritas en el mismo.

2.2.1 Definición del MCER y contexto histórico. Una descripción muy general del MCER puede ser que es precisamente eso, un marco de referencia, el cual sirve a diferentes usuarios del mismo para comprender las competencias del aprendiz o usuario de un lenguaje. Dicha comprensión está guiada por una serie de descripciones hechas por el MCER con respecto a las cuatro habilidades generales del lenguaje, expresión oral y escrita, así como comprensión lectora y auditiva, las cuales se ubican en 6 niveles de dominio, siendo los niveles A1 y A2 descritos como pertenecientes a usuarios básicos, el B1 y B2 a usuarios independientes, y el C1 y C2 de usuarios competentes. (Universidad de Cambridge, 2011).

Taylor y Jones (2006) mencionan que el origen del MCER data desde principios de los años 70, “cuando el Consejo de Europa promovía el trabajo dentro de su Proyecto de Lenguajes Modernos para desarrollar los niveles *Waystage* y *Threshold* como un conjunto de objetivos de aprendizaje especificados para propósitos de enseñanza del lenguaje.” (p.2).

La Universidad de Cambridge (2011) expone que el MCER surge de una necesidad por contar con un marco de referencia internacional en el contexto del aprendizaje de lenguajes, esto con la finalidad de alinear el entendimiento del nivel de dominio que los usuarios de distintos lenguajes poseen.

La Universidad de Cambridge (2011) da un panorama de lo que siguió después de los años 70, mismo que será descrito a continuación. Para los años 80, el enfoque comunicativo se establece y con esto la enseñanza y aprendizaje de lenguajes sufre un cambio al enfatizar el desarrollo de competencias productivas. A principios de los 90 se establece, por medio del simposio intergubernamental “Transparencia y Coherencia en el Aprendizaje de Lenguas en Europa”, un grupo de autoría así como un partido de trabajo internacional para el MCER. El grupo de autoría comprende al jefe de la División de Política del Lenguaje, Joe Shiels, así como a John Trim, Brian North y Daniel Coste. Los objetivos clave del grupo de autoría del MCER son:

establecer una herramienta útil para la comunicación que permitirá a los practicantes en contextos diversos hablar sobre objetivos y niveles de lenguaje de una forma más coherente; y motivar a los practicantes a reflexionar sobre su práctica actual en cuanto al establecimiento de objetivos y en el rastreo del progreso en aprendices con una mira a mejorar la enseñanza y evaluación del lenguaje a través del continente.”
(p.6).

Para el 2001, se publica el borrador final del MCER, así como el lanzamiento del Portafolio Europeo de Lenguaje. En este mismo año se traduce el MCER a alrededor de 37 lenguajes.

Esta mirada al contexto histórico del MCER permite la apreciación de la evolución del mismo, y cómo ha sido el trabajo detrás de la creación del mismo, de lo que podemos resaltar, a partir de la necesidad de distintos países de Europa por contar con una herramienta que les permita comprender las competencias del lenguaje desde una misma perspectiva, y no sólo eso, sino tener un punto de referencia tanto para aprender y enseñar lenguajes como para evaluar el dominio que se tiene de las cuatro habilidades generales de cualquier idioma.

2.2.2 Funciones del MCER. El Consejo de Europa (2001) y la Universidad de Cambridge (2013) argumentan que el MCER puede ser empleado como base para la creación y elaboración de programas, currículos, exámenes, libros, y una infinidad de materiales y elementos pertenecientes a la enseñanza, aprendizaje y evaluación de lenguajes. Es posible llegar a la producción de los diferentes elementos antes mencionados ya que el MCER “describe de una forma comprensiva lo que los aprendices necesitan hacer para usar un lenguaje para comunicarse y que conocimientos y competencias necesitan desarrollar para ser capaces de desempeñarse efectivamente.” (Consejo de Europa, 2001, p.1).

Otra de las finalidades del MCER, que ya se señalaba en párrafos anteriores, es superar las barreras que puedan suscitarse entre profesionales e instituciones dedicados a la enseñanza de lenguajes, pues éste provee “criterios objetivos para describir el dominio de lenguajes que facilitarán el reconocimiento mutuo de cualificaciones alcanzadas en diferentes contextos de aprendizaje, a su vez ayudarán a la movilidad europea.” (Consejo de Europa, 2001, p.1).

Noijons, Béresová, Breton y Szabó (2011), añaden a los usos anteriormente mencionados, el de “permitir que diferentes exámenes sean relacionados uno con otro indirectamente sin ninguna afirmación de que los dos exámenes sean exactamente equivalentes.” (p.18). Lo anterior nos quiere decir que los exámenes pueden ser, y seguramente son, distintos, pero esto no significa que no cubran tanto las categorías como los niveles indicados por el marco de referencia que representa el MCER.

2.2.3 Niveles del Marco Común. Como ya se apuntaba en los primeros párrafos de este apartado, el MCER cuenta con 6 niveles que describen lo que se desearía que un usuario de cada uno de ellos pudiera hacer. La siguiente tabla, tomada de la publicación del MCER por parte del Consejo de Europa (2001, p.24) muestra una descripción general de los mismos. La misma ha sido traducida para este proyecto de investigación.

Tabla 2

Tabla de niveles del Marco Común Europeo de Referencia para las Lenguas (Consejo de Europa, 2001, p. 24)

<p>Usuario Competente</p>	<p>C2</p>	<p>Puede entender con facilidad todo lo que es escuchado o leído virtualmente. Puede resumir información desde diferentes fuentes habladas o escritas, reconstruyendo argumentos y reportes en una presentación coherente. Puede expresarse espontáneamente, de forma muy fluida y precisa, diferenciando tonos más finos de significado incluso en situaciones de mayor complejidad.</p>
-------------------------------	-----------	---

	C1	<p>Puede entender una gran variedad de textos largos y demandantes, y reconocer significado implícito. Puede expresarse con fluidez y espontáneamente sin una búsqueda de expresiones obvia. Puede usar el lenguaje de manera flexible y efectiva para propósitos sociales, académicos, y profesionales. Puede producir textos claros, bien estructurados y detallados sobre temas complejos, mostrando un uso controlado de patrones organizacionales, conectores y recursos de cohesión.</p>
<p>Usuario Independiente</p>	B2	<p>Puede entender las ideas principales en textos complejos tanto en temas concretos como abstractos, incluyendo discusiones técnicas en su campo de especialización. Puede interactuar con un grado de fluidez y espontaneidad que hace bastante posible una interacción con hablantes nativos, esto sin mayor esfuerzo por ninguna de las dos partes. Puede producir textos claros y detallados sobre una variedad amplia de temas y explicar puntos de vista sobre un tema de discusión dando las ventajas y desventajas de varias opciones.</p>
	B1	<p>Puede entender los puntos principales de aportes claros y estándar sobre temas familiares regularmente encontrados en el trabajo, la escuela, ocio, etc. Puede lidiar con la mayoría de</p>

		<p>situaciones posibles de suceder mientras viaja en un área donde el lenguaje es hablado. Puede producir textos conectados de forma simple sobre temas que le son familiares o de interés personal. Puede describir experiencias y eventos, sueños, esperanzas, y ambiciones y dar razones y explicaciones de forma breve para opiniones y planes.</p>
<p>Usuario Básico</p>	<p>A2</p>	<p>Puede entender oraciones y expresiones usadas con frecuencia relacionadas a áreas de relevancia inmediata (por ejemplo, información personal y de familia muy básica, compras, geografía local, empleo). Puede comunicarse forma simple tareas rutinarias que requieren un intercambio de información simple y directa sobre temas familiares y rutinarios. Puede describir en términos simples aspectos de sus antecedentes, entorno inmediato y temas en áreas de necesidad inmediata.</p>
	<p>A1</p>	<p>Puede entender y usar expresiones cotidianas y frases muy básicas centradas en la satisfacción de necesidades de tipo concreto. Puede presentarse a sí mismo y a otros y puede pedir y responder preguntas sobre detalles personales tales como donde vive, la gente que conoce y cosas que tiene. Puede interactuar de forma simple previsto que la otra persona hable lenta y claramente y esté dispuesta a ayudar.</p>

2.2.4 Habilidades de lenguaje. En la enseñanza de idiomas el lenguaje es visto como la integración de cuatro habilidades frecuentemente referidas como “macro-habilidades”. Estas habilidades son a) la expresión oral, b) la expresión escrita, c) la comprensión lectora y d) la comprensión auditiva.

Como Aydoğan (2014) lo discute, estas cuatro habilidades tienden a ser entendidas o diferenciadas entre sí tanto por la forma en que la comunicación toma lugar como por la dirección en que ésta sucede. Así, podemos entender que cuando hablamos de forma, nos referimos a que la comunicación puede ser oral o escrita. La expresión escrita y la comprensión lectora son las habilidades involucradas en la comunicación escrita, mientras que las habilidades que toman parte en la comunicación oral son la expresión oral y la comprensión auditiva.

Ahora, si se prefiere diferenciar a las habilidades por la dirección en que la comunicación sucede, hablamos de habilidades receptoras y productivas. De esta forma, y de acuerdo con Aydoğan (2014), “la comprensión auditiva es la habilidad receptiva en la forma oral [de la comunicación] (...) La expresión oral es la habilidad productiva [en esta misma forma de comunicación]” (p. 673). Al referirnos a la forma escrita de la comunicación, la lectura es la habilidad receptiva mientras que la escritura adquiere el rol de habilidad productiva.

Por razones de interés en este estudio, ya que el examen de colocación PLET sólo evalúa las habilidades receptoras del idioma inglés, junto con la sub-habilidad de estructura escrita (gramática), a continuación sólo se abordará la descripción de la comprensión lectora y la comprensión auditiva, así como la de estructura escrita o gramática.

2.2.4.1 Comprensión. Para poder abordar tanto la comprensión auditiva como la comprensión lectora es necesario definir qué se entiende por comprensión en el contexto del lenguaje.

De acuerdo con Snowling, Cain, Nation y Oakhill (2009), la comprensión es la meta de las dos habilidades receptoras, y para que esta sea considerada como ideal, debe “permitir que los lectores (o escuchas) adquieran información, experimenten y estén al tanto de otros mundos (incluyendo ficticios), para comunicarse satisfactoriamente, y alcanzar el éxito académico.” (p. 1). Una vez establecida la definición de comprensión en el contexto lingüístico, es momento de explorar la comprensión tanto auditiva como lectora.

2.2.4.2 Comprensión auditiva. De acuerdo con Buck (2001), la comprensión auditiva es un proceso bastante complejo, que bien puede ser entendido como un proceso de “abajo hacia arriba en el sentido de que los varios tipos de conocimiento que son involucrados al momento de entender un lenguaje, no son aplicados en un orden fijo” (p.3), pues el momento en que se emplean puede ser cualquiera e incluso alternadamente. Aunado a la idea de complejidad del proceso de comprensión auditiva, Osada (2004), se refiere a esta habilidad como una dimensión crítica en el aprendizaje de algún lenguaje, pero que permanece como “uno de los procesos menos comprendidos.” (p. 53). Lo anterior, explica, se debe a que en la enseñanza de idiomas se atiende esta habilidad de una forma que deja mucho que desear.

Buck (2001) reconoce a la comprensión auditiva como “el resultado de una interacción entre un número de fuentes de información (...) y los escuchas usan cualquier información que tengan disponible, o cualquier información que les parezca relevante para

ayudarles a interpretar lo que el hablante está diciendo.” (p.3). Así, podemos aterrizar lo dicho al contexto de la evaluación, en el que para evaluar la comprensión auditiva de los sustentantes de un examen es necesario tener en cuenta el proceso por el cual pasa la comprensión de lo que es escuchado, así como el rol que adquiere el que escucha, además de las herramientas que le permitan satisfacer dicho proceso.

En sus seis niveles de dominio el MCER describe qué se espera que un usuario de cada nivel pueda hacer con respecto a la comprensión auditiva.

La siguiente tabla ha sido traducida de los descriptores que el MCER ha establecido para cada nivel en cuanto a la comprensión auditiva.

Tabla 3

Tabla de comprensión auditiva por niveles (Consejo de Europa, 2001, p. 66)

COMPRESIÓN AUDITIVA EN GENERAL	
C2	No tiene dificultad en entender cualquier tipo de lenguaje hablado, ya sea en vivo o transmitido, comunicado a una velocidad rápida y perteneciente a nativos del mismo.
C1	Puede entender suficiente para seguir un discurso extendido sobre temas abstractos y complejos más allá de su campo de conocimiento, aunque podría necesitar la confirmación de detalles ocasionales, especialmente si el acento no le es familiar. Puede reconocer un amplio rango de expresiones idiomáticas y coloquialismos, apreciando cambios de registro. Puede seguir discursos extensos incluso cuando no son estructurados claramente y cuando las relaciones son implícitas y no señaladas

	explícitamente.
	Puede entender lenguaje estándar hablado, en vivo o transmitido, sobre temas tanto que le son familiares como los que no lo son, normalmente encontrados en la vida personal, social, académica o vocacional. Sólo ruido extremo de fondo, estructura inadecuada del discurso y/o uso idiomático del lenguaje puede influir en la habilidad para entender.
B2	Puede entender las ideas principales de un discurso por silogismos y complejo lingüísticamente sobre temas tanto abstractos como concretos comunicados en un dialecto estándar, incluyendo discusiones técnicas de su campo de especialización. Puede seguir discursos extensos y líneas complejas de un argumento siempre que el tema sea razonablemente familiar, y la dirección de la plática sea señalizada con marcas explícitas.
B1	Puede entender información factual directa sobre temas comunes del día a día o relacionados con el trabajo, identificando tanto mensajes generales como detalles específicos, siempre que el discurso sea claramente articulado en un acento generalmente familiar.
	Puede entender los puntos principales de un discurso estándar sobre temas que le sean familiares regularmente encontrados en el trabajo, la escuela, ocio, etc., incluyendo narrativas cortas.
A2	Puede entender suficiente como para ser capaz de alcanzar necesidades de un tipo

	concreto siempre que el discurso sea clara y lentamente articulado.
	Puede entender frases y expresiones relacionadas a áreas de mayor prioridad inmediata (por ejemplo, información personal muy básica y de familia, compras, geografía local, empleo) siempre que el discurso sea clara y lentamente articulado.
A1	Puede seguir discursos que son muy lentos y cuidadosamente articulados, con pausas largas para que pueda asimilar el significado.

Puede apreciarse que en el MCER la comprensión auditiva es referida principalmente a la capacidad que el estudiante tiene tanto para reconocer o identificar ideas principales, la esencia de un texto, información específica, entendimiento detallado, entre otras cuestiones que conforman el proceso de escuchar y comprender lo que se escucha.

Así mismo, el Consejo de Europa (2001) sugiere ciertos escenarios o contextos en los que se desarrolla una situación en que el estudiante prueba su comprensión auditiva. Tales escenarios van desde grabaciones o audios en medios de comunicación, anuncios e instrucciones, siendo parte de una audiencia en vivo, hasta la interacción entre dos hablantes nativos, siendo todos estos escenarios los marcadores o elementos que ayudan a separar por nivel de dificultad la práctica de la comprensión auditiva.

2.2.4.3 Comprensión lectora. Así como la comprensión auditiva es referida como una habilidad receptiva, la comprensión lectora forma parte de esa misma categoría, y también es considerada un proceso.

Snow (2002), define a la comprensión lectora como “el proceso de extraer y construir significado simultáneamente a través de la interacción e involucramiento con el lenguaje escrito.” (p. 11). Woolley (2011) también se refiere a la comprensión lectora como un proceso cuyo producto es la construcción de significado, de forma que el lector se haga de un entendimiento general de lo que es descrito en un texto. Referente al término de interacción, Alyousef (2005), concuerda en considerar a la lectura como un proceso que pone al lector y a un texto como dos agentes que interactúan “dinámicamente (...) mientras [el lector] trata de obtener el significado y donde varios tipos de conocimiento están siendo usados: conocimiento lingüístico o sistemático así como conocimiento esquemático.” (p. 144).

De esta forma, y para intereses de esta investigación, la comprensión lectora ha sido entendida como un proceso interactivo entre lector y texto, así como un proceso mediante el cual se trata de producir un significado a partir de lo leído, ya sea como conjunto o por partes.

En sus seis niveles de dominio el MCER describe qué se espera que un usuario de cada nivel pueda hacer con respecto a la comprensión lectora.

La siguiente tabla ha sido traducida de los descriptores que el MCER ha establecido para cada nivel en cuanto a la comprensión auditiva.

Tabla 4

Tabla de comprensión lectora por niveles (Consejo de Europa, 2001, p. 69)

	COMPRESIÓN LECTORA EN GENERAL
C2	Puede entender e interpretar críticamente casi todas las formas de lenguaje escrito incluyendo escritos literarios y no literarios abstractos, de estructura compleja, o muy coloquiales. Puede entender una gran variedad de textos largos y complejos, apreciando distinciones sutiles de estilo y significado tanto implícito como explícito.
C1	Puede entender a detalle textos extensos y complejos, ya sea que esté o no relacionado a su área de especialidad, siempre que pueda releer secciones difíciles.
B2	Puede leer con un alto grado de independencia, adaptándose al estilo y velocidad de lectura de diferentes textos y propósitos, y usando selectivamente fuentes de referencia apropiadas. Tiene un amplio vocabulario de lectura activo, pero puede experimentar algo de dificultad con expresiones idiomáticas usadas con poca frecuencia.
B1	Puede leer textos factuales de forma directa sobre temas relacionados a su campo de interés con un nivel de comprensión satisfactorio.

A2	Puede entender textos cortos y simples sobre temas que le sean familiares de tipo concreto que consista de lenguaje de alta frecuencia en el día a día o relacionado al trabajo.
	Puede entender textos cortos y simples que contengan el vocabulario usado con mayor frecuencia, incluyendo una proporción de ítems de vocabulario compartidos internacionalmente.
A1	Puede entender textos muy cortos y simples, una frase a la vez, reconociendo nombres, palabras y frases básicas que le sean familiares y releendo lo que sea requerido.

Por medio de la tabla anterior, puede resumirse que la comprensión lectora es entendida por el MCER como un proceso que el lector o usuario del lenguaje desempeña con el fin de encontrar la esencia de un texto, sus ideas principales, información específica y entendimiento detallado, entre otros elementos que hacen de la comprensión una meta para el proceso de lectura.

Como en la comprensión auditiva, la dificultad de la comprensión lectora está regida con base en distintos ejes, los cuales van desde leer instrucciones, información general y argumentos, leer para orientarse, hasta leer correspondencia, escenarios en los que la extensión del texto, el vocabulario empleado y la finalidad del mismo, juegan un papel

importante en los niveles de complejidad a los que un usuario del lenguaje podría enfrentarse en su aprendizaje y adquisición de una nueva lengua.

2.2.4.4 Estructura escrita (Gramática). Entendida más bien como una sub-habilidad, la estructura escrita, término que se usa en esta investigación para referirse a la construcción gramatical del lenguaje, es otro proceso que es evaluado en diversos exámenes de lenguaje, con la finalidad de conocer la competencia o competencias que el usuario del lenguaje posee para identificar errores en una oración, como para completar adecuadamente oraciones en las que hace falta una parte que dé sentido a las mismas.

El Handbook de gramática de la Universidad Capella (s.f.), define a la construcción básica de una oración en inglés como una estructura que es aprendida al inicio de cursos de inglés. Como estructura, ésta se ve compuesta de diferentes elementos y partes con un significado individual que en conjunto otorgan a la oración un significado colectivo.

Como elementos, se entienden cuestiones como los bloques de construcción de oraciones (frases, cláusulas, oraciones simples, oraciones compuestas, oraciones complejas), los tipos de oración (declarativa, imperativa, interrogativa, exclamatoria), y la expansión de la oración (oraciones combinadas, oraciones de corrido (sin el uso de comas) coordinación de oraciones, subordinación de oraciones).

Clifton y Duffy (2001), sugieren diferentes tipos de estructura que una oración y un texto pueden llegar a tener. Estos tipos de estructura son a) la estructura semántica, b) la estructura léxica, c) la estructura prosódica, y d) la estructura semántica, todas ellas descritas en los siguientes puntos.

a) Estructura sintáctica

La estructura sintáctica de una oración se refiere a la composición de la misma basada en principios gramaticales, limitada a la interpretación semántica de la oración dicha o escrita, y cuyo proceso analítico y de comprensión se centra en cómo se relacionan una y otra palabra en la oración.

b) Estructura léxica

En este tipo de composición de oraciones, la comprensión o el análisis de las mismas se desarrolla bajo la consideración de que el lenguaje está compuesto por distintas dimensiones y cada dimensión es regida por sus propias reglas. De esta forma, el análisis de la estructura de oraciones a nivel léxico se refiere a comprender una oración por partes, dichas partes vendrían a representar frases compuestas por dos o más palabras. La interpretación de las mismas, así como con la estructura sintáctica, se limita al significado semántico.

c) Estructura prosódica

Puesto de una forma muy simple, la estructura prosódica de oraciones es más bien parte de oraciones habladas y que le exigen al usuario del lenguaje analizarlas por medio de la habilidad auditiva, de tal forma que el análisis o la comprensión de la oración expresada de forma oral se ayuda de la interpretación del ritmo y la melodía empleadas por quien dice la oración.

d) Estructura semántica

La estructura semántica de una oración es analizada por medio de las relaciones entre los elementos del habla (verbos, sujetos, adjetivos, artículos, etc.) y por medio de la relación que existe entre ellos construir un significado de la oración. Puede decirse que si bien requiere de la comprensión sintáctica, ésta, la semántica, va un paso más allá, pues busca comprender la denotación de la oración como un conjunto, y no por sus partes.

Entonces, después de este recorrido lingüístico sobre la construcción de oraciones, se puede decir que la estructura de oraciones tiene como fin un análisis por medio del cual el usuario del lenguaje pueda detectar el significado de una oración así como de sus partes, teniendo en cuenta las cuestiones gramaticales que ayuden a la construcción y comprensión de la misma.

Si bien el MCER no cuenta con una descripción sobre el uso del lenguaje, a continuación se ha traducido una tabla sobre contenidos gramaticales que el usuario del lenguaje debe conocer para comprender, por medio del reconocimiento de errores y la compleción de oraciones, el significado y lógica de una oración. La construcción de esta tabla se hizo en función de los intereses de este estudio, que giran en torno a cómo es abordada la estructura de oraciones en el examen PLET. Por la misma razón, sólo se describen los contenidos gramaticales de los primeros cuatro niveles del MCER.

Tabla 5

Tabla de contenidos gramaticales por nivel del MCER. (Exam English.com, 2011; Tabla de Contenidos PLET, 2015)

	ESTRUCTURA DE ORACIONES (GRAMÁTICA)
B2	Adjetivos y adverbios; Futuro perfecto; Futuro perfecto continuo; Condicionales mixtos; Revisión de tiempos gramaticales; Especificidad en temas vistos en niveles inferiores.
B1	Adverbios; Condicionales segundo y tercero; Futuro continuo; Modales de deducción; Pasado perfecto; resente perfecto continuo; Habla reportada; Voz pasiva.
A2	Adjetivos comparativos y superlativos; Frases adverbiales de tiempo, lugar y frecuencia; Adverbios de frecuencia; Artículos; Sustantivos contables y no contables; Futuro simple; Gerundios; Going to; Modales de obligación y permiso; Pasado continuo; Pasado simple; Verbos frasales o compuestos; Presente continuo para futuro; Presente Perfecto; Patrones de verbo infinitivos y gerundios; Condicionales cero y primero.
A1	Adjetivos; Adverbios de frecuencia; Comparativos; Going to; Sustantivos contables y no contables comunes; Imperativos; Verbos modales; Pasado simple; Adjetivos y pronombres posesivos; Preposiciones de lugar y tiempo; Presente continuo;

	Presente simple; Pronombres; Preguntas; Verbo to be.
--	--

La tabla anterior debe entenderse como aquellos temas que son generalmente vistos en cursos de inglés por cada nivel, entendiendo que son de carácter seriado, pues lo que se ve en el nivel más básico es reforzado en niveles superiores, de tal forma que un nivel no podría ser sin el anterior. La complejidad, como se puede apreciar, radica principalmente en la especificidad con que se abordan los temas y la combinación que se hace de algunos de ellos.

Capítulo 3. METODOLOGÍA DE LA INVESTIGACIÓN

El diseño de esta investigación es de tipo metodológico y su propósito es la validación de un instrumento de medición para la colocación en niveles de inglés de acuerdo con el dominio que los sustentantes muestran con respecto a habilidades del lenguaje, tales como la comprensión lectora y auditiva, además de la estructura escrita o gramática. Este estudio recopiló evidencia de validez basada en relación con otras variables a partir de los puntajes del examen institucional de la UAA para la colocación de estudiantes en cursos de inglés, PLET. Se evaluaron, pues, los puntajes del mencionado examen al compararlos con los juicios que los docentes de los cursos del PILE generaron sobre la colocación de los estudiantes, así como la percepción que los estudiantes tienen del nivel de dominio de las habilidades de lenguaje anteriormente señaladas. El juicio emitido por los docentes, y a su vez, la autoevaluación que los estudiantes hicieron sobre sus competencias comunicativas, fueron registrados en instrumentos elaborados a partir de las competencias señaladas por nivel en el MCE. El desempeño demostrado por los estudiantes en las secciones de comprensión auditiva, gramática, y comprensión lectora en su examen

intermedio, también se tomó como una fuente de evidencia de validez, y fueron analizadas y evaluadas por medio de una rúbrica.

3.1 Participantes

Los participantes de este estudio son docentes y alumnos de la UAA, de diferentes niveles de los cursos de inglés que el PILE ofrece. La muestra de estudiantes fue elegida a partir de 109 grupos que estaban a cargo del PILE en el periodo de agosto – diciembre de 2016, grupos que a su vez estaban divididos en los ocho niveles que ofrece el programa.

Se utilizó una fórmula de muestreo por conglomerados para determinar el tamaño de la muestra, el cual terminó siendo de un total de 14 grupos. Una vez determinado el tamaño de la muestra, se utilizó una fórmula de muestreo estratificado por proporción para determinar cuántos grupos de cada nivel serían necesarios, y como resultado se obtuvo un grupo de Nivel 1, dos de Nivel 2, tres de Nivel 3, cuatro de Nivel 4, y para los niveles 5, 6, 7, y 8, un grupo. La selección de los grupos fue por medio de un muestreo aleatorio. Al final, se trabajó con 12 de los 14 grupos seleccionados, pues dos de los docentes se negaron a participar en el estudio. Es de esta forma que entonces fueron dos grupos de Nivel 3 en lugar de tres, y quedó en tres grupos de Nivel 4 en lugar de cuatro.

Los alumnos, quienes fueron los primeros informantes del estudio, son estudiantes de diferentes carreras en la universidad, así como de diferentes semestres. Su edad varía de los 18 a los 25 años de edad. Los grupos estuvieron compuestos tanto por hombres como por mujeres. Al ser grupos con tantas diferencias, lo único que compartían en común independientemente de su carrera, semestre, edad, o género, es el nivel de inglés en el que fueron colocados.

Los docentes, quienes fungieron como los segundos informantes en este estudio, son profesionales en la enseñanza del idioma inglés. De acuerdo con los requisitos para trabajar en el PILE, en su totalidad, estos docentes son egresados de la licenciatura en Enseñanza del Inglés. La mayoría de ellos cuentan con experiencia frente a grupo de al menos dos años, y varios de ellos han trabajado o trabajan en distintos niveles educativos. Una gran parte de los docentes del PILE imparten diferentes niveles de inglés. Todos cuentan con acreditación vigente del idioma inglés por parte del TOEFL. Todos estos elementos contribuyen al perfil de un maestro cuya experiencia es suficiente para emitir juicios sobre el nivel de competencia de sus estudiantes. Si bien no existe un periodo de tiempo que sea universalmente aceptado en la literatura como adecuado para emitir un juicio, tomando como fundamento el trabajo realizado por Papageorgiou y Cho (2014), podemos establecer que después de un mes de trabajo el maestro ya tiene suficiente información sobre el estudiante para emitir un juicio sobre su desempeño. Además, el hecho de que estos maestros sean quienes trabajan con grupos de estudiantes que han sido colocados en cierto nivel de inglés a partir de sus resultados en el PLET, los hace los participantes ideales para el estudio que aquí se desarrolla.

En la siguiente tabla se muestra el número de participantes por instrumento.

Tabla 6

Cantidad de Participantes por Instrumento

Instrumento	Número de Participantes
Encuesta para la Autopercepción de Habilidades en Inglés.	158 estudiantes del PILE.

Encuesta para el Juicio Docente.	12 docentes valorando a 48 estudiantes, ambos del PILE.
Rúbrica de Evaluación de Desempeño.	34 exámenes de estudiantes del PILE valorados por el investigador.

3.2 Instrumentos

Se utilizaron tres instrumentos para la recolección de información en este estudio, además del PLET.

3.2.1 Encuesta para la Autopercepción de Habilidades en Inglés

3.2.1.1 Propósito. El primer instrumento fue una encuesta para los estudiantes a la que se le llamó “Encuesta para la Autopercepción de Habilidades en Inglés”. El propósito del instrumento fue recoger información sobre la percepción que los estudiantes tienen sobre su nivel de dominio en los contenidos que previamente fueron evaluados por el PLET, para de esta forma comparar el nivel que ellos creen poseer (derivado de las respuestas que otorgaron en la encuesta) y el nivel al que fueron colocados.

3.2.1.2 Estructura. Esta encuesta está construida por diferentes enunciados que buscan explorar la percepción que los estudiantes tienen sobre su uso del idioma inglés.

Los enunciados se formulan todos a partir del verbo “Puedo”, de tal forma que el estudiante registre cuáles son aquellas actividades que considera puede llevar a cabo con el nivel de dominio del idioma que posee.

TESIS TESIS TESIS TESIS TESIS

Todos los enunciados exploran el uso del idioma por medio de actividades del día a día en diferentes contextos que resultan comunes para cualquier usuario del lenguaje o en los que se esperaría que el usuario pudiera desempeñarse.

Los enunciados para todos los niveles de comprensión lectora y comprensión auditiva son una traducción hecha a partir de los enunciados que North (2008) ha creado precisamente con la intención de que los sustentantes de pruebas estandarizadas puedan explorar su nivel de dominio de cualquier lengua. Los enunciados para gramática de todos los niveles son una construcción por parte del investigador con base en los descriptores del MCER y los programas de los cursos de inglés del PILE.

Cada enunciado fue ponderado por el estudiante por medio de una escala Likert, cuyos niveles son “Muy de acuerdo”, “De acuerdo”, “Indeciso”, “Desacuerdo”, y “Muy en desacuerdo”, mismos que pretenden reflejar la postura del estudiante con respecto a lo expresado en cada uno de los enunciados.

Es importante mencionar que la encuesta fue dividida en ocho versiones (una por nivel del PILE). La estructura de todas las versiones de encuesta es similar, pues se incluyen todos los enunciados generados para el nivel en que el estudiante se encuentra además de enunciados que atienden al nivel inmediato inferior y al nivel inmediato superior. Con respecto a los otros niveles inferiores y niveles superiores no se incluye todo el paquete de enunciados, pero sí se incluye uno o dos enunciados representativos de la habilidad por cada uno de esos niveles. Así que, por ejemplo, la escala para el nivel cuatro tiene todo el paquete de enunciados generados para atender a ese nivel, así como todos los enunciados para el nivel tres y todos los enunciados para el nivel cinco, mientras que para atender los niveles uno, dos, seis, siete y ocho, sólo se

TESIS TESIS TESIS TESIS TESIS

incluye un enunciado representativo por nivel para la habilidad de comprensión lectora, otro por nivel para gramática, y dos por nivel para la comprensión auditiva.

3.2.1.3 Contenido. Todos los enunciados que tienen que ver con el uso de las habilidades de comprensión lectora y auditiva fueron traducidos del inventario de North(2008), mientras que aquellos enunciados que pretenden explorar la estructura de oraciones fueron construidos personalmente a partir de los temas que fueron señalados en el sub-apartado de estructura de oraciones en el apartado dedicado al MCER en este documento, en la tabla 7.

Como ya se aclaraba anteriormente, los enunciados son desarrollados de forma que expresen actividades del día a día en diferentes contextos, como el académico, el laboral, de relaciones sociales y/o familiares, entre otros. Este instrumento se inspira en la metodología de Powers, Kim, y Weng (2008).

3.2.2 Encuesta para el Juicio Docente

3.2.2.1 Propósito. El segundo instrumento es otra encuesta que se denomina “Encuesta para el Juicio Docente”. El instrumento tiene como propósito explorar el juicio que el docente ha construido sobre las habilidades de comprensión auditiva y comprensión lectora, así como la sub-habilidad de gramática de sus estudiantes. Esta encuesta resultó en la aprobación o desacuerdo con la colocación de algunos de los estudiantes del grupo y nivel que el docente imparte.

3.2.2.2 Estructura. La escala está compuesta por una traducción de los descriptores del MCER, además de contar con ejemplos de lo que se espera que un estudiante del nivel que el docente imparte deba lograr como usuario del lenguaje en las habilidades de comprensión lectora y auditiva, y en la sub-habilidad de estructura de oraciones.

TESIS TESIS TESIS TESIS TESIS

También cuenta con un espacio para comentarios donde el docente tiene la libertad de registrar observaciones cuando sienta que sea pertinente justificar o añadir algo que la encuesta no permita expresar.

Los enunciados se formulan todos a partir del verbo en tercera persona singular “Puede”, de tal forma que el docente registre cuales son aquellas actividades que considera su alumno puede llevar a cabo con el nivel de dominio del idioma que posee.

Todos los enunciados exploran el uso del idioma por medio de actividades del día a día en diferentes contextos que resultan comunes para cualquier usuario del lenguaje o en los que se esperaría que el usuario pudiera desempeñarse; sin embargo, para este instrumento los enunciados evalúan estas actividades de una forma más general que los enunciados del primer instrumento, de tal suerte que los enunciados de este instrumento sean una especie de resumen de todos los enunciados que en el instrumento para los estudiantes exploraban un nivel. Por ejemplo, si en el Inventario para la Autopercepción de Habilidades en Inglés para explorar el nivel B1.1 se tienen cinco enunciados que deberán ser respondidos individualmente, en este instrumento, el Inventario para el Juicio Docente, el enunciado es una suma del contenido que exploran esos cinco enunciados, ya que fueron traducidos a partir de los descriptores del MCER por nivel y habilidad.

Cada enunciado fue ponderado por el maestro por medio de una escala Likert, cuyos niveles fueron “Dos o más niveles inferiores al esperado”, “Un nivel inferior al esperado”, “De acuerdo al nivel esperado”, “Un nivel superior al esperado”, y “Dos niveles o más superiores al esperado”.

En “De acuerdo al nivel esperado” se adjuntó la descripción de ese nivel, así como para “Un nivel inferior al esperado” y “Un nivel superior al esperado”. Para “Dos o más niveles inferiores al esperado” y “Dos o más niveles superiores al esperado”, se incluyeron descripciones genéricas como “Presenta dificultades serias para la realización de actividades correspondientes a X habilidad” y “Las actividades correspondientes a este nivel no presentan reto alguno para el estudiante”, respectivamente.

3.2.2.3 Contenido. Así como con el primer instrumento, en la Encuesta para Juicio Docente se exploraron las habilidades de comprensión lectora y comprensión auditiva, así como la sub-habilidad de gramática.

Los enunciados contienen un descriptor del MCER correspondiente a una habilidad y a un nivel. Los descriptores para los niveles de comprensión lectora y comprensión auditiva son tomados directamente del MCER; los de estructura de oraciones son una construcción personal a partir del contenido de los niveles del MCER y el contenido de los programas de los cursos de PILE.

La encuesta está inspirada en la forma utilizada por Papageorgiou y Cho (2014).

3.2.3 Rúbrica de Evaluación de Desempeño

3.2.3.1 Propósito. El tercer instrumento es una rúbrica, la cual lleva por nombre “Rúbrica de Evaluación de Desempeño”. El propósito de la rúbrica es el de evaluar el desempeño mostrado por algunos de los estudiantes con respecto a algunas tareas y/o actividades relacionadas con la comprensión lectora y auditiva, así como el dominio en la estructura de oraciones.

3.2.3.2 Estructura y contenido. La rúbrica se conformó de la siguiente manera:

Primero se indicó la habilidad de la que fueron registrados los resultados del examen, ya sea comprensión auditiva, comprensión lectora, o gramática.

Después se indicó el nivel del PILE y el nivel del MCER, el cual se acompaña de una descripción sobre lo que se espera que un estudiante de dicho nivel deba de lograr.

Luego, se tiene una casilla que indica el número de reactivo perteneciente al examen intermedio que fue registrado.

Posterior a esta casilla se encuentra otra donde se registra el tipo de actividad de cada reactivo. Por ejemplo, para comprensión auditiva se tienen actividades tales como identificación de información específica, identificación de detalles, e identificación de ideas principales. Para comprensión lectora los tipos de actividad son los mismos que para comprensión auditiva. En el caso de gramática, los tipos de actividad son completar una oración a partir de dos o tres opciones de respuesta, completar una oración conjugando un verbo, estructuración de oraciones, y ordenamiento de oraciones.

Antes de registrar si el alumno tuvo una respuesta correcta o no en un reactivo, se decidió dar un valor a cada uno de los reactivos en cada una de las habilidades para cada una de las rúbricas. De esta forma, el valor “1” expresa que el reactivo y tipo de actividad presentan un nivel de dificultad adecuado para el nivel que cursa el estudiante; mientras que el valor “2” quiere decir que la dificultad y el tipo de actividad en el reactivo son adecuados para el nivel que se cursa pero exigen una demanda cognitiva mayor que los otros reactivos en la sección o habilidad que se evalúa. La decisión sobre la dificultad de cada reactivo viene de dos fuentes que el investigador consideró, las cuales son los descriptores del MCER y el programa de cada

TESIS TESIS TESIS TESIS TESIS

nivel de los cursos del PILE. De esta forma, para comprensión auditiva, por ejemplo, se escucharon los audios utilizados para evaluar la habilidad en cada uno de los niveles, y se consideraron cuestiones en el audio tales como velocidad del habla, uso de lenguaje estándar, ruidos de fondo que incrementaran la dificultad, la cantidad de información que se daba, etc.

En el caso de la sección de gramática la dificultad fue designada dependiendo del tipo de actividad, pues no representa la misma dificultad para un estudiante el que, por ejemplo, le den opciones para completar una oración a cuando se le pide estructurar una oración completa. Además, también se tomaba en cuenta, en las actividades donde se tenía que conjugar un verbo, qué tan común era éste para el nivel y el tipo de curso que se lleva.

Para la sección de comprensión lectora la dificultad está dada por el tipo de actividad así como el tipo de preguntas que se hacían. Lo anterior quiere decir que para una actividad que por ejemplo evalúa la habilidad para identificar información específica en un texto, puede que la información aparezca tal cual o más bien aparezca por el uso de sinónimos; en este caso se evaluaba lo común del vocabulario para los estudiantes, no del texto en sí, si no de lo que pedía el reactivo. De igual forma, con las actividades sobre la identificación de ideas principales había algunas en las que se presentaba con opciones de respuesta, mientras que otras empleaban ejercicios de “verdadero o falso” donde en algunas ocasiones el alumno tenía que inferir información para determinar si una idea principal sobre un párrafo o sobre el texto era verdadera o falsa. Sucede lo mismo en algunas actividades de identificación de información específica y detalles.

Por medio de la siguiente tabla se expresa quienes fueron las fuentes de información para cada instrumento, así como los constructos que cada uno exploró:

Tabla 7

Instrumentos de Obtención de Información e Informantes.

Instrumento	Informante	Constructo
Encuesta para la Autopercepción de Habilidades en Inglés	Estudiantes del PILE	Comprensión lectora Comprensión auditiva Estructura escrita (gramática)
Encuesta para el Juicio Docente	Docentes del PILE	Comprensión lectora Comprensión auditiva Estructura escrita (gramática) Colocación de estudiantes
Rúbrica de Evaluación de Desempeño	Exámenes intermedios de cursos del PILE (Exámenes) Investigador	Comprensión lectora Comprensión auditiva Estructura escrita (gramática)

3.3 Procedimiento

3.3.1 Construcción y Validación de Instrumentos. Para la creación de los instrumentos de investigación, previamente se realizó una revisión de la literatura disponible al respecto de las temáticas que se pretenden explorar con los mismos. Este proceso de lectura resultó en la definición de los constructos a investigar por medio del juicio docente, la autoevaluación de habilidades por parte de estudiantes, y los resultados que arroje el uso de la rúbrica.

Se tomó la definición del MCER para la conceptualización de habilidades de comprensión lectora y comprensión auditiva. Además, la revisión de programas de los cursos de inglés y el contenido de los libros de esos cursos, ayudaron a la creación de contextos y delimitación de cuestiones y temas gramaticales por cada nivel en el MCER.

Las encuestas fueron herramientas de obtención de la información de carácter autoadministrado. Consisten de enunciados sobre actividades que el estudiante puede hacer y que evidencian su dominio sobre los constructos explorados, así como, en el caso de la escala para el docente, enunciados donde se exprese lo que se espera que un estudiante pueda hacer en un nivel determinado. En la literatura inglesa, estos inventarios son conocidos como *'can-do' checklists*.

3.3.2 Construcción de Marcos de Referencia. La revisión de literatura permitió la posibilidad de construir un marco de referencia en el cual los ejes de la investigación son reflejados. Para esto, el MCER fue la base principal para la creación del marco de referencia para los instrumentos. De esta forma, se atiende la validez relacionada con otras variables, el juicio docente, la autoevaluación estudiantil, la comprensión lectora, la comprensión auditiva, y la estructura de oraciones o gramática. La conceptualización de cada uno de estos elementos de la

investigación, a partir del MCER, ayudó a aclarar cuál será el foco de cada uno de los instrumentos, y a su vez ayudó a la construcción de la operacionalización de constructos.

Así, la fuente de información más importante ha sido el Consejo de Europa y su documento del MCER del 2001 de donde se ha obtenido más soporte con respecto a la construcción de varios conceptos; esto debido a que el examen PLET emplea los descriptores de éste para la creación de su contenido y es el marco de referencia utilizado por instituciones como el ETS, Cambridge, entre otros, para el diseño de pruebas estandarizadas como el PLET.

3.3.3 Operacionalización de los Constructos. La operacionalización de los constructos a explorar tuvo como base primordial los descriptores del MCER así como información con respecto de los contextos y la gramática que es vista en los distintos niveles que ofrece el PILE; esto se logró por medio del análisis del contenido en los libros y programas de los cursos de inglés de la institución.

Los pasos para la operacionalización fueron de la siguiente forma:

- a) Construcción conceptual de comprensión lectora, comprensión auditiva, y estructura del lenguaje por niveles tanto del MCER como del PILE.
- b) Delimitación de dimensiones a explorar.
- c) Extracción de indicadores a partir de la construcción conceptual de cada una de las habilidades y la sub-habilidad referidas en el inciso A.

Una vez que se construyó la operacionalización de variables, se dio paso a la selección de enunciados que fueron traducidos del inventario de North (2008) para la creación del Inventario para la Autopercepción de Habilidades en Inglés, y la selección de descriptores para el diseño del

TESIS TESIS TESIS TESIS TESIS

Inventario para el Juicio Docente y para la Rúbrica para la Evaluación de Producciones Académicas. Esta selección se dio al contrastar las dimensiones e indicadores en la operacionalización con el contenido de cada enunciado, de manera que se rescataron aquellos enunciados que en verdad exploran las dimensiones e indicadores que forman parte del contenido de la operacionalización de variables para los instrumentos.

3.3.4 Traducción de Enunciados para las Encuestas. Como ya se ha referido anteriormente, para explorar la competencia de los estudiantes en el uso de las habilidades de comprensión lectora y auditiva, se tradujeron varios de los enunciados que North (2008) desarrolló a partir de los descriptores del MCER para la Encuesta para la Autopercepción de Habilidades en Inglés, y también se hizo una traducción de estos descriptores para la Encuesta para el Juicio Docente.

El proceso de traducción de instrumentos requirió de una serie de pasos que fueron llevados a cabo cuidadosamente de manera que la pragmática y semántica del contenido no fueran alteradas tras el proceso de traducción. Como Gaité, Ramírez, Herrera, y Vázquez-Barquero (1997) mencionan, el principal aspecto que se debe tener en cuenta al traducir es la equivalencia de las versiones, misma que se divide en equivalencia técnica, equivalencia semántica, y equivalencia conceptual. (p. 8). También ofrecen la explicación de distintas metodologías de traducción.

De esta manera, tras haber seleccionado los enunciados que resultaron adecuados de acuerdo con la operacionalización, se tradujeron por dos personas, y como resultado de la comparación de traducciones se llegó a la primera versión de los instrumentos.

3.3.5 Validación de Instrumentos por Jueces Expertos. Una de las fases de mayor importancia en el diseño de instrumentos es el jueceo por parte de personas con experiencia en el campo donde los constructos explorados tienen presencia. Basado en la información recopilada para esta investigación y el objetivo de la misma, se diseñó de manera personal un perfil del juez experto para la actividad de valoración de los instrumentos, y este fue:

- a) Persona con dominio del idioma inglés.
- b) Persona que conozca de forma detallada el contenido del MCER.
- c) Persona que haya tenido experiencia con el diseño de evaluaciones del idioma inglés.
- d) Persona que tenga experiencia con la impartición de cursos relacionados al desarrollo de competencias comunicativas.

La tarea de jueceo demanda a los responsables a evaluar los reactivos y la operacionalización previa al diseño de los mismos. Esta evaluación se enfocó en que los reactivos respetaran los siguientes criterios:

1. Claridad: El enunciado refleja claramente una actividad que requiere el dominio, en cierto nivel, de una de las habilidades o sub-habilidad a explorar en la investigación, y no permite más de una interpretación.

Problemas a evitar:

2. Consistencia: El enunciado es apropiado de acuerdo con lo expresado en la operacionalización, pues respeta el contenido de lo que se espera en la habilidad y el nivel.
3. Coherencia: Entre las definiciones conceptuales, las dimensiones, y los indicadores.

4. Adecuación: El enunciado es adecuado para valorar el constructo a explorar.
5. Suficiencia: Lo expresado en el enunciado es exhaustivo en cuanto a indicadores de la variable a explorar, pues comprende cuestiones de cantidad y calidad.

3.3.6 Pilotaje. Se seleccionó una muestra por medio de un muestreo probabilístico aleatorio simple tanto para la selección de maestros como para la selección de estudiantes.

El pilotaje se llevó a cabo después de la selección de la muestra y una vez que los instrumentos fueron expuestos al jueceo de expertos y a las adecuaciones sugeridas a partir de este proceso.

El pilotaje tuvo lugar en la UAA, y se llevó a cabo durante los cursos intensivos del verano de 2016, pues fueron grupos de estudiantes y docentes con las mismas características que la población a quienes posteriormente se aplicaron los instrumentos definitivos.

Se aplicó la Encuesta para la Autopercepción de Habilidades en Inglés en la segunda semana del curso, mismo que tuvo una duración de tres semanas. Se les dio la instrucción a los estudiantes sobre lo que tenían que hacer en cada enunciado, respondiendo de manera honesta lo que se les pidió en los mismos. Además, en el instrumento de pilotaje se les proporcionó un espacio para que hicieran comentarios sobre enunciados que les resultaran confusos o difíciles de contestar. La aplicación de la encuesta tomó alrededor de 25 minutos. Se aplicó a todos los estudiantes para de ahí seleccionar una cantidad de ellos que hayan expresado estar en el nivel adecuado, algunos que hayan expresado estar en un nivel más alto, y algunos que hayan indicado encontrarse en un nivel inferior.

La Encuesta para el Juicio Docente se aplicó al cabo de la segunda semana del curso, cuando el docente ya había construido una opinión y juicio sólidos sobre la colocación y desempeño en el idioma inglés de sus estudiantes. Se le indicó al docente de igual forma que sus comentarios y retroalimentación sobre el instrumento serían agradecidos. Sólo se le pidió emitir su juicio sobre una muestra de todos sus estudiantes, procurando que se cubran todos los casos posibles, tanto los que hayan expresado estar en el nivel adecuado como los que se perciban en un nivel que no es consistente con sus habilidades, ya sea porque es un nivel más alto o más bajo del que creen tener.

Una vez que se recolectaron estos datos, se analizaron la validez y confiabilidad de los instrumentos, para posteriormente aplicar la Rúbrica para la Evaluación de Producciones Académicas.

La rúbrica se empleó después de que terminaran los cursos de inglés. Se le pidió con anterioridad al maestro el solicitar a los alumnos seleccionados todas aquellas actividades donde el estudiante haya trabajado su comprensión lectora y auditiva, además del uso de la gramática. Aunque la respuesta por parte de maestros fue en general acorde a lo que se les solicitó, una revisión previa de los materiales y producciones académicas de los estudiantes permitió descartar lo que fueron libros de trabajo y mini exámenes, para sólo enfocarse en los denominados “exámenes intermedios” que se presentan después del primer mes y medio del curso en la modalidad de curso extensivo. La decisión de descartar el libro de trabajo fue que, a pesar de evaluar las tres habilidades de interés para este estudio, eran actividades que el alumno llevaba a cabo como tareas que podía contestar en casa, lo que hizo de esta fuente de información un elemento no muy confiable al momento de ofrecernos datos sobre el desempeño verdadero del estudiante, pues así como podría poner los audios para las actividades de comprensión auditiva

una sola vez, podría hacerlo dos o más veces, y de forma similar con los ejercicios de gramática y comprensión lectora, pues podría hacer uso de herramientas como la Internet o traductores que facilitarían su obtención de respuestas correctas en los ejercicios, dejando muy poco que ver sobre sus habilidades de comprensión auditiva y lectora, y el uso de gramática. En el caso de los mini exámenes, se descartaron pues ninguno evaluaba las tres habilidades; había casos en que sí evaluaban gramática, pero no comprensión lectora ni comprensión auditiva, o viceversa, el punto es que ninguno evaluaba las tres habilidades en conjunto para estos mini exámenes.

Tras haber terminado la aplicación de pilotaje, se hicieron las modificaciones pertinentes a los instrumentos, cuya aplicación definitiva se dio durante los cursos extensivos del semestre Agosto-Diciembre 2016 en la UAA.

3.3.7 Aplicación Definitiva. La información fue recabada de la siguiente manera:

Para dar respuesta a la primera pregunta de investigación, relacionada con la autopercepción que los estudiantes tienen sobre su competencia en las habilidades y sub-habilidad mencionadas anteriormente, se aplicó el Inventario para la Autopercepción de Habilidades en Inglés durante las primeras dos semanas del curso extensivo, de manera que los estudiantes aún no hubieran desarrollado ningún conocimiento significativo sobre el nivel al que fueron colocados. Los estudiantes fueron previamente seleccionados a partir del muestreo aleatorio simple.

Se aprovechó para pedirle al maestro que mantenga un registro sobre las actividades donde evalúe la comprensión lectora y auditiva, así como la estructura de oraciones.

Una vez analizados los resultados de la aplicación del instrumento, se dejaron pasar tres semanas más hasta que el segundo instrumento pudo ser aplicado. De acuerdo con Papageorgio y

Cho (2014), cuatro semanas es el tiempo adecuado para que el docente construya un juicio sobre los estudiantes que pertenecen y sobre aquellos que no pertenecen al nivel que se está impartiendo. Se volvió con los grupos que fueron seleccionados y se le pidió al maestro completar la Encuesta para el Juicio Docente, la cual fue contestada con respecto de una muestra pequeña de estudiantes por cada grupo.

Una vez que los datos fueron analizados, se procedió al último paso de la investigación, en que se le solicitó a los profesores evidencias sobre algunos de sus estudiantes, las cuales fueron limitadas a la recuperación de exámenes intermedios, mismos que se aplican al término de las primeras tres unidades del curso. Estos exámenes fueron valorados por medio del uso de la Rúbrica para la Evaluación de Producciones Académicas.

En conjunto, los resultados que se obtuvieron de cada instrumento, y una vez que fueran puestos a un análisis de correlación con la colocación obtenida a partir del examen PLET, sirvieron para dar respuestas a las preguntas de investigación, lo cual ayudó a cumplir con el objetivo de la investigación, referido a la aportación de evidencias de validez basadas en relación con otras variables con que cuenta el PLET.

El análisis realizado fue de asociación entre dos variables, pues se pretendió encontrar relaciones entre los puntajes y la colocación del PLET, los juicios emitidos por maestros y la percepción de los estudiantes sobre su colocación y desempeño en el idioma inglés.

3.3.8 Plan de Análisis

3.3.81 Encuesta para la Autopercepción de Habilidades en Inglés. Para la Encuesta para la Autopercepción de Habilidades en Inglés se corrieron análisis de fiabilidad en SPSS. También

se sacaron frecuencias de respuesta para el reporte. Para visualizar los porcentajes obtenidos en cada reactivo de la encuesta, vaya al Anexo 1.

En esta misma encuesta se sacaron puntajes utilizando el Modelo de Rasch, de manera que el análisis de correlación con el puntaje de PLET pudiera realizarse.

3.3.8.2 Modelo de Rasch. El Modelo de Rasch es un modelo de medida desarrollado por el matemático Georg Rasch en 1960. El Modelo de Rasch nace como una alternativa a la Teoría Clásica de los Test (TCT), pues a diferencia de éste, el primero permite estimar el nivel del constructo de interés a partir del patrón de respuesta de cada sustentante. (Rojas & Montero, 2012). De acuerdo con Prieto y Delgado (2013), el modelo encuentra su fundamentación en distintos supuestos tales como que:

“1. El atributo que se desea medir puede representarse en una única dimensión en la que se situarían conjuntamente las personas y los ítems.” y “2. El nivel de la persona en el atributo y la dificultad del ítem determinan la probabilidad de que la respuesta sea correcta.” (p. 94)

De esta forma, el análisis de resultados de una prueba obtiene beneficios a nivel interpretativo y práctico.

La ecuación que Rasch utiliza es la siguiente:

$$\ln (P_{is} / 1 - P_{is}) = (\theta_s - \beta_i)$$

Donde $(P_{is} / 1 - P_{is})$ representan la posibilidad de ambas, una respuesta correcta y una respuesta incorrecta, en función de la diferencia del constructo tanto en el nivel que posee la persona (θ_s) y en el nivel del reactivo (β_i) . De esta forma, si el nivel de la persona es mayor al del reactivo, la posibilidad de que la respuesta sea correcta será evidentemente mayor, y lo

TESIS TESIS TESIS TESIS TESIS

opuesto sucede si el nivel del reactivo es mayor al de la persona, entonces la probabilidad mayor será la de dar una respuesta incorrecta.

Así, esta fórmula permite establecer una relación entre la probabilidad de respuesta, el nivel del constructo y de la persona, y la dificultad que posee el reactivo.

Rojas y Montero (2012) rescatan cuatro propiedades del Modelo, las cuales se enlistan y explican a continuación:

1. Medición Conjunta: Básicamente, el modelo da la posibilidad de poner en una escala los niveles tanto de los reactivos como del constructo a medir, haciendo posible la relación entre los niveles de dificultad de los reactivos y el nivel de desempeño de la persona, todo en función del o los constructos que se desean explorar.

2. Objetividad Específica: No hay una dependencia de las medidas de niveles tanto en reactivos como en personas en las condiciones en que se han obtenido. Es decir, no hay variación en la dificultad de los reactivos en función de los sustentantes, pues esa dificultad no se determina en base a quien contesta el reactivo, sino en el nivel mismo del reactivo y cómo mide el constructo de interés.

3. Propiedades de Intervalo: La probabilidad de respuestas correctas e incorrectas es una diferencia constante entre reactivos y personas.

4. Especificidad del Error Típico de Medida: Dentro de los niveles del constructo a explorar se permiten diferentes errores en función de esos mismos niveles, pues la prueba no tiene una fiabilidad constante en todo el constructo sometido a medición.

Posterior a la conversión de puntajes en la escala de la encuesta se procedió a realizar una segmentación en la misma, esto para diferenciar a los estudiantes encuestados que sí presentaron el PLET de aquellos que no. Una vez que se tuvo la segmentación se hicieron pruebas de normalidad para determinar el tipo de correlación a usar en el análisis y de esta forma se corrieron los análisis.

3.3.8.3 Encuesta para el Juicio Docente. Para este instrumento también se corrió un análisis de fiabilidad y también se sacaron porcentajes de respuesta que ayudaron a la construcción del apartado de análisis descriptivo de este documento referente a la encuesta a la que este párrafo refiere.

En esta encuesta también se procedió a realizar una segmentación en la misma, esto para diferenciar aquellos estudiantes valorados por el docente que sí presentaron el PLET de aquellos que no. Una vez que se tuvo la segmentación se hicieron pruebas de normalidad para determinar el tipo de correlación a usar en el análisis y de esta forma se corrieron los análisis.

3.3.8.4 Rúbrica de Evaluación de Desempeño. De igual forma que con los dos instrumentos anteriores, se corrió un análisis de fiabilidad para los resultados obtenidos por medio de la rúbrica y para el análisis descriptivo también se sacaron porcentajes de respuesta que serán atendidos en el siguiente apartado.

Aquí también se realizó una segmentación en el archivo de SPSS para diferenciar los resultados de estudiantes que sí presentaron el PLET de aquellos que no. Posterior a esto se realizó una prueba de normalidad que determinó el tipo de correlación a utilizar.

Capítulo 4. RESULTADOS

A continuación se presenta el análisis descriptivo de los resultados obtenidos a partir del trabajo de campo de la presente investigación. Los resultados fueron obtenidos a partir de tres momentos; el primero fue la aplicación de la Encuesta para la Autopercepción de Habilidades en Inglés a estudiantes de 12 grupos que van del nivel 1 al nivel 8 de los cursos de inglés ofrecidos por el PILE; el segundo momento fue la aplicación de la Encuesta para el Juicio Docente a profesores de esos 12 grupos; y por último, se recolectaron producciones académicas que constaron de exámenes intermedios.

Como se menciona en párrafos anteriores en la introducción de este capítulo, este primer acercamiento al análisis de los datos fue descriptivo, y se enfoca en analizar las medias y porcentajes de los valores de respuesta tanto en el instrumento aplicado a los estudiantes como el instrumento aplicado a profesores.

Lo anterior fue hecho con la intención de identificar la tendencia de respuestas y a su vez conocer de manera global la autopercepción del estudiante sobre su colocación, el juicio hecho por el docente con respecto a la misma, y los resultados en los exámenes intermedios.

A continuación se presentan los porcentajes encontrados en cada encuesta tanto por habilidad como por nivel.

4.1 Resultados de Encuesta para la Autopercepción de Habilidades en Inglés

Las tablas 8, 9, y 10, que se presentan a continuación, contienen información sobre los promedios de respuesta que obtuvieron cada paquete de reactivos por nivel en cada una de las ocho encuestas. Cabe hacer la aclaración que estas tablas presentan el promedio de los

porcentajes de respuesta obtenidos en las encuestas para las etiquetas “De acuerdo” y “Muy de acuerdo”.

En otras palabras, lo que aquí se presenta es el nivel de acuerdo que los estudiantes mostraron con las actividades que se describían en cada uno de los reactivos por habilidad y por encuesta. Como ya se mencionó en el párrafo anterior, los porcentajes que se pueden apreciar son en realidad promedios de los porcentajes de respuesta por reactivo en cada una de las encuestas. Para conocer el porcentaje por reactivo puede referirse al Anexo 1.

De esta forma, y por mencionar un ejemplo, en la Tabla 8, en la fila “N 1 y Nivel MCER A1.1” se presentan el promedio final de los porcentajes de respuesta obtenidos para los reactivos de comprensión lectora que describen actividades que se espera una persona con nivel 1 de PILE (equivalente a nivel A1.1 del MCER) pueda llevar a cabo. El porcentaje es entonces el grado de acuerdo que cada grupo (ordenado en cada una de las filas de la tabla) siente con respecto de poder llevar a cabo esas actividades.

4.1.1 Comprensión lectora. Con respecto de la habilidad de comprensión lectora, los resultados obtenidos de la Encuesta para la Autopercepción de Habilidades en Inglés muestran diferentes tendencias en cada nivel.

En el nivel uno 94% de los estudiantes encuestados no sólo muestran acuerdo con llevar a cabo actividades de esta habilidad pertenecientes al nivel que cursan, sino que también muestran un alto acuerdo (de 82.18%) con respecto de las actividades pertenecientes a un nivel arriba. Incluso en el nivel tres y cuatro un 55% muestran un mayor porcentaje de acuerdo que de desacuerdo con la capacidad para realizar actividades de tres niveles arriba del propio, aunque los porcentajes mostrados no sean tan altos. A partir del nivel cinco los estudiantes comienzan a

mostrar indecisión y desacuerdo (en más de 65%) con poder llevar a cabo las actividades descritas en la encuesta.

En el nivel dos sucede algo similar que en el nivel uno, pues los estudiantes muestran acuerdo con la realización de actividades propias del nivel que se cursa (a un 78.04%), así como en actividades pertenecientes a dos niveles superiores (nivel tres a 67.50% y nivel cuatro a 79.10%). Es a partir del nivel cinco que la indecisión y el desacuerdo se hacen presentes en porcentajes superiores al 50%.

Para estudiantes de nivel tres el acuerdo para realizar actividades pertenecientes al nivel que se cursa es de alrededor de 75%, mientras que para actividades de niveles inferiores es superior al 90%. Lo que sucede con el nivel cuatro, inmediato superior al nivel que se cursa, es que también presenta un nivel de acuerdo alto, de 74.07%. A diferencia de la claridad con que se notaba en las respuestas de las encuestas de nivel uno y dos, aquí el porcentaje de lo que opinan los estudiantes sobre la capacidad que perciben en sí mismos para realizar actividades pertenecientes incluso a dos niveles superiores, ya no parece ser tan clara aunque sigue presente, pues un 50.00% de los encuestados de este nivel cree poder realizar aquellas actividades de nivel cinco, mientras que el otro 50.00% tiende al desacuerdo y la indecisión.

En el nivel cuatro los estudiantes presentan acuerdo a más del 85% en niveles inferiores al que se cursa, mientras que para el propio nivel se tiene un 84.35%. También en este nivel los estudiantes parecen percibirse capaces de realizar actividades a un nivel superior al suyo, donde el acuerdo es de un 66.33%. Lamentablemente la impresión de los reactivos de esta habilidad tuvo algunos problemas y los ítems de dos niveles superiores o más no aparecieron en la impresión.

También en el nivel cinco el nivel de acuerdo que los estudiantes presentan con respecto de la capacidad para realizar actividades de niveles inferiores al suyo supera el 85%, pues incluso en los tres primeros niveles se puede apreciar un 100% de acuerdo. Con respecto del nivel que se cursa existe un 73.33% de acuerdo, e incluso para el nivel seis, el inmediato superior, se presenta un 76.67%. En la impresión de esta encuesta tampoco fue posible evaluar dos niveles o más superiores al que se cursa.

En el caso del nivel seis, así como ha pasado en niveles anteriores, el nivel de acuerdo para la capacidad que el estudiante percibe en sí para la realización de actividades pertenecientes a niveles inferiores, es bastante claro, pues en esta encuesta los resultados permiten apreciar que los cuatro primeros niveles (dos o más inferiores al que se cursa) tienen más del 90% de acuerdo, mientras que el inmediato inferior, el nivel cinco, tiene un 77%. Sobre el nivel que se cursa, el acuerdo es de 66%, y aquí también se puede apreciar que los estudiantes se colocan un nivel superior al cursado; en este caso el acuerdo es de 62%. El reactivo que evalúa el nivel ocho, que en este caso sería dos niveles superiores al que se cursa, no apareció en la impresión de la encuesta.

En el nivel siete se puede apreciar que para los primeros cuatro niveles, el porcentaje de acuerdo con la realización de actividades pertenecientes a ese nivel es de 100%. El nivel que funge como inmediato inferior en este caso, el nivel seis, presenta un acuerdo de 80.00%, mientras que en el nivel propio es de 78%. La indecisión y desacuerdo se reparten de igual forma con el acuerdo en un nivel superior.

En el nivel ocho, el nivel de acuerdo en actividades de niveles inferiores es superior al 90% (es del 100% en cinco niveles inferiores, de hecho), mientras que el porcentaje para el nivel que cursan es del 81.43%.

Tabla 8

Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de comprensión lectora.

Habilidad: Comprensión Lectora	Grupos del PILE encuestados							
Reactivos agrupados por nivel de MCER	N1 n = 9	N2 n = 24	N3 n = 26	N4 n = 49	N5 n = 10	N6 n = 20	N7 n = 10	N8 n = 10
Nivel MCER A1.1	94%	92%	96%	98%	100%	90%	100%	100%
Nivel MCER A1.2	82%	78%	90%	95%	100%	95%	100%	100%
Nivel MCER A2.1	55%	67%	76%	87%	100%	95%	100%	100%
Nivel MCER A2.2	55%	79%	74%	84%	65%	95%	100%	100%
Nivel MCER B1.1	22%	38%	50%	66%	73%	78%	60%	90%
Nivel MCER B1.2	33%	46%	46%	*	77%	66%	80%	100%
Nivel MCER B2.1	33%	21%	19%	*	*	62%	78%	93%
Nivel MCER B2.2	11%	21%	15%	*	*	*	50%	81%

Nota.

*Información no disponible debido a que los reactivos pertenecientes a estos niveles no aparecieron en la impresión de las encuestas para niveles 4, 5, y 6 de PILE.

En suma, es posible decir que la comprensión lectora es una habilidad en la que los estudiantes se sienten con una notoria confianza al momento de llevar a cabo actividades donde se trabaje con esta habilidad. Es tal su percepción que en todos los niveles los estudiantes se perciben como capaces de lograr actividades de un nivel superior, y en algunos casos hasta de dos niveles superiores al que cursan.

4.1.2 Comprensión auditiva. Así como en la sección de la Encuesta para la Autopercepción de Habilidades en Inglés que atiende a la comprensión lectora, la información obtenida sobre la comprensión auditiva muestra diferentes tendencias en cada nivel.

En el nivel uno los estudiantes muestran un nivel de acuerdo de 87.48%, mientras que para el inmediato superior es de 71.40%. En los reactivos del nivel tres, que funge aquí como dos niveles superiores al cursado, el acuerdo es apenas del 50%. Para niveles superiores, el nivel de indecisión y desacuerdo es superior al 60%.

Para el nivel dos, se presentan resultados similares al nivel uno. El nivel de acuerdo para la realización de actividades pertenecientes al nivel es de 73%, mientras que para el nivel inmediato superior es de alrededor de 60%. Algo interesante sucede aquí, pues si bien la indecisión y el desacuerdo se presentan en casi un 55% en el nivel cuatro (dos niveles superiores al cursado), el acuerdo vuelve a aparecer en un 54% en las actividades de nivel cinco. Se retoma la tendencia por la indecisión y el desacuerdo a partir del nivel seis.

En el nivel tres los estudiantes se perciben como capaces de realizar actividades pertenecientes a su nivel e inferiores con un acuerdo mayor al 64%. Par el inmediato superior, así

TESIS TESIS TESIS TESIS TESIS

como para el nivel que funge como dos niveles superiores, es del 57%, mientras que la indecisión y el desacuerdo empiezan a ser tendencia a partir del nivel seis.

Para el nivel cuatro, el nivel de acuerdo para la realización de actividades pertenecientes al nivel cursado e inferiores es superior al 70%, y sigue siendo así hasta un nivel superior, donde se presenta un 75.12% de acuerdo. Incluso en el nivel seis hay un 57% de acuerdo con la capacidad por realizar actividades de ese nivel. Es hasta el nivel siete donde la indecisión y el desacuerdo se hacen presentes.

En el nivel cinco, si bien los porcentajes en el nivel de acuerdo para la realización de actividades pertenecientes al nivel siete y ocho no supera el 64% en ambos casos, en todos los niveles el acuerdo fue tendencia clara, resultando en la conclusión de que los estudiantes que cursan el nivel cinco se perciben como capaces no sólo de realizar actividades de su nivel e inferiores, sino también de un nivel o más superiores al propio.

En el nivel seis los estudiantes se perciben como capaces de realizar actividades pertenecientes a su nivel y niveles inferiores, mientras que la indecisión y el desacuerdo se reparten de igual forma que en el acuerdo para el caso de actividades pertenecientes al nivel inmediato superior. En el nivel ocho, la indecisión y el desacuerdo son bastante claros, con un 70%.

Para el caso del nivel siete, los estudiantes muestran un claro acuerdo con poder realizar actividades pertenecientes a niveles inferiores, pues es superior al 75% en todos los casos, mientras que el nivel de acuerdo para actividades del nivel propio es de 68%. La indecisión en el nivel inmediato superior es de 63.33%.

En el nivel ocho, como era de esperarse, el nivel de acuerdo con la realización de actividades propias del nivel y niveles inferiores es claro, pues supera el 70% en todos los casos y en la mayoría supera el 90% incluso.

Tabla 9

Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de comprensión auditiva.

Habilidad: Comprensión Auditiva	Grupos del PILE encuestados							
Reactivos agrupados por nivel de MCER	N1 n = 9	N2 n = 24	N3 n = 26	N4 n = 49	N5 n = 10	N6 n = 20	N7 n = 10	N8 n = 10
Nivel MCER A1.1	87%	92%	92%	94%	100%	90%	100%	100%
Nivel MCER A1.2	71%	73%	79%	72%	95%	88%	100%	100%
Nivel MCER A2.1	50%	58%	64%	77%	75%	70%	85%	90%
Nivel MCER A2.2	33%	46%	57%	72%	92%	78%	95%	100%
Nivel MCER B1.1	39%	54%	57%	75%	88%	78%	85%	75%
Nivel MCER B1.2	33%	42%	42%	56%	72%	55%	76%	90%
Nivel MCER B2.1	27%	42%	28%	44%	58%	50%	68%	81%
Nivel MCER B2.2	17%	29%	8%	28%	63%	30%	37%	72%

La comprensión auditiva, así como la comprensión lectora, parece ser una habilidad en la que los estudiantes se sienten cómodos y con cierta confianza, pues la tendencia en todos los niveles es a percibirse como capaces de realizar actividades que pertenecen a un nivel superior, e

TESIS TESIS TESIS TESIS TESIS

incluso, en algunos niveles los estudiantes se perciben como capaces de realizar actividades de dos niveles o superiores.

4.1.3 Gramática. La habilidad gramatical también presenta ciertas tendencias por nivel, aunque difiere de aquellas que fue posible apreciar en los reactivos de comprensión lectora y comprensión auditiva.

En el nivel uno, los estudiantes se sienten capaces de realizar actividades de su propio nivel e incluso un nivel superior, aunque el porcentaje de acuerdo sea apenas del 55%. Para actividades de nivel tres o superiores, el desacuerdo y la indecisión se presentan como una tendencia clara de respuesta.

En el nivel dos, la indecisión y el desacuerdo son una tendencia clara en todos los niveles, incluso en el nivel propio, pues supera el 50% incluso desde el nivel inmediato inferior, por lo que en este caso los estudiantes se perciben como poco o nada capaces de llevar a cabo actividades de su propio nivel e incluso uno inferior.

En el nivel tres las cosas cambian un poco, pues en niveles inferiores al que se cursa, la tendencia por el acuerdo supera el 55%. El nivel de acuerdo para la realización de actividades al nivel que se cursa es baja, de 52%. Para niveles superiores al que se cursa el nivel de indecisión y desacuerdo es superior al 55% y alcanza en algunos casos porcentajes de casi 80%.

Para el nivel cuatro se presenta algo muy diferente a lo que se veía en niveles anteriores, pues la tendencia por el acuerdo no sólo se presenta en el nivel cursado e inferiores, también lo hace en el inmediato superior e incluso tres niveles superiores, en porcentajes que rebasan el 53%. Sólo en el nivel ocho la indecisión y el desacuerdo se presentan como tendencia.

En el nivel cinco los estudiantes se perciben como capaces de realizar actividades pertenecientes a niveles inferiores al que cursan, como sería de esperarse en caso que la colocación del PLET haya sido adecuada. En el caso de los porcentajes que refieren al nivel cursado, se reparten en 50% tanto en la indecisión y el desacuerdo como en el acuerdo. También es interesante ver que en un nivel inmediato superior la tendencia es hacia el desacuerdo y la indecisión en un 59%, pero el acuerdo se retoma en el nivel siete, y lo hace en un 60%, mientras que en el nivel ocho, tres niveles superiores al cursado, la indecisión y el desacuerdo es de un 80%.

Para el nivel seis, los niveles de acuerdo para la realización de actividades que pertenecen al nivel cursado e inferiores superan el 70%, e incluso en un nivel superior el acuerdo sigue siendo tendencia en un 68%. La tendencia cambia a indecisión y desacuerdo hasta el nivel ocho.

En el nivel siete, el acuerdo es tendencia en todos los reactivos tanto de niveles inferiores al cursado como en el nivel superior, de tal forma que los estudiantes se perciben como un nivel superior al que tienen en cuestiones gramaticales.

En el caso del nivel ocho, el acuerdo surge como tendencia de respuesta en casi todos los niveles, pues, curiosamente, en actividades correspondientes al nivel seis, dos niveles inferiores al cursado, la indecisión y el desacuerdo tienen el 60%, pero el acuerdo se retoma como tendencia en el nivel siete y ocho.

Tabla 10

Promedios de porcentajes de respuesta “De acuerdo” y “Muy de acuerdo” por reactivo para niveles de gramática.

Habilidad: Gramática	Grupos del PILE encuestados							
Reactivos agrupados por nivel de MCER	N1 n = 9	N2 n=24	N3 n = 26	N4 n = 49	N5 n = 10	N6 n =20	N7 n = 10	N8 n = 10
Nivel MCER A1.1	54%	47%	76%	76%	70%	95%	90%	100%
Nivel MCER A1.2	56%	42%	59%	73%	60%	80%	80%	70%
Nivel MCER A2.1	22%	33%	52%	63%	70%	80%	70%	60%
Nivel MCER A2.2	33%	45%	43%	67%	60%	70%	70%	60%
Nivel MCER B1.1	11%	17%	21%	59%	50%	76%	60%	50%
Nivel MCER B1.2	11%	17%	38%	53%	41%	73%	64%	40%
Nivel MCER B2.1	33%	46%	38%	70%	60%	68%	67%	87%
Nivel MCER B2.2	22%	13%	21%	31%	20%	40%	63%	68%

De acuerdo con los resultados reportados en las tablas correspondientes a la Encuesta para la Autopercepción de Habilidades en Inglés, en cuanto a cuestiones gramaticales, puede decirse que en general los estudiantes se sienten capaces de realizar actividades que correspondan a su nivel e inferiores, y en algunos niveles incluso se perciben como capaces de realizar actividades de un nivel superior al suyo. Sin embargo, a pesar de que las conclusiones sobre gramática podrían parecer similares a las de comprensión lectora y comprensión auditiva con respecto de que los estudiantes no sólo se sienten capaces de realizar actividades de su nivel e inferiores, sino de un nivel superior incluso, los porcentajes de acuerdo no son tan altos, pues oscilan entre el 50% y el 70% en la mayoría de los casos, de manera que la confianza que sienten en esta sub-habilidad no es igual a la que sienten en las otras dos habilidades.

Los estadísticos de fiabilidad en cada una de las encuestas estuvieron por encima de .9, como se puede apreciar en la siguiente tabla:

Tabla 11

Resultados de fiabilidad en Encuestas para la Autopercepción de Habilidades en Inglés

Estadísticas de fiabilidad para Encuestas de Autopercepción de Habilidades en Inglés	
Encuesta por nivel	Alfa de Cronbach
NIVEL 1	.967
NIVEL 2	.970
NIVEL 3	.931
NIVEL 4	.951
NIVEL 5	.974
NIVEL 6	.951
NIVEL 7	.978
NIVEL 8	.919

4.2 Resultados de la Encuesta para el Juicio Docente

A continuación se presentan los porcentajes de respuesta en la Encuesta para el Juicio Docente y posteriormente una interpretación de los mismos.

4.2.1 Comprensión lectora. En el caso de la comprensión lectora, los resultados de la Encuesta para el Juicio Docente de nivel uno, nivel dos, nivel cuatro, y nivel ocho, nos dicen que la mayoría de los estudiantes de estos niveles están bien colocados, mientras que en la encuestas

de nivel tres, nivel cinco, y nivel seis, los estudiantes, en la opinión de los docentes de esos niveles, deberían de ir a un nivel inferior en esta habilidad. Sólo en el caso del nivel siete la mayoría de los estudiantes son percibidos por el docente como capaces de realizar actividades de un nivel superior al suyo.

Tabla 12

Porcentajes de respuesta para Comprensión Lectora en las Encuestas para el Juicio

Docente.

Habilidad: Comprensión Lectora	Grupos							
	N1 n = 4/9	N2 n = 8/24	N3 n = 8/26	N4 n = 12/49	N5 n = 4/10	N6 n = 4/20	N7 n = 4/10	N8 n = 4/10
Dos o más niveles inferiores al esperado	0%	12.5%	0%	12.5%	0%	0%	0%	0%
Un nivel inferior al esperado	0%	25%	50%	37.5%	75%	75%	0%	0%
De acuerdo al nivel esperado	75%	62.5%	37.5%	50%	25%	25%	0%	75%
Un nivel superior al esperado	25%	0%	12.5%	0%	0%	0%	75%	25%
Dos o más niveles superiores al esperado	0%	0%	0%	0%	0%	0%	25%	0%

4.2.2 Comprensión auditiva. En los resultados reportados con respecto del juicio que los docentes tienen sobre sus estudiantes y su desempeño y habilidades de comprensión auditiva nos dice que en el nivel uno y ocho, la mayoría de estudiantes podrían estar un nivel arriba al que se

curso. En los niveles, dos, tres, cuatro, y siete, los docentes perciben a la mayoría de sus estudiantes como pertenecientes al nivel de comprensión auditiva del nivel que se cursa. Es en el nivel cinco donde la mayoría de los estudiantes son percibidos por sus docentes como pertenecientes a un nivel inferior. En el nivel seis la mayoría son percibidos como pertenecientes a dos niveles o más inferiores al esperado.

Tabla 13

Porcentajes de respuesta para Comprensión Auditiva en las Encuestas para el Juicio Docente.

Habilidad: Comprensión Auditiva	Grupos							
	N1 n = 4/9	N2 n = 8/24	N3 n = 8/26	N4 n = 12/49	N5 n = 4/10	N6 n = 4/20	N7 n = 4/10	N8 n = 4/10
Dos o más niveles inferiores al esperado	0%	12.5%	12.5%	12.5%	25%	50%	0%	0%
Un nivel inferior al esperado	0%	12.5%	12.5%	37.5%	50%	25%	25%	0%
De acuerdo al nivel esperado	25%	75%	62.5%	50%	25%	25%	50%	25%
Un nivel superior al esperado	75%	0%	12.5%	0%	0%	0%	25%	50%
Dos o más niveles superiores al esperado	0%	0%	0%	0%	0%	0%	0%	25%

4.2.3 Gramática. En el caso de la sub-habilidad gramatical, los resultados nos dicen que en los niveles dos, tres, cuatro y ocho, la mayoría de los estudiantes sobre los que los docentes

emitieron su juicio están en el nivel adecuado, mientras que en los niveles uno y cinco, los docentes creen que la mayoría de sus estudiantes necesitan ayuda de un nivel inferior. En el nivel seis la mayoría de los estudiantes son percibidos como pertenecientes a dos niveles inferiores o más en cuestiones gramaticales. Para el nivel siete, los resultados se reparten de igual forma en nivel inferior al esperado, de acuerdo al nivel esperado, un nivel superior al esperado, y dos niveles superiores o más al esperado.

Tabla 14

Porcentajes de respuesta para Gramática en las Encuestas para el Juicio Docente.

Habilidad: Gramática	Grupos							
	N1 n = 4/9	N2 n = 8/24	N3 n = 8/26	N4 n = 12/49	N5 n = 4/10	N6 n = 4/20	N7 n = 4/10	N8 n = 4/10
Etiquetas								
Dos o más niveles inferiores al esperado	0%	12.5%	12.5%	12.5%	25%	50%	0%	0%
Un nivel inferior al esperado	0%	25%	0%	25%	50%	25%	25%	0%
De acuerdo al nivel esperado	50%	62.5%	87.5%	62.5%	25%	25%	25%	75%
Un nivel superior al esperado	50%	0%	0%	0%	0%	0%	25%	0%
Dos o más niveles superiores al esperado	0%	0%	0%	0%	0%	0%	25%	25%

4.2.4 Valoración global. En el juicio global que los docentes emitieron sobre la colocación de sus estudiantes, en el caso de los niveles dos, tres, y siete, consideran que la mayoría de sus estudiantes fueron colocados apropiadamente. En el nivel uno se considera que la

mitad fueron colocados adecuadamente y la otra mitad podría estar en un nivel superior. En el nivel cuatro se considera que la mayoría de los estudiantes deberían de estar en un nivel superior. Para los niveles cinco y ocho, la consideración de los docentes es de que la mayoría de los estudiantes estarían bien en un nivel inferior al asignado, mientras que sólo en el nivel seis se presenta la situación de que el docente considera que la mayoría de los estudiantes deberían haber sido colocados dos niveles o más inferiores en el que están cursando.

Tabla 15

Porcentajes de respuesta para la Valoración Global en las Encuestas para el Juicio Docente.

Valoración Global	Grupos							
	N1 n = 4/9	N2 n = 8/24	N3 n = 8/26	N4 n = 12/49	N5 n = 4/10	N6 n = 4/20	N7 n = 4/10	N8 n = 4/10
Dos o más niveles inferiores al esperado	0%	0%	12.5%	8.3%	25%	50%	0%	25%
Un nivel inferior al esperado	0%	25%	12.5%	8.3%	50%	25%	25%	50%
De acuerdo al nivel esperado	50%	75%	62.5%	50%	25%	25%	50%	25%
Un nivel superior al esperado	50%	0%	12.5%	66.7%	0%	0%	25%	0%
Dos o más niveles superiores al esperado	0%	0%	0%	33.3%	0%	0%	0%	0%

4.3 Resultados de la Rúbrica de Evaluación de Desempeño

Para los resultados recabados por medio del tercer instrumento, se trabajó con ocho rúbricas distintas. Se analizaron nueve paquetes de “Exámenes Intermedios” (un examen que se aplica al cabo de un mes y medio de trabajo donde se evalúan todas las habilidades del idioma inglés aplicadas a contenidos de las primeras tres unidades del curso). Cada uno de los nueve paquetes analizados proviene de un grupo distinto. Se esperaba trabajar con 12 paquetes de exámenes pero tres de ellos no fueron entregados (un grupo de nivel dos, otro de nivel tres, y el grupo de nivel ocho). Se analizaron alrededor de cuatro casos por nivel en cada una de las tres habilidades de interés en este estudio (comprensión auditiva, comprensión lectora, y gramática), a excepción del nivel tres donde sólo se analizaron tres casos, y del nivel cuatro, donde, por la cantidad de grupos, se trabajó con 11 casos.

Tabla 16

Distribución de exámenes intermedios analizados con la rúbrica.

Nivel	Número de casos analizados con la rúbrica	Versiones de examen
1	4	1
2	4	1
3	3	1
4	11	2
5	4	1

6	4	1
7	4	1
8	No se entregó ningún examen de este nivel.	

4.3.1 Comprensión lectora. En la sección designada para evaluar el desempeño del estudiante en la habilidad de comprensión lectora, se contó en promedio con 15 reactivos que se distribuían en diferentes tipos de tarea, que iban desde la identificación de detalles y la identificación de información específica, hasta la identificación de ideas principales en un texto. La rúbrica ayudó a identificar qué tipo de actividad o tarea evaluaba cada reactivo de la sección, así como el nivel de demanda en cada uno de ellos, encontrando que todos los reactivos pertenecían al nivel que los estudiantes cursaban, pues en su totalidad eran preguntas y textos disponibles en el material exclusivo del profesor, provisto por la editorial de los libros con que se trabaja.

Con respecto de los reactivos considerados como “apropiados para el nivel y habilidad”, se encontró que la calificación promedio dentro de todos los casos fue de 7.96.

Con respecto a reactivos denominados como “adecuados para el nivel y habilidad pero de mayor demanda cognitiva” el promedio de calificación obtenida fue de 5.86. Cabe señalar que de los 34 casos, fueron en 20 de ellos donde se encontraron este tipo de reactivos. La desviación estándar, justo como en la comprensión auditiva, está por encima del tipo de reactivos reportados en el párrafo anterior.

En promedio general para la habilidad de comprensión lectora, los estudiantes tienen una calificación de 7.07, lo que permite apreciar que la muestra tiene un desempeño esperado de acuerdo al nivel al que fueron colocados, pues estamos hablando de que los estudiantes alcanzan una calificación del 70% de respuestas correctas en sus exámenes respecto de esta sub-habilidad del idioma inglés.

En la Tabla 17 se puede apreciar que un 97% de la muestra es capaz de realizar actividades de su nivel en cuanto a comprensión lectora, de acuerdo con el desempeño visto en la sección del examen intermedio que atiende a esta habilidad.

Tabla 17

Porcentajes de alumnos que pueden realizar actividades de Comprensión Lectora de su nivel.

Comprensión Lectora	Grupos							
	N1	N2	N3	N4	N5	N6	N7	N8
Etiquetas	n = 4/9	n = 4/24	n = 3/26	n = 11/49	n = 4/10	n = 4/20	n = 4/10	n = 0/10
No es capaz de realizar actividades de su nivel	0%	0%	0%	18.2%	0%	0%	0%	
Es capaz de realizar actividades de su nivel	100%	100%	100%	81.8%	100%	100%	100%	
<p>Nota</p> <p>La forma en que se asignaron estos porcentajes a cada etiqueta fue sacando un promedio de aciertos en la sección de comprensión lectora. Posterior a esto, se determinaba que quien obtuviera un promedio >5 + 1 entraba en “Es capaz de realizar actividades de su nivel”.</p>								

4.3.2 Comprensión auditiva. En la sección encargada de evaluar el desempeño del estudiante en la habilidad de comprensión auditiva, se contaba en promedio con 15 reactivos que se distribuían en diferentes tipos de tarea, que iban desde la identificación de detalles y la identificación de información específica, hasta la identificación de ideas principales en un audio. La rúbrica ayudó a identificar qué tipo de actividad o tarea evaluaba cada reactivo de la sección, así como el nivel de demanda en cada uno de ellos, encontrando que todos los reactivos pertenecían al nivel que los estudiantes cursaban, pues en su totalidad eran preguntas y audios disponibles en el material exclusivo del profesor, provisto por la editorial de los libros con que se trabaja.

Con respecto de los reactivos considerados como “apropiados para el nivel y habilidad”, se encontró que la calificación promedio dentro de todos los casos fue de 7.8 en una escala de 0 a 10.

Con respecto a reactivos denominados como “adecuados para el nivel y habilidad pero de mayor demanda cognitiva” presentes en 11 de los 34 casos, el promedio de calificación obtenida fue de 4.99.

En promedio general para la habilidad de comprensión lectora, los estudiantes tienen una calificación de 7.38, lo que permite apreciar que la muestra tiene un desempeño esperado de acuerdo al nivel al que fueron colocados, pues estamos hablando de que los estudiantes alcanzan una calificación de más del 70% de respuestas correctas en sus exámenes respecto de la habilidad de comprensión auditiva.

Hablando en porcentaje, es posible apreciar en la Tabla 19 que casi el 90% de la muestra es capaz de llevar a cabo actividades de comprensión auditiva de su nivel, de acuerdo con el

desempeño visto en el examen intermedio. Cabe señalar que los porcentajes más altos se encuentran en el nivel cuatro y superiores.

Tabla 18

Porcentajes de alumnos que pueden realizar actividades de Comprensión Auditiva de su nivel.

Comprensión Auditiva	Grupos							
	N1	N2	N3	N4	N5	N6	N7	N8
Etiquetas	n = 4/9	n = 4/24	n = 3/26	n = 11/49	n = 4/10	n = 4/20	n = 4/10	n = 0/10
No es capaz de realizar actividades de su nivel	25%	25%	33.3%	0%	0%	0%	0%	
Es capaz de realizar actividades de su nivel	75%	75%	66.7%	100%	100%	100%	100%	
Nota	La forma en que se asignaron estos porcentajes a cada etiqueta fue sacando un promedio de aciertos en la sección de comprensión auditiva. Posterior a esto, se determinaba que quien obtuviera un promedio >5 + 1 entraba en “Es capaz de realizar actividades de su nivel”.							

4.3.3 Gramática. En la sección diseñada para evaluar el desempeño del estudiante en la sub-habilidad de gramática el promedio de ejercicios era de 20, mismos que se distribuían en diferentes tipos de tarea que iban desde la elección de una opción de respuesta para completar una oración, la conjugación de verbos dependiendo del tiempo gramatical, el ordenamiento de oraciones, hasta la estructuración de distintos tipos oraciones. La rúbrica ayudó a identificar qué tipo de actividad o tarea evaluaba cada reactivo de la sección, así como el nivel de demanda en

TESIS TESIS TESIS TESIS TESIS

cada uno de ellos, encontrando que todos los reactivos pertenecían al nivel que los estudiantes cursaban, pues en su mayoría eran ejercicios disponibles en el material exclusivo del profesor, provisto por la editorial de los libros con que se trabaja. Una pequeña cantidad de reactivos fue diseñada por algunos profesores, pero aun así respetaban los contenidos y el tipo de actividad adecuada para el nivel.

Con respecto de los reactivos considerados como “apropiados para el nivel y habilidad”, se encontró que la calificación promedio dentro de todos los casos fue de 7.07.

Con respecto a reactivos denominados como “adecuados para el nivel y habilidad pero de mayor demanda cognitiva”, presentes en 23 de los 34 casos, el promedio de calificación obtenida fue de 6.59.

En promedio general para la sub-habilidad de gramática, los estudiantes tienen una calificación de 6.93, lo que permite apreciar que la muestra tiene un desempeño esperado de acuerdo al nivel al que fueron colocados, pues estamos hablando de que los estudiantes alcanzan una calificación casi del 70% de respuestas correctas en sus exámenes respecto de esta sub-habilidad del idioma inglés.

En la Tabla 18 se puede observar que cerca de un 87% de la muestra es capaz de llevar a cabo actividades de gramática de su nivel, de acuerdo con el desempeño visto en el examen intermedio para esta sección.

Tabla 19

Porcentajes de alumnos que pueden realizar actividades de Gramática de su nivel.

Gramática	Grupos							
Etiquetas	N1	N2	N3	N4	N5	N6	N7	N8
	n = 4/9	n = 4/24	n = 3/26	n = 11/49	n = 4/10	n = 4/20	n = 4/10	n = 0/10
No es capaz de realizar actividades de su nivel	0%	0%	33.3%	9.1%	25%	0%	25%	
Es capaz de realizar actividades de su nivel	100%	100%	66.7%	90.9%	75%	100%	75%	
<p>Nota</p> <p>La forma en que se asignaron estos porcentajes a cada etiqueta fue sacando un promedio de aciertos en la sección de gramática. Posterior a esto, se determinaba que quien obtuviera un promedio $>5 + 1$ entraba en “Es capaz de realizar actividades de su nivel”.</p>								

Al promediar las medias de las calificaciones obtenidas por los estudiantes en cada una de las habilidades, se obtienen los siguientes datos:

Tabla 20

Porcentajes de alumnos que pueden realizar actividades de las tres habilidades de su nivel, de acuerdo con lo visto en cada habilidad.

Global	Grupos							
Etiquetas	N1	N2	N3	N4	N5	N6	N7	N8
	n = 4/9	n = 4/24	n = 3/26	n = 11/49	n = 4/10	n = 4/20	n = 4/10	n = 0/10
No es capaz de realizar actividades de su nivel	0%	0%	66.7%	0%	0%	0%	0%	

Es capaz de realizar actividades de su nivel	100%	100%	33.7%	100%	100%	100%	100%	
--	------	------	-------	------	------	------	------	--

El promedio global de los 34 casos nos dice que tienen un desempeño aceptable para el nivel que cursan los estudiantes. Si el promedio global nos dijera que están sacando un 90% ó 100% de respuestas correctas en el examen, se pueden hacer hipótesis tales como que a) los estudiantes se prepararon adecuadamente para el examen, b) es el curso y la instrucción del docente quien les ayudó a obtener buenas calificaciones, o c) los estudiantes están siendo colocados en niveles que no representan ningún tipo de exigencia para sus habilidades; en otras palabras, están siendo colocados en niveles inferiores al que pertenecen. Por otro lado, si el promedio global nos dejara ver que los estudiantes están sacando sólo un 40% o menos, las hipótesis van en sentido de que a) los estudiantes no se prepararon para el examen, b) el curso y la instrucción de los profesores están resultando inapropiados, o c) los estudiantes están siendo colocados en niveles donde el desempeño requerido dista del desempeño con que cuentan y han llegado al nivel donde fueron colocados.

De esta forma, las evidencias revisadas en el examen intermedio de distintos niveles y grupos, nos permite ver que, en promedio, se cumplen las hipótesis, o al menos alguna de ellas, que explican por qué los estudiantes son capaces de llevar a cabo actividades correspondientes al nivel que cursan, mismas que representan una exigencia adecuada para las habilidades del estudiante.

4.4 Análisis de Correlación

A continuación se presentan los análisis de correlación por cada uno de los instrumentos. Este es un acercamiento a las correlaciones que existen entre el puntaje obtenido en PLET y los

resultados para la Encuesta para la Autopercepción de Habilidades en Inglés, y entre la ubicación en los cursos de PILE, para la Encuesta para el Juicio Docente y la Rúbrica de Evaluación de Desempeño. Si bien en primera instancia se tenía intencionado correlacionar los resultados de los instrumentos con el puntaje que los estudiantes obtuvieron en el PLET, esto sólo se hizo con la Encuesta para la Autopercepción de Habilidades en Inglés, pues la cantidad de la muestra de 158 alumnos y la cantidad de reactivos por encuesta (alrededor de 70) permitió hacer un análisis por medio del Modelo de Rasch, y así sacar un puntaje que fuera comparable con aquél del PLET, pues como Linacre (1994) lo sugiere, “así como el tamaño de una muestra incrementa, las diferencias se hacen más pequeñas. (párr. 4)” De esta forma, se siguió la recomendación de aplicar el Modelo de Rasch (Linacre, 1994) en la Encuesta para la Autopercepción de Habilidades en Inglés, el único instrumento para el que la muestra fue mayor a 100 participantes, cosa que ni en la Encuesta para el Juicio Docente ni en la Rúbrica de Evaluación de Desempeño llegó a cumplirse, pues la muestra para el primero era de 48 estudiantes y la del otro de 34, respectivamente. Lo anterior aunado a que por problemas con la selección de participantes (no todos los encuestados presentaron el PLET), terminó en la decisión de segmentar las bases de datos en dos grupos: aquellos que no presentaron el PLET (asignándoles un valor de .00) y aquellos que sí lo presentaron (cuyo valor asignado fue 1.00).

Para asignar una etiqueta al grado de correlación se utilizaron las sugerencias por rango de Bisquerra (1987), las cuales son: a) Mayor a 0.80: muy alta; b) entre 0.60 y 0.79: alta; c) entre 0.40 y 0.59: moderada; d) entre 0.20 y 0.39: baja; y e) menor a 0.20: muy baja.

4.4.1 Correlación entre Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje PLET

4.4.1.1 Correlación en Comprensión Lectora. De acuerdo con lo visto en la tabla 23 es posible decir que la correlación entre el puntaje en PLET de 69 estudiantes y el puntaje de estos mismos en la encuesta es positiva, al ser de 0.51. Esto nos indica que la asociación entre la percepción que estos 69 estudiantes tienen es moderada en función de sus resultados en el examen PLET.

Tabla 21

Correlación de Pearson entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Comprensión Lectora.

Correlaciones				
Segmentación			Encuesta de Comprensión Lectora	Puntaje de Comprensión Lectora en PLET
1.00	Encuesta de Comprensión Lectora	Correlación de Pearson	1	.512**
		Sig. (bilateral)		.000
		N	69	69
	Puntaje de Comprensión Lectora en PLET	Correlación de Pearson	.512**	1
		Sig. (bilateral)	.000	
		N	69	69

** . La correlación es significativa en el nivel 0,01 (bilateral).

4.4.1.2 Correlación en comprensión auditiva. Debido al valor de la prueba de normalidad (.043), para la correlación se decidió utilizar Rho de Spearman. De acuerdo con lo

visto en la tabla 21 es posible decir que la correlación entre el puntaje en PLET de 69 estudiantes y el puntaje de estos mismos en la encuesta es positiva, al ser de 0.50 en Spearman. Esto nos indica que la asociación entre la percepción que estos 69 estudiantes tiene es moderada considerando los valores en Spearman, en función de sus resultados en el examen PLET.

Tabla 22

Correlación de Spearman entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Comprensión Auditiva.

Correlaciones					
Segmentación			Encuesta de Comprensión Auditiva	Puntaje de Comprensión Auditiva en PLET	
1.00	Rho de Spearman	Encuesta de Comprensión Auditiva	Coeficiente de correlación	1.000	.501**
			Sig. (bilateral)	.	.000
			N	69	69
		Puntaje de Comprensión Auditiva en PLET	Coeficiente de correlación	.501**	1.000
			Sig. (bilateral)	.000	.
			N	69	69
**. La correlación es significativa en el nivel 0,01 (bilateral).					

4.4.1.3 Correlación en gramática De acuerdo con el contenido en la tabla 22 es posible decir que la correlación entre el puntaje en PLET de 67 estudiantes y el puntaje de estos mismos en la encuesta es positiva, al ser de 0.40. Esto nos indica que la asociación entre la percepción que estos 67 estudiantes tiene es moderada en función de sus resultados en el examen PLET.

Tabla 23

Correlación de Pearson entre los resultados de la Encuesta para la Autopercepción de Habilidades en Inglés y el puntaje en el PLET, para la habilidad de Gramática.

Correlaciones				
Segmentación			Encuesta de Gramática	Puntaje de Gramática en PLET
1.00	Encuesta de Gramática	Correlación de Pearson	1	.400**
		Sig. (bilateral)		.001
		N	67	67
	Puntaje de Gramática en PLET	Correlación de Pearson	.400**	1
		Sig. (bilateral)	.001	
		N	67	69
**. La correlación es significativa en el nivel 0,01 (bilateral).				

En relación con lo observado en las tablas 20 a 22, puede decirse que la asociación que existe entre la percepción que los estudiantes tienen con respecto de sus habilidades de comprensión lectora, comprensión auditiva y gramática, y los resultados que obtuvieron en el PLET en estas mismas habilidades, es moderada creciente. La interpretación que puede hacerse a partir de este resultado es que este grado de asociación se debe en general a que una cantidad considerable de los estudiantes que contestaron la encuesta se perciben capaces de llevar a cabo tareas y actividades correspondientes tanto a las tres habilidades anteriormente mencionadas y al nivel al que fueron asignados por el PLET. Podría decirse que si los resultados de correlación no

son más altos se debe a que, de acuerdo a lo que ya se veía en el análisis descriptivo de los resultados de la encuesta, también en promedio eran bastantes los estudiantes que se percibían capaces de llevar a cabo actividades y tareas, de estas tres habilidades, correspondientes a uno e incluso dos niveles superiores al que fueron asignados.

4.4.2 Correlación entre Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en cursos de PILE

La prueba de normalidad sugiere que para la correlación de los resultados en todas las habilidades y la valoración general del nivel de los estudiantes se utilice la correlación de Spearman, al ser valores de significancia por debajo del 0.05.

4.4.2.1 Correlación en comprensión lectora. De acuerdo con lo observado en los resultados de correlación, es posible apreciar que la asociación entre el juicio o valoración que los docentes emitieron sobre 21 alumnos y el nivel de ubicación en cursos de PILE es positiva, con un valor de 0.62. Se considera entonces la asociación entre la opinión que los docentes tienen sobre el desempeño de estos 21 estudiantes en la habilidad de comprensión lectora es alta en función del nivel de ubicación en cursos de PILE como resultado de su puntaje en el PLET.

Tabla 24

Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Comprensión Lectora.

Correlaciones					
Segmentación				Encuesta de Comprensión Lectora	Nivel de ubicación del MCE en Comprensión Lectora
1.00	Rho de Spearman	Encuesta de Comprensión Lectora	Coefficiente de correlación	1.000	.619**
			Sig. (bilateral)	.	.003

			N	21	21
		Nivel de ubicación del MCE en Comprensión Lectora	Coefficiente de correlación	.619**	1.000
			Sig. (bilateral)	.003	.
			N	21	21
**. La correlación es significativa en el nivel 0,01 (bilateral).					

4.4.2.2 Correlación en Comprensión Auditiva. De acuerdo con lo observado en los resultados de correlación, es posible apreciar que la asociación entre el juicio o valoración que los docentes emitieron sobre 21 alumnos y el nivel de ubicación en cursos de PILE es positiva, con un valor de 0.59. Se considera entonces que la asociación entre la opinión que los docentes tienen sobre el desempeño de estos 21 estudiantes en la habilidad de comprensión auditiva es moderada en función del nivel de ubicación en cursos de PILE como resultado de su puntaje en el PLET.

Tabla 25

Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Comprensión Auditiva.

Correlaciones					
Segmentación			Encuesta de Comprensión Auditiva	Nivel de ubicación del MCE en Comprensión Auditiva	
1.00	Rho de Spearman	Encuesta de Comprensión Auditiva	Coefficiente de correlación	1.000	.597**
			Sig. (bilateral)	.	.004
			N	21	21

	Nivel de ubicación del MCE en Comprensión Auditiva	Coefficiente de correlación	.597**	1.000
		Sig. (bilateral)	.004	.
		N	21	21
**. La correlación es significativa en el nivel 0,01 (bilateral).				

4.4.2.3 Correlación en gramática. De acuerdo con lo observado en los resultados de correlación, es posible apreciar que la asociación entre el juicio o valoración que los docentes emitieron sobre 21 alumnos y el nivel de ubicación en cursos de PILE es positiva, con un valor de 0.68. Se considera entonces que la asociación entre la opinión que los docentes tienen sobre el desempeño de estos 21 estudiantes en la habilidad de gramática es alta en función del nivel de ubicación en cursos de PILE como resultado de su puntaje en el PLET.

Tabla 26

Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la habilidad de Gramática.

Correlaciones					
Segmentación			Encuesta de Gramática	Nivel de ubicación del MCE en Gramática	
1.00	Rho de Spearman	Encuesta de Gramática	Coefficiente de correlación	1.000	.680**
			Sig. (bilateral)	.	.001
			N	21	21
	Nivel de ubicación del MCE en	Coefficiente de correlación	.680**	1.000	

	Gramática	Sig. (bilateral)	.001	.
		N	21	21

** . La correlación es significativa en el nivel 0,01 (bilateral).

4.4.2.4 Correlación en la valoración general del desempeño del estudiante. De acuerdo con lo observado en los resultados de correlación, es posible apreciar que la asociación entre el juicio o valoración que los docentes emitieron sobre 21 alumnos y el nivel de ubicación en cursos de PILE es positiva, con un valor de 0.75. Se considera entonces que la asociación entre la opinión que los docentes tienen sobre el desempeño de estos 21 estudiantes en su valoración general del desempeño de los estudiantes es alta en función del nivel de ubicación en cursos de PILE como resultado de su puntaje en el PLET.

Tabla 27

Correlación de Pearson entre los resultados de la Encuesta para el Juicio Docente y el Nivel de Ubicación del MCE en PILE, para la Valoración General del Desempeño del Estudiante.

Correlaciones					
Segmentación			Encuesta General sobre las Tres Habilidades	Nivel de ubicación en los cursos del PILE	
1.00	Rho de Spearman	Encuesta General sobre las Tres Habilidades	Coefficiente de correlación	1.000	.754**
			Sig. (bilateral)	.	.000
			N	21	21
	Nivel de ubicación en los cursos del PILE	Coefficiente de correlación	.754**	1.000	
			Sig. (bilateral)	.000	.
			N	21	21

** . La correlación es significativa en el nivel 0,01 (bilateral).

En relación con los valores de significancia en la correlación de Spearman tanto para cada habilidad como para la valoración general que los docentes hacen sobre los 21 estudiantes es posible observar que la asociación entre la opinión que los docentes tienen, siendo un juicio que han construido con el trabajo y convivencia diaria con sus estudiantes, y la ubicación que les fue dada a partir de su desempeño en el PLET es alta. La interpretación que se puede hacer a partir de estos resultados y de lo visto en el análisis descriptivo es que el juicio del docente parece no distar mucho de la asignación de nivel que se hace a partir del puntaje en el PLET, por lo que podría decirse que una cantidad considerable de los docentes que participaron en el estudio consideran que la mayoría de los estudiantes sobre quienes emitieron una valoración han sido colocados apropiadamente. Esto es incluso más fácil de apreciar si enfocamos nuestra atención a la tabla 27, sobre la valoración general que hacen los docentes, donde el valor de significancia es de 0.75.

4.4.3 Correlación entre Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCE en cursos de PILE

Los niveles de significancia en la prueba de normalidad nos indican que la correlación de Pearson es la indicada para las cuatro variables con que se trabaja en este instrumento, ya que todas son mayores a 0.05.

4.4.3.1 Correlación en comprensión lectora

De acuerdo con el valor encontrado en la correlación de Pearson se puede decir que la asociación entre el desempeño mostrado por los 16 estudiantes en su examen intermedio, en la sección de comprensión lectora, y la ubicación de nivel que les es asignada para esta habilidad es positivo con un valor de 0.82. Lo anterior nos dice que esta asociación entre lo que el estudiante es capaz de hacer con respecto de la comprensión lectora de su nivel y la ubicación a ese mismo

nivel a partir del puntaje obtenido en el PLET, es muy alta para estos 16 estudiantes y lo que demostraron en su sección de comprensión lectora en el examen intermedio de su curso.

Tabla 28

Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Comprensión Lectora.

Correlaciones				
Segmentación			Resultados de Rúbrica para Comprensión Lectora	Nivel de ubicación del MCE en Comprensión Lectora
1.00	Resultados de Rúbrica para Comprensión Lectora	Correlación de Pearson	1	.828**
		Sig. (bilateral)		.000
		N	16	16
	Nivel de ubicación del MCE en Comprensión Lectora	Correlación de Pearson	.828**	1
		Sig. (bilateral)	.000	
		N	16	16
**. La correlación es significativa en el nivel 0,01 (bilateral).				

4.4.3.2 Correlación en comprensión auditiva. De acuerdo con el valor encontrado en la correlación de Pearson se puede decir que la asociación entre el desempeño mostrado por los 16 estudiantes en su examen intermedio, en la sección de comprensión auditiva, y la ubicación de nivel que les es asignada es positivo con un valor de 0.82. Lo anterior nos dice que esta asociación entre lo que el estudiante es capaz de hacer con respecto de la comprensión auditiva

de su nivel y la ubicación a ese mismo nivel es muy alta para estos 16 estudiantes y lo que demostraron en su sección de comprensión auditiva en el examen intermedio de su curso.

Tabla 29

Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Comprensión Auditiva.

Correlaciones				
Segmentación			Resultado de Rúbrica para Comprensión Auditiva	Nivel de ubicación del MCE en Comprensión Auditiva
1.00	Resultado de Rúbrica para Comprensión Auditiva	Correlación de Pearson	1	.825**
		Sig. (bilateral)		.000
		N	16	16
	Nivel de ubicación del MCE en Comprensión Auditiva	Correlación de Pearson	.825**	1
		Sig. (bilateral)	.000	
		N	16	16
**. La correlación es significativa en el nivel 0,01 (bilateral).				

4.4.3.3 Correlación en gramática. De acuerdo con el valor encontrado en la correlación de Pearson se puede decir que la asociación entre el desempeño mostrado por los 16 estudiantes en su examen intermedio, en la sección de gramática, y la ubicación de nivel que les es asignada para esta habilidad es positivo con un valor de 0.39. Lo anterior nos dice que esta asociación entre lo que el estudiante es capaz de hacer con respecto de la gramática de su nivel y la ubicación a ese mismo nivel a partir del puntaje obtenido en el PLET, es baja para estos 16

estudiantes y lo que demostraron en su sección de gramática en el examen intermedio de su curso.

Tabla 30

Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Gramática.

Correlaciones				
Segmentación			Resultado de la Rúbrica para Gramática	Nivel de ubicación del MCE en Gramática
1.00	Resultado de la Rúbrica para Gramática	Correlación de Pearson	1	.397
		Sig. (bilateral)		.128
		N	16	16
	Nivel de ubicación del MCE en Gramática	Correlación de Pearson	.397	1
		Sig. (bilateral)	.128	
		N	16	16

*. La correlación es significativa en el nivel 0,05 (bilateral).

4.4.3.4 Correlación en Valoración General de los Resultados del Examen Intermedio en las Tres Habilidades. De acuerdo con el valor encontrado en la correlación de Pearson se puede decir que la asociación entre el desempeño mostrado por los 16 estudiantes en su examen intermedio, en las tres secciones (Comprensión auditiva, gramática, y comprensión lectora), y la ubicación de nivel que les es asignada para la a partir de su desempeño en el PLET es positiva con un valor de 0.88. Lo anterior nos dice que esta asociación entre lo que el estudiante es capaz

de hacer con respecto de las tres habilidades de su nivel y la ubicación a ese mismo nivel a partir del puntaje obtenido en el PLET, es muy alta para estos 16 estudiantes y lo que demostraron en las tres secciones en el examen intermedio de su curso.

Tabla 31

Correlación de Pearson entre los resultados de la Rúbrica de Evaluación de Desempeño y el Nivel de Ubicación del MCER en PILE, para Valoración General de Capacidad para Actividades del Nivel Asignado.

Correlaciones				
Segmentación			Resultado de Rúbrica para las Tres Habilidades	Nivel de ubicación en los cursos del PILE
1.00	Resultado de Rúbrica para las Tres Habilidades	Correlación de Pearson	1	.885**
		Sig. (bilateral)		.000
		N	16	16
	Nivel de ubicación en los cursos del PILE	Correlación de Pearson	.885**	1
		Sig. (bilateral)	.000	
		N	16	16

** . La correlación es significativa en el nivel 0,01 (bilateral).

Similar a los resultados y valores de correlación en la Encuesta para el Juicio Docente, en lo observado para la Rúbrica de Evaluación de Desempeño, se puede decir que la asociación entre el desempeño que demuestran para las tres habilidades en el examen intermedio de su curso y el desempeño demostrado en el PLET es creciente muy alta, con excepción de gramática, donde es baja. Por esto, puede decirse que lo que los estudiantes fueron capaces de hacer en el

examen intermedio de su curso con respecto de comprensión auditiva y comprensión lectora de su nivel está bastante asociado con el desempeño que muestran en estas habilidades en el PLET. En el caso de gramática, el que la asociación sea positiva baja podría deberse tanto a que el examen PLET está siendo laxo al evaluar esta habilidad como a que se está evaluando con severidad. En la valoración general, también presentando un resultado similar al de la Encuesta para el Juicio Docente, el valor es muy alto y nos dice que los estudiantes son capaces de llevar a cabo actividades correspondientes al nivel que fueron ubicados como resultado de su desempeño en el PLET.

4.4.4 Interpretación de resultados con base en el MCER. Para dar una interpretación más amplia de lo que el análisis de correlación nos quiere decir, se retomará lo expresado en la tabla 1 de este documento sobre los descriptores del MCER para cada nivel.

El argumento principal para el que se aportó evidencia de validez en esta investigación fue el de aseverar que el PLET, como examen de colocación y acreditación, es capaz de ubicar a sus sustentantes de manera apropiada en un nivel de desempeño del lenguaje que refleje su nivel de dominio del idioma inglés lo más cercano a la realidad.

Con sustento en los valores de correlación presentados en apartados anteriores, podemos dar validez a los argumentos referentes a que a) los alumnos colocados en el nivel A1, como usuarios básicos del idioma, aún no cuentan con habilidades del idioma inglés lo suficientemente desarrolladas como para interactuar a un nivel simple con otros hablantes del idioma; b) que los alumnos colocados en el nivel A2, son capaces de comunicarse de formas simples con otros hablantes, alcanzando un grado de interacción que les permite expresar ideas sencillas, así como solicitar información básica para conocer personas y lugares, todo ello por medio del uso de

frases y expresiones que reflejan cotidianeidad; c) que los estudiantes ubicados en el nivel B1, como usuarios independientes, son capaces de expresarse con propiedad en temas cotidianos e incluso son capaces de hacer cambios de tiempos de manera que su comunicación sea más rica en detalles que la de un usuario básico; d) y que los sustentantes colocados en el nivel B2, cuentan con la capacidad para comprender y comunicar ideas de mediana complejidad, haciendo uso apropiado de expresiones y frases cotidianas, siendo capaz de responder acertadamente a una variedad de situaciones en el ámbito laboral y académico.

Capítulo 5. DISCUSIÓN DE RESULTADOS

En el presente capítulo se discuten los resultados obtenidos en el proceso de recolección de datos después de que han sido analizados. Específicamente se abordarán los resultados del análisis de correlación y lo que la literatura existente informa al respecto. Así mismo, se irá dando respuesta a cada una de las preguntas de investigación que fueron planteadas para llegar al objetivo de la misma, que fue aportar evidencias de validez con relación a otras variables para las interpretaciones y usos del examen institucional PLET.

5.1 Autopercepción de Habilidades en Inglés

La primera pregunta de investigación de este trabajo es: ¿Cuál es el grado de asociación que existe entre el resultado del examen PLET y la percepción de los estudiantes sobre su nivel de dominio en las habilidades de comprensión lectora y auditiva, además de la sub-habilidad de gramática, del idioma inglés? De acuerdo con los valores obtenidos en el análisis de correlación de los datos para la Encuesta para la Autopercepción de Habilidades en Inglés y los puntajes en el PLET, se llegó a la conclusión de que el grado de asociación es moderado positivo, pues en

gramática se obtuvo un valor de .40, en comprensión auditiva .50, y en comprensión lectora .51, obteniendo una media de .47 entre las tres habilidades.

Con respecto de otras investigaciones, Powers, Kim, y Weng (2008) encontraron un grado de asociación similar entre los puntajes obtenidos por alrededor de 5,400 estudiantes en el examen TOEIC y en las autoevaluaciones *Can-Do* que esa cantidad de estudiantes respondió. El valor para la correlación de variables en comprensión auditiva fue de .57, mientras que en el caso de comprensión lectora fue de .50 (es difícil determinar cuánto de este valor corresponde a gramática, pues si bien el TOEIC no tiene una sección determinada como “gramática”, esta sub-habilidad se evalúa por medio de algunos ejercicios de completar oraciones en la sección de comprensión lectora). Estos valores se encuentran en el mismo rango “moderado” que los del presente proyecto de investigación, aunque para la comprensión auditiva el valor en el estudio de Powers et. al (2008) superó a los obtenidos en esta investigación, mientras que los obtenidos en cuanto a comprensión lectora son similares.

En otro estudio realizado por Yoshizawa (2009) se encontraron valores de correlación por debajo de los encontrados en el presente estudio. Para la correlación de comprensión lectora se obtuvo un valor de .358, mientras que la sección de comprensión auditiva obtuvo un .26 valores que, en palabras del autor de este estudio y en acuerdo con las etiquetas sugeridas por Bisquerra (1987), son considerados como bajos. De acuerdo con Yoshizawa (2009), “Esto presenta algo de evidencia para decir que los reactivos de comprensión lectora en la autoevaluación *can-do* tienen validez concurrente (...) En contraste, el estudio de correlación no presentó ningún soporte para los reactivos de comprensión auditiva en la autoevaluación *can-do* en términos de validez.” (p. 79).

TESIS TESIS TESIS TESIS TESIS

Yoshizawa justifica estos valores para la comprensión lectora debido a que “Uno de los factores podría ser el número limitado de reactivos en la autoevaluación de comprensión lectora. Se tenían diez reactivos, pero tres fueron borrados. (...) no son suficiente para describir lo que los aprendices del inglés como lengua extranjera en diferentes niveles de desempeño pueden llevar a cabo.” (p. 80).

Mientras que para el valor en la habilidad de comprensión lectora argumenta que puede deberse a “la discrepancia entre tareas descritas en la autoevaluación y las habilidades de los sustentantes” (p. 80), el cual parece ser un argumento aplicable a todo este tipo de investigaciones y el por qué, si bien se obtienen valores positivos, no se obtienen valores de correlación más altos, ya sea porque las habilidades descritas en la autoevaluación sean más difíciles o más fáciles de lo que el estudiante se percibe como capaz de realizar.

Podría ser ésta la misma justificación por la que en el estudio de Cotton y Conrow (1998), donde se pidió a dos grupos de estudiantes de distinto semestre valorar su desempeño académico, se encontraron valores bajos e incluso negativos tras la correlación hecha con las diferentes secciones del IELTS. Por ejemplo, para comprensión lectora se encontró una correlación de -0.25 en los 32 estudiantes de primer semestre y de 0.46 en los 22 estudiantes de segundo semestre, mientras que para la comprensión auditiva los valores fueron de -0.31 y 0.16 respectivamente. Aunque Cotton y Conrow argumentan la posibilidad de que estos valores negativos y en el rango de bajo puedan deberse a que en algunas culturas esté mal visto en algunos grupos culturales de la muestra, el cual no parece ser un factor aparente en nuestro estudio. También argumentan que las percepciones de los estudiantes podrían verse afectadas por lo que ellos comprendan como constituyente de éxito. “Altas expectativas podrían resultar en

valoraciones bajas ya que los estudiantes sienten que no están cumpliendo con sus propias expectativas.” (p. 97).

Ito, Kawaguchi, y Ohta (2005) realizaron un estudio donde se aplicaron instrumentos similares para explorar la asociación entre la percepción del estudiante sobre sus habilidades y lo demostrado en el TOEIC. Los valores de correlación observados para comprensión lectora fueron de .70, mientras que para la habilidad de comprensión auditiva fue de .65, mayores que los del presente, y considerados altos a diferencia de los moderados obtenidos aquí. La similitud entre este estudio y el propio proyecto de investigación se encuentra en que en ambos la comprensión lectora obtuvo un valor de asociación superior al de la comprensión auditiva, y la explicación de Ito et. al (2005) se divide en dos argumentos; el primero es que “puede atribuirse a que la mayoría de nuestros sujetos aprendieron inglés en el sistema educativo japonés donde la comprensión lectora ha sido más enfatizada. (...) la segunda razón para esto es que las habilidades lectoras podrían no amenazar la identidad propia o el ego tanto como las otras habilidades podrían.” (p. 16). De estos argumentos, el que podría dar una panorámica más general sobre lo observado con respecto de la habilidad de comprensión lectora y el valor que obtuvo en los análisis de correlación, siempre más altos que los otros constructos, es el del posible énfasis que se da en los cursos de inglés para esta habilidad.

En un estudio con similar metodología a la discutida en investigaciones previamente mencionadas, Powers, Roever, Huff, y Trapani (2003), aplicaron un par de escalas de autoevaluación y analizaron su correlación con las puntuaciones en la evaluación por medio del software LanguEdge, los valores de correlación en la primer escala fueron de .50 para comprensión auditiva y .43 para comprensión auditiva, mientras que en la segunda escala fueron

de .51 para comprensión auditiva y de .46 para comprensión lectora. En suma, similares en cuanto al grado moderado de los valores de la presente investigación.

Tabla 32

Valores de Correlación entre Autoevaluaciones de Habilidades en Inglés y Exámenes de Colocación encontrados en otros Estudios

Estudio	Valor de correlación en comprensión lectora	Valor de correlación en comprensión auditiva	N
Powers, Kim, y Weng (2008)	.50	.57	5,400 aproximadamente
Yoshizawa (2009)	.36	.26	
Cotton y Conrow (1998)	-.25	-.31	32
Cotton y Conrow (1998)	.46	.16	22
Ito, Kawaguchi, y Ohta (2005)	.70	.65	8,386
Powers, Roever, Huff, y Trapani (2003)	.43	.50	2,703
Powers, Roever, Huff, y Trapani (2003)	.46	.51	2,703

Así como se puede apreciar en la tabla 42, los valores de correlación encontrados en otros estudios para comprensión lectora van desde $-.25$ hasta $.70$, mientras que el propio se encuentra en $.51$, por lo que, basado en los estudios citados en este apartado, puede decirse que este valor aporta evidencia de validez en relación con otras variables para las interpretaciones del PLET con respecto de esta habilidad. En el caso de comprensión auditiva, los valores de correlación en los estudios incluidos en la tabla 42 van desde $-.31$ a $.65$, mientras que en el presente se alcanzó un valor de $.50$, lo que nos permite apreciar que la asociación entre los puntajes del PLET y los resultados de la encuesta para estudiantes en cuanto a esta habilidad también tiene un comportamiento esperado, tomando en cuenta los valores en las otras investigaciones. Incluso para la sub-habilidad de gramática puede extenderse esta aserción, pues el valor de la correlación ($.40$) es considerado moderado y entra en el rango de los valores vistos en otras investigaciones para las habilidades de comprensión auditiva y comprensión lectora.

5.1.1 Posibles factores y argumentos para valores de correlación. Es posible hacer interpretaciones del porqué los valores de correlación se comportan de la forma en que fueron discutidos en el sub-apartado anterior al asociar autoevaluaciones sobre las habilidades de los estudiantes y los resultados en exámenes estandarizados. Yoshizawa (2009) declara que uno de los factores se determina como “factor de experiencia, es decir, si la tarea [descrita] en la autoevaluación está relacionada a la experiencia que estudiantes de una segunda lengua con el lenguaje ya sea a través de la instrucción o el contacto con el lenguaje.” (p. 81). Lo anterior parece un factor considerable para el contexto de nuestro estudio, pues a pesar de que se procuró que las tareas descritas en las habilidades de interés en la Encuesta para la Autopercepción de Habilidades en Inglés estuvieran relacionadas al contexto de los estudiantes como aprendices del

TESIS TESIS TESIS TESIS TESIS

inglés como una lengua extranjera, la exposición que tienen a dichas habilidades varía de sujeto a sujeto, y al no vivir en un país donde la lengua oficial es el inglés, limita el contacto con este idioma a cursos de inglés y otras cuestiones académicas. Ross (1998 en Yoshizawa, 2009) expresa que la experiencia con comprensión auditiva de estudiantes en contextos donde el inglés es una lengua extranjera es menor a la que se tiene con comprensión lectora, lo cual puede ser tomado como un hecho, pues a no ser que la institución educativa donde los alumnos estudien sea bilingüe, hay una mayor posibilidad de que el contacto que los estudiantes tienen con el idioma inglés sea por medio de libros y artículos que contengan información que les apoye a sus tareas y formación en la carrera o materias que estudian, mientras que la comprensión auditiva podría no suceder en este tipo de contextos, pues la comunicación entre personas dentro y fuera de la institución es muy poco probable que suceda en inglés. Esto podría explicar la situación de aquellos sujetos en la muestra que perciben su habilidad de comprensión auditiva como menor que las otras. Ito et. al (2005), apoyan esta idea sobre la importancia de la experiencia, pues “la autoevaluación es más acertada cuando las preguntas son relevantes a la experiencia del estudiante.” (p. 6), además agregan que al tener menos experiencia y exposición ante el idioma inglés hace que no puedan tener una percepción apropiada de sus habilidades.

Sin embargo, lo anterior no es un argumento en contra de las autoevaluaciones en inglés, pues como se ha evidenciado en los estudios citados para este apartado sobre autopercepción y desempeño en exámenes de colocación, la mayoría de los resultados apuntan a correlaciones positivas y de un grado de asociación considerado entre moderado y alto. Además, como ya se discutía en el marco teórico en referencia con las autoevaluaciones estudiantiles, mientras que no exista un impacto a la calificación de los estudiantes durante su curso o efectos similares a partir del ejercicio de autoevaluación sobre sus habilidades, lo cual es apoyado por Ito et. al (2005),

TESIS TESIS TESIS TESIS TESIS

pues “las sobreestimaciones son más posibles de ser encontradas si la autoevaluación contribuye a la calificación del estudiante en el curso”, (p. 3), lo cual fue aclarado y se enfatizó en que no sucedería cuando se aplicó la Encuesta para la Autopercepción de Habilidades en Inglés a cada uno de los grupos. Aun así, y con base en lo declarado por Ito et. al (2005), se recomendaría que los estudiantes tuvieran experiencia o un entrenamiento sobre cómo ejercer una autoevaluación de sus habilidades, factor desconocido en los estudiantes del presente estudio.

A manera de conclusión sobre lo encontrado en la asociación de los resultados para la Encuesta para la Autopercepción de Habilidades en Inglés y los puntajes obtenidos por los estudiantes en el PLET, se puede decir lo siguiente:

a) Los valores de correlación entran en la etiqueta de moderados y son positivos, por lo que se puede argumentar que hay evidencia para decir que las interpretaciones hechas a partir de los puntajes del PLET cuentan con validez basada en la relación con la variable de Autopercepción de Habilidades en Inglés de los estudiantes.

b) Una alternativa para incrementar los valores de asociación entre estas dos variables podría ser instruir a los estudiantes en cuanto a cómo realizar una autovaloración de sus habilidades, aunque la demanda cognitiva y de introspección en la encuesta de este estudio no sea tan alta debido a que el estudiante sólo tenía que llevar a cabo un ejercicio de acuerdo o desacuerdo con afirmaciones sobre su desempeño en el idioma.

c) A pesar de haber obtenido valores de correlación favorables al objetivo del estudio, es difícil determinar si los estudiantes están dando estimaciones muy altas o bajas sobre su desempeño en el idioma o si es que el PLET está siendo severo o laxo en la evaluación del idioma, esto sobre todo en gramática, el área donde existe una menor asociación (comparada con

las otras dos) entre lo que percibe el estudiante sobre su desempeño en la sub-habilidad y el desempeño mostrado en el PLET para esta sección.

5.2 Juicio Docente

La segunda pregunta de investigación en el proyecto es: ¿Cuál es el grado de asociación que existe entre el resultado del examen PLET y el juicio de los maestros sobre la colocación de estudiantes en un determinado nivel de dominio del idioma inglés, basado en el desempeño visto de sus estudiantes en el curso?, cuya respuesta es que la asociación es alta, esto con base en los valores obtenidos en la correlación de los resultados de la Encuesta para el Juicio Docente y la colocación de estudiantes en cursos del PILE a partir de los resultados en el PLET, pues se obtuvo un .59 en comprensión auditiva, .62 en comprensión lectora, .68 en gramática, y en la valoración general de desempeño .75, obteniendo una media de .66 entre los cuatro rubros.

En un estudio realizado por Coladarci (1986), se explora la certeza con que los docentes valoran el desempeño de sus estudiantes en áreas de comprensión lectora y matemáticas, concluyendo, a partir de la correlación de ambas variables, que los docentes tienden a otorgar valoraciones muy acertadas sobre el desempeño mostrado por sus estudiantes, pues los valores del juicio docente van del .78 al .92, lo cual habla sobre grados muy altos de correlación. En el caso de comprensión lectora, la correlación del desempeño académico del estudiante y el juicio del maestro fue de .90. Perry y Meisels (1996) argumentan en su revisión de investigaciones sobre juicios docentes que los docentes tienden a ser más certeros en sus juicios en áreas que tienen que ver con comprensión lectora, lenguaje en las artes y en matemáticas, lo que puede ser apreciado en los valores de correlación del estudio de Coladarci (1986) y en el estudio presente. En esta revisión de Perry y Meisels (1996) sobre investigaciones en el juicio docente se ha

llegado incluso a la conclusión de que el juicio del docente ha demostrado ser mejor predictor de desempeño que otras medidas estandarizadas.

Hoge y Coladarci (1989), a partir de una revisión de literatura sobre juicios de logro académico basado en los docentes, logran distinguir una división sobre el tipo de juicio docente, teniendo por un lado un juicio directo y por el otro uno indirecto. Básicamente la diferencia entre uno y otro de estos tipos de juicios se encuentra en si el docente tiene que emitir un juicio sobre el desempeño de sus estudiantes en función de criterios especificados (directo) o si tiene que emitir un juicio sin ningún tipo de guía sobre los criterios de interés. Al respecto, los autores de dicha investigación encontraron que las correlaciones entre el juicio del docente y el desempeño de sus estudiantes en pruebas estandarizadas que pueden ser sobre lectura, matemáticas, inglés en general, y estudios sociales, en el caso de juicios directos los valores van desde .48 a .92, y los de juicios indirectos van de .28 a .86. En esta revisión de literatura hecha por los autores puede apreciarse que en estudios donde se ha explorado la asociación del juicio docente con pruebas que miden la habilidad lectora, los valores no descienden del .55 y no superan el .86 (Hoge y Coladarci, 1989, pp. 304 – 305), lo cual nos deja ver que el valor obtenido para comprensión lectora en el presente estudio (.62), queda dentro de este rango y da fundamento no sólo para decir que el juicio docente en la encuesta fue acertado, sino que apoya el argumento de colocación apropiada por el PLET, y no solamente en comprensión lectora, sino en las otras dos áreas del lenguaje también.

En el estudio realizado por Papageorgiou y Cho (2014), se exploró la correlación entre los juicios de maestros sobre el desempeño de estudiantes en dos escuelas distintas tras haber sido colocados en un nivel de inglés como resultado de haber presentado el TOEFL® Junior™. Los resultados de dicha correlación para la escuela A fueron de .72 y para la escuela B de .83.

TESIS TESIS TESIS TESIS TESIS

No se comunicaron resultados por habilidad. Estos resultados apoyan a los estudios y revisiones literarias mencionadas en párrafos anteriores, pues se encuentran dentro de los rangos de alto y muy alto, a los cuales se inscribe el resultado de correlación encontrado en el presente estudio entre la valoración general que hicieron los docentes sobre el desempeño de sus estudiantes y el nivel de ubicación emitido por el PLET, el cual obtuvo un .75.

En el estudio de Powers, Roever, Huff y Trapani (2003), también se explora el juicio docente y su asociación con el examen del software LanguEdge, encontrando valores de correlación moderados. En la habilidad de comprensión auditiva la correlación da .52, mientras que para comprensión lectora queda en .48. Cotton y Conrow (1998) también encontraron valores por debajo de los considerados altos y muy altos en estudios referidos en párrafos anteriores. En cuanto a la puntuación obtenida en la sección de comprensión lectora del IELTS y su asociación con la calificación emitida por los profesores fue de .36, mientras que para comprensión auditiva fue apenas de .07; y la valoración global emitida en ese estudio marca un .15. Los autores explican estos valores debido a que el juicio que el maestro hacía era con base en distintas mediciones que los profesores utilizaron para elaborar su opinión. Por ejemplo “si los estudiantes no participan, entonces no es probable que el personal [profesores] les den valores altos” (p. 95). Sin embargo, la razón de más peso para que estos resultados sean tan bajos parece ser, en palabras de Cotton y Conrow (1998), “los profesores fueron encuestados casi diez meses después de que el IELTS fuera presentado por los estudiantes, haciendo imposible establecer cualquier relación”. (p. 96). En el presente estudio se descarta ese problema, pues los docentes fueron encuestados poco después del primer mes de haber estado trabajando con sus estudiantes, y como puede apreciarse en los resultados, ninguno de ellos se considera bajo, razón que sirve para apoyar el factor sobre la influencia del tiempo; Cotton y Conrow (1998) no

ahondan en esto pero probablemente los estudiantes, tras esos diez meses de haber presentado el IELTS, adquirieron nuevas habilidades que los hicieron alejarse de los puntajes obtenidos en el mencionado examen.

5.2.1 Posibles factores y argumentos para valores de correlación. Para el caso de juicios no adecuados por parte de docentes podría exponerse el argumento que expresan Lukša, Radanović, y Garašić, (2014) sobre las expectativas poco realistas por parte de los profesores, lo cual evidencia que el docente podría no ser consciente de los problemas que ciertos estudiantes se encuentran en su aprendizaje. Esto podría ser una posibilidad por la que la correlación entre el juicio docente y la ubicación en niveles de PILE no llega a ser mayor, aunque ya sea alta; podría suceder que aquellos estudiantes no están siendo apoyados por sus docentes en su aprendizaje como lo demanda su nivel, pues estos últimos podrían tener una opinión lejana a la realidad del estudiante y su desempeño.

Después de haber discutido los resultados en otros estudios donde el juicio docente era una variable cuya asociación con puntajes de exámenes estandarizados, de colocación o no, fue explorada, es posible decir que la ubicación del PLET para los 21 sujetos sobre los cuales los docentes hicieron su valoración, es una interpretación válida, tomando en cuenta el grado de correlación positivo alto que existe tanto en cada una de las habilidades como en la valoración global.

También es posible decir que podría trabajarse más en la encuesta de ser necesario para incrementar el valor de asociación, pues probablemente en algunos casos la descripción y criterios que sirvieron como guía para que el docente emitiera su valoración eran o muy

específicos, resultando como un problema si el docente no conoce bien al estudiante y su desempeño.

5.3 Desempeño de los Estudiantes en su Examen Intermedio

La tercera pregunta de investigación es: ¿Cuál es la asociación que existe entre el resultado del examen PLET y el desempeño demostrado por parte de los estudiantes en las secciones de comprensión lectora, comprensión auditiva y gramática en el examen intermedio del curso? Dando respuesta a esta pregunta, basado en los resultados de la correlación entre el desempeño de los estudiantes mostrado en el examen intermedio de su curso y la ubicación PILE a partir del PLET, el grado de asociación se considera alto, pues los valores de correlación fueron de .82 en comprensión auditiva, .39 para gramática, .83 para comprensión lectora, .88 de valoración general de las tres habilidades, todos estos valores con una media de .73.

En relación con esta variable sobre el desempeño académico del estudiante durante el curso al que fue asignado a partir de un examen de colocación, no se ha encontrado literatura sobre estudios que hayan utilizado una metodología donde se incluya una variable de esta naturaleza para discutir la validez basada en relación con otras variables.

Sin embargo, existen estudios en los cuales pueden rescatarse ciertas similitudes y de esta forma argumentar la discusión de los resultados obtenidos en el presente con respecto de una variable sobre desempeño académico de características similares con la que aquí se trabajó, pues no se trata de un examen estandarizado con el que se compararán los resultados del PLET, ni es un promedio final sobre el desempeño que los estudiantes tuvieron en sus carreras, se trata de una variable que aborda el desempeño académico del estudiante en un curso de inglés en un momento en que dicho curso no ha terminado y aún no tienen un impacto tan alto en el

estudiante cuestiones más bien relacionadas a efectos del contenido del curso, la calidad en la instrucción del docente, el material de trabajo, etc.

Rumsey (2013), exploró los puntajes del examen de colocación CEPA y su correlación con variables como los resultados en el examen final del curso de inglés, las calificaciones finales del trabajo en clase del curso de inglés, y el promedio final del curso. La autora del estudio encontró una correlación de .48 entre el examen final de inglés y el puntaje en CEPA, de .46 entre la calificación final del trabajo en clase, y un .47 entre la calificación final del curso. Si bien en el estudio citado se buscaba ver el poder predictivo del CEPA para el éxito académico tanto en cursos de inglés como de matemáticas, las variables con que se hace la correlación son similares a la variable que se discute en este apartado, donde la correlación entre los constructos en el PLET y los constructos del Examen Intermedio en cursos del PILE fue en todos los casos alta, a excepción de gramática, donde, al igual que las variables en el estudio de Rumsey, el valor de correlación permanece como moderado, similar a lo encontrado por Pugh y Lowther (2004), autores de un estudio donde se exploró la correlación entre los puntajes de un examen de colocación de matemáticas (AMATH) y la calificación final obtenida en dicho curso. Ellos encontraron una correlación de .33 entre estas dos variables.

Como ya se advertía al principio de este sub-apartado, por medio de la revisión de literatura se encontró con que existe una escases de información disponible sobre estudios donde el desempeño del estudiante sea abordado en un momento durante el curso, y no al final donde varios factores relacionados con cuestiones independientes de la colocación podrían afectar los resultados. Los dos estudios aquí referidos fueron considerados pues se asocian las calificaciones finales de un curso y de exámenes de dicho curso en los que se exploran las mismas áreas de conocimiento que se miden en el examen de colocación que les ubicó en el nivel y curso donde

estudian, lo que los hace diferentes a numerosos estudios donde la correlación se explora entre el examen de colocación y variables como la puntuación de un examen estandarizado o el promedio final de un estudiante en el primer semestre no de una, sino de todas las materias que lo conformaron.

5.3.1 Posibles factores y argumentos para valores de correlación. Una posible explicación del porqué el valor de correlación en gramática no es tan alto como los encontrados para comprensión auditiva y comprensión lectora podría deberse a que en el PLET los reactivos diseñados para medir el desempeño en gramática están siendo severos en la dificultad con que evalúan esta sub-habilidad del idioma, esto se puede apreciar pues en promedio, como se puede ver en las tablas 20 y 21, los estudiantes obtenían un siete como calificación, dentro de una escala de cero a 10, lo que, si bien no habla de un desempeño perfecto de esta habilidad, permite ver que el desempeño de por parte de los estudiantes es aceptable para el nivel al que son asignados. Esto puede apreciarse de forma más clara en la tabla 23, donde, tras promediar los porcentajes de estudiantes capaces de realizar actividades gramaticales esperadas para su nivel, se obtiene que un 86.8% de ellos son capaces de llevarlas a cabo satisfactoriamente.

Otro factor para este valor en gramática puede deberse a los puntos de corte que se precisaron en el PLET para la sub-habilidad gramatical. Es decir, la dificultad no se encuentra en los reactivos, sino en lo demandante que resulta para el sustentante alcanzar la puntuación que ha sido especificada, a través del proceso de selección de puntos de corte del examen, como deseable para cubrir con cierto nivel. En otras palabras, la toma de decisiones para la precisión de los puntos de corte fue severa, aumentando la dificultad para que un sustentante alcanzara un nivel determinado, ya que las marcas que fueron precisadas para indicar la puntuación límite de un nivel y el comienzo de otro, exigen un puntaje alto.

Este posible factor cobra sentido sobre todo si se presta atención al desempeño que los estudiantes tuvieron en su examen intermedio del curso, en el que, de acuerdo con lo visto en la tabla X, puede decirse que fueron capaces de llevar a cabo actividades gramaticales de su nivel de una forma más o menos satisfactoria.

En la asociación del desempeño demostrado por los estudiantes en su examen intermedio y la ubicación PILE a partir del PLET, se encuentran valores positivos, por lo que se puede confirmar que el PLET coloca a los estudiantes de manera apropiada en función de su nivel de desempeño en el inglés, recordando que esto es considerando lo visto en el desempeño académico de los estudiantes en su examen intermedio.

La correlación en el caso de la sub-habilidad de gramática podría ser mayor si se revisará la dificultad en los reactivos del PLET para este constructo o los puntos de corte.

Capítulo 6. Conclusiones

6.1 Conclusiones Generales

Con apoyo de lo visto en el análisis de resultados y la discusión de los mismos con lo encontrado en otros estudios es posible decir que de acuerdo con a) la autopercepción de habilidades en inglés por parte de los estudiantes, b) el juicio docente sobre el desempeño observado de algunos de sus estudiantes, y c) el desempeño mostrado por esos mismos estudiantes en el examen intermedio, las interpretaciones del PLET sobre la colocación de estudiantes en niveles de inglés cuentan con evidencias a favor de su validez en relación con otras variables.

TESIS TESIS TESIS TESIS TESIS

Después de lo visto en los análisis de correlación se puede apreciar que de las tres variables con que se asocia el puntaje del PLET y la ubicación de estudiantes a partir del mismo, es el desempeño mostrado en el examen intermedio del curso de inglés la que cuenta con valores más altos de correlación, lo que hace de esta variable el estándar de oro.

Sin embargo, lo anterior no quiere decir que se le resta importancia a las otras dos variables, pues los valores en el juicio docente también son altos y a pesar de que en la autopercepción del estudiante son moderados cumplen con un comportamiento similar al visto en otros estudios que son referidos en el sub-apartado para esta variable dentro del capítulo de discusión de resultados.

En las tres variables también se puede apreciar un comportamiento similar en cuanto a cómo responden los valores de correlación en cada habilidad, pues en la autopercepción de habilidades en inglés y en el desempeño mostrado en el examen intermedio, la sub-habilidad de gramática obtuvo los valores más bajos en comparación con las otras dos habilidades, las cuales, dentro de cada variable, se distanciaban una de otra con diferencias mínimas en sus valores. Por ejemplo en la correlación con respecto de la autopercepción de habilidades en inglés, así como en las correlaciones respecto del desempeño en el examen intermedio, la comprensión lectora apenas supera en valor de correlación a la comprensión auditiva con un .01, mientras que en el juicio docente lo hace con apenas un .03.

Lo anterior hace que se abran interrogantes con respecto de la forma y dificultad con que se está evaluando la sub-habilidad de gramática en el PLET, pues además de ser la habilidad con los valores más bajos en dos de las variables comparada con los valores de comprensión auditiva y comprensión lectora dentro de esas mismas variables, se puede notar un patrón que sí se

TESIS TESIS TESIS TESIS TESIS

presenta en las tres variables, y es que los valores de correlación para gramática se alejan con a) .08 puntos en el caso del juicio docente (la única variable donde gramática obtuvo el valor más alto de correlación con la ubicación del PLET de entre las tres habilidades), b) .10 puntos (inferiores a los valores en las otras habilidades) en la autopercepción de habilidades en inglés del estudiante, y c) con .43 puntos (inferiores a los valores en las otras habilidades) en el desempeño mostrado en el examen intermedio.

Como ya se discutía en el capítulo de discusión de resultados, se sugiere al PILE revisar cuestiones tanto de dificultad en reactivos del PLET destinados a evaluar el área de gramática así como en los puntos de corte del examen para esta sub-habilidad.

6.2 Limitaciones del Estudio y Recomendaciones para Futuros Estudios

La principal deficiencia del estudio radica en la muestra. A pesar de que se llegó a la determinación de la cantidad de grupos, estudiantes, y profesores por medio de técnicas de muestreo probabilístico tales como el muestreo por conglomerados y el muestreo estratificado, el detalle que, inesperadamente, hizo más pequeña la muestra fue que no todos los estudiantes a quienes se les aplicó la Encuesta de Autopercepción para las Habilidades del Inglés, ni todos los estudiantes cuyo desempeño en clase fue valorado por los profesores a través de la Encuesta para el Juicio Docente, mismos sobre quienes se evaluó su desempeño en el examen intermedio de su curso de inglés, aplicaron la versión del examen PLET cuyos usos e interpretaciones se buscaron validar con la relación de variables en este estudio. Es decir, no todos los estudiantes que se tenían contemplados para la recuperación de información en el proyecto, fueron ubicados por la versión 2016 del PLET en el curso que tomaban.

De esta forma, la recomendación fundamental para futuros estudios parte de la deficiencia referida en el párrafo anterior. Es necesario tener un control más meticuloso sobre quiénes son los participantes, pues si bien el PLET se ha aplicado desde el 2015, ha contado con versiones distintas, que a pesar de seguir el mismo proceso de diseño, cada una de éstas requiere la validación de usos e interpretaciones a partir de los resultados específicos de la versión de examen con que se está trabajando. Así, es necesario contar con una muestra apropiada que esté compuesta por participantes que tengan este criterio en común.

Como se ha venido precisando desde el primer capítulo de la investigación, sólo se exploraron las habilidades de comprensión auditiva y comprensión lectora, además de la sub-habilidad de gramática, ya que son las áreas de desempeño que evalúa el PLET; sin embargo, resultaría interesante que para futuros estudios que se incluyera la exploración sobre las habilidades de producción, tanto oral como escrita, y detectar el grado de correlación que guarda la ubicación que otorga el PLET y el desempeño de estas habilidades, visto desde el juicio docente y desde varias evidencias de producción y trabajo en clase, así como una autovaloración del mismo estudiante. Esto abriría temas de discusión interesantes sobre cómo la evaluación de las habilidades receptivas puede o no guardar correlaciones significativas con la competencia en habilidades productivas.

Una última recomendación en torno a lo referente con la variable sobre el desempeño del estudiante a través de exámenes y actividades en clase, es que podría resultar eficiente el diseño de exámenes diagnósticos por parte el investigador, mismos que se apliquen en la primera semana de la impartición de los cursos a que son ubicados los estudiantes. Esto podría resultar en una obtención de información sobre desempeño más enfocada a criterios y elementos de interés

exclusivo de la investigación, esto es, hacer de estos exámenes diagnósticos por nivel la evidencia necesaria para evaluar el desempeño del estudiante en las habilidades de interés.

Es importante mencionar que, a pesar de encontrar valores positivos y que van de moderados a altos en la asociación de variables, el estudio no apoya que cualquiera de las variables, de manera aislada, pueda dar validez a las interpretaciones de un examen de colocación, pues esto significaría limitarse a la panorámica que nos puede ofrecer una sola variable. Tener información a partir de más variables enriquece la discusión sobre la validez de interpretaciones y usos de una prueba y sus resultados, pues se abre paso a más de una opinión, a más de una visión, ya que puede ayudar a que algún argumento sea consistente en todas las variables, y de no serlo, da oportunidad a la identificación de áreas de oportunidad.

Así mismo, ninguna de las variables con que se asociaron los puntajes y la ubicación emitida por el PLET, puede sustituir en cuestiones de colocación a un examen con el proceso de diseño que tiene dicha prueba; pueden argumentar a favor o en contra de las interpretaciones y usos del mismo, pero los resultados en esta investigación no apoyan de ninguna forma que alguno de los tres instrumentos con que se exploró la correlación con el PLET pueda tener la función de ubicar a estudiantes en cursos de inglés.

A pesar de que el desempeño mostrado en el examen intermedio por parte de estudiantes haya sido determinado como el estándar de oro en esta investigación tanto por los valores de correlación encontrados ahí como por el tipo de evidencia que es (una prueba de desempeño a diferencia de las otras dos que son opiniones), valdría la pena incluir evidencia sobre desempeño académico con reactivos o producciones donde se pueda explorar el dominio de los estudiantes en estas habilidades pero de una forma abierta, y no con reactivos de respuesta cerrada como los

del examen intermedio, además que, como se tenía pensado, pero por problemas con la respuesta por parte de docentes no se logró, lo ideal sería tener más de una fuente de evidencia para la variable de desempeño, tales como tareas, diferentes tipos de evaluaciones, actividades en clase, etc.



Referencias

- Alnahdi, G.H. (2015). Aptitude tests and successful college students: The predictive validity of the General Aptitude Test (GAT) in Saudi Arabia. *International Education Studies*, 8 (4), 2 – 6.
- Alyousef, H.S. (2005). Teaching reading comprehension to ESL/EFL learners. *The Reading Matrix*, 5(2), 143 – 154.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Estados Unidos, WA: American Educational Research Association.
- Armstrong, W.B. (2001). Explaining student course outcomes by analyzing placement test scores, student background data, and instructor effects. San Diego, CA. Recuperado de: <http://eric.ed.gov/?id=ED454907>
- Armstrong, W.B. (1994). English placement testing. Analysis of the criterion- and content-related validity evidence for the Reading & Writing Placement Tests in the San Diego Community College District. San Diego, CA: Research and Planning office for San Diego Community College.
- Aydoğan, H. (2014). The four basic language skills, whole language & intergrated skill approach in mainstream university classrooms in Turkey. *Mediterranean Journal of Social Sciences*, 5(9), 672 – 680.
- Bayliss, A. (2006). IELTS as a predictor of academic language performance. (Trabajo no publicado). Universidad de Melbourne.
- Bisquerra, R. (1987). *Introducción conceptual al análisis multivariable: un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*, Barcelona: PPU.
- Blue, G.M. (1994). Self-assessment of foreign language skills: Does it work? *CLE Working Papers*, 3, University of Southampton. U.K., 18-35.
- Borsboom, D. (2015). Zen and the art of validity theory. *Assessment in Educatio Principles Policy and Practice*, DOI: 10.1080/0969594X.2015.1073479

- Buck, G. (2001). An overview of listening comprehension. En *Assessing Listening* (Capítulo I). Recuperado de <http://www.cambridge.org/gb/cambridgeenglish/catalog/teacher-training-development-and-research/assessing-listening/assessing-listening-1/resources>
- Capella University. (s.f.). *Grammar handbook*. Recuperado de <http://www.capella.edu/interactivemedia/onlinewritingcenter/downloads/grammar.pdf>
- Clifton, C. y Duffy, S.A. (2001). Sentence and text comprehension: Roles of linguistic structure. *Annual Reviews*, 52, 167 – 190.
- Coladarci, T. (1986). Accuracy of teacher judgments of student responses to standardized test items. *Journal of Educational Psychology*, 78(2), 141 – 146.
- Consejo de Europa. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.
- Cotton, F. y Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a group of international students studying at the University of Tasmania. *IELTS*, 1 (4), 72 – 115.
- Daugherty, R. (1995). *National Curriculum Assessment: A review of policy, 1987 – 1994*. Reino Unido, LDN: The Falmer Press.
- Denicolo, P.M. y Kompf, M. (2005). *Teacher thinking and profesional action*. Estados Unidos, NY: Routledge.
- DuPaul, G.J. y Rapport, M.D. (1991). Teacher ratings of academic skills: The development of the academic performance rating scale. *School Psychology Review*, 20, (2), 284 – 300.
- Exam English. (2011). *Grammar by CEF level*. Recuperado de http://www.examenglish.com/CEFR/cefr_grammar.htm
- Facultad de química de la UNAM. (2013). ¿Cuál es la importancia del aprendizaje del idioma inglés? *El Filogisto Ilustrado*, 2 (93), 1. – 20.
- Fulcher, G. (1997). An English Language Placement Test: Issues in reliability and validity. *Language Testing* 14 (2), 113 – 139.

- Fuller, M.L. (2000). Teacher judgment as formative and predictive assessment of student performance on Ohio's fourth and sixth grade proficiency tests. Documento presentado en la Junta Anual de la Asociación Americana de Investigación Educativa, Nueva Orleans.
- García-Medina, A.M. y Pedroza-Zúñiga, L.H. (2015). *Elaboración del examen de competencias de la lengua inglesa. Diseño del instrumento. Versión preliminar.* (Trabajo no publicado). Universidad Autónoma de Aguascalientes, México.
- Gardner, D. (1999). Self-assessment for autonomous language learners. *Links & Letters*, 7, 49 – 60
- Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *Research Notes*, 12, 8 – 10.
- Hamilton, C. y Shinn, M.R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, 32 (2), 228 – 240.
- Hoge, R.D. y Coladarci, T. (1989). Teacher-Based Judgments of Academic Achievement: A Review of Literature. *Review of Educational Research*, 59(3), 297 – 313.
- Ito, T., Kawaguchi, K., & Ohta, R. (2005). A Study of the Relationship between TOEIC Scores and Functional Job Performance: Self-assessment of Foreign Language Proficiency. *TOEIC Research Report*, 1-40.
- Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38 (4), 319 – 342.
- Kane, M.T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50 (1), 1 – 73.
- Karakaya, I. y Tavsancil, E. (2008). The predictive validity of the university student selection examination. *Kuram ve Uygulamada Eğitim Bilimleri / Educational Sciences: Theory & Practice*, 8 (3), 1011 – 1019.

- Kokhan, K. (2012). Investigating the possibility of using TOEFL scores for university ESL decision-making: Placement trends and effect of time lag. *Language Testing*, 29 (2), 291 – 308.
- Kokhan, K. (2013). An argument against using standardized test scores for placement of international undergraduate students in English as a Second Language (ESL) courses. *Language Testing*, 30 (4), 467 – 489.
- Le, H. (2011). *Developing a validity argument for the English Placement Listening Fall 2010 test at Iowa State University*. (Tesis no publicada). Universidad Estatal de Iowa, Ames, Iowa.
- LeBlanc, R. y Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, 19 (4), 673 – 687
- Lee, Y. y Greene, J. (2007). The predictive validity of an ESL placement test: A mixed methods approach. *Language Testing*, 1 (4), 366 – 389.
- Linacre, J.M. (1994). Sample Size and Item Calibration Stability. *Rasch Measurement Transactions*, 7(4), 328. Recuperado de: <https://www.rasch.org/rmt/rmt74m.htm>
- Linn, R.L. (2010). Validity. En *International Encyclopedia of Education*. (Vol. 3, pp. 181 – 785). P. Peterson, E. Baker and B. McGraw. Oxford: Elsevier.
- Lukša, Ž., Radanović, I, y Garašić, D. (2014). Relationship of teachers' perception of students' knowledge and real students' knowledge acquisition. *European Science Education Research Association*, 317 – 326.
- Márquez-García, C.L., Medina-Escorcia, M.M., y Pérez-Sarmiento, D.B. (2013). Diseño de un procedimiento de autoevaluación del proceso de comprensión lectora en estudiantes de educación media. *Escenarios*, 11 (2), 57 – 68
- Martínez, J.F., Stecher, B. y Borko, H. (2009). Classroom assessment practices, teacher judgements, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment*, 14, 78 – 102.

- Martínez-Figueira, E., Tellado-González, F., y Rasposo-Rivas, M. (2013). La rúbrica como instrumento para la autoevaluación: un estudio piloto. *REDU. Revista de Docencia Universitaria*, 11 (2), 373 – 390
- Meiers, M., Ozolins, C., y McKenzie, P. (2007). Improving consistency in teacher judgments: An investigation for the Department of Education, Victoria. *ACEReSearch*, 7, 1 – 89.
- Mendoza-Ramos, A. (2015). La validez en los exámenes de alto impacto. Un enfoque desde la lógica argumentativa. *Perfiles Educativos*, 37 (149), 169 – 186.
- Moritoshi, Paul. (2001). The Test of English for International Communication (TOEIC®): necessity, proficiency levels, test score utilisation and accuracy. The University of Birmingham (Internal publication) Recuperado de: <http://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/testing/PaulMoritoshisM5Testingassignmentgraded72.pdf>
- Muyskens, P., Betts, J., Lau, M.Y. & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *The California School Psychologist*, 14, 11 – 21.
- Newton, P., & Shaw, S. (2014). *Validity in Educational and Psychological Assessment*. SAGE.
- Niojons, J., Bérešová, J., Breton, G., y Szabó, G. (2011). Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR). *RelEx Relating Language Examinations*, 1 – 93.
- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3, 53 – 66.
- Oscarson, M. (2013). The challenge of student self-assessment in language education. *Voices in Asia*, 1 (1), 1 – 14.
- Panahi, A. (2014). Threats to Validity: construct-irrelevant variances contributing to performance under-representation on Graduate Record Exam (GRE). *Journal of Education & Human Development*, 3(1), 327 – 346.

- Papageorgiou, S. y Cho, Y. (2014). An investigation of the use of TOEFL® Junior™ Standard scores for ESL placement decisions in secondary education. *Language Testing*, 31 (2), 223 – 239.
- Perry, N.E. y Meisels, S.J. (1996). *How accurate are teacher judgments of students' academic performance?* (National Center for Education Statistics No. 96-08) Washington D.C.: NCES.
- Powers, D.E., Kim, H-J, Weng, V.Z. (2008). *The redesigned TOEIC (Listening and Reading) test: Relations to test-taker perceptions of proficiency in English*. Recuperado de: <https://www.ets.org/Media/Research/pdf/TC-10-06.pdf>
- Powers, D.E., Roever, C., Huff, K.H., y Trapani, C.S. (2003). Validating LanguEdge courseware scores against faculty rating and student self-assessments. *Educational Testing Service: Research Reports*, 1 – 29
- Prieto, G. y Delgado, A.R. (2013). Análisis de un test mediante el modelo de Rasch. *Piscotherma*, 15(1), 94 – 100.
- Programa Institucional de Lenguas Extranjeras. (2015). Informe de evaluación, examen de colocación/acreditación PLET. Aguascalientes: PILE.
- Pugh, C.M. y Lowther, S. (2004). College Math Performance and Last High School Math Course. Recuperado de:
- Riazi, M. (2014). Concurrent and predictive validity of Pearson Test of English Academic (PTE Academic). *Papers in Language Testing and Assessment*, 2(2), 1 – 8.
- Rojas, S. y Montero, E. (2012). El modelo de Rasch: una herramienta esencial para la evaluación educativa de gran escala. En Cuarto Informe del Estado de la Educación.
- Rumsey, L.W. (2013). *Shedding Light on the predictive validity of English proficiency tests in predicting academic success*. Recuperado de:
- Snow, C. (2002). Reading for understanding: Toward an R&D program in reading comprehension. Santa Monica, CA: RAND Corporation.

- Snowling, M., Cain, K., Nation, K. y Oakhill, J. (2009). *Reading comprehension_ nature, assessment, and teaching*. Recuperado de <http://psyweb.psy.ox.ac.uk/lcd/files/ESRCcomprehensionbooklet.pdf>
- Südkamp, A., Kaiser, J., y Möller, J. (2014). Teacher's judgements of students' academic achievement. En Krolak-Schwerdt, S., Glock, S., y Böhmer, M. (Eds.), *Teachers' professional development. Assessment, training, and learning*.
- Taras, M. (2015). Student self-assessment: what have we learned and what are the challenges. *RELIEVE*, 21 (1), art. ME8. DOI: 10.7203/relieve.21.1.63944
- Taylor, L. y Jones, N. (2006). Cambridge ESOL exams and the Common European Framework of Reference (CEFR). *Research Notes*, 24, 1 – 24.
- Universidad Autónoma de Aguascalientes (s.f.) Egreso y titulación. En *Departamento de Control Escolar*. Recuperado de: <http://www.uaa.mx/direcciones/sg/ce>
- University of Cambridge. (2011). *Using the CEFR: Principles of good practice*. Recuperado de <http://www.cambridgeenglish.org/images/126011-using-cefr-principles-of-good-practice.pdf>
- University of Cambridge. (2013). *Introductory guide to the Common European Framework of Reference (CEFR) for English Language Teachers*. Recuperado de <http://www.englishprofile.org/images/pdf/GuideToCEFR.pdf>
- Wall, D., Clapham, C., & Alderson, J. C. (1994). Evaluating a placement test. *Language Testing*, 11(3), 321–344.
- Wang, J., Niemi, D., & Wang, H. (2007). Predictive validity of an English language arts performance assessment. *CRESST*, 729, 1 – 27.
- Wooley, G. (2011). Reading comprehension. En *Reading Comprehension: Assisting children with learning difficulties* (14 – 34). Germany: Springer.

Yoshizawa, K. (2009). To what extent can self-assessment of language skills predict language proficiency of EFL learners in school context in Japan? *Foreign Language Education Research (外国語教育研究)*, 17, 65-82.



ANEXO 1

Porcentajes de respuesta por reactivo en la Encuesta para la Autopercepción de Habilidades del Inglés

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
Nivel 1	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	0.0%	44.4%	55.6%	100.0%
A1.2R	0.0%	0.0%	11.1%	33.3%	55.6%	100.0%
A1.3R	0.0%	11.1%	0.0%	22.2%	66.7%	100.0%
A1.4R	0.0%	0.0%	0.0%	33.3%	66.7%	100.0%
A1.5R	0.0%	0.0%	11.1%	55.6%	33.3%	100.0%
A1.6R	0.0%	0.0%	22.2%	55.6%	22.2%	100.0%
A1.7R	0.0%	0.0%	22.2%	44.4%	33.3%	100.0%
A1.8R	0.0%	0.0%	22.2%	44.4%	33.3%	100.0%
A1.9R	0.0%	0.0%	11.1%	44.4%	44.4%	100.0%
A1.10R	0.0%	0.0%	44.4%	11.1%	44.4%	100.0%
A1.11R	0.0%	11.1%	33.3%	33.3%	22.2%	100.0%
A1.12R	11.1%	22.2%	44.4%	22.2%	0.0%	100.0%
A1.13R	11.1%	22.2%	33.3%	22.2%	11.1%	100.0%
A1.14R	22.2%	33.3%	11.1%	22.2%	11.1%	100.0%
A1.15R	22.2%	11.1%	55.6%	11.1%	0.0%	100.0%
A2.1L	0.0%	0.0%	0.0%	11.1%	88.9%	100.0%
A2.2L	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
A2.3L	0.0%	0.0%	11.1%	33.3%	55.6%	100.0%

A2.4L	0.0%	0.0%	11.1%	44.4%	44.4%	100.0%
A2.5L	0.0%	0.0%	0.0%	11.1%	88.9%	100.0%
A2.6L	0.0%	0.0%	0.0%	33.3%	66.7%	100.0%
A2.7L	0.0%	0.0%	22.2%	44.4%	33.3%	100.0%
A2.8L	0.0%	0.0%	55.6%	11.1%	33.3%	100.0%
A2.9L	0.0%	0.0%	44.4%	22.2%	33.3%	100.0%
A2.10L	0.0%	0.0%	22.2%	33.3%	44.4%	100.0%
A2.11L	0.0%	0.0%	0.0%	77.8%	22.2%	100.0%
A2.12L	0.0%	0.0%	11.1%	55.6%	33.3%	100.0%
A2.13L	0.0%	0.0%	33.3%	33.3%	33.3%	100.0%
A2.14L	0.0%	0.0%	44.4%	44.4%	11.1%	100.0%
A2.15L	0.0%	11.1%	33.3%	55.6%	0.0%	100.0%
A2.16L	11.1%	33.3%	33.3%	22.2%	0.0%	100.0%
A2.17L	11.1%	0.0%	11.1%	77.8%	0.0%	100.0%
A2.18L	11.1%	22.2%	33.3%	33.3%	0.0%	100.0%
A2.19L	11.1%	0.0%	55.6%	22.2%	11.1%	100.0%
A2.20L	11.1%	11.1%	33.3%	44.4%	0.0%	100.0%
A2.21L	11.1%	11.1%	44.4%	22.2%	11.1%	100.0%
A2.22L	11.1%	11.1%	55.6%	22.2%	0.0%	100.0%
A2.23L	11.1%	22.2%	22.2%	33.3%	11.1%	100.0%
A2.24L	11.1%	11.1%	33.3%	33.3%	11.1%	100.0%
A2.25L	11.1%	22.2%	55.6%	0.0%	11.1%	100.0%
A2.26L	22.2%	22.2%	44.4%	11.1%	0.0%	100.0%
A2.27L	33.3%	22.2%	22.2%	22.2%	0.0%	100.0%
A3.1G	11.1%	22.2%	0.0%	66.7%	0.0%	100.0%
A3.2G	0.0%	33.3%	22.2%	44.4%	0.0%	100.0%

A3.3G	11.1%	11.1%	33.3%	44.4%	0.0%	100.0%
A3.4G	11.1%	11.1%	22.2%	44.4%	11.1%	100.0%
A3.5G	11.1%	22.2%	44.4%	22.2%	0.0%	100.0%
A3.6G	11.1%	11.1%	33.3%	33.3%	11.1%	100.0%
A3.7G	11.1%	22.2%	33.3%	33.3%	0.0%	100.0%
A3.8G	11.1%	11.1%	22.2%	55.6%	0.0%	100.0%
A3.9G	11.1%	22.2%	22.2%	44.4%	0.0%	100.0%
A3.10G	11.1%	0.0%	33.3%	55.6%	0.0%	100.0%
A3.11G	11.1%	11.1%	33.3%	44.4%	0.0%	100.0%
A3.12G	11.1%	11.1%	44.4%	33.3%	0.0%	100.0%
A3.13G	11.1%	11.1%	55.6%	22.2%	0.0%	100.0%
A3.14G	11.1%	22.2%	33.3%	22.2%	11.1%	100.0%
A3.15G	0.0%	44.4%	44.4%	11.1%	0.0%	100.0%
A3.16G	11.1%	22.2%	55.6%	11.1%	0.0%	100.0%
A3.17G	11.1%	22.2%	33.3%	22.2%	11.1%	100.0%
A3.18G	11.1%	22.2%	44.4%	22.2%	0.0%	100.0%

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
Nivel 2	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	4.2%	33.3%	62.5%	100.0%
A1.2R	0.0%	4.2%	8.3%	25.0%	62.5%	100.0%
A1.3R	4.2%	0.0%	0.0%	45.8%	50.0%	100.0%
A1.4R	0.0%	0.0%	8.3%	37.5%	54.2%	100.0%
A1.5R	0.0%	0.0%	25.0%	54.2%	20.8%	100.0%

A1.6R	0.0%	4.2%	12.5%	66.7%	16.7%	100.0%
A1.7R	0.0%	8.3%	8.3%	37.5%	45.8%	100.0%
A1.8R	8.7%	4.3%	21.7%	43.5%	21.7%	100.0%
A1.9R	0.0%	0.0%	16.7%	45.8%	37.5%	100.0%
A1.10R	0.0%	4.2%	29.2%	41.7%	25.0%	100.0%
A1.11R	0.0%	0.0%	37.5%	50.0%	12.5%	100.0%
A1.12R	0.0%	4.2%	33.3%	45.8%	16.7%	100.0%
A1.13R	0.0%	12.5%	16.7%	45.8%	25.0%	100.0%
A1.14R	4.2%	8.3%	12.5%	33.3%	41.7%	100.0%
A1.15R	0.0%	4.2%	16.7%	58.3%	20.8%	100.0%
A1.16R	0.0%	8.3%	54.2%	25.0%	12.5%	100.0%
A1.17R	4.2%	4.2%	45.8%	29.2%	16.7%	100.0%
A1.18R	4.2%	29.2%	45.8%	16.7%	4.2%	100.0%
A1.19R	25.0%	29.2%	25.0%	8.3%	12.5%	100.0%
A2.1L	0.0%	0.0%	0.0%	8.3%	91.7%	100.0%
A2.2L	0.0%	0.0%	0.0%	16.7%	83.3%	100.0%
A2.3L	0.0%	0.0%	8.3%	29.2%	62.5%	100.0%
A2.4L	0.0%	4.2%	4.2%	29.2%	62.5%	100.0%
A2.5L	0.0%	0.0%	0.0%	8.3%	91.7%	100.0%
A2.6L	0.0%	0.0%	0.0%	29.2%	70.8%	100.0%
A2.7L	4.2%	4.2%	8.3%	66.7%	16.7%	100.0%
A2.8L	0.0%	8.3%	25.0%	62.5%	4.2%	100.0%
A2.9L	0.0%	4.2%	12.5%	54.2%	29.2%	100.0%
A2.10L	0.0%	8.3%	12.5%	50.0%	29.2%	100.0%
A2.11L	0.0%	4.2%	16.7%	37.5%	41.7%	100.0%
A2.12L	0.0%	4.2%	20.8%	41.7%	33.3%	100.0%

A2.13L	4.2%	0.0%	20.8%	45.8%	29.2%	100.0%
A2.14L	0.0%	4.2%	20.8%	54.2%	20.8%	100.0%
A2.15L	0.0%	25.0%	29.2%	37.5%	8.3%	100.0%
A2.16L	0.0%	8.3%	29.2%	41.7%	20.8%	100.0%
A2.17L	4.2%	8.3%	20.8%	41.7%	25.0%	100.0%
A2.18L	8.3%	8.3%	16.7%	37.5%	29.2%	100.0%
A2.19L	0.0%	25.0%	33.3%	37.5%	4.2%	100.0%
A2.20L	0.0%	16.7%	20.8%	54.2%	8.3%	100.0%
A2.21L	0.0%	4.2%	50.0%	33.3%	12.5%	100.0%
A2.22L	4.2%	8.3%	50.0%	25.0%	12.5%	100.0%
A2.23L	0.0%	8.3%	37.5%	41.7%	12.5%	100.0%
A2.24L	0.0%	8.3%	41.7%	41.7%	8.3%	100.0%
A2.25L	0.0%	16.7%	25.0%	54.2%	4.2%	100.0%
A2.26L	4.2%	12.5%	50.0%	33.3%	0.0%	100.0%
A2.27L	0.0%	8.7%	39.1%	39.1%	13.0%	100.0%
A2.28L	4.2%	12.5%	33.3%	41.7%	8.3%	100.0%
A2.29L	8.3%	20.8%	37.5%	33.3%	0.0%	100.0%
A2.30L	12.5%	20.8%	29.2%	37.5%	0.0%	100.0%
A2.31L	12.5%	33.3%	33.3%	20.8%	0.0%	100.0%
A4.1G	0.0%	8.3%	29.2%	58.3%	4.2%	100.0%
A4.2G	0.0%	0.0%	54.2%	41.7%	4.2%	100.0%
A4.3G	0.0%	8.3%	50.0%	37.5%	4.2%	100.0%
A4.4G	0.0%	4.2%	58.3%	29.2%	8.3%	100.0%
A4.5G	0.0%	0.0%	58.3%	33.3%	8.3%	100.0%
A4.6G	0.0%	4.2%	41.7%	41.7%	12.5%	100.0%
A4.7G	0.0%	0.0%	58.3%	29.2%	12.5%	100.0%

A4.8G	0.0%	4.2%	45.8%	45.8%	4.2%	100.0%
A4.9G	0.0%	8.3%	58.3%	29.2%	4.2%	100.0%
A4.10G	0.0%	12.5%	54.2%	33.3%	0.0%	100.0%
A4.11G	4.2%	12.5%	45.8%	37.5%	0.0%	100.0%
A4.12G	0.0%	12.5%	33.3%	50.0%	4.2%	100.0%
A4.13G	0.0%	12.5%	33.3%	45.8%	8.3%	100.0%
A4.14G	0.0%	8.3%	62.5%	29.2%	0.0%	100.0%
A4.15G	0.0%	8.3%	54.2%	37.5%	0.0%	100.0%
A4.16G	0.0%	16.7%	58.3%	25.0%	0.0%	100.0%
A4.17G	4.2%	16.7%	41.7%	20.8%	16.7%	100.0%
A4.18G	4.2%	20.8%	37.5%	29.2%	8.3%	100.0%
A4.19G	4.2%	25.0%	45.8%	16.7%	8.3%	100.0%
A4.20G	4.2%	29.2%	45.8%	20.8%	0.0%	100.0%
A4.21G	4.2%	25.0%	25.0%	45.8%	0.0%	100.0%
A4.22G	4.2%	20.8%	58.3%	8.3%	8.3%	100.0%
A4.23G	4.2%	29.2%	50.0%	12.5%	4.2%	100.0%
A4.24G	0.0%	29.2%	25.0%	33.3%	12.5%	100.0%
A4.25G	20.8%	12.5%	54.2%	8.3%	4.2%	100.0%

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	12	Total
Nivel 3	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	3.8%	46.2%	50.0%	0.0%	100.0%
A1.2R	0.0%	0.0%	3.8%	42.3%	53.8%	0.0%	100.0%

A1.3R	0.0%	0.0%	7.7%	61.5%	30.8%	0.0%	100.0%
A1.4R	0.0%	0.0%	7.7%	38.5%	53.8%	0.0%	100.0%
A1.5R	0.0%	0.0%	30.8%	30.8%	38.5%	0.0%	100.0%
A1.6R	0.0%	0.0%	0.0%	42.3%	57.7%	0.0%	100.0%
A1.7R	0.0%	0.0%	15.4%	57.7%	26.9%	0.0%	100.0%
A1.8R	0.0%	3.8%	19.2%	69.2%	7.7%	0.0%	100.0%
A1.9R	0.0%	0.0%	23.1%	46.2%	30.8%	0.0%	100.0%
A1.10R	0.0%	0.0%	34.6%	34.6%	30.8%	0.0%	100.0%
A1.11R	0.0%	0.0%	15.4%	53.8%	26.9%	3.8%	100.0%
A1.12R	0.0%	0.0%	38.5%	42.3%	19.2%	0.0%	100.0%
A1.13R	0.0%	0.0%	44.0%	44.0%	12.0%	0.0%	100.0%
A1.14R	0.0%	0.0%	23.1%	57.7%	19.2%	0.0%	100.0%
A1.15R	0.0%	3.8%	19.2%	53.8%	23.1%	0.0%	100.0%
A1.16R	0.0%	0.0%	15.4%	65.4%	19.2%	0.0%	100.0%
A1.17R	0.0%	0.0%	11.5%	65.4%	23.1%	0.0%	100.0%
A1.18R	0.0%	0.0%	50.0%	46.2%	3.8%	0.0%	100.0%
A1.19R	0.0%	0.0%	53.8%	38.5%	7.7%	0.0%	100.0%
A1.20R	0.0%	23.1%	57.7%	15.4%	3.8%	0.0%	100.0%
A1.21R	3.8%	23.1%	57.7%	11.5%	3.8%	0.0%	100.0%
A2.1L	0.0%	0.0%	3.8%	34.6%	61.5%	0.0%	100.0%
A2.2L	0.0%	0.0%	11.5%	38.5%	50.0%	0.0%	100.0%
A2.3L	0.0%	0.0%	15.4%	46.2%	38.5%	0.0%	100.0%
A2.4L	0.0%	0.0%	3.8%	57.7%	38.5%	0.0%	100.0%
A2.5L	0.0%	0.0%	19.2%	65.4%	15.4%	0.0%	100.0%
A2.6L	0.0%	3.8%	15.4%	57.7%	23.1%	0.0%	100.0%
A2.7L	0.0%	0.0%	30.8%	50.0%	19.2%	0.0%	100.0%

A2.8L	0.0%	0.0%	26.9%	57.7%	15.4%	0.0%	100.0%
A2.9L	0.0%	0.0%	30.8%	57.7%	11.5%	0.0%	100.0%
A2.10L	0.0%	0.0%	26.9%	53.8%	19.2%	0.0%	100.0%
A2.11L	0.0%	0.0%	11.5%	73.1%	15.4%	0.0%	100.0%
A2.12L	0.0%	4.0%	24.0%	48.0%	24.0%	0.0%	100.0%
A2.13L	0.0%	15.4%	46.2%	30.8%	7.7%	0.0%	100.0%
A2.14L	0.0%	8.0%	24.0%	52.0%	16.0%	0.0%	100.0%
A2.15L	0.0%	12.0%	44.0%	40.0%	4.0%	0.0%	100.0%
A2.16L	0.0%	0.0%	48.0%	48.0%	4.0%	0.0%	100.0%
A2.17L	0.0%	0.0%	28.0%	64.0%	8.0%	0.0%	100.0%
A2.18L	0.0%	4.0%	44.0%	52.0%	0.0%	0.0%	100.0%
A2.19L	0.0%	4.0%	36.0%	48.0%	12.0%	0.0%	100.0%
A2.20L	0.0%	0.0%	43.5%	43.5%	13.0%	0.0%	100.0%
A2.21L	0.0%	0.0%	52.0%	48.0%	0.0%	0.0%	100.0%
A2.22L	0.0%	8.3%	20.8%	66.7%	4.2%	0.0%	100.0%
A2.23L	0.0%	8.0%	48.0%	40.0%	4.0%	0.0%	100.0%
A2.24L	0.0%	16.0%	44.0%	36.0%	4.0%	0.0%	100.0%
A2.25L	0.0%	4.0%	52.0%	40.0%	4.0%	0.0%	100.0%
A2.26L	0.0%	16.0%	40.0%	36.0%	8.0%	0.0%	100.0%
A2.27L	4.0%	20.0%	64.0%	8.0%	4.0%	0.0%	100.0%
A2.28L	12.0%	12.0%	64.0%	12.0%	0.0%	0.0%	100.0%
A2.29L	24.0%	24.0%	48.0%	0.0%	4.0%	0.0%	100.0%
A4.1G	0.0%	4.0%	20.0%	52.0%	24.0%	0.0%	100.0%
A4.2G	0.0%	4.0%	24.0%	60.0%	12.0%	0.0%	100.0%
A4.3G	0.0%	4.0%	44.0%	40.0%	12.0%	0.0%	100.0%
A4.4G	0.0%	4.2%	25.0%	58.3%	12.5%	0.0%	100.0%

A4.5G	0.0%	4.0%	52.0%	40.0%	4.0%	0.0%	100.0%
A4.6G	4.0%	0.0%	40.0%	44.0%	12.0%	0.0%	100.0%
A4.7G	0.0%	4.0%	44.0%	48.0%	4.0%	0.0%	100.0%
A4.8G	4.0%	0.0%	52.0%	40.0%	4.0%	0.0%	100.0%
A4.9G	0.0%	4.0%	24.0%	68.0%	4.0%	0.0%	100.0%
A4.10G	0.0%	4.0%	20.0%	68.0%	8.0%	0.0%	100.0%
A4.11G	0.0%	4.0%	28.0%	60.0%	8.0%	0.0%	100.0%
A4.12G	0.0%	12.0%	48.0%	32.0%	8.0%	0.0%	100.0%
A4.13G	0.0%	16.7%	41.7%	37.5%	4.2%	0.0%	100.0%
A4.14G	0.0%	8.0%	68.0%	24.0%	0.0%	0.0%	100.0%
A4.15G	0.0%	8.0%	40.0%	48.0%	4.0%	0.0%	100.0%
A4.16G	0.0%	4.0%	32.0%	60.0%	4.0%	0.0%	100.0%
A4.17G	0.0%	4.0%	20.0%	48.0%	28.0%	0.0%	100.0%
A4.18G	0.0%	12.0%	40.0%	48.0%	0.0%	0.0%	100.0%
A4.19G	4.0%	16.0%	68.0%	12.0%	0.0%	0.0%	100.0%
A4.20G	4.0%	4.0%	52.0%	32.0%	8.0%	0.0%	100.0%
A4.21G	4.0%	16.0%	68.0%	12.0%	0.0%	0.0%	100.0%
A4.22G	4.2%	20.8%	54.2%	20.8%	0.0%	0.0%	100.0%
A4.23G	0.0%	16.7%	45.8%	37.5%	0.0%	0.0%	100.0%
A4.24G	0.0%	0.0%	62.5%	33.3%	4.2%	0.0%	100.0%
A4.25G	8.3%	25.0%	45.8%	20.8%	0.0%	0.0%	100.0%

Nivel 4	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	23	Total
---------	----------------	------------	----------	------------	----------------	----	-------

	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	2.0%	24.5%	73.5%	0.0%	100.0%
A1.2R	0.0%	0.0%	4.1%	16.3%	79.6%	0.0%	100.0%
A1.3R	0.0%	0.0%	2.1%	33.3%	64.6%	0.0%	100.0%
A1.4R	0.0%	0.0%	20.4%	42.9%	36.7%	0.0%	100.0%
A1.5R	2.0%	0.0%	4.1%	55.1%	38.8%	0.0%	100.0%
A1.6R	2.0%	2.0%	14.3%	42.9%	38.8%	0.0%	100.0%
A1.7R	0.0%	2.1%	12.5%	41.7%	43.8%	0.0%	100.0%
A1.8R	0.0%	0.0%	16.3%	46.9%	36.7%	0.0%	100.0%
A1.9R	0.0%	0.0%	26.5%	40.8%	32.7%	0.0%	100.0%
A1.10R	0.0%	2.0%	16.3%	42.9%	38.8%	0.0%	100.0%
A1.11R	0.0%	2.0%	14.3%	40.8%	42.9%	0.0%	100.0%
A1.12R	0.0%	0.0%	8.2%	46.9%	44.9%	0.0%	100.0%
A1.13R	0.0%	2.0%	6.1%	44.9%	46.9%	0.0%	100.0%
A1.14R	0.0%	0.0%	40.8%	40.8%	18.4%	0.0%	100.0%
A1.15R	0.0%	0.0%	26.5%	44.9%	28.6%	0.0%	100.0%
A1.16R	0.0%	0.0%	28.6%	51.0%	20.4%	0.0%	100.0%
A1.17R	0.0%	0.0%	24.5%	38.8%	36.7%	0.0%	100.0%
A1.18R	0.0%	6.1%	40.8%	44.9%	8.2%	0.0%	100.0%
A1.19R	0.0%	2.0%	32.7%	38.8%	26.5%	0.0%	100.0%
A2.1L	0.0%	0.0%	2.0%	12.2%	85.7%	0.0%	100.0%
A2.2L	0.0%	0.0%	10.2%	28.6%	61.2%	0.0%	100.0%
A2.3L	0.0%	0.0%	18.4%	34.7%	46.9%	0.0%	100.0%
A2.4L	0.0%	4.1%	32.7%	38.8%	24.5%	0.0%	100.0%
A2.5L	0.0%	2.0%	10.2%	55.1%	32.7%	0.0%	100.0%

A2.6L	0.0%	4.1%	12.2%	51.0%	32.7%	0.0%	100.0%
A2.7L	0.0%	2.0%	12.2%	49.0%	36.7%	0.0%	100.0%
A2.8L	4.1%	8.2%	36.7%	38.8%	12.2%	0.0%	100.0%
A2.9L	0.0%	0.0%	14.3%	44.9%	40.8%	0.0%	100.0%
A2.10L	0.0%	6.1%	28.6%	34.7%	30.6%	0.0%	100.0%
A2.11L	2.0%	6.1%	24.5%	42.9%	24.5%	0.0%	100.0%
A2.12L	2.0%	4.1%	22.4%	44.9%	26.5%	0.0%	100.0%
A2.13L	2.0%	4.1%	28.6%	42.9%	22.4%	0.0%	100.0%
A2.14L	4.1%	0.0%	22.4%	38.8%	34.7%	0.0%	100.0%
A2.15L	0.0%	0.0%	18.4%	53.1%	28.6%	0.0%	100.0%
A2.16L	0.0%	2.0%	22.4%	49.0%	26.5%	0.0%	100.0%
A2.17L	0.0%	8.2%	12.2%	59.2%	20.4%	0.0%	100.0%
A2.18L	0.0%	4.1%	14.3%	63.3%	16.3%	2.0%	100.0%
A2.19L	0.0%	2.1%	8.3%	58.3%	31.3%	0.0%	100.0%
A2.20L	0.0%	6.1%	26.5%	40.8%	26.5%	0.0%	100.0%
A2.21L	2.0%	2.0%	28.6%	46.9%	20.4%	0.0%	100.0%
A2.22L	0.0%	2.0%	30.6%	51.0%	16.3%	0.0%	100.0%
A2.23L	2.0%	16.3%	34.7%	28.6%	18.4%	0.0%	100.0%
A2.24L	2.1%	2.1%	29.2%	50.0%	16.7%	0.0%	100.0%
A2.25L	2.1%	8.3%	35.4%	29.2%	25.0%	0.0%	100.0%
A2.26L	8.2%	16.3%	40.8%	28.6%	6.1%	0.0%	100.0%
A2.27L	8.2%	22.4%	36.7%	26.5%	6.1%	0.0%	100.0%
A2.28L	14.3%	24.5%	38.8%	16.3%	6.1%	0.0%	100.0%
A4.1G	0.0%	2.0%	22.4%	46.9%	28.6%	0.0%	100.0%
A4.2G	0.0%	2.0%	24.5%	51.0%	22.4%	0.0%	100.0%
A4.3G	0.0%	4.1%	36.7%	42.9%	16.3%	0.0%	100.0%

A4.4G	0.0%	2.0%	18.4%	38.8%	40.8%	0.0%	100.0%
A4.5G	0.0%	0.0%	16.3%	57.1%	26.5%	0.0%	100.0%
A4.6G	0.0%	2.0%	26.5%	46.9%	24.5%	0.0%	100.0%
A4.7G	2.0%	0.0%	38.8%	42.9%	16.3%	0.0%	100.0%
A4.8G	0.0%	8.2%	40.8%	36.7%	14.3%	0.0%	100.0%
A4.9G	0.0%	4.1%	55.1%	32.7%	8.2%	0.0%	100.0%
A4.10G	2.0%	2.0%	40.8%	26.5%	28.6%	0.0%	100.0%
A4.11G	0.0%	4.1%	22.4%	51.0%	22.4%	0.0%	100.0%
A4.12G	0.0%	2.0%	12.2%	46.9%	38.8%	0.0%	100.0%
A4.13G	0.0%	4.1%	40.8%	36.7%	18.4%	0.0%	100.0%
A4.14G	0.0%	8.2%	44.9%	40.8%	6.1%	0.0%	100.0%
A4.15G	2.1%	2.1%	23.4%	48.9%	23.4%	0.0%	100.0%
A4.16G	2.0%	6.1%	20.4%	51.0%	20.4%	0.0%	100.0%
A4.17G	2.0%	4.1%	28.6%	53.1%	12.2%	0.0%	100.0%
A4.18G	2.0%	6.1%	49.0%	36.7%	6.1%	0.0%	100.0%
A4.19G	2.0%	0.0%	22.4%	51.0%	24.5%	0.0%	100.0%
A4.20G	0.0%	2.0%	40.8%	46.9%	10.2%	0.0%	100.0%
A4.21G	0.0%	2.0%	26.5%	55.1%	16.3%	0.0%	100.0%
A4.22G	0.0%	8.2%	44.9%	36.7%	10.2%	0.0%	100.0%
A4.23G	0.0%	10.2%	36.7%	42.9%	10.2%	0.0%	100.0%
A4.24G	0.0%	6.1%	24.5%	51.0%	18.4%	0.0%	100.0%
A4.25G	2.0%	26.5%	40.8%	22.4%	8.2%	0.0%	100.0%

Nivel 5	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
---------	----------------	------------	----------	------------	----------------	-------

	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A1.2R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.3R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.4R	0.0%	0.0%	10.0%	30.0%	60.0%	100.0%
A1.5R	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A1.6R	0.0%	0.0%	20.0%	10.0%	70.0%	100.0%
A1.7R	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A1.8R	0.0%	0.0%	20.0%	20.0%	60.0%	100.0%
A1.9R	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A1.10R	0.0%	0.0%	20.0%	40.0%	40.0%	100.0%
A1.11R	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A1.12R	0.0%	10.0%	10.0%	50.0%	30.0%	100.0%
A1.13R	0.0%	0.0%	30.0%	30.0%	40.0%	100.0%
A1.14R	0.0%	10.0%	40.0%	30.0%	20.0%	100.0%
A1.15R	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A1.16R	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A1.17R	0.0%	0.0%	30.0%	30.0%	40.0%	100.0%
A1.18R	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A1.19R	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A1.20R	0.0%	20.0%	10.0%	40.0%	30.0%	100.0%
A1.21R	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A2.1L	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A2.2L	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A2.3L	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%

A2.4L	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A2.5L	0.0%	0.0%	40.0%	50.0%	10.0%	100.0%
A2.6L	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A2.7L	0.0%	0.0%	0.0%	50.0%	50.0%	100.0%
A2.8L	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A2.9L	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A2.10L	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A2.11L	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A2.12L	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A2.13L	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A2.14L	0.0%	0.0%	10.0%	70.0%	20.0%	100.0%
A2.15L	0.0%	0.0%	20.0%	40.0%	40.0%	100.0%
A2.16L	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A2.17L	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A2.18L	0.0%	0.0%	30.0%	20.0%	50.0%	100.0%
A2.19L	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A2.20L	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A2.21L	0.0%	0.0%	10.0%	70.0%	20.0%	100.0%
A2.22L	0.0%	0.0%	40.0%	40.0%	20.0%	100.0%
A2.23L	0.0%	0.0%	30.0%	30.0%	40.0%	100.0%
A2.24L	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A2.25L	0.0%	0.0%	50.0%	30.0%	20.0%	100.0%
A2.26L	0.0%	20.0%	20.0%	50.0%	10.0%	100.0%
A2.27L	0.0%	10.0%	40.0%	50.0%	0.0%	100.0%
A4.1G	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A4.2G	0.0%	0.0%	40.0%	40.0%	20.0%	100.0%

A4.3G	0.0%	10.0%	20.0%	50.0%	20.0%	100.0%
A4.4G	0.0%	10.0%	20.0%	50.0%	20.0%	100.0%
A4.5G	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A4.6G	0.0%	10.0%	30.0%	40.0%	20.0%	100.0%
A4.7G	0.0%	0.0%	60.0%	10.0%	30.0%	100.0%
A4.8G	0.0%	10.0%	10.0%	50.0%	30.0%	100.0%
A4.9G	0.0%	0.0%	50.0%	10.0%	40.0%	100.0%
A4.10G	0.0%	0.0%	60.0%	20.0%	20.0%	100.0%
A4.11G	0.0%	20.0%	60.0%	10.0%	10.0%	100.0%
A4.12G	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A4.13G	0.0%	10.0%	60.0%	20.0%	10.0%	100.0%
A4.14G	0.0%	0.0%	0.0%	70.0%	30.0%	100.0%
A4.15G	10.0%	10.0%	50.0%	20.0%	10.0%	100.0%
A4.16G	10.0%	10.0%	40.0%	30.0%	10.0%	100.0%
A4.17G	10.0%	10.0%	50.0%	20.0%	10.0%	100.0%
A4.18G	0.0%	10.0%	60.0%	10.0%	20.0%	100.0%
A4.19G	0.0%	20.0%	40.0%	10.0%	30.0%	100.0%
A4.20G	0.0%	10.0%	30.0%	40.0%	20.0%	100.0%
A4.21G	10.0%	20.0%	30.0%	30.0%	10.0%	100.0%
A4.22G	0.0%	10.0%	50.0%	20.0%	20.0%	100.0%
A4.23G	0.0%	0.0%	50.0%	30.0%	20.0%	100.0%
A4.24G	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A4.25G	10.0%	10.0%	60.0%	10.0%	10.0%	100.0%

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
Nivel 6	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	10.0%	0.0%	0.0%	15.0%	75.0%	100.0%
A1.2R	5.0%	0.0%	0.0%	20.0%	75.0%	100.0%
A1.3R	5.0%	0.0%	0.0%	25.0%	70.0%	100.0%
A1.4R	5.0%	0.0%	0.0%	40.0%	55.0%	100.0%
A1.5R	5.0%	10.0%	10.0%	50.0%	25.0%	100.0%
A1.6R	5.0%	0.0%	10.0%	55.0%	30.0%	100.0%
A1.7R	5.0%	0.0%	20.0%	35.0%	40.0%	100.0%
A1.8R	5.0%	0.0%	10.0%	35.0%	50.0%	100.0%
A1.9R	5.0%	10.0%	15.0%	65.0%	5.0%	100.0%
A1.10R	0.0%	5.0%	20.0%	50.0%	25.0%	100.0%
A1.11R	5.0%	0.0%	25.0%	30.0%	40.0%	100.0%
A1.12R	10.0%	0.0%	30.0%	40.0%	20.0%	100.0%
A1.13R	5.3%	10.5%	47.4%	31.6%	5.3%	100.0%
A1.14R	0.0%	5.0%	10.0%	45.0%	40.0%	100.0%
A1.15R	5.0%	5.0%	25.0%	55.0%	10.0%	100.0%
A1.16R	10.0%	0.0%	10.0%	65.0%	15.0%	100.0%
A1.17R	5.0%	10.0%	35.0%	30.0%	20.0%	100.0%
A1.18R	10.0%	5.0%	25.0%	55.0%	5.0%	100.0%
A1.19R	0.0%	10.0%	40.0%	35.0%	15.0%	100.0%
A1.20R	10.0%	5.0%	15.0%	50.0%	20.0%	100.0%
A1.21R	10.0%	0.0%	10.0%	70.0%	10.0%	100.0%
A2.1L	5.0%	5.0%	0.0%	0.0%	90.0%	100.0%

A2.2L	5.0%	5.0%	0.0%	15.0%	75.0%	100.0%
A2.3L	5.0%	5.0%	0.0%	20.0%	70.0%	100.0%
A2.4L	5.0%	5.0%	5.0%	50.0%	35.0%	100.0%
A2.5L	0.0%	10.0%	30.0%	45.0%	15.0%	100.0%
A2.6L	0.0%	5.0%	15.0%	40.0%	40.0%	100.0%
A2.7L	10.0%	0.0%	20.0%	20.0%	50.0%	100.0%
A2.8L	10.0%	0.0%	5.0%	55.0%	30.0%	100.0%
A2.9L	5.0%	5.0%	5.0%	60.0%	25.0%	100.0%
A2.10L	10.0%	0.0%	0.0%	50.0%	40.0%	100.0%
A2.11L	5.0%	5.0%	10.0%	40.0%	40.0%	100.0%
A2.12L	0.0%	10.0%	20.0%	45.0%	25.0%	100.0%
A2.13L	5.0%	10.0%	20.0%	55.0%	10.0%	100.0%
A2.14L	0.0%	15.0%	5.0%	45.0%	35.0%	100.0%
A2.15L	5.0%	10.0%	15.0%	50.0%	20.0%	100.0%
A2.16L	0.0%	10.0%	50.0%	30.0%	10.0%	100.0%
A2.17L	0.0%	15.0%	45.0%	20.0%	20.0%	100.0%
A2.18L	0.0%	15.0%	25.0%	45.0%	15.0%	100.0%
A2.19L	0.0%	10.0%	20.0%	40.0%	30.0%	100.0%
A2.20L	5.0%	20.0%	25.0%	35.0%	15.0%	100.0%
A2.21L	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A2.22L	0.0%	20.0%	40.0%	20.0%	20.0%	100.0%
A2.23L	0.0%	25.0%	30.0%	35.0%	10.0%	100.0%
A2.24L	0.0%	20.0%	30.0%	40.0%	10.0%	100.0%
A2.25L	0.0%	15.0%	40.0%	30.0%	15.0%	100.0%
A2.26L	0.0%	10.0%	25.0%	55.0%	10.0%	100.0%
A2.27L	10.0%	20.0%	25.0%	35.0%	10.0%	100.0%

A2.28L	15.0%	20.0%	35.0%	20.0%	10.0%	100.0%
A4.1G	0.0%	0.0%	5.0%	30.0%	65.0%	100.0%
A4.2G	5.0%	10.0%	5.0%	35.0%	45.0%	100.0%
A4.3G	5.0%	0.0%	15.0%	50.0%	30.0%	100.0%
A4.4G	5.0%	5.0%	20.0%	40.0%	30.0%	100.0%
A4.5G	0.0%	10.0%	15.0%	50.0%	25.0%	100.0%
A4.6G	0.0%	0.0%	10.0%	60.0%	30.0%	100.0%
A4.7G	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A4.8G	0.0%	0.0%	20.0%	60.0%	20.0%	100.0%
A4.9G	0.0%	10.0%	35.0%	50.0%	5.0%	100.0%
A4.10G	0.0%	0.0%	20.0%	45.0%	35.0%	100.0%
A4.11G	0.0%	0.0%	35.0%	40.0%	25.0%	100.0%
A4.12G	0.0%	0.0%	30.0%	55.0%	15.0%	100.0%
A4.13G	0.0%	5.0%	15.0%	50.0%	30.0%	100.0%
A4.14G	0.0%	0.0%	20.0%	60.0%	20.0%	100.0%
A4.15G	0.0%	0.0%	35.0%	40.0%	25.0%	100.0%
A4.16G	0.0%	0.0%	25.0%	55.0%	20.0%	100.0%
A4.17G	0.0%	15.0%	15.0%	50.0%	20.0%	100.0%
A4.18G	0.0%	0.0%	25.0%	55.0%	20.0%	100.0%
A4.19G	0.0%	0.0%	15.0%	35.0%	50.0%	100.0%
A4.20G	0.0%	0.0%	35.0%	40.0%	25.0%	100.0%
A4.21G	0.0%	0.0%	40.0%	40.0%	20.0%	100.0%
A4.22G	0.0%	10.0%	50.0%	35.0%	5.0%	100.0%
A4.23G	0.0%	10.0%	5.0%	70.0%	15.0%	100.0%
A4.24G	0.0%	25.0%	35.0%	30.0%	10.0%	100.0%

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
Nivel 7	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
A1.2R	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A1.3R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.4R	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A1.5R	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A1.6R	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A1.7R	0.0%	10.0%	10.0%	50.0%	30.0%	100.0%
A1.8R	0.0%	0.0%	40.0%	40.0%	20.0%	100.0%
A1.9R	0.0%	10.0%	0.0%	30.0%	60.0%	100.0%
A1.10R	0.0%	10.0%	10.0%	70.0%	10.0%	100.0%
A1.11R	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A1.12R	0.0%	10.0%	0.0%	40.0%	50.0%	100.0%
A1.13R	0.0%	10.0%	10.0%	70.0%	10.0%	100.0%
A1.14R	0.0%	20.0%	50.0%	20.0%	10.0%	100.0%
A1.15R	0.0%	0.0%	20.0%	70.0%	10.0%	100.0%
A1.16R	0.0%	0.0%	0.0%	70.0%	30.0%	100.0%
A1.17R	0.0%	10.0%	0.0%	50.0%	40.0%	100.0%
A1.18R	0.0%	0.0%	60.0%	20.0%	20.0%	100.0%
A1.19R	10.0%	10.0%	30.0%	30.0%	20.0%	100.0%
A1.20R	10.0%	0.0%	30.0%	40.0%	20.0%	100.0%
A1.21R	20.0%	0.0%	40.0%	30.0%	10.0%	100.0%

A1.22R	10.0%	10.0%	20.0%	50.0%	10.0%	100.0%
A2.1L	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
A2.2L	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A2.3L	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A2.4L	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A2.5L	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A2.6L	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A2.7L	0.0%	0.0%	0.0%	50.0%	50.0%	100.0%
A2.8L	0.0%	0.0%	10.0%	10.0%	80.0%	100.0%
A2.9L	0.0%	10.0%	0.0%	50.0%	40.0%	100.0%
A2.10L	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A2.11L	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A2.12L	0.0%	10.0%	20.0%	30.0%	40.0%	100.0%
A2.13L	0.0%	20.0%	10.0%	30.0%	40.0%	100.0%
A2.14L	10.0%	0.0%	20.0%	40.0%	30.0%	100.0%
A2.15L	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A2.16L	0.0%	20.0%	30.0%	40.0%	10.0%	100.0%
A2.17L	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A2.18L	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A2.19L	0.0%	0.0%	40.0%	50.0%	10.0%	100.0%
A2.20L	0.0%	10.0%	20.0%	50.0%	20.0%	100.0%
A2.21L	0.0%	10.0%	40.0%	40.0%	10.0%	100.0%
A2.22L	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A2.23L	0.0%	30.0%	40.0%	10.0%	20.0%	100.0%
A2.24L	0.0%	20.0%	40.0%	30.0%	10.0%	100.0%
A2.25L	10.0%	20.0%	60.0%	0.0%	10.0%	100.0%

A2.26L	0.0%	30.0%	40.0%	20.0%	10.0%	100.0%
A2.27L	0.0%	20.0%	30.0%	30.0%	20.0%	100.0%
A2.28L	0.0%	10.0%	30.0%	40.0%	20.0%	100.0%
A4.1G	10.0%	0.0%	0.0%	40.0%	50.0%	100.0%
A4.2G	0.0%	10.0%	10.0%	40.0%	40.0%	100.0%
A4.3G	10.0%	0.0%	20.0%	20.0%	50.0%	100.0%
A4.4G	10.0%	0.0%	20.0%	40.0%	30.0%	100.0%
A4.5G	10.0%	20.0%	10.0%	20.0%	40.0%	100.0%
A4.6G	10.0%	10.0%	10.0%	20.0%	50.0%	100.0%
A4.7G	0.0%	10.0%	30.0%	20.0%	40.0%	100.0%
A4.8G	10.0%	0.0%	40.0%	20.0%	30.0%	100.0%
A4.9G	0.0%	10.0%	0.0%	50.0%	40.0%	100.0%
A4.10G	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A4.11G	10.0%	0.0%	30.0%	30.0%	30.0%	100.0%
A4.12G	10.0%	0.0%	50.0%	10.0%	30.0%	100.0%
A4.13G	0.0%	10.0%	20.0%	40.0%	30.0%	100.0%
A4.14G	10.0%	0.0%	40.0%	10.0%	40.0%	100.0%
A4.15G	0.0%	0.0%	30.0%	20.0%	50.0%	100.0%
A4.16G	0.0%	0.0%	40.0%	20.0%	40.0%	100.0%
A4.17G	10.0%	0.0%	10.0%	30.0%	50.0%	100.0%
A4.18G	10.0%	20.0%	20.0%	20.0%	30.0%	100.0%
A4.19G	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A4.20G	10.0%	0.0%	20.0%	30.0%	40.0%	100.0%
A4.21G	10.0%	0.0%	20.0%	20.0%	50.0%	100.0%
A4.22G	10.0%	0.0%	20.0%	40.0%	30.0%	100.0%
A4.23G	10.0%	10.0%	40.0%	10.0%	30.0%	100.0%

	Muy desacuerdo	Desacuerdo	Indeciso	De acuerdo	Muy de acuerdo	Total
Nivel 8	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila	% del N válido de fila
A1.1R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.2R	0.0%	0.0%	0.0%	10.0%	90.0%	100.0%
A1.3R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.4R	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A1.5R	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A1.6R	0.0%	0.0%	0.0%	20.0%	80.0%	100.0%
A1.7R	0.0%	0.0%	20.0%	10.0%	70.0%	100.0%
A1.8R	0.0%	0.0%	10.0%	30.0%	60.0%	100.0%
A1.9R	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A1.10R	0.0%	0.0%	0.0%	50.0%	50.0%	100.0%
A1.11R	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A1.12R	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A1.13R	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A1.14R	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A1.15R	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A1.16R	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A1.17R	0.0%	0.0%	20.0%	60.0%	20.0%	100.0%
A1.18R	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A1.19R	0.0%	0.0%	10.0%	30.0%	60.0%	100.0%
A2.1L	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%

A2.2L	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A2.3L	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A2.4L	0.0%	0.0%	0.0%	30.0%	70.0%	100.0%
A2.5L	0.0%	0.0%	20.0%	30.0%	50.0%	100.0%
A2.6L	0.0%	0.0%	0.0%	50.0%	50.0%	100.0%
A2.7L	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A2.8L	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A2.9L	0.0%	10.0%	30.0%	30.0%	30.0%	100.0%
A2.10L	0.0%	0.0%	10.0%	50.0%	40.0%	100.0%
A2.11L	0.0%	0.0%	20.0%	50.0%	30.0%	100.0%
A2.12L	0.0%	0.0%	0.0%	66.7%	33.3%	100.0%
A2.13L	0.0%	10.0%	30.0%	40.0%	20.0%	100.0%
A2.14L	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A2.15L	0.0%	20.0%	20.0%	60.0%	0.0%	100.0%
A2.16L	0.0%	0.0%	10.0%	70.0%	20.0%	100.0%
A2.17L	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A2.18L	0.0%	0.0%	20.0%	20.0%	60.0%	100.0%
A2.19L	0.0%	0.0%	20.0%	20.0%	60.0%	100.0%
A2.20L	0.0%	0.0%	30.0%	40.0%	30.0%	100.0%
A2.21L	0.0%	10.0%	20.0%	40.0%	30.0%	100.0%
A2.22L	0.0%	0.0%	40.0%	30.0%	30.0%	100.0%
A2.23L	0.0%	0.0%	40.0%	60.0%	0.0%	100.0%
A2.24L	0.0%	0.0%	10.0%	60.0%	30.0%	100.0%
A2.25L	0.0%	0.0%	20.0%	40.0%	40.0%	100.0%
A4.1G	0.0%	0.0%	0.0%	40.0%	60.0%	100.0%
A4.2G	10.0%	10.0%	10.0%	30.0%	40.0%	100.0%

A4.3G	10.0%	10.0%	20.0%	40.0%	20.0%	100.0%
A4.4G	0.0%	10.0%	30.0%	20.0%	40.0%	100.0%
A4.5G	10.0%	20.0%	20.0%	30.0%	20.0%	100.0%
A4.6G	10.0%	30.0%	20.0%	30.0%	10.0%	100.0%
A4.7G	0.0%	0.0%	0.0%	80.0%	20.0%	100.0%
A4.8G	0.0%	0.0%	10.0%	40.0%	50.0%	100.0%
A4.9G	0.0%	0.0%	20.0%	20.0%	60.0%	100.0%
A4.10G	0.0%	0.0%	0.0%	60.0%	40.0%	100.0%
A4.11G	10.0%	0.0%	30.0%	20.0%	40.0%	100.0%
A4.12G	0.0%	0.0%	10.0%	80.0%	10.0%	100.0%
A4.13G	0.0%	0.0%	30.0%	20.0%	50.0%	100.0%
A4.14G	0.0%	0.0%	20.0%	30.0%	50.0%	100.0%
A4.15G	0.0%	0.0%	30.0%	50.0%	20.0%	100.0%
A4.16G	0.0%	0.0%	50.0%	30.0%	20.0%	100.0%