

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo



M A E S T R Í A

“Desarrollo y validación de una prueba diagnóstica del idioma inglés con propósitos de publicación científica para estudiantes de posgrado”

Tesis
Que para obtener el grado de
Maestra en Ciencias Educativas

Presenta

Liliana Parra González

Director de tesis

Dr. Juan Carlos Rodríguez Macías

Ensenada, Baja California, México

Diciembre, 2024



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo
Maestría en Ciencias Educativas



**“Desarrollo y validación de una prueba
diagnóstica del idioma inglés con propósitos de
publicación científica para estudiantes de
posgrado”**

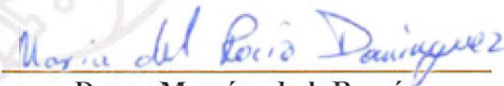
Tesis
Que para obtener el grado de
Maestra en Ciencias Educativas


Presenta
Liliana Parra González

Aprobado por:


Dr. Juan Carlos Rodríguez Macías
Director de tesis


Dr. Carlos David Díaz López
Sinodal


**Dra. María del Rocío
Dominguez Gaona**
Sinodal


Dra. Maday A. Coronel Santos
Sinodal



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 26 de noviembre de 2024

ASUNTO: Voto aprobatorio al trabajo
de tesis para el grado de Maestra en Ciencias Educativas

Dra. Rubí Surema Peniche Cetzal
Coordinadora de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la C. **Liliana Parra González**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Maestría en Ciencias Educativas, con el trabajo titulado:

“Desarrollo y validación de una prueba diagnóstica del idioma inglés con propósitos de publicación científica para estudiantes de posgrado”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dr. Juan Carlos Rodríguez Macías



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 26 de noviembre de 2024

ASUNTO: Voto aprobatorio al trabajo de tesis para el grado de Maestra en Ciencias Educativas

Dra. Rubí Surema Peniche Cetzal
Coordinadora de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Liliana Parra González**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Maestría en Ciencias Educativas, con el trabajo titulado:

“Desarrollo y validación de una prueba diagnóstica del idioma inglés con propósitos de publicación científica para estudiantes de posgrado”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dr. Carlos David Díaz López



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 26 de noviembre de 2024

ASUNTO: Voto aprobatorio al trabajo
de tesis para el grado de Maestra en Ciencias Educativas

Dra. Rubí Surema Peniche Cetzal
Coordinadora de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Liliana Parra González**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Maestría en Ciencias Educativas, con el trabajo titulado:

“Desarrollo y validación de una prueba diagnóstica del idioma inglés con propósitos de publicación científica para estudiantes de posgrado”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dra. María del Rocío Domínguez Gaona



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 26 de noviembre de 2024

ASUNTO: Voto aprobatorio al trabajo de tesis para el grado de Maestra en Ciencias Educativas

Dra. Rubí Surema Peniche Cetzal
Coordinadora de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Liliana Parra González**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Maestría en Ciencias Educativas, con el trabajo titulado:

“Desarrollo y validación de una prueba diagnóstica del idioma inglés con propósitos de publicación científica para estudiantes de posgrado”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dra. Maday Alicia Coronel Santos

Agradecimiento a CONAHCYT

Agradezco al Consejo Nacional de Humanidades, Ciencia y la Tecnología por su apoyo para dedicarme de tiempo completo como estudiante y poder culminar este trabajo de investigación.

Agradecimientos

A Dios, mi maestro que me acompaña en cada etapa.

A Hazael, mi esposo, compañero y motivación. Gracias por creer en mí y animarme a seguir cumpliendo sueños, nuestros sueños. Tu amor es mi motor.

A mi familia que siempre estuvieron al pendiente de esta etapa, acompañándome y dándome ánimos. Son un gran pilar en mi vida.

A mis compañeros de generación, su dedicación me inspiró. Su compañerismo trajo calidez a la vida en el posgrado, haciendo la experiencia mucho más placentera.

A mi tutor, Dr. Juan Carlos Rodríguez Macías, por su acompañamiento desde el primer día. Su disposición, paciencia y dedicación me dieron motivación para continuar y mejorar.

A mis sinodales, Dra. María del Rocío Domínguez Gaona, Dr. Carlos David Díaz López y Dra. Maday Alicia Coronel Santos, por sus comentarios tan puntuales, que me ayudaron en mi formación y siempre con la mira en la mejora. A la Dra. Edna Luna Serrano por sus valiosas aportaciones.

Al Instituto de Investigación y Desarrollo Educativo por la oportunidad de formar parte del programa de Maestría. No solo los aprendizajes se quedan conmigo, también todo el cuerpo de docentes y administrativos por su trato y calidez.

Índice

Agradecimiento a CONAHACYT	2
Agradecimientos.....	3
Resumen.....	6
Índice de tablas.....	7
Índice de figuras.....	8
Siglas y acrónimos.....	9
Introducción	11
Capítulo I. Planteamiento de la investigación.....	13
Antecedentes	13
Planteamiento del problema.....	19
Objetivos de la investigación	21
Justificación.....	22
Alcances	23
Capítulo II. Marco Referencial	24
2.1 Marco teórico	24
2.1.1 Lengua inglesa.....	25
2.1.2 Modelos y pruebas para evaluar el dominio del inglés.....	28
2.2 Marco Contextual	39
2.2.1 Enseñanza del inglés en el Sistema Educativo Mexicano	39
2.2.2 Contexto en UABC.....	43
2.3 Nociones sobre medición y validez.....	44
2.3.1 Taxonomía de Anderson y Krathwohl	47
2.3.2 Inteligencia Artificial Generativa.....	48
Capítulo III. Método	50
Diseño de la investigación.....	50
Fase I. Revisión Sistemática de la Literatura.....	51
Preguntas de orientación a la revisión sistemática.	51
Estrategia de búsqueda.	51
Criterios de elegibilidad.....	52
Selección de estudios.....	53
Estrategia de análisis de datos	53
Fase II. Diseño y desarrollo de la prueba.....	54
Etapa I. Marco General.....	55
Etapa II Definición de la variable de medida.....	55
Etapa III. Elaboración de Especificaciones.....	58
Etapa IV. Elaboración de los ítems.....	59

Etapa V. Validación de la prueba	61
Etapa VI. Edición	63
Etapa VII. Pilotaje.....	63
Fase III. Análisis de psicométricos y descriptivos de la prueba	67
Aspectos éticos	69
Capítulo IV. Resultados	70
Fase I. Revisión Sistemática de la Literatura.....	70
¿Cuáles son las propuestas para la evaluación del dominio del inglés en nivel superior?.....	70
¿En qué países se han aplicado?	71
¿Qué características tiene la población objetivo de las propuestas?.....	72
¿Qué metodología utilizan para el diseño y desarrollo de pruebas?.....	72
¿Qué pruebas evalúan el inglés con fines académicos?.....	73
¿Cuáles son los marcos o modelo más utilizados?	73
¿Qué habilidades o conocimientos evalúan las pruebas?.....	74
¿Cuáles son los fines del uso de las pruebas?	74
Fase II. Diseño y desarrollo de la prueba.....	75
Etapa 1. Marco general	75
Etapa 2. Definición de la variable de medida	76
Etapa 3. Especificaciones.....	78
Etapa 4. Elaboración de ítems	82
Etapa 5. Validación de contenido de la prueba.....	83
Etapa 6. Edición	86
Etapa 7. Pilotaje	87
Fase III. Análisis psicométricos y descriptivos de la prueba	89
Comportamiento de la prueba.....	89
Comportamiento de examinados ante la prueba.....	93
Uso de la Inteligencia Artificial para la evaluación de escritos.....	96
Capítulo V. Discusión y conclusiones	99
Operacionalizar los niveles de dominio del EAP	99
Desarrollar una prueba diagnóstica de inglés con propósitos de publicación científica.....	101
Conclusiones.....	102
Recomendaciones.....	103
Futuras líneas de investigación.	104
Referencias.....	105
Anexos.....	122

Resumen

El objetivo de este trabajo de investigación fue diseñar, desarrollar y obtener evidencias de validez de contenido de una prueba diagnóstica del dominio de inglés con propósitos de publicación de textos científicos para estudiantes de posgrado. La metodología empleada permitió la operacionalización de niveles de dominio del inglés académico con propósitos específicos y la integración de la inteligencia artificial como herramienta de evaluación de textos. Este estudio psicométrico contó con la participación de jueces expertos en los procesos de cribado, validación y evaluación; y 11 estudiantes de posgrado que contestaron a la prueba en el pilotaje, 7 a lápiz y papel y 2 en computadora.

El método se dividió en tres etapas: 1) revisión sistemática de la literatura, 2) desarrollo del instrumento, y 3) recopilación de evidencias de validez del aspecto de contenido.

Como parte de los resultados, se obtuvo una tabla de operacionalización de los niveles de dominio del Inglés con Propósitos Académicos (EPA), tabla de especificaciones y una prueba diagnóstica para evaluar el dominio de estudiantes de posgrado para escribir textos científicos. Además, se encontró que al emplear Inteligencia Artificial (IA) para evaluar los textos producidos por los examinados, ésta tiende a emitir evaluaciones más severas y homogéneas, lo que sugiere su potencial como herramienta complementaria en tareas de evaluación y retroalimentación de los estudiantes.

El principal resultado de la investigación fue el obtener una prueba que permite identificar áreas de oportunidad en la escritura científica, proporcionando insumos valiosos para el diseño de estrategias pedagógicas y didácticas. Asimismo, se propone la integración de inteligencia artificial para optimizar procesos de evaluación, reducir costos y fomentar investigaciones sobre la interacción entre juicios humanos y automáticos. Este trabajo tiene implicaciones significativas en el ámbito académico y social, al contribuir al fortalecimiento de las competencias de publicación científica y, con ello, al aumento del volumen y calidad de los artículos publicados por estudiantes de posgrado. Finalmente, el estudio abre líneas de investigación futuras en áreas como el diseño de pruebas especializadas para otros contextos, la creación de materiales didácticos en inglés académico, y la exploración de nuevas aplicaciones de la inteligencia artificial en la evaluación educativa.

Palabras clave: Inglés con Propósitos de Publicación Científica, inglés académico, evaluación educativa, psicometría.

Índice de tablas

- Tabla 1** *Comparativa de autores de habilidades de lectura y escritura académica*
- Tabla 2** *Tabla de competencias por categorías y habilidades*
- Tabla 3** *Tabla de descriptores generales por nivel en el MCER*
- Tabla 4** *Comparación de los niveles de dominio entre los diferentes modelos de evaluación.*
- Tabla 5** *Índices y cadena de búsqueda para la Revisión Sistemática de la Literatura*
- Tabla 6** *Refinamiento de la búsqueda para la Revisión Sistemática de la Literatura*
- Tabla 7** *Criterios y participantes del cribado de indicadores*
- Tabla 8** *Actividades en el análisis de datos*
- Tabla 9** *Cronograma de actividades para el desarrollo de la prueba*
- Tabla 10** *Resultados del cribado en la Revisión Sistemática de la Literatura*
- Tabla 11** *Marco general del instrumento*
- Tabla 12** *Características de los modelos y pruebas de referencia para el desarrollo de la prueba*
- Tabla 13** *Proceso de cribado de indicadores*
- Tabla 14** *Categorización de indicadores*
- Tabla 15** *Tabla resumen de habilidad de lectura con sus niveles taxonómicos por nivel de MCERL*
- Tabla 16** *Tabla de alineación de indicador con criterio evaluado en las rúbricas*
- Tabla 17** *Categorización de ítems*
- Tabla 18** *Resumen de coeficiente de Aiken y Tristán por subcriterio*
- Tabla 19** *Resumen de coeficientes de Aiken y Tristán por ítem*
- Tabla 20** *Concentrado de observaciones realizadas por jueces*
- Tabla 21** *Concentrado de entrevistas en pilotaje a lápiz y papel*
- Tabla 22** *Muestra de evaluación realizada por ChatGPT al examinado 5 en criterios de Main Idea e Introduction*
- Tabla 23** *Similitud entre criterios evaluados en la sección de escritura e ítems de lectura.*
- Tabla 24** *Comparación de evaluaciones por criterio en rúbrica de lenguaje entre humano e inteligencia artificial*
- Tabla 25** *Comparación de evaluaciones por criterio en rúbrica de lenguaje entre humano e inteligencia artificial*

Índice de figuras

Figura 1 *Etapas para el desarrollo del instrumento*

Figura 2 *Ejemplo de tabla de evaluación de indicadores*

Figura 3 *Ejemplo de tabla de ítems y taxonomía*

Figura 4 *Ejemplo de formato para la especificación de ítems e ítems muestra del examen de inglés con propósito de publicación científica en posgrado*

Figura 5 *Ejemplo de formato de evaluación para evidencias de validez de contenido*

Figura 6 *Resumen de rúbrica de evaluación de la lengua*

Figura 7 *Resumen de rúbrica de evaluación del resumen*

Figura 8 *Extracto de instrucciones en Chat GPT para la evaluación de escritos con rúbrica de convenciones y usos de la lengua*

Figura 9 *Instrucciones en Chat GPT para la evaluación de escritos con rúbrica de elaboración de un resumen de artículo científico*

Figura 10 *Instrucciones para contestar la prueba*

Figura 11 *Número de respuestas correctas por individuo y por ítem*

Figura 12 *Correlación Biserial Puntual y Valor P de Facilidad*

Figura 13 *Porcentaje de elección de cada alternativa para ítem 16*

Figura 14 *Porcentaje de elección de cada alternativa para ítem dos*

Figura 15 *Distribución de puntajes por tipo de programa*

Figura 16 *Distribución por tipo de cursos tomados*

Figura 17 *Distribución por años de estudio*

Figura 18 *Comparación de resultados reportados en TOEFL y en la sección de lectura en la prueba*

Figura 19 *Comparación de juicio de expertos e Inteligencia Artificial en rúbrica de lenguaje*

Figura 20 *Comparación de juicio de expertos e Inteligencia Artificial en rúbrica de resumen*

Figura 21 *Comparación de resultados en sección de escritura de aspecto de lenguaje evaluado por humanos y por la inteligencia artificial*

Figura 22 *Comparación de resultados en sección de escritura de aspecto de resumen evaluado por humanos y por la inteligencia artificial*

Siglas y acrónimos

ACERT - Academic Certificate of Competence in a Foreign Language
ACTFL - American Council on the Teaching of Foreign Languages
AERA - American Educational Research Association
ANUIES - Asociación Nacional de Universidades e Instituciones de Educación Superior
APA - American Psychological Association
CBI - Content-Based Instruction
CEMAI - Capacitación para la Escritura de Manuscritos en Inglés
CENNI - Certificación Nacional del Nivel de Inglés
CertAcles - Certificados de la Asociación de Centros de Lenguas en la Enseñanza Superior
CET - College English Test
CET-SET - College English Test–Spoken English Test
CLIL - Content and Language Integrated Learning
COMEPO - Consejo Mexicano de Estudios de Posgrado
CONAHCYT - Consejo Nacional de Humanidades, Ciencias y Tecnologías
DGAIR - Dirección General de Acreditación, Incorporación y Revalidación
DGDC - Dirección General de Desarrollo Curricular
DOF - Diario Oficial de la Federación
EAL- English as an Additional Language
EAP - English for Academic Purposes
EFL - English as a Foreign Language
EMI - English Medium Instruction
EOP - English for Occupational Purposes
ERPP - English for Research Publication Purposes
ESL - English as a second language
ESP - English for Specific Purposes
GAI – Generative Artificial Intelligence
GATE - Graduate Admission Test of English
GBI - Genre-based Instruction
GSE - Global Scale of English
IA - Inteligencia Artificial
IELTS - International English Language Testing System
IESOL - International English for Speakers of Other Languages

IIDE - Instituto de Investigación y Desarrollo Educativo
IMCO - Instituto Mexicano para la Competitividad
IPA - Inglés con Propósitos Académicos
IPE - Inglés con Propósitos Específicos
IVC - Índice de Validez de Contenido
KEPT - Khon Kaen University English Proficiency Test
MCERL - Marco Común Europeo de Referencia para las Lenguas
MMD - Miles de Millones de dólares
MUET - Malaysian University English Test
NCME - National Council on Measurement in Education
NESB - Non-English-Speaking Background
PAWT - Productive Academic Word Test
PELA - Post-Entry Language Assessment
PET Academic - Preliminary English Test Academic
PNIEB - Programa Nacional de Inglés en Educación Básica
PRONI - Programa Nacional de Inglés
RAE - Real Academia Española
RSL - Revisión Sistemática de la Literatura
RVC - Razón de Validez de Contenido
SEE - Sistema Educativo Estatal
SEP - Secretaría de Educación Pública
TBI - Task-based instruction
TCT - Teoría Clásica de Test
TOEFL - Test of English as a Foreign Language
TOEFL iBT - TOEFL Internet Based Test
TOEFL ITP - TOEFL Institutional Testing Program
UABC - Universidad Autónoma de Baja California
WELT - Warwick English Language Test

Introducción

De acuerdo con el Consejo Mexicano de Estudios de Posgrado, A. C. (COMEPO) (2015), solo el 59% de los programas de posgrado consideran el nivel de inglés como un requisito de ingreso, y el 25% como requisito de egreso. La importancia de la investigación que se presenta en este documento radica en aportar a los programas de posgrado en investigación de una prueba para evaluar de manera diagnóstica los conocimientos y habilidades de escritura en inglés de sus estudiantes, con el fin de identificar áreas de oportunidad para mejorar la escritura de textos científicos y la difusión de conocimiento. El desarrollo de una prueba propia para los estudios de posgrado permite responder a las necesidades del contexto de los estudiantes mexicanos y busca atender a las necesidades de las Instituciones de Educación Superior y los centros e institutos de investigación.

El propósito de la investigación que se presenta en este documento fue desarrollar y obtener evidencias de validez de una prueba diagnóstica que evalúe el dominio del idioma inglés con propósitos de publicación científica para estudiantes de posgrado. Durante el proceso de desarrollo de la prueba se realizaron aportaciones metodológicas para futuros desarrollos como la categorización de indicadores sobre las habilidades de lectura y escritura científica, la cual se realizó mediante la revisión de los principales referentes teóricos y modelos metodológicos predominantes. Esto se efectuó mediante una Revisión Sistemática de la Literatura, la cual también contribuyó en la identificación de las necesidades de la prueba y en la operacionalización de los niveles de dominio del Inglés con Propósitos Académicos.

Para cumplir con dicho objetivo se ha organizado el escrito en cinco capítulos. El primer capítulo describe el planteamiento de la investigación, el cual contiene 5 subapartados: antecedentes, planteamiento del problema, objetivos, justificación y alcances. En este apartado se exponen los antecedentes a la problemática que enfrentan tanto estudiantes de posgrado como algunos académicos al publicar textos científicos, en lo cual se sustenta la propuesta de investigación. Por otra parte, se describen los objetivos de investigación, los cuales se justifican al señalar los beneficios que el desarrollo de este tipo de prueba abona al campo social y educativo.

En el segundo capítulo se muestra el marco referencial en que se sustenta el estudio, el cual está dividido en 3 apartados: el marco teórico, el marco contextual y nociones de validación.

En el tercer capítulo se describe el método el cual se divide en dos principales

apartados: la Revisión Sistemática de la Literatura (RSL) y el diseño y validación del instrumento. El primer apartado presenta las preguntas para iniciar la RSL, las estrategias de búsqueda, los criterios de elegibilidad, selección de estudios y las estrategias de análisis de datos. El segundo apartado presenta la población para la cual fue diseñada la prueba, y el desarrollo del instrumento.

Como cuarto capítulo se presentan los hallazgos obtenidos de la revisión sistemática de la literatura, además se reportan los productos obtenidos del desarrollo y validación de la prueba. Como parte de la validez de contenido se presentan los resultados de los juicios que emitieron los expertos. También se presentan los resultados del pilotaje y los análisis psicométricos de los ítems y la prueba en su conjunto.

Finalmente, el quinto capítulo de conclusiones y discusión, presenta los apartados finales. Dentro de este apartado se discuten los principales hallazgos y aportaciones, además de las recomendaciones finales y las futuras líneas de investigación. En suma, el esfuerzo académico realizado permitió obtener una primera versión de una prueba diagnóstica que evalúa el dominio del idioma inglés con propósitos de publicación científica para estudiantes de posgrado, además de documentar el proceso de desarrollo y validación. Cabe destacar, que durante dicho proceso se exploró el uso de la inteligencia artificial para la evaluación de textos. Esto sin duda es una línea de investigación abierta que requiere mayor exploración e investigación, pero el estudio que se documenta presenta un primer avance en esta materia.

Capítulo I. Planteamiento de la investigación

Antecedentes

El inglés ha cobrado una mayor relevancia, colocándose en el 2022 como la lengua extranjera con mayor número de usuarios (Eberhard et al., 2022). Uno de los principales factores que ha contribuido en este proceso es la globalización. Para Carrillo-Beltrán et al. (2021, p.2) “la globalización es, sin lugar a duda, hoy en día uno de los principales factores que detonan el desarrollo económico, social, educativo en innovación tecnológica lo que hace que este idioma [el inglés] tenga relevancia en todas las naciones del mundo”.

La globalización la define Steger como “la expansión e intensificación de las relaciones sociales y de la conciencia a través del tiempo mundial y del espacio mundial” (2019, p.19), ante tal expansión e intensidad de relaciones a través del mundo, el inglés continuará asumiendo un rol preponderante en tales procesos, pero en la medida en que otros países asuman un papel protagónico en la economía mundial, su idioma ganará importancia en detrimento del uso del inglés como lengua de mayor expansión mundial.

Por ejemplo, China en el transcurso de las últimas 4 décadas, aumentó su relevancia global, de manera que es considerada como la segunda potencia económica y tecnológica mundial (Haro y Liaudat, 2021). A la par, el chino mandarín también ha crecido en número de usuarios. De acuerdo con el Instituto Cervantes (2022a), el chino mandarín se coloca como la lengua con más hablantes nativos, y la segunda en total de usuarios y utilizada en internet. Aunque Estados Unidos sigue siendo el principal productor de artículos científicos en Scopus, con aproximadamente 16 millones de documentos según *Scientifica Journals Ranking* (2024); China ha presentado un crecimiento ascendiendo a la segunda posición, con una producción de 10 millones de documentos.

Para Steger (2019, p.56) “la globalización de las lenguas puede entenderse como un proceso en virtud del cual algunos idiomas se utilizan cada vez más en la comunicación internacional mientras otros pierden importancia o incluso desaparecen por falta de hablantes”. De esta manera, se hace evidente el vínculo que existe entre la economía mundial y las lenguas. Además, presentan cinco puntos clave en el proceso de globalización de las lenguas. Primeramente, es la cantidad de lenguas, esto se debe a la influencia de culturas dominantes donde ciertas lenguas van perdiendo su posición en el mundo. El segundo, se refiere a la circulación de las personas en el mundo, lo cual promueve la difusión de los idiomas, esto en gran medida se debe a los patrones de migración o turismo, porque las

personas llevan la lengua consigo. En tercer lugar, está el aprendizaje de lenguas extranjeras, ya que igualmente propician la difusión de los idiomas. El cuarto punto se asocia al uso de las Tecnologías de la Información y Comunicación, ya que facilita el posicionamiento de las lenguas como un medio instantáneo, global y de acceso rápido a la información y la comunicación entre las personas. Por lo anterior, la comunicación internacional a través de medios digitales o virtuales es clave al momento de realizar un análisis de la influencia lingüística. El último punto clave se refiere a las publicaciones científicas internacionales, a decir de los mismos autores, estas “contienen los idiomas del discurso intelectual global, por lo que impactan de forma decisiva en las comunidades intelectuales implicadas en la producción, reproducción y circulación de conocimiento por todo el mundo” (2019, p.56).

El uso de la lengua inglesa ha permeado en distintos niveles, desde locales hasta internacionales, e incluso en distintas disciplinas como en los entornos académicos y universitarios, y de manera particular en la comunidad científica. Li y Flowerdew (2020) mencionan que, tanto académicos como estudiantes de investigación de todo el mundo, se ven cada vez más presionados a publicar en revistas internacionales donde se requiere un nivel alto de dominio del inglés. Para Cargill y Burgess (2008), la mayoría de las revistas de mayor reconocimiento, están centradas en dicho idioma, lo cual representa un reto para los científicos y estudiantes que tienen al Inglés como Lengua Adicional (EAL, por sus siglas en inglés).

La globalización también ha influido en la importancia que tiene el aprendizaje de las lenguas extranjeras en los Sistemas Educativos de los países, tal como lo menciona Canale (2011, p. 51), “la promoción de lenguas extranjeras en los sistemas educativos responde en gran medida a los discursos de globalización y a la conformación de bloques político-económicos regionales y continentales”. Por lo anterior, algunos países emprenden acciones para introducir estas ideas en sus sistemas educativos, por ejemplo, crean políticas públicas, planes o programas que respondan a la tendencia mundial, y lograr posicionar al inglés como una segunda lengua. Lo cual contribuye a que el inglés se posicione como la lengua extranjera con mayor número de estudiantes.

De igual manera, ante la influencia ejercida por la lengua inglesa, diversas disciplinas se han adaptado a las demandas que impone la globalización. Explica Hyland y Jiang (2021b) que en la década de 1960 el campo del Inglés con Propósitos Específicos (IPE), surgió como respuesta a la globalización de los mercados y a la expansión del Inglés como Lengua Franca en el ámbito comercial. Por lo tanto, ante dicho contexto, una gran cantidad de estudiantes y

profesionales de todo el mundo buscaron ganar fluidez comunicativa (verbal y escrita) en inglés. El IPE fue definido en sus inicios como una alineación de las necesidades de los alumnos con los métodos de enseñanza apropiados (Stevens, 1977, como se cita en Hyland y Jiang, 2021b), los cuales tienen sus inicios en la metodología basada en textos y sus características gramaticales. Además, el IPE se desprendió como una rama de la lingüística enfocada en el desarrollo tanto de la teoría como de la práctica innovadora en la enseñanza de idiomas. Así, se convirtió en un enfoque que busca dotar a los alumnos con todos los conocimientos lingüísticos, y sobre todo con las habilidades comunicativas necesarias para desenvolverse en ambientes particulares (sea el comercio, el turismo, política, etc.).

Las diversas visiones sobre las necesidades del alumno, son las que rigen el enfoque del inglés para propósitos específicos (las características del discurso y las estructuras del texto) y lo que impulsó su expansión a más áreas, ramas o géneros. Entre las ramas más destacadas se encuentran el Inglés con Propósitos Académicos (IPA) y el Inglés con Propósitos Ocupacionales, ambas igual de prolíficas en subdivisiones (Hyland y Jiang, 2021b).

Según Hyland y Jiang (2021a) el concepto de Inglés con Propósitos Académicos (IPA) fue acuñado por Tim Johns en 1974, y lo definieron como un enfoque de la lingüística que busca identificar las características específicas del lenguaje, las prácticas discursivas y habilidades comunicativas de los grupos académicos. Por lo tanto, pretende responder a las demandas de la academia y universidades, tanto en la lengua como el área de estudio. Su importancia crece en la medida en que las demandas en educación superior involucran la publicación de artículos en revistas bien posicionadas e investigación conjunta entre investigadores de diferentes países. Además, como ya se mencionó, la globalización e internacionalización ofrecen mayores oportunidades para aumentar el diálogo académico mediante la ampliación del corpus de la literatura académica, proporcionando nuevas vías para la investigación y colaboración, y la apertura de más canales de comunicación que informen sobre investigaciones específicas (Hyland, 2016).

Hyland (2016) argumenta que la creciente publicación internacional representa un desarrollo positivo, tanto para académicos como para los países en desarrollo que buscan formar parte de la "economía del conocimiento". Al respecto de la economía del conocimiento, Gaona et al. (2017, p. 66) mencionan que dentro de la economía del conocimiento "el conocimiento se reconoce ahora como el motor de la productividad y del crecimiento económico, dando lugar a un nuevo enfoque sobre el papel de la información, la

tecnología y el aprendizaje en el rendimiento económico.” Igualmente, los autores mencionan cuatro pilares de la economía del conocimiento: I. Incentivos económicos y regímenes institucionales, II. Educación y recursos humanos, III. Sistemas de innovación, y IV. Infraestructura de información.

Sin embargo, existe un reto dentro del ámbito académico, esto es dominar la lengua. Dentro de estas razones se encuentra la complejidad que demanda el inglés académico. Este se puede observar en un estudio realizado por Pearson (2015), en donde analizó y concentró todos los indicadores publicados por el Marco Común Europeo de Referencias (MCER) y otros autores, después cribó aquellos indicadores que estuvieran dirigidos a los objetivos del inglés académico. Como resultado observó que para las 4 habilidades (lectura, escucha, escritura y habla), la mayoría de los indicadores inician en un nivel intermedio, esto es, un B1 de acuerdo con el MCER, y con mayor concentración en niveles avanzados como el C1 y C2.

De esta misma manera, lo anterior se ve reflejado en las subramas del inglés académico como lo es el Inglés con Propósitos de Publicación de Investigación (ERPP, por sus siglas en inglés) (McDowell y Liardet, 2019). En ella se encuentran todos los componentes lingüísticos que corresponden a la publicación de textos científicos como lo son el lenguaje y la estructura, que son una particularidad de los textos, discursos, materiales y la comunidad en general.

De acuerdo con la revisión de la literatura realizada por Li y Flowerdew (2020), los enfoques pedagógicos emergentes en el ERPP están concentrados principalmente en la habilidad de la escritura. De manera colateral, se incluyen en algunos enfoques la habilidad de la lectura, como recurso de apoyo para los estudiantes por medio de artículos redactados por nativos de la lengua inglesa. En cuanto a las otras dos habilidades (escucha y habla), son consideradas solamente en universidades/instituciones con el inglés como medio de instrucción (EMI, por sus siglas en inglés). En estas modalidades, se utiliza la lengua para impartir las clases sin utilizar el idioma nativo de los estudiantes. Sin embargo, aún en estos casos, las habilidades dominantes, tanto en la instrucción como en la evaluación son la lectura y, sobre todo, la escritura.

Por parte de la escritura, Corcoran y Englander (2016), resaltan que en general la escritura en inglés es más difícil, y, por lo tanto, menos efectiva como segundo idioma para una persona. No sólo implica el desafío de producir datos científicos rigurosos, sino también redactar artículos que cumplan con las expectativas en inglés de los editores y revistas. Adicionalmente, escribir un artículo para una revista científicas de alto impacto, representa

un reto aún mayor para algunas personas, dado que les genera ansiedad y menos satisfacción cuando no es aceptado para su publicación, lo cual desmotiva la escritura de futuros productos científicos en inglés.

Henshall (2018), encontró que el inglés tiene tal impacto en la cultura científica, ya que se considera un componente importante en la internacionalización, no solo por la comunidad científica sino también dentro de revistas y bases de datos. Incluso, ha llegado a ser catalogado como una lengua hegemónica en la comunidad científica (Ferguson, 2007; Henshall, 2018). Sin embargo, a pesar de las críticas, es evidente que el inglés tiene sumo peso en la publicación y acceso de la información. Esta relevancia encontró eco en el Instituto Cervantes (2022b) y realizó una investigación al respecto. Los resultados demostraron que la producción científica en inglés concentra mayor impacto en relación con la producida en cualquier otro idioma e incluso, las revistas generadas en Estados Unidos son mayormente citadas con relación a las de otros países. Este hecho influye en las decisiones de los científicos que están interesados en publicar sus investigaciones, y los enfrenta a la necesidad de conocer el idioma y dominarlo para tales fines.

López-Leyva et al. (2018) señalan que el inglés representa un obstáculo importante para los docentes universitarios en México a la hora de publicar. Los principales obstáculos que identificaron son el dominio del idioma, ya que es difícil ser un experto para publicar en inglés; y los apoyos limitados que ofrecen las instituciones a las que pertenecen para capacitarse en la redacción en inglés, recibir asistencia de revisores en este idioma o brindarles servicios de traducción especializada.

En México, para atender a las necesidades de acreditar el idioma inglés en educación superior se ha recurrido a pruebas estandarizadas como el TOEFL INT, TOEFL IBT, IELTS. De entre las pruebas que se pueden mencionar está la Pearson Test of English Academic. La cual surgió por la necesidad de atender a las demandas de cursos específicos del idioma inglés y explora otras áreas específicas como el de profesionales, adultos y niños (Pearson, 2015)

Por otra parte, la prueba de TOEFL iBT presenta a los examinados, reactivos que simulan a los requeridos por las universidades que usan el inglés como medio de comunicación. A esta prueba se le han realizado estudios sobre evidencias de validez, y se ha utilizado como herramienta en la toma de decisiones, como la admisión de estudiantes (Malone y Montee, 2014).

Una de las pruebas más utilizadas a lo largo del mundo para medir el nivel de

competencia en el inglés es el International English Language Test System (IELTS), sobre todo para las personas interesadas en estudiar o trabajar en un país angloparlante. Hashemi y Daneshfar (2018) realizaron un estudio a 90,000 aplicantes y produjeron alrededor de 13 versiones para evaluar las habilidades de escucha y lectura, lo cual también sirvió para mostrar evidencias de validez de la prueba.

Ante la necesidad de pruebas contextualizadas y focalizadas que respondan a necesidades concretas, Zainal y Jamil (2015) desarrollaron una prueba contextualizada para Malasia, la cual pretende responder a demandas de estudiantes de posgrado que, a pesar de tener altos puntajes en pruebas como TOEFL o IELTS, no concuerda con su desempeño en clases, el cual era menor al esperado. Esta prueba, busca ajustarse al contexto académico, y al nivel superior, además de responder a las condiciones particulares del país y a las necesidades de los estudiantes de posgrado.

Por otra parte, Jin (2019) presenta la prueba College English Test (CET) de China. Esta prueba surgió ante la necesidad de incorporar al inglés como primera lengua extranjera en el país. En la actualidad esta prueba es obligatoria y es un requisito de egreso de la educación superior. Esta prueba incluye las cuatro habilidades mencionadas (lectura, escritura, habla y escucha), además de un apartado adicional de traducción.

Planteamiento del problema

La globalización es una serie de procesos sociales que cambian la situación nacional o local, y produce un estado global (Steger, 2019). En consecuencia, implica la transformación de las interacciones humanas y los diferentes ámbitos: social, cultural, económico y político. Estas interacciones han utilizado como medio de conexión una lengua que sirva como punto de encuentro. En la actualidad, el idioma inglés ha tomado ese lugar de conexión e interacción en todo el mundo, como lo menciona el Instituto Mexicano para la Competitividad (IMCO, 2015).

En México el inglés ha tomado también un rol decisivo. Esto se puede explicar en cierta medida por la cercanía con los Estados Unidos de Norteamérica, la frontera compartida y, su influencia a nivel mundial. Estos factores han influido en algunos cambios políticos que han repercutido en el ámbito educativo, y de manera particular en la inclusión del inglés en el currículum nacional.

La implementación de la lengua inglesa en el currículum mexicano demuestra la necesidad de formar ciudadanos que sean capaces de competir en un mundo cada vez más global. Sin embargo, el reporte realizado por el IMCO (2015), mostró que el dominio del inglés es bajo. En el caso de la educación superior, encontró que poco más del 50 % de las universidades no incluyen al inglés como materia obligatoria, y en 60 % no se exige como segundo idioma para terminar la carrera. No obstante, el reporte menciona que el 85 % de las universidades lo consideran como un requisito para obtener el grado. Por ejemplo, para obtener un título de licenciatura, el 70 % de las universidades exigen al menos un nivel B1 o B2.

Dicho informe planteó propuestas para las instituciones educativas respecto al inglés como segunda lengua, las principales fueron: a) articular el currículum de todos los niveles educativos, sobre todo en niveles de media superior y superior, ya que los contenidos cuentan con una articulación limitada ; b) establecer entre las Instituciones de Educación Superior (IES) “una agenda curricular que permita a los estudiantes obtener mayores competencias en el idioma” (p.42) para la formación y certificación del idioma entre los alumnos; y c) instaurar como requisito obligatorio el nivel B1 o B2 de acuerdo con el MCER. Dichas propuestas tenían como objetivo incrementar el dominio del inglés y con ello, la competitividad, la movilidad social, la empleabilidad y mejorar los sueldos de los estudiantes al egresar de su formación académica e incorporarse al mercado laboral. Además, el IMCO considera al inglés como una herramienta de acceso al conocimiento y reconoce la relevancia

que tiene en el mundo de los negocios.

De manera particular, esta problemática de bajo dominio del inglés afecta a quienes deciden estudiar un posgrado en investigación, y en mayor medida a los interesados en publicar artículos científicos. El 21.8% de los libros publicados en el mundo están redactados en inglés y poco más de 45% de las publicaciones académicas también están en esta lengua (IMCO, 2015). Por otra parte, egresar del nivel licenciatura con un bajo nivel de inglés, muchas veces se traduce en estudiantes que llegan al posgrado con un bajo nivel de dominio, y en palabras de Diaz-Castelazo “Ante este escenario, resulta evidente que el idioma inglés es fundamental para adquirir conocimiento” (2018, p.61). Así, las posibilidades de redactar artículos que respondan a las demandas de escribir en inglés se reducen o se convierte en una tarea difícil o imposible, lo cual puede generar estrés, frustración y desmotivación.

Sin embargo, de acuerdo con Encinas et al. (2020) en la actualidad “...en el contexto internacional, aún existen pocas propuestas pedagógicas”, y solo existen algunos estudios sobre las prácticas pedagógicas relacionadas con la publicación en inglés utilizadas en posgrados en México. Esto constituye una dificultad, sobre todo cuando se trata de superar el rezago que presentan los estudiantes en el dominio del inglés y mejorar sus condiciones de aprendizaje en el nivel de educación superior. Si a esta situación se le agrega la falta de políticas educativas y la poca atención a la población de estudiantes universitarios, resulta la necesidad de investigaciones que aborden estos problemas y aporten elementos para superar los retos. El superar los retos representa un beneficio no solo para la formación de estudiantes competitivos en un mundo global, sino también en la generación de conocimiento sobre la evaluación y la enseñanza del inglés. En ambos casos, el resultado final aporta al desarrollo del país y su participación en el mundo global. Otros beneficios pueden ser la internacionalización de las universidades y la movilidad de sus estudiantes, docentes e investigadores a países no hispanoparlantes para enriquecer su formación, actualización o aceleración de ciertos procesos de investigación al compartir laboratorios, tecnología y equipo especializado que no se dispone en México.

Si bien, el Consejo Nacional de Humanidades, Ciencia y Tecnologías (CONAHCYT), ante la necesidad del dominio de la lengua y aumentar la producción científica, diseñó el programa de Capacitación para la Escritura de Manuscritos Académicos en Inglés (CEMAI, s.f.), con la finalidad de capacitar a estudiantes y tutores para publicar en la lengua. Este esfuerzo no es suficiente, se requiere de una estrategia que considere la evaluación como elemento central en el proceso de enseñanza del inglés, al aportar la información relevante

sobre las áreas de oportunidad que tienen los estudiantes para lograr los aprendizajes que demanda la publicación de artículos científicos en revistas de alto impacto.

De acuerdo con Karimpour y Mazlum (2024) desde las últimas tres décadas han incrementado los intentos por aportar los fundamentos teóricos en la evaluación del Inglés con Propósitos Específicos y Académicos (ESP y EAP, por sus siglas en inglés). Sin embargo, en su mayoría se han enfocado en pruebas internacionales a gran escala centradas en competencias, dejando marginadas a las pruebas EAP con funciones importantes como logro, ubicación y diagnóstico.

Por otra parte, la gran mayoría de las pruebas utilizadas no están localizadas, como lo mencionan Zainal y Jamil (2015), es decir, no responden al contexto, población, necesidades, objetivos y dominio en los que será utilizada la prueba. Esto concuerda con el amplio uso de las pruebas estandarizadas, con funciones de alto impacto, como ingreso o egreso a programas educativos. Esto refleja el rezago en el desarrollo de pruebas diagnósticas centradas en el proceso de enseñanza del inglés, que aporten información sobre las áreas de mejora que tienen los estudiantes para lograr las habilidades lingüísticas que demanda la publicación de artículos científicos.

Objetivos de la investigación

Diseñar, desarrollar y validar una prueba diagnóstica del dominio del idioma inglés con propósitos de publicación científica en estudiantes de posgrado.

Para el logro de dicho objetivo se definieron tres objetivos específicos:

1. Operacionalizar los niveles de dominio IPA con enfoque en la publicación científica.
2. Obtener evidencias de validez de contenido y calidad técnica de la prueba.
3. Analizar las puntuaciones que obtienen los examinados en la prueba.

Justificación

México, con el fin de responder a las demandas globales, ha diseñado diversos planes, estrategias y programas en distintas administraciones federales. Sin embargo, Toledo y García (2020) señalan que cerca del 36% de los estudiantes que egresan de las universidades cumplen con el requisito mínimo del dominio del inglés (leerlo, hablarlo, escribirlo y comprenderlo). A través de la prueba diagnóstica que se desarrolló y validó en el presente estudio, se aporta un instrumento de evaluación que sirve para orientar las decisiones formativas de los futuros investigadores en un dominio muy importante para su desarrollo como lo es el inglés, lo cual contribuye a visibilizar su producción científica en el mundo y generar mayor conocimiento. Además, las evaluaciones diagnósticas son importantes porque aportan información para emprender acciones en el diseño y aplicación de políticas públicas en beneficio de la educación, la investigación y la difusión, divulgación, generación y transferencia de conocimiento. El estudio permitió revisar las acciones y esfuerzos realizados por las organizaciones, consejos e incluso las instituciones de educación superior que tienen el fin de aumentar la producción científica de investigadores mexicanos y su visibilidad a nivel internacional.

Los resultados de la aplicación de la prueba diagnóstica del dominio del idioma inglés con propósitos de publicación científica en estudiantes de posgrado, aporta información para diseños instruccionales más eficaces y centrados en las necesidades de los estudiantes mexicanos; además de dar cuenta del grado o nivel de dominio del inglés que tienen los estudiantes que cursan un posgrado con orientación hacia la investigación. Esto, a su vez, contribuye en cierta medida, al cumplimiento de los objetivos de CONAHCYT de internacionalizar los programas de posgrado en México.

Por otra parte, aunque este proyecto está contextualizado en la Universidad Autónoma de Baja California (UABC), el estudio pretende aportar una prueba diagnóstica para las diversas instituciones de educación superior a nivel nacional. De manera que este instrumento aporte a los procesos de evaluación del inglés en los posgrados y sea de utilidad para el diseño y aplicación de estrategias instruccionales que contribuyan en la formación de los estudiantes en torno al inglés como segunda lengua, particularmente en el ámbito académico y científico.

Por ejemplo, en el caso de UABC, el Plan de Desarrollo Institucional 2019-2023, promueve fortalecer esquemas para el aprendizaje y dominio del inglés (UABC, 2019b). De la misma manera, el Plan de Desarrollo Institucional 2023-2027 (UABC, 2023) presenta

estas estrategias en sus prioridades institucionales de desarrollo regional e internacionalización. Dichas estrategias están igualmente relacionadas a la colaboración con instituciones académicas internacionales, mejorar la visibilidad a nivel internacional, y consolidar la participación en conferencias internacionales y publicación en revistas científicas de prestigio (estrategias cinco y seis).

Además, en el Modelo Educativo de UABC establece el dominio de un idioma extranjero dentro de la normatividad (UABC, 2018), y otorga a las unidades académicas, las modalidades ofertadas y el nivel requerido de inglés al egreso del trayecto formativo (UABC, 2019a). Sin embargo, al dejar la instrucción de un segundo idioma bajo la coordinación de las unidades académicas (UABC, 2019a) ocasiona la heterogeneidad de exámenes, cursos, e incluso el nivel de competencia de egreso.

Con el uso de un instrumento de evaluación diagnóstica de la lengua es posible aportar en diferentes aspectos a las necesidades de alumnos en posgrados. Primeramente, conocer las necesidades de los alumnos, reconociendo la diferencia de contextos educativos. Además de aportar a objetivos de internacionalización, visibilidad y participación en revistas internacionales que, como ya se ha discutido, están principalmente en inglés. Finalmente, en aportar una base para el diseño de otras pruebas que estén personalizadas a las necesidades institucionales y de los alumnos.

Alcances

El presente trabajo se realizó como parte del proceso de acreditación para obtener el grado de maestro, por lo que, su diseño estuvo delimitado a un espacio temporal de dos años, del año 2022 al 2024, tiempo de duración del programa académico. El marco de dos años permitió solamente el pilotaje y el análisis de los datos recolectados con el fin de cumplir con las etapas de la ruta crítica del programa. Durante este periodo, la investigación se desarrolló en 3 fases: operacionalizar los niveles de dominio, desarrollar la prueba y obtener evidencias de validez.

Capítulo II. Marco Referencial

En este capítulo se presentan los referentes que sustentan el desarrollo del trabajo, sirviendo como la base conceptual para analizar, comprender y estructurar los diversos elementos implicados en el estudio. Se abordan temáticas clave relacionadas con el idioma inglés y su relevancia global, los enfoques pedagógicos para su enseñanza y evaluación, y los modelos teóricos que respaldan el diseño y validez de pruebas de medición en el ámbito académico.

El primer apartado explora el inglés como lengua franca, sus raíces históricas y su consolidación como idioma predominante en la comunicación internacional, académica y científica. Asimismo, se discuten las competencias lingüísticas específicas para la escritura científica, integrando referencias al Marco Común Europeo de Referencia para las Lenguas (MCER) y la Escala Global de Inglés (GSE). Estos marcos se analizan tanto desde su perspectiva general como desde su aplicación particular en habilidades de lectura y escritura.

Posteriormente, se introducen teorías psicométricas como la Teoría Clásica de los Tests y la Teoría de Respuesta al Ítem, fundamentales para evaluar la validez y confiabilidad de las herramientas diseñadas para medir competencias lingüísticas. En este contexto, se resalta la importancia de la validez de contenido y se presenta el uso de la Taxonomía de Anderson y Krathwohl como herramienta esencial para estructurar objetivos de aprendizaje y evaluaciones dentro del diseño del instrumento propuesto.

Este marco referencial no solo guía el análisis de los resultados, sino que también fundamenta las decisiones metodológicas tomadas a lo largo del estudio. Al integrar aspectos lingüísticos, pedagógicos y de evaluación, este capítulo establece una plataforma sólida para entender y abordar las complejidades del aprendizaje y evaluación del inglés en contextos académicos y de investigación.

2.1 Marco teórico

En el siguiente apartado se encuentran los referentes teóricos que sustentan este trabajo. Dentro de este primer marco encontraremos el apartado de la lengua inglesa donde se abordan sus orígenes como lengua franca, sus ramificaciones en la lingüística, y las diferencias entre lengua e idioma. Además, se discuten las competencias propuestas para la escritura científica, primeramente, de manera general en dos marcos (Marco Común Europeo y la Global Scale of English, propuesta por PEARSON), y luego de manera particular en las

habilidades de escritura y lectura. Por otra parte, se recapitulan los referentes teóricos en las pruebas, así como en las evidencias de validez, como la de contenido y las principales teorías que sustentarán el análisis de resultados como la Teoría de Respuesta al Ítem y la Teoría Clásica de los Test. Finalmente, se presenta la taxonomía de Anderson y Krathwohl, la cual fue relevante para el diseño del instrumento propuesto.

2.1.1 Lengua inglesa

2.1.1.1 El inglés como lengua franca. El término de lengua franca surgió como resultado de la globalización de las lenguas hace miles de años. Desde la Edad Media, los árabes nombraron a todos los europeos como Francos (tribu que es la hoy Francia) ya que, existía gran afluencia de intercambio económico árabe-europeo (Moscoso et al., 2019). De modo que para facilitar la comunicación se dio un proceso de globalización de las lenguas en el que los comerciantes desarrollaron una lengua común producto del francés, italiano, el árabe y otros idiomas, con el fin de ser entendidos de manera sencilla, la cual nombraron Lingua Franca o lengua franca.

Aunado a lo anterior, Brosch (2015) establece que el concepto de lengua franca tiene sus orígenes en los *pidgins*. Acorde con la Real Academia Española (RAE) *pidgins* se define como “lengua mixta, creada sobre la base de una lengua determinada y con la aportación de numerosos elementos de otra u otras, que usan especialmente en enclaves comerciales hablantes de diferentes idiomas para relacionarse entre sí” (s.f., definición 2). Es importante destacar que los *pidgins*, se constituyen por la mezcla de lenguas sin desplazar a ninguna otra, pero, como se rescata a lo largo de la historia, el dominio o procesos imperialistas posicionan a uno u otro idioma como la lengua franca. En el caso del griego, fue extendido por las conquistas de Alejandro Magno, e incluso el latín, traído por los romanos (Serrano, 2016).

En el caso del inglés, obtuvo gran alcance debido a la colonización británica a finales del siglo XVI y más adelante sería reforzada a nivel global por la revolución industrial, los avances tecnológicos y científicos. Tiempo después, sería Estados Unidos a finales del siglo XVIII quien aumentaría su economía al incrementar su poder hegemónico y con ello el uso del inglés; además del liderazgo industrial de Gran Bretaña en el siglo XX (Serrano, 2016). De esta manera, como apunta Brosch (2015), el término Lengua Franca ha sido fuertemente ligado al inglés y, como consecuencia, un nuevo campo de investigación. Dada la gran proliferación del idioma inglés se ha denominado como la *lingua franca* del siglo XXI, debido al poder que ejerce la industria cultural angloamericana en el mundo (Steger 2009,

como se cita en Garrido, 2010).

Lo anterior se refleja no solo en la demografía del uso de la lengua inglesa, sino también en la importancia que le dan las instituciones universitarias. Por ejemplo, casi la mitad de la población mundial de estudiantes extranjeros está matriculada en instituciones de países angloamericanos (Steger, 2019). Además, tiene una gran relevancia en el ámbito académico y en la comunidad científica, particularmente en la publicación de artículos científicos en revistas de gran relevancia mundial. De acuerdo con Garrido (2010, p.85) “aunque el volumen de traducción es inmenso, el flujo principal de la comunicación y directamente la investigación científica se producen en inglés”.

Dentro de las discusiones que se pueden presentar en el campo de la lingüística es la diferencias entre lengua e idioma. De acuerdo con el Diccionario del Español de México (DEM) (2024) la lengua se define como un sistema de signos fónicos o gráficos con el que se comunican los miembros de una comunidad humana. Por otra parte, el idioma lo definen como la lengua de un pueblo o nación, o lengua que lo caracteriza: idioma español, idiomas modernos. Si bien en la práctica son intercambiados ambos términos, existe una distinción, sobre todo en lenguas como el inglés que se ha conformado en distintos países como el idioma oficial de la región.

Dentro de este último fenómeno resalta el inglés americano o estadounidense y el británico, ya que derivado de aspectos propios de la región pueden existir diferencias léxicas entre ambos. Sin embargo, como muestra el estudio realizado por Gu y So (2015), existe una preferencia por el uso de variedades estándar de inglés para preguntas escritas. De esta manera y para fines de esta prueba, se optó por un inglés estándar, el cual no hace diferencia entre ambos, ya que en textos académicos se busca la universalidad del idioma y el fácil acceso de hablantes no nativos.

2.1.1.2 Ramas pedagógicas del inglés. Dentro del amplio abanico de posibilidades que representa la enseñanza del inglés, pasando por el inglés general, se encuentra el Inglés con Propósitos Específicos (ESP, por sus siglas en inglés. Como lo menciona González (2015), el ESP tiene su enfoque centrado en el estudiante (learner-centred), ya que tiene como prioridad atender las necesidades específicas de los estudiantes y las demandas profesionales o vocacionales. De modo que, esta rama se ha diversificado de acuerdo con las necesidades de los estudiantes.

Si bien el ESP muestra un gran auge en la actualidad por su diversificación y alcance, el estudio y práctica de este enfoque de enseñanza del inglés no es algo nuevo. Ma'mura et al.

(2020) describen que los cursos de lenguas con propósitos específicos datan de la primera década del siglo XX, siendo Carl Rogers, el padre de la teoría de enfoque centrado en el cliente, quien aplicó este conocimiento al ámbito educativo, ya que sentó las nociones para contenidos centrados en el alumno (O'Neill y McMahon, 2005).

Este nuevo enfoque ganó fuerza no solo por las demandas globales sino también debido al modelo pedagógico centrado en el alumno. De acuerdo con los autores, el estudio de curso personal, participó en la expansión de esta teoría en la pedagogía, aunque también destacan el papel que tuvo Piaget y Malcolm Knowles en los diseños centrados en los estudiantes. Entre 1930 y 1940 el aprendizaje de lenguas tomó una ruta centrada en niños, sería a finales de 1940 cuando se retomaría la pedagogía en adultos, y de manera más preponderante en 1960, con el ESP.

El ESP puede clasificarse en 2 grupos grandes, o subramas, que son el Inglés con propósitos Académicos (EAP, por sus siglas en inglés) y el Inglés con Propósitos Ocupacionales (EOP, por sus siglas en inglés). El EAP está mayormente asociado con la academia y estudios en general, y el EOP con la práctica profesional. Hyland y Jiang (2021b) representan esta diferencia con el ejemplo del Inglés con Propósitos Legales Académicos e Inglés con Propósitos Legales (ambos híbridos de EAP y EOP); el primero va asociado a estudiar derecho y el segundo a practicarlo.

Referente a la enseñanza del EAP, Hyland y Jiang (2021a) mencionan que el crecimiento de la investigación de la enseñanza del EAP es el resultado de dos grandes cambios sociales. El primero es la creciente complejidad y diversidad de contextos derivados de las redes globalizadoras que han llevado el inglés a todos los rincones de la vida académica, lo cual ha conformado una base de estudiantes diversa y heterogénea. La segunda razón de la expansión de la investigación en EAP, es el resultado de un esfuerzo sostenido y consciente de profesionalización en el campo. Sin embargo, reconoce Bell (2022) la necesidad de apuntar las nuevas investigaciones en el “cómo” del EAP en vez del “qué”, esto es, en la pedagogía o metodologías que lo sostienen.

Por otro lado, el inglés se puede clasificar según el contexto de los estudiantes y uso que le den a esta lengua. De acuerdo con Iwai (2011), el Inglés como Segunda Lengua (ESL) se refiere a aquellos alumnos que aprenden la lengua en contextos donde el inglés es utilizado como medio de comunicación, además de tener igual importancia como lo es la lengua materna; un ejemplo de esto son los hispanos en Estados Unidos aprendiendo inglés. Por otra parte, el Inglés como lengua Extranjera (EFL) se atribuye en aquellos países que no son hablantes del inglés; para este ejemplo menciona al alumnado del inglés en Japón, donde el

idioma no figura como lengua oficial ni medio de comunicación local. Otro ejemplo, es el Inglés como Medio de Instrucción (EMI, por sus siglas en inglés) donde las instituciones de educación superior utilizan el inglés como lengua para impartir el contenido, con el fin de mejorar no solo en contenido sino también en el dominio de la lengua (Barrios y Lopez-Gutierrez, 2021).

2.1.2 Modelos y pruebas para evaluar el dominio del inglés

De acuerdo con North (2000), las escalas de dominio del idioma se les ha nombrado de manera diferente, y lo que todas las definiciones tienen en común es que intentan proporcionar de manera ascendente una serie de niveles de competencia lingüística. En otras palabras, las escalas de dominio del idioma dan imágenes de niveles sucesivos de logro de aprendizaje de idiomas. Por otra parte, explica que las definiciones de las escalas de competencia lingüística en la literatura dependen de la perspectiva del escritor.

De manera extendida, la lengua se ha enseñado y evaluado a través de las cuatro habilidades (lectura, escritura, escucha y habla), sin embargo, en el ámbito académico y de investigación destacan la lectura y la escritura. El impacto de estas habilidades en el desempeño es significativo, ya que estas competencias se entrelazan con el desarrollo del pensamiento crítico y la construcción del conocimiento. La lectura efectiva potencia la comprensión y las habilidades analíticas, mientras que la escritura competente facilita la síntesis de información y la comunicación de ideas (Morales, 2020; Castillo-Martínez y Ramírez-Montoya, 2021).

Además, la lectura y la escritura son componentes fundamentales del proceso de investigación, particularmente en entornos académicos. La investigación realizada por Prihandoko (2024) en 162 estudiantes reveló que la lectura académica y la competencia de investigación explican juntos aproximadamente 58% de la varianza en la escritura académica. Este trabajo destaca la importancia de las habilidades de lectura e investigación para mejorar las capacidades de escritura.

Partiendo de la importancia de ambas habilidades, diversos autores plantean competencias dentro de estas dos habilidades. En la Tabla 1 se presenta un comparativo de las competencias lingüísticas para lectura y escritura que han utilizado diferentes autores al momento de crear pruebas estandarizadas para evaluar el dominio del inglés.

Tabla 1*Comparativa de autores de habilidades de lectura y escritura académica*

<i>Autores</i>	<i>Lectura</i>	<i>Escritura</i>
<i>Prapinwong y Koolsriroj (2019)</i>	General English Skills	
	Academic Writing skills	
	Academic English Skills	
	Research skills	
<i>Nakamura (2022)</i>	Vocabulary and Word usage	Summary
	Comprehending main ideas	Essay
	Comprehending details	
	Comprehending charts/graphs	
<i>Song y Zhou (2022)</i>	Identifying key ideas	Expressing ideas clearly/logically
	Reading quickly to find information	Expressing ideas in correct English
	Reading carefully to understand a text	Writing references/bibliographies
	Identifying supporting ideas/examples	Summarizing/paraphrasing
	Using own words in note taking	Writing conclusions
	Reading quickly to get overall meaning	Planning written assignment
	Understanding organization of a text	Synthesizing information/ideas
	Taking brief, relevant notes	Linking sentences smoothly
	Understanding specialist vocabulary	Using appropriate academic style
	Working out meaning of difficult words	Writing introduction
	Evaluating an academic journal	Proof-reading written assignments
		Referring to sources
		Writing body sections
		Revising written work
	Writing coherent paragraphs	
<i>Grabe y Zhang (2013)</i>	Taking notes from a text	
	Summarizing text information	
	Paraphrasing textual resources	
	Combining information from multiple text sources in a synthesis task	
	Comparing multiple points of view from written texts and producing a critical synthesis	
	Answering essay exam questions in writing	
	Writing an extended research paper or literature review	
<i>Qayyum (2024)</i>	Responding to assigned text	
	Content	
	Reader/writer interaction	
	Accuracy	
	Fluency	
	Organization	
Mechanics		

<i>Autores</i>	<i>Lectura</i>	<i>Escritura</i>
<i>Zainal y Jamil (2015)</i>	Making inferences and drawing conclusion based on evidence or information in texts	Summarizing a text
	Determining author's purpose and point of view	Writing an argumentative essay
	Using deductive or inductive reason to determine author's argument/ interpreting graphic aids, which authors use to present information	
	Evaluating author's argument to determine objectivity, completeness, validity, and credibility	
	Identifying errors and making corrections Paraphrasing sentences given first words	
<i>Sata y Karakaya (2021)</i>		Organization
		Content
		Coherence
		Cohesion
		Grammar
		Vocabulary
		Mechanics

Si bien existen autores que delimitan las competencias dependiendo de la habilidad, lectura o escritura, autores resaltan la integración de ambas habilidades a través de ciertas competencias (Prapinwong y Koolsriroj, 2019; Grabe y Zhang, 2013; Zainal y Jamil, 2015). En palabras de Grabe y Zhang (2013), la integración de lectura/escritura es crítico para el éxito académico. De entre las competencias que los autores mencionan está el resumen, la síntesis, toma de notas entre otras. Estas tareas podrán ser retadoras, pero los autores destacan la importancia de integrar ambas habilidades tanto en la enseñanza como en la evaluación.

Similarmente, Plakans et al., (2018) y O'Grady y Taşkesen (2022), comentan que las evaluaciones integradas, requieren que los estudiantes utilicen el contenido de lectura para producir respuestas escritas, evaluando así tanto las habilidades de comprensión como de composición. Esto sugiere que las evaluaciones de lectura pueden servir como medidas indirectas del dominio de la escritura dentro del contexto de las pruebas de lenguaje.

Por otra parte, como se puede apreciar, algunos autores clasifican las competencias por tareas, otros por aspectos concretos de la lengua o por la función que cumplen. De estos autores, resaltan Prapinwong y Koolsriroj (2019), quienes clasifican de acuerdo con la función, es decir, habilidades del inglés general, habilidades de escritura y lectura académica,

habilidades del inglés académico, y habilidades de investigación. Estas mismas categorías podrían englobar a otras competencias mencionadas como se muestra a continuación en la Tabla 2. Esta tabla presenta estas cuatro grandes categorías que agrupan competencias, tanto para las habilidades de escritura como lectura.

Tabla 2*Tabla de competencias por categorías y habilidades*

<i>Competencias</i>	<i>Lectura</i>	<i>Escritura</i>
<i>General English Skills</i>	Comprehending main, key and supporting ideas and details Working out meaning of difficult words Reading carefully to understand a text Identifying errors	Grammar Vocabulary and expressions usage Coherence Cohesion Mechanics (spelling and punctuation) Making corrections
<i>Academic Reading and Writing skills</i>	Comprehending charts/graphs Understanding organization Proofreading written assignments Revising written work	Organization, Idea development Paraphrasing Summary Essay Writing introductions, body sections and conclusions Writing references/bibliographies and Citation Synthesizing Content Reader/writer interaction Taking brief, relevant notes
<i>Academic English Skills</i>	Specialized vocabulary, Genre-based analysis	Specialized vocabulary,
<i>Research skills</i>	Critical reading and analysis Searching skills Breadth and depth of knowledge in the discipline Evaluating an academic journal Using deductive or inductive reason	Critical analysis, Breadth and depth of knowledge in the discipline

Nota. Generado a partir de Prapinwong y Koolsriroj (2019), Nakamura (2022), Song y Zhou (2022), Grabe y Zhang (2013), Qayyum (2024), Zainal y Jamil (2015), Sata y Karakaya (2021)

El Marco Común Europeo de Referencia (MCER) ha sido uno de los desarrollos más influyentes en los campos de la enseñanza de idiomas y pruebas en la última década (Figueras et al., 2005, como se cita en Cephe y Toprak, 2014). El MCERL es un proyecto del Consejo de Europa, que fomenta la armonización y la transparencia entre las instituciones transnacionales y promueve la ciudadanía europea (Fulcher, 2004, como se cita en Cephe y Toprak, 2014). De acuerdo con el Consejo de Europa (2020, p.35) el MCERL “Su objetivo es mejorar la transparencia y la coherencia entre el currículo, la enseñanza y la evaluación dentro de una institución, así como la transparencia y coherencia entre instituciones, sectores educativos, regiones y países”.

Desde su creación en el 2001 el Marco Común Europeo de Referencia para las Lenguas ha perseguido distintos objetivos, entre ellos “proporciona un metalenguaje que permite abordar la complejidad del dominio lingüístico para todos los ciudadanos en una Europa multilingüe e intercultural y que los responsables de las políticas lingüísticas puedan trabajar con objetivos y resultados de aprendizaje coherentes y transparentes”. El Marco Común Europeo de Referencia para las Lenguas (MCERL) es una convención utilizada para el diseño de cursos, guías, libros, entre otros alrededor de Europa. Sin embargo, este marco es ampliamente utilizado a niveles internacionales ya que provee los parámetros necesarios para describir a un usuario de la lengua. Por otra parte, aporta descriptores de aprendizaje específico que sirven de referencia para que los estudiantes, con el fin de alcanzar cierto nivel de dominio, como se muestra de manera resumida en la Tabla 3, de acuerdo con la información proporcionada por el Consejo de Europa (2020). De esta manera, se pueden diseñar exámenes o pruebas que permitan medir habilidades o competencias lingüísticas de manera precisa (Council of Europe, 2001).

Tabla 3*Tabla de descriptores generales por nivel en el MCERL*

Habilidad	A1	A2	B1	B2	C1	C2
<i>Comprensión de lectura</i>	Comprendo palabras/ signos y nombres conocidas/os y frases muy sencillas, por ejemplo, las que hay en avisos, letreros, carteles, pósteres y catálogos.	Leo textos muy breves y sencillos. Encuentro información específica y predecible en material sencillo y cotidiano, como anuncios publicitarios, prospectos, cartas de restaurantes, y comprendo cartas personales breves y sencillas.	Comprendo textos elaborados con un lenguaje de uso muy frecuente relativo a asuntos cotidianos o al trabajo. Comprendo la descripción de acontecimientos, sentimientos y deseos en cartas personales.	Leo artículos e informes relativos a problemas contemporáneos en los que los autores adoptan posturas o puntos de vista concretos. Comprendo la prosa literaria contemporánea.	Comprendo textos largos y complejos de carácter literario o de carácter fáctico, apreciando distinciones de estilo. Comprendo artículos especializados e instrucciones técnicas largas, aunque no estén relacionados con mi especialidad.	Leo con facilidad prácticamente todas las formas de lengua escrita/signada, incluidos textos abstractos estructural o lingüísticamente complejos, como manuales, artículos especializados y obras literarias.

Habilidad	A1	A1	B1	B2	C1	C2	
<i>Expresión escrita</i>	Elaboro frases y oraciones sencillas y aisladas.	Elaboro una serie de frases y oraciones sencillas enlazadas con conectores sencillos tales como <i>y, pero</i> y <i>porque</i> .	Elaboro textos sencillos y cohesionados sobre temas que me son conocidos o de interés personal.	Elaboro textos claros y detallados sobre una amplia variedad de temas relacionados con mis intereses. Elaboro redacciones o informes transmitiendo información o proporcionando razones a favor o en contra de un punto de vista concreto.	Elaboro textos claros y detallados sobre una amplia variedad de temas relacionados con mis intereses. Elaboro redacciones o informes transmitiendo información o proporcionando razones a favor o en contra de un punto de vista concreto.	Me expreso en textos claros y bien estructurados, exponiendo puntos de vista con cierta extensión y profundidad. Elaboro exposiciones detalladas sobre temas complejos en una redacción o un informe, resaltando lo que considero que son aspectos importantes. Elaboro diferentes tipos de texto en un estilo adecuado al/la lector/a que tengo en mente.	Elaboro textos claros y fluidos en un estilo adecuado. Elaboro cartas, informes o artículos complejos que presentan una argumentación con una estructura lógica y eficaz que ayuda al/la destinatario/a a fijarse en las ideas importantes y a recordarlas. Elaboro resúmenes y reseñas de obras profesionales o literarias.

Nota. Extraído de Consejo de Europa (2020)

Sin embargo, al ser una herramienta de amplio uso, no es específico en las habilidades necesarias en áreas de enseñanza y evaluación como lo es el inglés con propósitos académicos. En estos casos cada institución establece el nivel de dominio necesario para los programas, como en el caso de Reino Unido que es el B1 (Pearson, 2022) e igualmente en otros casos.

Otra de las escalas a considerar es la Escala Global de Inglés (GSE, por sus siglas en inglés). Es una escala granular estandarizada que mide el dominio del idioma inglés y, a diferencia de otros marcos, describe el rendimiento en bandas anchas, ya que identifica lo que un alumno puede hacer en cada punto de la escala a través de las habilidades de hablar, escuchar, leer y escribir (Pearson, 2015). La escala fue diseñada para demostrar con precisión el progreso paso a paso de los alumnos por medio de indicadores, que son competencias concretas que puede hacer el estudiante, como lo hace el MCERL con sus descriptores. A diferencia de otras escalas como la del MCERL, está dividida en seis niveles y va de puntuaciones del 10 al 90. Esta escala considera los lineamientos sugeridos por el MCERL y otros autores sobre las habilidades que deben tener los estudiantes para dominar el inglés.

La GSE surgió a partir de la necesidad de proporcionar un punto en la escala que declarara de manera precisa el nivel de dominio en una prueba, además de responder a necesidades específicas de los alumnos, sean con necesidades académicas, alumnos adultos, entre otros. Los creadores de la escala realizaron estudios en los que se determinaron de manera particular las características de los principales dominios de la lengua.

Pearson (2015) de manera sistemática operacionaliza el dominio del inglés académico y sus indicadores. Además, considera que el campo educativo se caracteriza por una dimensión de contextualización que determina diferencias importantes en el uso del idioma, a saber, la edad y el desarrollo cognitivo. Por lo tanto, los objetivos que persigue en el campo académico abordan las necesidades de los estudiantes en el nivel superior. En la Tabla 4 se presenta un comparativo de los niveles de dominio de los diferentes modelos de evaluación del dominio del inglés, y se pueden apreciar las diferencias en cuanto a su rango de evaluación, siendo el MCERL y el GSE (junto a la prueba PTE Academic), los más amplios.

Tabla 4*Comparación de los niveles de dominio entre los diferentes modelos de evaluación.*

MCER	GSE	PTE Academic	TOEFL iBT	TOEFL PBT	IELTS
C2	85 -90	85 -90	Sin puntaje	Sin puntaje	8.5 – 9.0
C1	76 -84	76 -84	110 -120	635 – 677	7.5 – 8.5
B2	59- 75	59- 75	87 – 109	569 – 634	6.5 – 7.5
B1	43- 58	43- 58	57 -86	490 – 568	5.5 – 6.5
A2	30-42	30-42	Sin puntaje	Sin puntaje	4.5 – 5.5
A1	22 – 29	22 – 29	Sin puntaje	Sin puntaje	Sin puntaje
<A1 Turista	13 – 21	13 – 21	Sin puntaje	Sin puntaje	Sin puntaje
< Turista	10-12	10-12	Sin puntaje	Sin puntaje	Sin puntaje

Nota. Traducido de De Jong y Benigno (2017)

Dentro de las pruebas que se identificaron como de bajo impacto está el Austest (Farnill y Hayes, 1996). Esta prueba fue diseñada con el propósito de identificar aquellos alumnos que pudieran presentar dificultades académicas debido a barreras del lenguaje. Esta prueba se aplicó en estudiantes recién ingresados a la educación superior en Australia, particularmente en este contexto existe un gran crecimiento en inmigración y como consecuencia, un aumento de alumnos de contextos de habla no inglesa (NESB, por sus siglas en inglés). Si bien esta prueba tiene como fin principal el cribar a los alumnos, también se hacen con la finalidad de orientar la construcción de talleres que fortalezcan la competencia en el inglés, como en el caso de las pruebas diagnósticas.

Por otra parte, se identificaron el uso de pruebas en cuasiexperimentos, que más adelante servirían para fortalecer ciertas áreas del proceso del dominio de la lengua. En la investigación de Altay et al., (2022) se utilizó el Cambridge Preliminary English Test para encontrar la correlación entre el dominio de la lengua y el éxito académico. En el caso de Chang (2023) se utilizaron las puntuaciones del TOEFL iBT para probar el éxito en el uso de las Tecnologías de la Información y la comunicación en el aula, y aportar a la mejora de las intervenciones pedagógicas.

Dentro de las pruebas enfocadas en el Inglés Académico se encuentran TOEFL iBT y el IELTS, las cuales son reconocidas internacionalmente, ya que expiden certificaciones especialmente útiles para aquellos alumnos interesados en movilidades o ingresar a instituciones EMI; estas pruebas evalúan las 4 habilidades (habla, escucha, escritura y lectura).

Otras pruebas encontradas son el WELT, ACERT y el CertAcles, las cuales son ampliamente utilizadas en sus países de origen, estas se caracterizan por ser con fines de ingreso a las instituciones de educación superior; el WELT se caracteriza por evaluar gramática, lectura y escritura, mientras que las otras dos incluyen la escucha.

Además, existen pruebas desarrolladas por investigadores que identificaron la necesidad de crear pruebas especializadas a un contexto específico, por ejemplo, la GATE es una prueba para posgrados en Malasia (Zainal y Jamil, 2015), la PAWT es para estudiantes de Canadá (Crossman, 2018) y la adaptación al CET realizada por Hu et al., (2022) es para China. En el caso del GATE, se creó con la finalidad de evitar el sesgo entre resultados de otras pruebas internacionales estandarizadas y evaluar el desempeño real de los alumnos en el aula, además de responder a las demandas lingüísticas de un posgrado en un país oriental. En cuanto al PAWT, busca atender a las necesidades de fortalecer el vocabulario académico de los alumnos de Canadá.

Finalmente, hay pruebas que son diseñadas con el fin de mejorar el uso del inglés en ámbitos académicos, como es en el caso de Kalnberzina y Karapetjana (2021), Simonova (2019) y Şata y Karakaya (2021). Estas están dirigidas a aspirantes y alumnos de educación superior o educación terciaria. Sin embargo, se identificó particularmente la prueba GATE que es dirigida al contexto de posgrado. Esta prueba fue diseñada en Malasia, de acuerdo con las necesidades identificadas por los autores Zainal y Jamil (2015) en cuanto a la redacción de tesis y de artículos. Un concepto fundamental para este diseño fue la localización, la cual parte de la idea de crear pruebas que estén personalizadas y respondan al contexto en el cual se han de aplicar. Dentro de las características que se consideran al contextualizar se encuentra el nivel académico al que va dirigido, país y localidad de los examinados, idioma objetivo, e incluso rama del conocimiento como la educación o la investigación.

2.2 Marco Contextual

2.2.1 Enseñanza del inglés en el Sistema Educativo Mexicano

En México desde 1926 el Estado decidió instaurar las escuelas secundarias como una extensión de la educación primaria y se incorporó la enseñanza del inglés dentro del plan de estudios con carácter obligatorio (Zorrilla, 2004). Para el año de 1992, se firma El Acuerdo Nacional para la Modernización de la Educación Básica bajo la dirección del secretario de educación Ernesto Zedillo Ponce de León y del presidente Carlos Salinas de Gortari donde uno de sus puntos más destacados fue la descentralización de la educación. Con este acuerdo se otorgó autoridad a los estados federativos en materia educativa, y de esta manera algunas entidades federativas implementaron programas de inglés a nivel primaria siendo el estado de Morelos el primero en 1992 y en cuanto a Baja California en 1997 (Fierro, 2015).

Es importante señalar que para 1993 se firmó el Tratado de Libre Comercio de América del Norte (TLCAN) lo cual reforzó las relaciones con los países angloparlantes aumentando el comercio, entre 1993 y 2017, casi 6 veces, esto es 522.2 MMD y con Canadá, casi 8 veces siendo de 21.1 MMD (Secretaría de Economía, 2019). Por lo anterior, en el 2009, 21 entidades federativas ya contaban con programas de inglés en escuelas primarias públicas (Fierro, 2015).

Por parte de Baja California, el Sistema Educativo Estatal (SEE, 2010, citado en Fierro, 2015) menciona la importancia de las gestiones realizadas por coordinadores, directores de escuelas, profesores y el apoyo de padres de familia en la implementación de los programas de inglés desde 1997. Además, de iniciar tales programas, sin presupuesto económico, ni formación docente.

En el año 2006, se propuso un programa piloto de Enciclomedia en inglés, el cual buscaba impartir inglés a estudiantes de sexto de primaria, donde el maestro no requería conocer la lengua ya que el material estaba diseñado para ser semi-autodidáctico. Sin embargo, como lo menciona Hernández “los maestros tuvieron problemas para dedicar las dos horas semanales previstas, no recibieron la asesoría esperada y se reportó que el material era un restringido apoyo audiovisual.” (2019).

En el Diario Oficial de la Federación (DOF) de 2015, cuando se reformó la educación secundaria del 2006 “... la asignatura de inglés también lo hace y dentro de los programas de estudio se detalla el nivel de logro de los estudiantes, el cual se estableció como nivel A2 alineándose al Marco Común Europeo de Referencia para las Lenguas (MCERL), que es el referente internacional obligado en términos de estándares para

aprender lenguas extranjeras”. En este momento se comienza a hablar de estándares en la enseñanza de dicha lengua.

Fue en el año de 2009 que entró en rigor el programa piloto PNIEB (Programa Nacional de Inglés en Educación Básica, por sus siglas) en un intento de homogeneizar la enseñanza del idioma en las 32 entidades federativas. De esta manera, se sustituyeron varios programas estatales y se procuró articular los aprendizajes desde tercero de preescolar hasta sexto de primaria y vincularlos a los tres grados de la educación secundaria. A pesar de esto, y como lo señaló Hernández (2019) “el principal problema del programa es que se limitó a acciones administrativas y fiscales de transferencia de recursos para contratación de maestros de inglés en los estados.”

El 19 de agosto de 2011 se publicó en el Diario Oficial de la Federación el Acuerdo número 592, la Articulación de la Educación Básica y el Plan de Estudios y enmarca el perfil de egreso y las habilidades que debe poseer de la lengua, como lo son comprensión oral, comprensión lectora, producción oral y producción escrita en los niveles de dominio B1 del MCERL (DOF, 2011). Para lo cual Rodríguez (2021) menciona la alineación que se dispuso para el ciclo 4 de PNIEB, en el cual se declaraba el nivel B1 según el MCERL, la competencia deseada a alcanzar durante los tres años de secundaria logrando competencias que señala de la siguiente manera “Este nivel se considera como el necesario para que un visitante se desenvuelva de manera eficiente en un país extranjero. Incluye dos características principales: la habilidad para interactuar y comunicarse en varios contextos y la habilidad de solucionar problemas de la vida cotidiana” (Rodríguez, 2021, p.36).

Igualmente, la Secretaría de Educación Pública (SEP) instauró programas que fomenten la evaluación de lenguas extranjeras a nivel nacional por medio de un marco, y que está “orientado a elevar la calidad educativa en la materia, especialmente en el caso de la enseñanza del idioma inglés como lengua extranjera” (SEP, 2023, p.3) la Certificación Nacional del Nivel de Inglés (CENNI) para establecerlo como estándar de acuerdo con MCERL y certificar el nivel de conocimientos de los alumnos (SEP, 2015). Esta certificación sirve a la Dirección General de Acreditación, Incorporación y Revalidación (DGAIR) y otras instancias y autoridades “evaluar, acreditar y certificar conocimientos y aptitudes en materia de lenguas extranjeras, así como del español como lengua extranjera o lengua adicional” (SEP, 2023, p.3). El CENNI se presenta con 3 distintas finalidades que son específicos (oficio o profesión), general y académico.

La SEP, en el 2015, redactó que el dominio del inglés incrementa la probabilidad de un mejor empleo, movilidad social y mejores sueldos. Siendo así

vemos que las autoridades educativas reconocen la importancia del inglés, y no solo en la actualidad ya que desde décadas atrás podemos ver cómo se ha ido perfilando la lengua en la educación básica.

Dentro de las reformas propuestas para el programa sectorial se encuentra el Programa Nacional de Inglés (PRONI), que inició su operación en el ejercicio fiscal 2016 siendo la Dirección General de Desarrollo Curricular (DGDC) la unidad responsable. El PRONI vino a sustituir a PNIEB aunque retomando ciertos aspectos de este.

La Secretaría de Gobernación en el DOF del 2015 describe el objetivo de PRONI el cual “busca apoyar a las Entidades Federativas con la finalidad de fortalecer y dar continuidad a las acciones que se han generado desde que inició la prueba piloto en el ciclo escolar 2009-2010 en escuelas públicas de preescolar y primaria, ...” En otras palabras, pretende auxiliar en cuestiones técnicas y pedagógicas de la enseñanza del inglés a cada entidad federativa. Lo anterior, de acuerdo con Plan Nacional de Desarrollo 2013-2018 bajo el sexenio de Enrique Peña Nieto con el lema México con Educación de Calidad.

De acuerdo con Hernández (2019) “El PRONI utiliza el marco curricular diseñado para el PNIEB. De acuerdo con este marco, todos los estudiantes de tercer grado de preescolar hasta sexto grado de primaria deben recibir de dos a dos horas y media de instrucción a la semana. No obstante, dado su carácter de programa nacional, el PRONI ha alcanzado sólo al 18% de la población en edad escolar básica.

Las acciones tomadas dentro de los planes y programas de cada sexenio impactan no solo en los trayectos formativos de la comunidad estudiantil, sino también en etapas terminales y laborales. “Los profesionales que dominan el inglés pueden asegurar un sueldo entre 30 y 35% por encima de personas que no lo hablan o pueden realizar estudios superiores o de posgrado en universidades reconocidas internacionalmente” (SEP, 2015, p.4). De acuerdo con el estudio realizado por EF English Proficiency Index (EPI, 2021), el nivel de dominio del inglés en México se encuentra en los niveles de Muy Bajo, posicionándose en el lugar 88 de 110 en el ranking de países. Esto representa un descenso en comparación al 2017, en el cual México se posicionó en un nivel bajo. Dentro de México, Baja California se encuentra en un nivel de dominio Muy Bajo, a pesar de sus condiciones de frontera.

El inglés representa un reto que dificulta el desarrollo y la innovación de países en los cuales no es su lengua oficial. Como se menciona en Salager-Meyer (2008) la lengua inglesa está fuertemente ligada tanto la producción científica como el acceso y la acumulación del

conocimiento al desarrollo de un país. Ante tal necesidad, la SEP ha planteado a lo largo de los sexenios planes y programas que respondan ante tales demandas sociales y globales.

De acuerdo con el estudio realizado por Toledo y García (2020), cerca del 36% de los estudiantes universitarios que egresan cumplen con el requisito mínimo del dominio del inglés (leerlo, hablarlo, escribirlo y comprenderlo). Por otra parte, el Instituto Mexicano para la Competitividad (IMCO, 2015, como se cita en Toledo y García, 2020) menciona que, a medida que aumenta el nivel de estudios, aumenta también el nivel de inglés, sin embargo, las universidades aún no logran capacitar mejor a sus egresados en el idioma. Esto se refleja en las cifras, donde solo 36,70% dice hablar inglés. Ante las estadísticas, el IMCO expresa las razones por las cuales promueven el inglés en el país y su relevancia en el mundo actual. Entre los motivos se encuentran la competitividad, la movilidad social, empleabilidad y mejora salarial, como herramienta de acceso al conocimiento y su relevancia en el mundo de los negocios.

Las cifras anteriores ponen en perspectiva los esfuerzos que diversas organizaciones, consejos e incluso las instituciones de educación superior han realizado por resolver esta problemática. Por ejemplo, entre los organismos del nivel de estudios superiores destaca la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES), la cual en sus líneas de acción para el objetivo específico 2.4 contempla “Desarrollar e implantar un programa para la enseñanza y el aprendizaje del idioma inglés en todas las instituciones asociadas a la ANUIES” (2016, p. 53) con el fin de adecuarse a un mundo globalizado por medio de la internacionalización.

Por otra parte, el Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) dentro de su Programa Institucional 2020-2024, considera no sólo la participación en espacios internacionales sino también la continuidad dentro de membresías internacionales (s.f., p.57). Igualmente, ante la necesidad del dominio de la lengua y de la producción científica, diseñó el programa de Capacitación para la Escritura de Manuscritos Académicos en Inglés (CEMAI), en el cual declara que “la escritura de manuscritos académicos en inglés es todavía imprescindible y deberá convertirse, en el mediano plazo, en uso y costumbre en nuestras instituciones” (s.f.). Por lo tanto, este programa se creó con la finalidad de capacitar a estudiantes y tutores para publicar en la lengua.

2.2.2 Contexto en UABC

De manera particular, la Universidad Autónoma de Baja California (UABC, 2019b), también se ha planteado acciones para cubrir la necesidad de un segundo idioma. En la estrategia 2.1.5. del Plan de Desarrollo Institucional 2019-2023 se propone, “Fortalecer los esquemas institucionales para el aprendizaje y dominio del idioma inglés” (p. 99), la cual se conecta con los propósitos de internacionalización y movilidad académica. Incluso, en la estrategia 5.1.5 se plantea “Impulsar procesos de formación y certificación en el dominio del idioma inglés en el personal académico” (p.107). Cabe también resaltar, que dentro del Modelo Educativo de UABC se menciona que el dominio de un idioma extranjero se encuentra dentro de la normatividad y que cada unidad académica establece el nivel requerido para su acreditación (UABC, 2018).

En el caso del Instituto de Investigación y Desarrollo Educativo (IIIDE, 2020), en su Plan de Desarrollo Institucional 2019-2023, incluye acciones dentro de las cuales está la 2.1.5. que busca “Fortalecer los esquemas institucionales para el aprendizaje y dominio del idioma inglés.” Esto se pretende lograrlo por medio de ofrecer a los estudiantes de posgrado unidades de aprendizaje optativas en inglés. Todo con la finalidad de promover la internacionalización y movilidad, además de responder a demandas sociales y mantener la competitividad.

La problemática de las Instituciones de Educación Superior no solo se ha abordado en México. Li y Flowerdew (2020) señalan que, desde principios de la década de 2000, las universidades de todo el mundo comenzaron cada vez más a instituir políticas que fomentan fuertemente la publicación en revistas de alto rango en inglés. Como resultado también de esta exigencia, ahora hay más instituciones superiores que buscan que sus estudiantes de posgrado publiquen también en revistas internacionales indexadas como condición para obtener el grado. Sin embargo, es relativamente reciente el desarrollo de intervenciones pedagógicas. John Swales (1990, como se cita en Li y Flowerdew, 2020) fue un precursor en este sentido, presentando intervenciones pedagógicas utilizadas por él para atacar las problemáticas que tenían sus estudiantes de doctorado de Inglés como Lengua Adicional (EAL, por sus siglas en inglés). Sin embargo, en los documentos revisados no se presentan con claridad los métodos de evaluación que utilizaron en sus intervenciones. En términos de evaluación diagnóstica, no existe la suficiente literatura que revele el nivel de conocimientos y habilidades de los estudiantes.

2.3 Nociones sobre medición y validez

La teoría Clásica de Test (TCT) es la teoría psicométrica más popular en el siglo XX (Muñiz, 2010) sobre la cual la mayoría de las teorías y prácticas de medición en las últimas décadas. La idea central de la TCT parte de que la puntuación de un examinado en una prueba corresponde a la suma de un componente de su puntuación verdadera y el error de medida.

En esta teoría existen dos conceptos principales. Primeramente, la puntuación verdadera hace referencia al promedio de la puntuación que se espera obtener si un examinado realiza formas paralelas de la misma prueba en un número infinito de ocasiones. En segundo lugar, el error de medida corresponde a la diferencia entre lo que observas en el promedio de implicaciones infinitas.

Entre otros conceptos claves de la TCT encontramos el error aleatorio y el error sistemático. De acuerdo con Finch y French (2019), el error aleatorio se refiere a los factores que son transitorios e idiosincráticos por naturaleza, afectando al desempeño de un solo individuo en la escala solamente en el momento en el que completan el instrumento. El error aleatorio es único de tiempo en tiempo y de persona a persona, significando que, si un individuo se les diera una prueba de matemáticas en diferentes puntos en el tiempo, el error diferiría cada vez.

Una de las grandes herramientas que posee la educación, es la evaluación y las distintas ramas y subramas del aprendizaje del inglés no es la excepción. Sobre todo, cuando se trata de la evaluación del aprendizaje, las actividades de medición y las pruebas han tenido un rol importante tanto en la calificación de un curso como en el diseño de materiales y unidades de aprendizaje

Dentro de las pruebas encontramos las estandarizadas, que de acuerdo con Jornet y Suarez (1996, p.142) existen otros instrumentos que hacen referencia a estas mismas como “...Tests Referidos al Dominio, Tests Referidos a Objetivos, Tests de Competencia, Tests de Certificación, Tests de Dominio, Tests referidos al Criterio, etc”. Por otra parte, estas pruebas nos ayudan a descartar la subjetividad por parámetros establecidos sobre las habilidades y conocimientos esperados de los alumnos (Hernández et al., 2018). Si bien las pruebas estandarizadas han tenido críticas, para Jornet y Suárez (1996) la problemática se encuentra en la descontextualización de estas pruebas y en los “usos e interpretaciones para los que normalmente no han sido construidas y, en ocasiones, se desarrollan con esquemas de elaboración que han sido diseñados metodológicamente para objetivos evaluativos diferentes” (p. 142). Adicionalmente, enuncian ciertas características que se deben tomar en

consideración como es la finalidad, la población a la que va dirigida, el objeto de medida y el uso de la prueba.

Buisán y Marín (2001, como se cita en Arriaga, 2015) identifican al proceso diagnóstico con función preventiva, predictiva y correctiva. Este mismo proceso es orientador ya que identifica “las posibilidades y limitaciones del sujeto, sus resultados servirán para definir el desarrollo futuro y la marcha del aprendizaje del objeto de estudio” (p.66), lo cual resulta muy útil en el aprendizaje de lenguas.

Con el fin de desarrollar la Prueba Diagnóstica del dominio del idioma inglés con propósitos de publicación científica en estudiantes de posgrado, se utilizó como referente las pruebas referidas a criterio. De acuerdo con Popham (1983, como se cita en Blanco, 1996, p. 12) “un test basado en criterios, se emplea para determinar la posición de un individuo con respecto a un dominio de la conducta perfectamente definido”.

Un pilar dentro del diseño de pruebas es la psicometría. De acuerdo con Meneses et al. (2013, p.38)

...es una rama de la psicología que, mediante teorías, métodos y técnicas vinculados al desarrollo y la administración de tests, se ocupa de la medida indirecta de los fenómenos psicológicos con el objetivo de hacer descripciones, clasificaciones, diagnósticos, explicaciones o predicciones que permitan orientar una acción o tomar decisiones sobre el comportamiento de las personas en el ejercicio profesional de la psicología.

Además, como lo mencionan Finch y French (2019) la medición de habilidades mentales representa un pilar tanto en la psicología como en la educación. Debido a la gran importancia de poder cuantificar estos rasgos intangibles del ser humano, la evaluación y medición han tomado gran relevancia. Por lo tanto, es importante reconocer el proceso de medición y llevarlo a cabo de manera correcta. Además de fiarse de instrumentos que realmente arrojen la información que se está buscando.

Entre las propiedades psicométricas que debe tener una prueba está la validez. De acuerdo con Tristán y Vidal (2006, p. 19) la definen como “el grado en que una prueba mide el propósito que se pretende medir”. Además, la describen como un atributo obligatorio, prioritario, ineludible de la prueba. En otras palabras, Messick (1995) lo define como un juicio compuesto por la evidencia empírica y los fundamentos teóricos que respaldan las

acciones basadas en puntajes de las pruebas. De esta manera, es recomendable contar con el mayor número de evidencias sobre su validez, para así garantizar su calidad y consistencia.

Otra propiedad psicométrica necesaria es la confiabilidad. En palabras de Reidl-Martínez (2013. p. 109) “La confiabilidad de una prueba se refiere a la consistencia de las calificaciones obtenidas por las mismas personas en ocasiones diferentes o con diferentes conjuntos de reactivos equivalentes”. Finch y French (2019) la describen como la medida de la consistencia del instrumento en las puntuaciones que arroja. Por lo anterior, la confiabilidad está ligada a la medición del error, esto se debe a la naturaleza de los datos utilizados.

Igualmente, es necesario definir el objeto de medida, sus interacciones con factores externos (psicológicos, biológicos, etc.) que puedan influir en las evidencias sobre el objeto, y, de manera muy importante el objetivo de la prueba. Tristán y Pedraza (2017) señalan que “junto con la validez y la confiabilidad, configuran los tres atributos fundamentales para el diseño, administración, interpretación y uso de las pruebas estandarizadas” (p.12).

Por último, está la imparcialidad, con ella se pretende asegurar que en la prueba estandarizada no existan sesgos. Esto es, que en la obtención de las evidencias se encuentren condiciones de equidad contextual. En este proceso de imparcialidad se debe considerar todas las variables que produzcan medidas equivocadas y apreciaciones injustas a personas de una muestra en particular (sea por género, nivel socioeconómico, etc.). Por tal motivo, el investigador debe asegurarse que los ítems no atenten contra la dignidad de las personas (Tristán y Pedraza, 2017).

Dentro del aspecto de validez, existen diferentes tipos de evidencias que se pueden recabar con el fin de aportar a la prueba. Entre estos, existen 4 principales: evidencia basada en contenido, evidencia basada en los procesos de respuesta, evidencia basada en estructura interna y evidencia basada en relación a otras variables (Cizek, 2020). Para efectos de este documento solo nos concentraremos en las evidencias basadas en contenido. “El contenido de la prueba hace referencia a los temas, la redacción y el formato de los ítems, tareas o preguntas de una prueba” (American Educational Research Association, American Psychological Association y National Council on Measurement in Education [AERA, APA y NCME], 2018, pp. 14-15).

Como lo menciona Meneses et al. (2013), la validez de contenido “hace referencia a la relación que existe entre los ítems que componen el test y lo que se pretende evaluar con él, prestando atención tanto a la relevancia como a la representatividad de los ítems”. Por lo

tanto, para obtener evidencias es necesario tener definido el dominio que se pretende evaluar, y después someter a juicio los ítems que conforman la prueba, según su relevancia.

2.3.1 Taxonomía de Anderson y Krathwohl

Uno de los procesos importantes dentro de la enseñanza- aprendizaje, así como de la evaluación, es la capacidad de organizar y gestionar los conocimientos y habilidades, ante esta necesidad surgieron las taxonomías. Por taxonomías se entiende como una “serie de ordenamientos sistemáticos y jerárquicos de los diferentes mecanismos generadores del aprendizaje” (Bayacán, 2013, p.110).

Dentro de las distintas taxonomías que se encuentran en la educación, existe la propuesta por Bloom, la de Anderson y Krathwohl, de Fink, de Simpson, de Dave, de Marzano (Benemérita Universidad Autónoma de Puebla, s.f.). Para el diseño de esta prueba se utilizó la taxonomía de Anderson y Krathwohl. Esta taxonomía fue diseñada en el 2001 como una extensión o revisión de la taxonomía propuesta por Bloom.

Como lo menciona Urgo et al. (2019) la propuesta de Anderson y Krathwohl fue creada con el fin de proporcionar a los docentes un marco que permitiera la discusión y articulación de los objetivos de aprendizaje, actividades educativas y evaluaciones. Esta taxonomía está estructurada por dos dimensiones: procesos cognitivos y tipo de conocimiento. De esta manera, se aprecia como una tabla de seis columnas para los procesos cognitivos, organizados de simple a complejo, esto es recordar, comprender, aplicar, analizar, evaluar y crear. Además, incluye cuatro filas para los tipos de conocimientos, que son factual, conceptual, procedimental y metacognitiva, dando como resultado una tabla de seis por cuatro.

Sin embargo, de acuerdo con Tuela y Palar (2022) existen otras clasificaciones para los procesos cognitivos las cuales han resultado útil en el diseño de materiales y evaluaciones. Estas clasificaciones se nombran de acuerdo al orden de habilidad de pensamiento. Siendo las primeras dos (recordar, comprender y aplicar) de orden inferior (LOTS, por sus siglas en inglés), y analizar, evaluar y crear como de orden alto (HOTS, por sus siglas en inglés). En las HOTS, algunos autores han identificado habilidades clave como lo es el pensamiento crítico, creativo, analítico, lógico, reflexivo, reflexivo y metacognitivo (Arif et al., 2020; Riefani y Utami, 2017).

2.3.2 Inteligencia Artificial Generativa

La inteligencia artificial ha representado un avance en la tecnología. Si bien, esta herramienta presenta debate actual sobre aspectos éticos, filosóficos y de integridad (Habibie y Starfield, 2023), es indudable su aporte sobre todo en el área de la educación y la investigación. Primeramente desde el aspecto digital, se encuentra que las pruebas digitales ofrecen nuevas formas de entregar y calificar tareas tradicionales (Cardwell et al., 2024). De esta manera, la automatización de la calificación de tareas es más eficiente. Además, que la gestión y análisis de datos es más sencillo desde las herramientas digitales.

Por otra parte, la automatización de los procesos es una gran ventaja. En el caso de la Inteligencia Artificial puede representar un recurso eficiente como lo menciona Cotos (2011) de retroalimentación inteligente para la escritura si (a) se individualiza de acuerdo con la información específica del alumno, (b) señala el tipo de error, (c) explica explícitamente el error y (d) conduce a la autocorrección. Además, no solo se considera rentable, sino que también brinda retroalimentación rigurosa y rápida (Hyland y Hyland, 2019, como se cita en Taskiran y Goksel, 2022).

De entre los distintos modelos creados de la inteligencia artificial se encuentra la generativa, que refiere a modelos de IA que crean contenido, como texto e imágenes, utilizando grandes conjuntos de datos (More, 2024). De acuerdo con Kaldaras et al. (2024), la GAI (Inteligencia Artificial Generativa, por sus siglas en inglés) tiene el potencial de transformar la educación debido a su programación que puede ser pre-entrenada para ciertas tareas, de manera que se ajusten a las necesidades del usuario. Un ejemplo de esto es Chat GPT (transformador pre-entrenado generativo, por sus siglas en inglés), esta GAI o GenAI se ha popularizado debido a su gran capacidad de generar respuestas gracias al acceso a una gran base de datos.

Si bien, los aspectos éticos de esta herramienta han causado controversia, recientes estudios apuntan a su incorporación en la educación y sobre todo en la evaluación (Chan, 2023; Kaldaras et al., 2024; Farzi y Dietz, 2024; Khlaif et al., 2024; Tobler, 2023). Este nuevo campo busca, en palabras de Farzi y Dietz (2024) integrar las fortalezas de ambas partes, esto es, el juicio humano y el juicio emitido por los modelos extensos de lenguaje (LLM, por sus siglas en inglés) de la IA.

En cuanto a las formas de integración, se encuentran distintos métodos, trabajos como el de Farzi y Dietz (2024), que propone el uso de rúbricas para la evaluación de textos por medio de rúbricas utilizadas por LLM. En este trabajo se descompone información en

criterios que pueden ser evaluados sistemáticamente. Otro método es el presentado por Khlaif et al. (2024), donde integra herramientas de GenAI con las pruebas adaptativas.

Capítulo III. Método

En el presente capítulo se detalla el enfoque metodológico adoptado para la investigación, describiendo con rigor cada una de las etapas que conforman el proceso de desarrollo de la prueba diagnóstica de inglés con propósitos de publicación científica. El capítulo tiene como objetivo proporcionar una descripción exhaustiva del diseño, las fases metodológicas, los análisis realizados, y los aspectos éticos considerados para garantizar la calidad y la validez del instrumento propuesto.

El capítulo está organizado en cuatro fases principales. La primera considera la Revisión Sistemática de la Literatura (RSL), destinada a identificar referentes teóricos y metodológicos clave, así como a delimitar los marcos conceptuales para la construcción de la prueba. En la segunda fase, se describen las actividades relacionadas con el diseño y desarrollo del instrumento, incluyendo la definición de las variables de medida, la elaboración de especificaciones, y la construcción inicial de los ítems. La tercera fase se centra en el análisis psicométrico y descriptivo de la prueba, incluyendo la validación mediante juicios de expertos y el pilotaje del instrumento en formatos análogos y digitales. Finalmente, se aborda la fase de análisis del comportamiento de la prueba y de los examinados, con un enfoque en la interpretación de resultados.

Cada sección incluye las técnicas y procedimientos específicos utilizados, así como las herramientas tecnológicas y estrategias implementadas para garantizar la calidad del proceso. Asimismo, se incorporan consideraciones éticas, subrayando el respeto a la confidencialidad de los datos, la participación voluntaria de los sujetos, y el uso responsable de herramientas de inteligencia artificial para la evaluación de los resultados. De este modo, el presente Capítulo ofrece una base metodológica sólida que respalda la validez del proceso de investigación, asegurando que los resultados obtenidos respondan de manera rigurosa a los objetivos planteados.

Diseño de la investigación

El presente proyecto se realizó bajo el diseño de investigación cuantitativo, ya que de acuerdo con Creswell y Creswell (2018) es un enfoque para probar teorías objetivas al examinar las relaciones entre variables, siendo estas variables medidas por medio de instrumentos para analizar sus datos numéricos en procedimientos estadísticos. Además, el diseño de esta investigación se caracteriza como no experimental y transversal, ya que en palabras de Muij (2011), no se utilizó un grupo de control y los datos fueron obtenidos en un

momento determinado. Finalmente, esta investigación es con enfoque psicométrico, ya que se hace uso de las teorías, métodos y técnicas que permiten, en este caso, diagnosticar los fenómenos psicológicos objeto de medida, que para este caso es el dominio del inglés como lo explica Meneses et al.:

mediante teorías, métodos y técnicas vinculados al desarrollo y la administración de tests, se ocupa de la medida indirecta de los fenómenos psicológicos con el objetivo de hacer descripciones, clasificaciones, diagnósticos, explicaciones o predicciones que permitan orientar una acción o tomar decisiones [...]. (2013, p.38)

Fase I. Revisión Sistemática de la Literatura

Con el objetivo de identificar las necesidades de la prueba, se realizó una Revisión Sistemática de la Literatura para reconocer los trabajos que servirán como referente, los estudios más importantes y hacer un mapeo del estado del arte en materia de pruebas de inglés en educación superior. Además, este ejercicio sirvió al primer objetivo de investigación que es el operacionalizar los niveles de dominio IPA con enfoque en la publicación científica, y reconocer las principales pruebas, sus características y su composición metodológica.

Preguntas de orientación a la revisión sistemática.

Antes de iniciar con el proceso de la Revisión Sistemática, se establecieron preguntas orientadoras con el fin de acotar el propósito de la misma. Estas preguntas fueron sobre todo de carácter metodológico, ya que se alinea con el perfil de este proyecto. Entre esas preguntas se pueden mencionar las siguientes: ¿Cuáles son las propuestas para la evaluación del dominio del inglés en nivel superior?, ¿dónde se han aplicado?, ¿cómo se caracteriza la población objetivo de las propuestas?, ¿En qué consiste el método?, ¿Cuáles son enfocadas a un inglés académico?, ¿Cuáles son los marcos o escalas utilizadas?, ¿Cuáles son las habilidades evaluadas? ¿cuáles son los fines del uso de las pruebas?

Estrategia de búsqueda.

En esta etapa se exploró con distintas cadenas de búsqueda con el fin de obtener aquellos textos que fueran relevantes a la temática. Al ser algunas palabras clave muy

específicas, la cantidad de resultados eran muy pocos, esto es, palabras clave como *ERPP* o similares; o por otra parte al generalizar con las palabras clave como *English* o *Test* los resultados crecían en número, pero decrecían en relevancia. Después de dicha exploración, se concluyó con una cadena, la cual, en conjunto con los criterios de inclusión y exclusión, arrojaron los resultados que mejor respondían a los objetivos de la revisión.

De esta manera se definieron las estrategias, palabras clave, operadores booleanos, periodo y tipo de búsqueda a considerar para generar la cadena. La búsqueda se realizó en inglés, el día 21 de febrero de 2023 en la base de datos Scopus con las siguientes palabras clave: english language, higher education, test y measurement.

La cadena extraída de la base de datos fue: TITLE-ABS-KEY ("english language" AND "higher education" AND ("test" OR "measurement"))

Solo se incluyeron artículos, capítulos de libros, libros y editoriales. En lo que respecta al idioma, se excluyeron aquellos documentos en español e inglés. Bajo estos criterios, Scopus arrojó 181 documentos y 50 de Web of science.

Criterios de elegibilidad.

Se presentan los criterios considerados para la inclusión y la exclusión de los estudios en la revisión sistemática. Los criterios a considerar para la inclusión de documentos fueron los siguientes:

Artículos en inglés y español, todos los años, artículos, capítulos de libros, libros y editoriales, trabajos con propuestas de pruebas de inglés, trabajos que analizan pruebas de inglés.

Se excluyeron aquellos documentos que:

No estuvieran situados en el nivel educativo superior, trabajos que no fueran de acceso abierto, resúmenes de libros, redactados en idiomas aparte del inglés y el español, proponen solamente escalas, cuasi- experimentos con pre test y/o post test, pruebas que no incluyan la habilidad de lectura y/o escritura.

Además, este proceso se realizó en los índices de Scopus y Web of science, ya que la mayoría de la literatura de interés se encuentra en estos índices. De esta segunda etapa se obtuvieron 181 resultados de Scopus y 50 de Web of science, siendo un total de 231. Sin embargo, se encontraron trabajos repetidos entre los resultados de ambos índices, así que se excluyeron en un segundo cribado, dejando 181 en Scopus y 13 en Web of science, con un

total de 194 resultados. De este segundo cribado resultaron un total de 32 artículos científicos.

Selección de estudios.

Antes de aplicar los criterios de inclusión y exclusión, la búsqueda de Scopus y Web of science con la cadena descrita anteriormente arrojó 231 documentos. Se descargó un documento con el autor, título, año, título de la fuente, volumen, asunto, artículo, páginas, doi, resumen, palabras clave, tipo de documento y acceso.

Primeramente, se tradujeron los títulos y resúmenes y se excluyeron aquellos documentos que no cumplieron con los criterios establecidos anteriormente. El criterio de acceso, se realizó hasta la siguiente etapa, donde se leyeron a profundidad los documentos. Después de esta etapa quedaron 35 artículos.

Posteriormente, se leyeron los documentos seleccionados al momento y se excluyeron aquellos que, después de la lectura, no cumplieron con los criterios establecidos. Después de esta etapa quedaron 31 artículos. En la tabla 5 y 6 se encuentra la información sintetizada de este proceso de revisión sistemática.

Estrategia de análisis de datos

Dentro de los resultados, se encontró que la mayoría de los artículos se refieren a pruebas utilizadas en instituciones que utilizan el inglés como medio de instrucción. Por otra parte, la mayoría de estas pruebas hacen referencia a la evaluación de las 4 habilidades en general (escritura, lectura, habla y escucha), además de un enfoque en el inglés general más que en un inglés focalizado.

Tabla 5

Índices y cadena de búsqueda para la Revisión Sistemática de la Literatura

Índices	Cadena de búsqueda:
<i>Scopus</i>	TITLE-ABS-KEY ("english language" AND
<i>Web of Science</i>	"higher education" AND ("test" OR
	"measurement"))

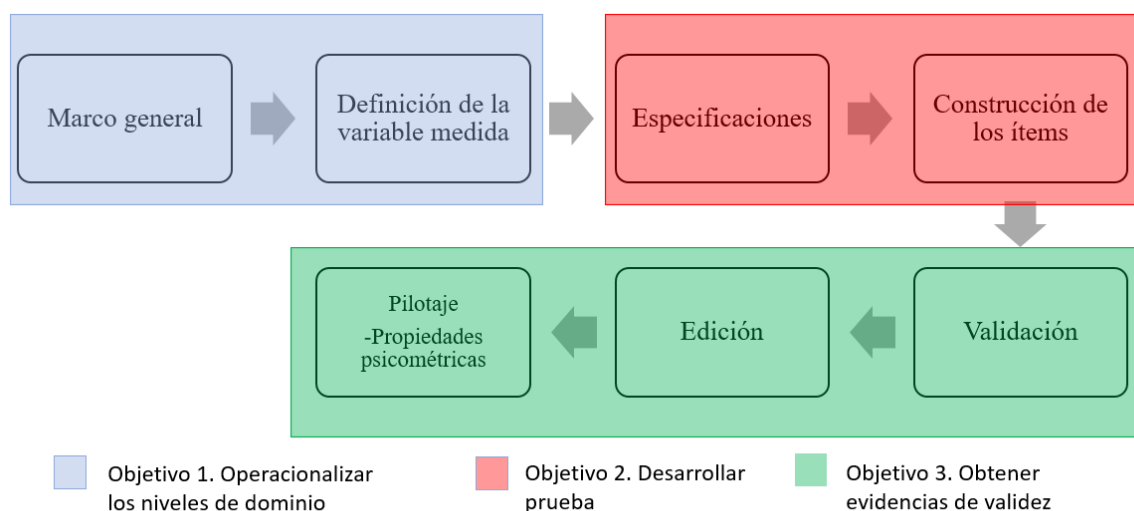
Tabla 6*Refinamiento de la búsqueda para la Revisión Sistemática de la Literatura*

Criterio	Inclusión	Exclusión
Producción del autor(es)	Nacional e internacional	No aplica
Nivel educativo	Educación superior	Todo lo que no se refiere a educación superior
Idioma	Inglés y español	Otros

Fase II. Diseño y desarrollo de la prueba

Con el propósito de alcanzar los objetivos establecidos en esta investigación, se decidió seguir los pasos propuestos por Muñiz y Fonseca-Pedrero (2019) para el diseño y desarrollo de la prueba diagnóstica. Se realizaron adaptaciones con el fin de optimizar el tiempo y cumplir con el cronograma de trabajo establecido, y de acuerdo con los alcances y limitaciones de este proyecto.

Dicho proceso se aprecia en la Figura 1, donde se muestran las etapas que buscan dar cumplimiento a los objetivos establecidos. En este apartado se describen cada una de las fases de manera cronológica de acuerdo con su procedimiento, participantes e instrumentos o materiales.

Figura 1*Etapas para el desarrollo del instrumento*

Nota. Creación propia, adaptado de Muñiz y Fonseca-Pedrero (2019)

Etapa I. Marco General

El objetivo general de esta etapa fue describir la justificación, el contexto de aplicación, el tipo de aplicación, el formato de administración, y el propósito (Muñiz y Fonseca-Pedrero, 2019) de la prueba diagnóstica.

Etapa II Definición de la variable de medida

El objetivo general de esta etapa fue delimitar de manera operativa, semántica y sintáctica la variable de medida (Muñiz y Fonseca-Pedrero, 2019).

Para cumplir con los objetivos de esta etapa se realizaron dos actividades. En primer lugar, como actividad 2.1 se definió la variable de acuerdo con los objetivos. Además, se realizó una revisión de las principales escalas del dominio de inglés, se aplicaron criterios para la selección del marco que se utilizaría como referente en el instrumento.

Posteriormente, en la actividad 2.2, como parte de la variable de medida, se seleccionaron los indicadores que corresponden a los objetivos de la prueba. Lo anterior se realizó con fin de evitar evaluar habilidades o conocimientos innecesarios y que la prueba final sea corta, ya que su uso final es el diagnóstico. Para llevar a cabo esta actividad, primeramente, se contactaron a tres especialistas en el desarrollo de pruebas sobre el inglés, a los cuales se les envió dos documentos: una presentación y un Excel de evaluación. La presentación para capacitación de expertos (Anexo 1), constaba de una descripción breve del propósito de la prueba, la escala a utilizar, y las actividades a realizar con un ejemplo de cómo llenar el formato de evaluación. La tabla de Excel de indicadores, como se muestra de ejemplo en la Figura 2 (para más información consultar Anexo 2), que fue utilizada como formato de evaluación, constaba de la lista de todos los indicadores encontrados por Pearson (2015), esto es 423 en total, además de una columna de evaluación, en la cual determinarían los expertos si el indicador es relevante (1) o no (0), y finalmente una columna de observaciones.

Figura 2

Ejemplo de tabla de evaluación de indicadores

MCER	GSE	READING	Habilidad requerida	Observaciones
		GSE 10-21/ Below A1	0-1	
A1-	10	10.1 Can recognise cardinal numbers up to 10. (P)		
		10.2 Can recognise the letters of the alphabet (P)		
	12	12 Can read and understand simple prices. (P)		
	14	14.1 Can identify very common food and drink on a menu. (P)		
		14.2 Can recognise basic plural forms of nouns (e.g cars, books). (P)		
15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)			
		GSE 22-29/ A1		
A1	26	26 Can follow short, simple written directions (e.g. to go from X to Y). (C)		
	27	27 Can understand short written notices, signs and instructions with visual support. (P)		
		GSE 30-35/ A2		
A2	31	31.1 Can understand simple questions in questionnaires on familiar topics. (P)		
		31.2 Can find specific, predictable information in everyday materials (e.g. menus, timetables). (CA)		
		31.3 Can understand short, simple messages on postcards, emails and social networks. (CA)		
	34	34.1 Can understand the general meaning of short, simple informational material and descriptions if there is visual support. (CA)		
		34.2 Can understand simple instructions on everyday equipment (e.g. cash machines). (CA)		
		GSE 36-42/ A2(+)		

Además, como parte de la actividad 2.2, una vez que se obtuvieron los juicios de los tres expertos (ejemplo en Anexo 3), se realizó un análisis de sus comentarios y se revisaron en función de la clasificación propuesta por Pearson (2015).

Como siguiente paso se seleccionaron las habilidades y conocimientos correspondientes a los objetivos de la prueba. En esta actividad participaron cinco expertos en el área de psicometría, manejo de la lengua inglesa y la publicación de artículos científicos. Uno de los expertos tenía un perfil en traducción de artículos científicos, dos expertos eran especialistas en la lengua, tenían experiencia docente y en la publicación de artículos científicos en inglés. El cuarto juez era experto en psicometría y en la publicación de artículos científicos. El quinto era un experto en el inglés.

La selección de indicadores que se realizó tuvo como fin determinar a la prueba lo más apegada a los propósitos planteados.

Dentro de la primera etapa se utilizaron diferentes recursos para el diseño y desarrollo de la prueba, desde documentos de capacitación de expertos hasta tablas de contenido y evaluación. La tabla de comparación y cribado (Anexo 4) concentraba los indicadores y las evaluaciones realizadas por los expertos. El resumen de los criterios y participantes involucrados en el proceso de cribado se muestra en la Tabla 7.

Tabla 7*Crterios y participantes del cribado de indicadores*

<i>¿Qué criterios se consideraron en la eliminación de los indicadores?</i>	<i>¿Quiénes participaron en el cribado?</i>
-Indicadores que tuvieran el acuerdo entre los expertos considerados para el primer cribado	3 personas expertas en la lengua inglesa y tuvieran experiencia con textos académicos o con artículos científicos. Además, se consideró la clasificación de Pearson de aquellos indicadores pertenecientes al Inglés con Propósitos Académicos.
-Indicadores relacionados con la producción de textos académicos. -Indicadores con niveles taxonómicos altos -Se excluyeron indicadores que estuvieran relacionados con la preparación para redactar textos.	2 personas expertas en la lengua y en el diseño de pruebas.
-Indicadores que incluyeran tipos de texto fuera del científico-académico.	2 personas expertas en la lengua y en el diseño de pruebas.
-Indicadores que de actividades previas a la escritura o indicadores que pueden ser evaluados en un mismo ítem	1 experto en el idioma inglés y en el diseño y aplicación de pruebas.
-Indicadores que evaluaran lo mismo (se conservó el más alto en dificultad) -Indicadores de escritura (se utilizarán criterios en una rúbrica)	1 experto en la lengua y otro en el diseño de pruebas.

Etapa III. Elaboración de Especificaciones

El objetivo general de esta etapa fue elaborar las especificaciones de los ítems que conformarían la prueba.

En la actividad 3.1 se definió el nivel taxonómico de cada indicador utilizando la taxonomía de Anderson y Krathwohl (2001), y elaboró un formato de especificaciones como se muestra de ejemplo en la Figura 3 (para más información consultar el Anexo 5) donde se señalaban la intención evaluativa, definición de precisión, demanda cognitiva, foco evaluativo del ítem, y criterios sobre la información a emplear e indicaciones para responder el o los ítems. Así como, las características de la base del ítem, la respuesta correcta y los distractores.

Figura 3
Ejemplo de tabla de ítems y taxonomía

Habilidad	MCE R	Pearson n	Descriptor	Nivel taxonómico *	Nº ítems	Tipo ítem
Lectura	B1	45	45.3 Can skim a simple text to identify key concepts. (P)	Recordar		
		48	48.5 Can identify key information in a simple academic text, if guided by questions. (P)	Recordar		
		49	49.2 Can scan a simple academic text to find specific information. (P)	Recordar		
		50	50.2 Can understand the main idea of a passage using textual clues. (P)	Comprender		
		52	52.2 Can recognise the organizational structure of a paragraph in a simple academic text. (P)	Recordar		
B1+		54	54.3 Can distinguish between fact and opinion in a simple academic text. (P)	Analizar		
			54.4 Can distinguish between the main idea and related ideas in a simple academic text in order to answer specific questions. (P)	Analizar		
		58	58.4 Can synthesize information from two or more basic texts, if guided by questions. (P)	Comprender		
			58.5 Can understand relationships between ideas in a simple academic text, if guided by questions. (P)	Comprender		
B2	61	61.5 Can critically evaluate the effectiveness of a simple problem-solution essay. (P)	Evaluar			
B2+	68		68.5 Can identify common features of an academic abstract. (P)	Recordar		
			68.8 Can distinguish between active and passive voice in a linguistically complex academic text. (P)	Analizar		

Para dar cumplimiento a la actividad 3.2 sobre la elaboración de las especificaciones, se programó una reunión virtual por Google Meet con el experto que realizaría la tarea de elaborar las especificaciones. Durante la reunión se realizó la presentación del proyecto y la capacitación para completar la tarea. A esta reunión le siguió una reunión más de seguimiento y sugerencias. La deliberación enviada por el experto fue sometida al jueceo de otros 2 expertos.

Etapa IV. Elaboración de los ítems

Esta etapa consistió en la construcción de los ítems bajo las recomendaciones realizadas por Haladyna et al. (2002). Estas recomendaciones incluyen aspectos de contenido del ítem, formato del ítem, estilo del ítem, redacción de la base del ítem y redacción de las opciones de respuesta. Para realizar la tarea de la elaboración de los ítems, un experto en la lengua cumplimentó el formato de especificaciones e ítems muestra (ejemplo en Figura 4, consultar en Anexo 6) y se siguieron los criterios y características que aparecían en cada una de las especificaciones que fueron elaboradas en la fase anterior. Adicionalmente, se ensambló la prueba con los ítems generados (Anexo 7), además del cuestionario de contexto (Anexo 8). El producto final de esta etapa es contar con una primera versión de la prueba misma que se sometió al proceso de validación.

Figura 4

Ejemplo de formato para la especificación de ítems e ítems muestra del examen de inglés con propósito de publicación científica en posgrado.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se evaluará en qué medida el estudiante puede inferir el contenido de un texto académico complejo lingüísticamente, por lo tanto, se elaborará una especificación para producir un ítem	
	Definición de precisión:	
	Se considerará un texto académico complejo lingüísticamente aquel que contenga las siguientes características:	Demanda cognitiva Comprender
Ítem		
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	71.3 Can predict the content of a linguistically complex academic text by reading introductory and summary statements	<ul style="list-style-type: none"> Se presentará un <i>abstract</i> de no más de 10 renglones (Arial,10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
<ul style="list-style-type: none"> Especificación de la base del ítem: Especificación de la respuesta correcta a. Especificación del distractor b. 	<p>Se le presentará los términos temática o contenido general, y se le solicitará que los identifique. O en su caso se le presentará la temática o contenido general de un texto y se le solicitará que identifique al <i>abstract</i> que corresponde de entre diferentes textos.</p> <p>Será la temática y contenido general del texto que corresponde al <i>abstract</i>. O en su defecto será el <i>abstract</i> que corresponde a la temática presentada.</p> <p>Será una temática o contenido del texto que no corresponde al tema principal pero que esté relacionado, o bien será un <i>abstract</i> que no corresponde a la temática y contenido general del texto.</p>	
<ul style="list-style-type: none"> Especificación del distractor c. 	Será una temática similar a la temática o contenido principal del texto, o bien un <i>abstract</i> sobre un tema similar a la temática o contenido principal.	
<ul style="list-style-type: none"> Ítem muestra (ítem 1, versión 1): 	<p>Plagiarism is a major problem for research. There are, however, divergent views on how to define plagiarism and on what makes plagiarism reprehensible. In this paper we explicate the concept of "plagiarism" and discuss plagiarism normatively in relation to research. We suggest that plagiarism should be understood as "someone using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own" and argue that this is an adequate and fruitful definition. We discuss a number of circumstances that make plagiarism more or less grave and the plagiariser more or less blameworthy. As a result of our normative analysis, we suggest that what makes plagiarism reprehensible as such is that it distorts scientific credit. In addition, intentional plagiarism involves dishonesty. There are, furthermore, a number of potentially negative consequences of plagiarism.</p> <p>What is the abstract about?</p> <p>a. Plagiarism in research</p> <p>b. Scientific credit</p> <p>c. Plagiarism as a research topic</p>	

Etapa V. Validación de la prueba

Para esta fase se recopilaron evidencias de validez de aspecto de contenido. Como lo mencionan Meneses et al. (2013), este tipo de evidencias busca demostrar la relación que existe entre el ítem y los que se pretende evaluar. Para esta etapa se retomaron los productos de la Etapa 3 y 4, los cuales sirvieron de insumos para realizar las actividades de esta etapa.

La validación se realizó mediante el juicio de tres expertos. Uno de los expertos, era un especialista en la lengua, contaba con experiencia docente en el área de la investigación y había publicado en revistas científicas. El segundo experto se desempeñaba como docente de la lengua, además de tener experiencia en el diseño de pruebas y en la publicación científica. Finalmente, el tercer experto se desempeña como docente, contaba con dominio de la lengua y experiencia en publicaciones científicas.

Para el proceso de emisión de juicios por parte de los jueces, se les enviaron dos documentos y un archivo de Excel. Un documento contenía todas las especificaciones e ítems (Anexo 6) y el segundo, la primera versión de la prueba (Anexo 7) que incluía portada, presentación, cuestionario de contexto (Anexo 8) y la prueba. El archivo de Excel para la validación de contenido (ejemplo en Figura 5). Este archivo contó con nueve criterios: Correspondencia de ítem con su respectiva especificación, correspondencia entre la especificación y el contenido esencial que evalúa, corrección conceptual, penalización injusta, ofensa o estereotipo, apego a las convenciones de redacción técnica, demanda lingüística apropiada, demanda cognitiva apropiada, y escritura. Además, cada criterio contenía subcriterios en forma de pregunta a los cuales el juez contestó marcando la casilla Sí o No para cada ítem. Adicionalmente, se agregó una casilla de observaciones donde los jueces justificaron su respuesta o en algunos casos aportaron correcciones.

Figura 5

Ejemplo de formato de evaluación para evidencias de validez de contenido

Sección	Ítem	Correspondencia de ítem con su respectiva especificación								
		¿El ítem presenta todos los elementos especificados?			¿El ítem presenta los elementos de la manera en que fueron especificados?			¿Existe otra falla de correspondencia entre el ítem y la especificación?		
		Si	No	Observación	Si	No	Observaciones	Si	No	Observaciones
Lectura	What is the abstract about?									
	Which of the following key concepts are related to the text?									
	Which is the key information of the paragraph?									
	Which criteria do the authors suggest for defining plagiarism?									
	What is the main idea of the paragraph?									
	What's the organizational structure of the paragraph?									
	What is the author's opinion about plagiarism?									
	What is the main idea of the paragraph?									
	Which is the best synthesis of the text?									
	What's the relationship between the ideas?									
	According to the author of the article, what is the best solution for scientists to publish in English?									
	Which are the common features of an abstract?									
	Which of the following sentences are in passive voice?									
	What is the main idea of the paragraph?									

Para la obtención de evidencias de validez, se calcularon el Índice de Validez de Contenido (IVC) de Aiken (1985) y Tristán (2008). Primeramente, para el cálculo del IVC de Aiken se transformó los puntajes del 0 al 1, y con base en los nuevos datos, se aplicó la fórmula propuesta por Aiken (1985):

$$IVC = \frac{\bar{X} - l}{k}$$

Donde se considera IVC = índice de validez de Aiken, \bar{X} como el promedio de las calificaciones de todos los jueces, l como la calificación mínima y k como la diferencia entre la calificación máxima y la mínima. Para este caso, se realizó la sustitución:

$$IVC = \frac{\bar{X} - 0}{1}$$

Con la sustitución anterior, se calcularon los valores de IVC en los criterios de evaluación y se promediaron los puntajes para obtener el IVC por Ítem. Además, se promedió el IVC global.

Por otra parte, se calculó la Razón de Validez de Contenido (RVC) de Lawshe (1975) modificada por Tristán (2008) (RVC'), debido a que la cantidad de jueces no permitió utilizar la RVC de Lawshe. Para el cálculo de la RVC' se recurrió a las siguientes formulas:

$$RVC = \frac{n_e - N/2}{N/2}$$

Lawshe (1975)

$$RVC' = \frac{RVC + 1}{2}$$

Tristán (2008)

Etapa VI. Edición

Dentro de la etapa de edición se diseñó una tabla para concentrar las observaciones que realizaron los expertos a los ítems en función de los aspectos y criterios evaluados. Además, se clasificaron de acuerdo con el tipo de observación realizada. Una vez que se atendieron los comentarios, observaciones y sugerencias realizadas por los expertos, se procedió a la edición de la prueba, donde se le dio el formato de diseño final para la versión escrita. Al terminar dicho proceso se obtuvo la segunda versión de la prueba.

Etapa VII. Pilotaje

El pilotaje tuvo por objetivo revisar el funcionamiento de la prueba en una muestra que posea las características de la población a la que va dirigida la prueba. Esto permitió “examinar las propiedades métricas de la versión preliminar del instrumento de medida y ha de llevarse a cabo con personas similares a las que va dirigida la prueba.” (Muñiz y Fonseca-Pedrero, 2019, p. 11)

La actividad 7.1 consistió en la aplicación de la segunda versión de la prueba, en formato lápiz y papel. Se aplicó de manera estandarizada a estudiantes de un programa de Maestría en Ciencias Educativas con orientación a la investigación. El proceso de aplicación estandarizada de la prueba, inició con dar las instrucciones de la prueba a los examinados y solicitarles su consentimiento informado de participación. Además, se proporcionó a los examinados el documento que contenía la prueba y hojas blancas para contestar el reactivo 23, el cual es un ítem de ensayo donde se le solicita a los examinados que redacten un resumen de su proyecto de investigación o intervención según sea el caso. Al momento en que cada examinado terminaba la prueba, se le entrevistaba para conocer su opinión sobre la dificultad de la prueba, para ello se emplearon cuatro preguntas: 1. ¿Cuál fue la pregunta más difícil? 2. ¿Fue complicado el texto? 3. ¿Existió algún término/palabra que no comprendiste? 4. ¿Cuál?

Dentro de la actividad 7.2 se realizó la evaluación de los examinados. En la sección de lectura se utilizó el tren de respuestas para la evaluación, ya que era en formato opción múltiple. Para la sección de escritura, que era una pregunta abierta, se recurrió a tres expertos,

a los cuales se les proporcionó por correo electrónico los textos escaneados que produjeron los examinados, en este caso, solo se les envió siete escritos, ya que solo siete contestaron el ítem 23.

Además, se enviaron dos rúbricas, una enfocada en las convenciones de la lengua y la otra en las partes del resumen científico. La rúbrica que evalúa las convenciones de la lengua (Figura 6, para rúbrica completa consultar Anexo) fue adaptada de Education Northwest (2021), con puntuaciones del 1 como mínimo y el 6 como máximo. Los criterios de la rúbrica original fueron seleccionados mediante un cribado y alineación con los 26 indicadores finales de escritura, esta alineación se encuentra en la Tabla 16 del capítulo de resultados. La rúbrica que evalúa el contenido, esto es las convenciones de un texto científico en formato de resumen, fue tomada de Office of Science de la U.S. Department of Energy (s.f.) (Figura 7, para rúbrica completa consultar Anexo), con puntuaciones de 0 como mínimo y 2 como máximo. Para el caso de esta rúbrica solo se conservaron los criterios principales del resumen: Introducción, método y resultados.

Figura 6

Resumen de rúbrica de evaluación de la lengua

Descriptor	Descripción
<i>Main idea</i>	Conveys a clear, focused, complex, and original main idea that drives the piece
<i>Details, evidence, support</i>	Supports the main idea convincingly, with highly specific, credible information and/or striking details that go beyond the obvious
<i>Reasoning/ thinking</i>	Conveys unique, useful insights using higher-order thinking and convincing reasoning. Clarifies and enhances ideas.
<i>Textual evidence</i>	Incorporates very smoothly integrated textual evidence
<i>Lead/ conclusion</i>	Creates an inviting lead that establishes clear purpose, draws reader in, and creates a strong sense of anticipation. Develops a satisfying conclusion that conveys a powerful and thoughtful sense of closure
<i>Sequencing</i>	Utilizes highly effective sequencing, making best choices for progression and enrichment of reader's understanding
<i>Pacing</i>	Skillfully uses pacing to compel the reader through the text and enhance its impact
<i>Purpose/ text structure</i>	Utilizes text structure that enhances understanding of purpose and flows very smoothly
<i>Tone</i>	Employs a tone that is consistently appropriate for purpose and audience, clarifying and validating the message. Consistently uses appropriate level of formality
<i>Word meaning</i>	Consistently features precise, creative, and powerful vocabulary, including exact domain-specific vocabulary if appropriate
<i>Grammar</i>	Chooses powerful parts of speech, including expressive, active verbs and polished nouns/modifiers, to best convey exact meaning.
<i>Sentence structure</i>	Controls strong sentence structure for maximum impact throughout text
<i>Connecting sentences</i>	Joins sentences and sentence parts throughout paragraphs with creative and varied transitional words or phrases that highlight relationships between sentences
<i>Spelling</i>	Exhibits correct spelling, including more difficult words
<i>Punctuation</i>	Employs correct end and internal punctuation; may use creative punctuation that enhances readability

Figura 7

Resumen de rúbrica de evaluación del resumen

Descriptor	Descripción
<i>Introduction</i>	The introduction succinctly describes and appropriately connects the subject and context/background to the purpose of the investigation.
<i>Methods</i>	The methods section succinctly identifies the methods used to study the subject of the investigation OR succinctly identifies the existing procedures, models or programs used.
<i>Results</i>	The results section provides a succinct and specific explanation of what was discovered, accomplished, collected or produced.

Con base en la experiencia de la aplicación de la prueba a lápiz y papel, se realizaron modificaciones a la prueba para su versión digital (actividad 7.3). Los cambios que se hicieron fueron una respuesta a las necesidades identificadas durante fase de las entrevistas breves al terminar la aplicación de la prueba. Como parte de las modificaciones se reorganizó el orden de los ítems, donde el primero pasó a ser número 15. Además, se optó por omitir la lectura del artículo completa con el fin de minimizar el tiempo de aplicación de la prueba y la fatiga de los examinados.

En la actividad 7.4 se realizó el pilotaje en formato digital, por medio de la plataforma de Google Forms. Dicha versión de la prueba se aplicó a una muestra reducida de estudiantes de un programa de Maestría en Ciencias de la Educación con orientación profesionalizante. Sin embargo, solo se obtuvieron respuesta de 2 examinados.

Finalmente, la actividad 7.5 consistió de la evaluación de los escritos por medio de la inteligencia artificial, en este caso la generativa debido a los beneficios que aporta este tipo, sobre todo sus usos en la evaluación. Particularmente, se recurrió a la herramienta de ChatGPT versión 4.0 mini, al cual se le dio instrucciones de evaluar los escritos de acuerdo con los criterios de ambas rúbricas. Los escritos proporcionados fueron los del pilotaje a lápiz y papel, los cuales fueron cuidadosamente transcritos para cuidar la integridad del contenido, y los escritos del pilotaje digital. Las instrucciones para la primera y segunda rúbrica están descritas en las Figuras 8 y 9. Este proceso se realizó el 7 de agosto de 2024, todos los escritos fueron evaluados en un mismo día con el fin de reducir fluctuaciones en las evaluaciones, sea por actualizaciones o cambios de versión. Las evaluaciones arrojadas por la inteligencia artificial aportaron no solo una puntuación como se solicitaba en las instrucciones, sino que además para cada criterio, proporcionaba observaciones, e incluso una puntuación y evaluación general.

Figura 8

Extracto de instrucciones en Chat GPT para la evaluación de escritos con rúbrica de convenciones y usos de la lengua

Evaluate this text " " from 1 to 6 (1 as Beginning, 2 as Emerging, 3 as Developing, 4 as Capable, 5 as Experienced, and 6 as Exceptional) according to the following 15 criteria:
Main idea: Conveys a clear, focused, complex, and original main idea that drives the piece.
Details, evidence, support: Supports the main idea convincingly, with highly specific, credible information and/or striking details that go beyond the obvious.
Reasoning/ thinking: Conveys unique, useful insights using higher-order thinking and convincing reasoning. Clarifies and enhances ideas.
...

Figura 9

Instrucciones en Chat GPT para la evaluación de escritos con rúbrica de elaboración de un resumen de artículo científico.

Evaluate this text " " from 0 to 2 according to the following 3 criteria:
Introduction: The introduction succinctly describes and appropriately connects the subject and context/ background to the purpose of the investigation.
Methods: The methods section succinctly identifies the methods used to study the subject of the investigation OR succinctly identifies the existing procedures, models or programs used.
Results: The results section provides a succinct and specific explanation of what was discovered, accomplished, collected or produced.

Fase III. Análisis de psicométricos y descriptivos de la prueba

Para calificar la prueba y realizar los análisis psicométricos básicos se utilizaron el Corrector 2.2 que es una extensión para Excel. Además, el paquete estadístico IBM SPSS Statistics 21 (Jöreskog y Sörbom, 1989) para realizar los análisis descriptivos que permitieron caracterizar los resultados que obtuvieron los examinados. La síntesis de estos análisis se encuentra en la tabla 8.

Tabla 8*Actividades en el análisis de datos*

<i>Etapas</i>	<i>Actividades</i>
<i>I. Análisis preliminares</i>	1.1. Eliminación de valores perdidos y datos erróneos.
	1.2. Análisis descriptivo de las puntuaciones totales de la prueba y por ítem
	1.3. Exploración de la normalidad y eliminación de casos con puntuaciones atípicas.
	1.4. Verificación de los supuestos de normalidad, linealidad y multicolinealidad de las puntuaciones.
	1.5. Análisis de dificultad, adivinación y discriminación mediante los supuestos de la TCT.

A manera de cierre de este capítulo, la prueba diagnóstica de inglés con propósitos de publicación científica para posgrado se realizó bajo el programa de actividades que se presentan en la Tabla 9.

Tabla 9*Cronograma de actividades para el desarrollo de la prueba*

<i>Objetivo</i>	<i>Fase</i>	<i>Actividad</i>
	1. Marco general	1.1 Determinación de características del instrumento
	2. Definición de la variable de medida	2.1 Definir el marco o escala bajo el cual se realizará la prueba.
		2.2 Seleccionar las habilidades y conocimientos correspondientes a los objetivos de la prueba.
<i>Desarrollo</i>	3. Especificaciones	3.1 Definir el nivel taxonómico
		3.2 Elaborar las especificaciones
	4. Construcción de los	4.1 Elaborar los ítems para la primera versión de la prueba
<i>Evidencias de validez</i>	5. Validación	5.1 Juicio de expertos
		5.2 Análisis de los datos
	6. Edición	6.1 Edición de la prueba
		7.1 Pilotaje de la versión de la prueba en papel y lápiz
		7.2 Evaluación de escritos por jueces
		7.3 Edición de la prueba
		7.4 Pilotaje de la versión digital
7.5 Evaluación de escritos por IA		
	Propiedades psicométricas	Análisis de datos

Aspectos éticos

Finalmente, con respecto a los aspectos éticos considerados para el presente trabajo de investigación, se garantizó a los participantes que la información solicitada es confidencial y de uso exclusivo para los fines de la investigación. También se les informó que la participación es voluntaria y sin ningún tipo de repercusión académica. Finalmente, en cuanto al uso de las inteligencias artificiales, se reportan los programas, versiones y *prompts* señalados, con el fin de dar claridad al proceso.

Capítulo IV. Resultados

El presente capítulo expone los hallazgos obtenidos a lo largo de las distintas fases metodológicas del estudio, organizados según los objetivos y las etapas planteadas en el diseño de investigación. Se sintetizan los resultados de la Revisión Sistemática de la Literatura (RSL), el proceso de diseño y desarrollo de la prueba diagnóstica, así como los análisis psicométricos y descriptivos realizados sobre el instrumento aplicado. Además, se incluyen los hallazgos derivados del uso de inteligencia artificial como herramienta de evaluación en la sección de escritura.

Este capítulo no solo sintetiza los hallazgos obtenidos, sino que también ofrece una visión crítica de los datos recolectados, identificando fortalezas y áreas de mejora del instrumento, con el propósito de contribuir al desarrollo de herramientas más precisas y adecuadas para evaluar el dominio del inglés en contextos académicos específicos.

Fase I. Revisión Sistemática de la Literatura

Dentro de este apartado se explora los resultados del cribado (Tabla 10) y los hallazgos encontrados en la Revisión Sistemática de la Literatura (RSL) y se organiza en subapartados conforme a las preguntas de orientación que se plantearon.

Tabla 10

Resultados del cribado en la Revisión Sistemática de la Literatura

Índice	Resultados	Relevantes	Seleccionados
<i>Scopus</i>	181	181	28
<i>Web of Science</i>	50	13	4
	231	194	32

¿Cuáles son las propuestas para la evaluación del dominio del inglés en nivel superior?

Entre las principales propuestas que se mencionan se encuentran pruebas utilizadas a gran escala, desde internacionalmente como nacionalmente, hasta pruebas desarrolladas por universidades en particular y pruebas diseñadas con propósito de pre test y post test.

Entre las pruebas más conocidas y utilizadas internacionalmente se mencionan el TOEFL iBT (Chang, 2023; Gu y So, 2015) y el IELTS (Charge y Taylor, 1997). Estas pruebas

se mencionan dentro de los artículos analizados, ya sea como una revisión histórica de su diseño y desarrollo, o como parte de investigaciones cuasiexperimentales. Las puntuaciones en estas pruebas suelen ser un referente para diseños de metodologías pedagógicas o como referentes para el diseño de otras pruebas (Zainal y Jamil, 2015).

Para las pruebas utilizadas nacionalmente, ya sea que son aprobadas nacionalmente como certificaciones o que son requisitos nacionales dentro de las universidades como el caso del CET o CET-SET en China. Se encontraron ejemplos como ACERT en Polonia y CertAcles en España los cuales se analizaron y compararon ya que existen esfuerzos por hacer equivalencias sobre todo para facilitar la movilización de alumnos. Además, se mencionan pruebas como GatesGintie Reading Test en Canadá y Warwick English Language Test en Reino Unido que son utilizados en las investigaciones como referentes para cuasiexperimentos. Por otra parte, el MUET en Malasia sirvió junto con el IELTS como referente en el diseño de una prueba especializada en aspirantes a posgrado.

Dentro de otras pruebas diseñadas y desarrolladas por universidades se mencionan el Cambridge Preliminary English Test, PELA, Oxford Placement Test, Austest, Language cert IESOL B1, GATE. Algunas de estas pruebas fueron referentes para la creación de pruebas más específicas como el Khon Kaen University English Proficiency Test y el Productive Academic Word Test.

¿En qué países se han aplicado?

Como ya se mencionó dentro de las pruebas revisadas se encontraron aquellas que son utilizadas y reconocidas internacionalmente, esto es, IELTS y TOEFL iBT, y dentro de los estudios son analizados desde su trayectoria de diseño, desarrollo y validación, o como estudio de caso en países particulares como Taiwan (Chang, 2023).

Por otra parte, existen países que presentaron varios estudios, como es el caso de España con 4 (Barrios y Lopez-Gutierrez, 2021; Pérez, 2010; Martínez-Hernández y Bellés-Fortuño, 2021; Zabala- Delgado y Sawicka, 2019). Turquía tuvo 3 investigaciones (Altay et al., 2022; Çelik y Baran, 2021; Şata y Karakaya, 2021) al igual que China (Hu et al., 2022; Jin, 2019; Zhang, 2022). Dentro de los países que obtuvieron 2 estudios están Ucrania (Devos et al., 2021; Kharchenko et al., 2021), Reino Unido (Pearson, 2022; Sharpling, 2010) y Australia (Baird y Dooley, 2017; Farnill y Hayes, 1996).

Finalmente, otros países encontrados en la RSL fueron Norte de Chipre (Banaeian y Gilanlioglu, 2021), Montenegro (Božović et al., 2022), Canadá (Crossman, 2018), Letonia (Kalnberzina y Karapetjana, 2021), Portugal (Lousado et al., 2011), Tailandia

(Nithideechaiwarachok et al., 2022), Indonesia (Said y Weda, 2018), Republica Checa (Simonova, 2019), Polonia (Zabala-Delgado y Sawicka, 2019) y Malasia (Zainal y Jamil, 2015).

Se puede observar del análisis realizado, la ausencia de estudios realizados en Latinoamérica y particularmente en México en materia de evaluación del inglés en educación superior, al menos bajo las características de esta RSL.

¿Qué características tiene la población objetivo de las propuestas?

En general, se observa que la población se caracteriza por pertenecer o ser aspirante a la educación superior o educación terciaria. Incluso como lo menciona Kalnberzina y Karapetjana (2021) para el cuerpo académico en educación superior. Además, se encontraron trabajos especializados en investigación para contextos de minorías lingüísticas (Baird y Dooley, 2017) y en aspirantes a posgrados (Zainal y Jamil, 2015).

Por otra parte, resalta que algunos contextos se caracterizan por el uso del Inglés como Medio de Instrucción (EMI, por sus siglas en inglés). Este concepto se refiere a instituciones de Educación Superior en las que el inglés no es el idioma oficial y, sin embargo, es utilizado para la enseñanza de contenido en las aulas. Para este tipo de contextos están trabajos realizados por Banaeian y Gilanlioglu (2021), Barrios y Lopez-Gutierrez (2021) y Çelik y Baran (2021) entre otros.

¿Qué metodología utilizan para el diseño y desarrollo de pruebas?

El objetivo central de la RSL y de la presente investigación era explorar las metodologías utilizadas en el ámbito del diseño y desarrollo de pruebas. Sin embargo, dentro de los documentos revisados no todos incluían una metodología ya que en algunos casos los artículos se centraban en la descripción de una prueba (Zabala-Delgado y Sawicka, 2019; Zhang, 2022; Pearson, 2021; Pearson, 2022).

Por otra parte, algunos documentos se centraban como cuasi experimentos, en los que utilizaban a las pruebas como referentes tipo pre test o pos test (Altay et al., 2022; Banaeian y Gilanlioglu, 2021; Barrios y Lopez-Gutierrez, 2021; Chang, 2023; Gu y So, 2015). Otras investigaciones realizaron sus propios pre test o pos test, sin embargo, el documento no da cuenta de la metodología que se utilizó en su elaboración, o solo se menciona el material de donde se tomaron las tareas para construir los ítems. En otros casos se hacían adecuaciones o sugerencias a las pruebas con el fin de adaptarse a las necesidades detectadas por los autores, o para cumplir con los objetivos de la investigación (Baird y Dooley, 2017; Crossman, 2018; Hu

et al., 2022; Sharpling, 2010).

Uno de los argumentos encontrados en los documentos revisados fue la necesidad de realizar un análisis de necesidad, ya que argumentan Nithideechaiwarachok et al. (2022) que, a pesar de tener la acreditación del idioma, los alumnos en educación superior no cumplen con las expectativas en el dominio del inglés. De igual manera, Božović et al. (2022) mencionan la importancia de atender a las demandas locales e internacionales, es importante las recomendaciones basadas en las necesidades en materia de políticas de lenguas extranjeras.

Finalmente, se encontraron artículos que detallaban el proceso de diseño y desarrollo de una prueba como fue el caso de la prueba GATE (Zainal y Jamil, 2015). Para el diseño de esta prueba, los autores establecen 3 fases: desarrollo, operación y análisis. A grandes rasgos cada fase se divide de la siguiente manera, en la primera fase determinan el marco, las especificaciones, escalas de calificación y los ítems, además se imparten talleres para los examinadores y calificadores, y se desarrolla la plataforma para aplicación en computadora. Para la segunda fase se realiza la aplicación del test a papel y computadora. En la tercera fase se hace el análisis de los resultados y la validación del test. Este test busca aportar una prueba que responda a las necesidades particulares del contexto de Malasia, y sobre todo que pueda participar con las competencias necesarias en los programas de posgrado.

Por parte de Kalnberzina y Karapetjana (2021) sugirieron el desarrollo de un sistema de evaluación de miembros del cuerpo académico en contextos de educación superior EMI. Para este caso, los autores conformaron un conjunto de ítems que fueron piloteados y validados. Además, los autores mencionan casos similares de diseño de pruebas para académicos en inglés como lengua extranjera. Entre ellos destacan para los fines de este trabajo aquellos enfocados en la escritura o publicación de artículos (Plews y Amos, 2020) y los que se enfocan en el inglés académico (Marsaulina, sf.).

¿Qué pruebas evalúan el inglés con fines académicos?

Entre las pruebas observadas en la RSL solo algunas estuvieron focalizadas en el inglés académico. Tal es el caso de TOEFL iBT, IELTS, PAWT, WELT, ACERT, CertAcles, GATE, incluyendo la adaptación al CET realizada por Hu et al., (2022). Además, algunas de las pruebas desarrolladas por los autores también estuvieron centradas en el inglés académico (Kalnberzina y Karapetjana, 2021; Simonova, 2019; Şata y Karakaya, 2021).

¿Cuáles son los marcos o modelo más utilizados?

Para la mayoría de las pruebas observadas el modelo utilizado fue el Marco Común

Europeo de Referencias (MCER). Aunque dependiendo la prueba puede que cambie las bandas en las que se puntúe a los examinados. Como es el caso de la prueba IELTS que otorga puntuaciones en bandas de 0 a 9, o el TOEFL iBT que tiene un rango de 0 120 puntos. Para ambas pruebas existen equivalencias validadas con el MCER. El caso particular del IESOL (Pearson, 2022) menciona el autor que las pruebas están focalizadas por nivel del MCER, es decir que para cada nivel (A1, A2, B1...) existe una prueba que es puntuada de 0 a 50.

Para la prueba WELT, se menciona que las bandas existen del A al E, según el dominio de la lengua. Por otra parte, la prueba CET se divide en 2 bandas (4 y 6) las cuales determinan el nivel de dominio de la lengua, sea básico (4) o avanzado (6).

¿Qué habilidades o conocimientos evalúan las pruebas?

Para la evaluación, la mayoría de las pruebas atienden a las 4 habilidades (lectura, escritura, habla y escucha). Sobre todo, en aquellas que son a nivel internacional o nacional, ya que en su mayoría buscan otorgar al examinado una certificación. La selección de las habilidades a medir depende de los objetivos que persigue la investigación, ya que existen casos donde el contexto tiene un papel relevante en la habilidad que se observa. Por ejemplo, en las instituciones EMI, las habilidades de producción y comprensión oral son necesarias en el trayecto académico.

En el caso de las pruebas que están enfocadas en el inglés académico, ya sea por contexto EMI, por posgrados (sobre todo de investigación), o enfocados a académicos las habilidades más recurrentes son la lectura y la escritura. Además, se destaca que algunas pruebas añaden como habilidad la gramática y el uso de la lengua.

¿Cuáles son los fines del uso de las pruebas?

Entre los principales usos que se les dan a las pruebas son de alto impacto. Por ejemplo, de admisión, ya que como se mencionó algunos contextos como el EMI requieren el dominio de la lengua como requisito de admisión (PELA, KEPT, TOEFL, IELTS, IESOL, WELT, ACERT, CertAcles, GATE). En otros casos, las pruebas son utilizadas como requisito de egreso, esto es común en contextos de inglés como lengua extranjera (EFL, por sus siglas en inglés), como es el caso de CERT.

Finalmente, dentro de los usos de bajo impacto está el Austest, que de acuerdo con Farnill y Hayes (1996) se diseñó como una forma de cribar a los alumnos de recién ingreso en educación superior e identificar aquellos con necesidades en el uso del inglés y poder dar

acompañamiento y asegurar el éxito académico. Por otra parte, en algunos casos se utilizaron como datos las puntuaciones como referentes en las investigaciones (éxito académico, éxito en el uso de las Tecnologías de la Información y la comunicación en el aula, entre otras estrategias pedagógicas). En estas pruebas se busca mejorar el desempeño académico, y la producción académica.

Fase II. Diseño y desarrollo de la prueba

Este apartado se presentan los resultados obtenidos en la fase de diseño y desarrollo de la prueba de inglés para posgrado con propósitos de publicación científica.

Etapa 1. Marco general

La RSL sirvió para determinar el marco general de la prueba y orientarla hacia el contexto de la UABC y de manera particular a las necesidades de los estudiantes de posgrado. De esta manera, se obtuvo una prueba de inglés que estuviera centrada en la redacción de textos académicos para alumnos de posgrado. En la Tabla 11 se muestran las características esenciales de la prueba contenidas en el marco general.

Tabla 11

Marco general del instrumento

<i>Justificación</i>	Ausencia de pruebas especializadas para la evaluación del dominio del inglés en el campo de la publicación científica para estudiantes de posgrado
<i>Contexto</i>	Universidad Autónoma de Baja California, campus Ensenada.
<i>Población objetivo</i>	Alumnos de maestría en educación de dos programas distintos, uno con orientación hacia la investigación y otro profesionalizante.
<i>Tipo y formato de aplicación</i>	Dos versiones, una lápiz y papel, y otra digital.
<i>Propósito</i>	Diagnóstico

Etapa 2. Definición de la variable de medida

Para el desarrollo de la prueba se consideraron solamente las habilidades de lectura y escritura de acuerdo con los diferentes marcos, modelos y pruebas que se revisaron. En la Tabla 12 se presenta un comparativo de las características de los modelos y pruebas que sirvieron como referencia para el desarrollo de la prueba.

Tabla 12

Características de los modelos y pruebas de referencia para el desarrollo de la prueba.

<i>Modelos y pruebas</i>	<i>Niveles de medida o escalas de puntuación</i>	<i>Indicadores</i>	<i>Clasificación Académica</i>
<i>MCER</i>	A1/Turista -C2	Descriptorios	No
<i>GSE</i>	10 -90	Objetivos de aprendizaje	Sí
<i>ACTFL</i>	Principiante bajo - Distinguido	Guías	No
<i>PTE Academic</i>	10 -90	-	Sí
<i>TOEFL iBT</i>	57 -120	-	Sí
<i>TOEFL ITP</i> (antes <i>PBT</i>)	490 - 677	-	Sí
<i>IELTS</i>	4.5 – 9.0	-	Sí

Para la selección del modelo o marco que se utilizó para el diseño de la prueba, se consideraron aspectos como los niveles de medida o las escalas de puntuación, el tipo de indicadores utilizados, y si aportaban información sobre el dominio del inglés académico. En cuanto a los niveles de medida o las escalas de puntuación, se encontró que el MCER utiliza siete niveles, e igualmente, el American Council on the Teaching of Foreign Languages (ACTFL). El resto utilizan un sistema de puntuación continuo, lo cual permite reportar información más precisa el dominio del idioma. En cuanto a los indicadores, solamente se encontró información en el MCER, GSE y ACTFL, sin embargo, la ACTFL solo describe de manera general las competencias por nivel, sin llegar a ser específico como los otros dos. Finalmente, Casi todos los materiales analizados se presentaban información sobre el inglés académico, a excepción del MCER y el ACTFL.

Para el desarrollo de la prueba se determinó usar el modelo de la GSE, ya que los niveles de medida son granulares, presenta indicadores específicos por nivel y ofrece una

clasificación de indicadores de acuerdo con los objetivos de inglés académico. Por otra parte, en el desarrollo de la GSE se recopiló distintos indicadores de otros marcos y ediciones (Pearson, 2015). Estos indicadores, o competencias concretas, los concentró en un solo documento, el cual permitió obtener un universo amplio de la variable, para ser cribado más adelante.

Para el cribado se realizaron cinco ejercicios de jueceo, de los cuales se obtuvieron 22 indicadores de lectura y 26 de escritura como se muestra en la Tabla 13, de acuerdo con los criterios y participantes mencionados en el capítulo de método. En cuanto a los indicadores obtenidos, se organizaron de acuerdo con las categorías encontradas en la literatura (Tabla 14), de esta manera se identificó una mayor presencia de indicadores en la categoría de lectura y escritura académica. Siguiendo las recomendaciones de expertos se seleccionaron criterios de evaluación de la escritura que estuvieran alineados a los criterios de la escala de Pearson.

Tabla 13
Proceso de cribado de indicadores

<i>Etapa</i>	<i>Total</i>	<i>Eliminados</i>
1	Lectura: 164 Escritura: 259 Total: 423	Lectura: 104 Escritura: 172 Total: 276
2	Lectura: 60 Escritura: 87 Total: 147	Lectura: 29 Escritura: 32 Total: 61
3	Lectura: 31 Escritura: 55 Total: 86	Lectura: 6 Escritura: 15 Total: 21
4	Lectura: 25 Escritura: 40 Total: 65	Escritura: 12
5	Lectura: 25 Escritura: 28 Total: 53	Lectura: 3 Escritura: 2
6	Lectura: 22 Escritura: 26	

Tabla 14*Categorización de indicadores*

<i>Categoría</i>	<i>Indicadores</i>
<i>General English Skills</i>	L45.3, L48.5, L49.2, L50.2, L54.4, L68.8, L70.3, E65.10, E74.7
<i>Academic Reading and Writing skills</i>	L.52.2, L.58.4, L58.5, L68.5, L70.4, L74.7, L77.1, E50.8, E56.3, E 58.6, E 58.8, E58.9, E59.10, E61.7, E61.8, E64.3, E70.4, E70.5, E70.6, E74.3, E75.2, E75.3, E75.5, E76.3, E76.4, E78.2
<i>Academic English Skills</i>	L54.3, L79.5, E66.6, E67.10,
<i>Research skills</i>	L61.5, L71.3, L74.2, L77.3, L79.2, L85.2, E70.7, E87.2

Etapa 3. Especificaciones

Con los resultados se diseñó un formato con las especificaciones de los ítems como se muestra en la Tabla 15. Este documento redacta de manera detallada las especificaciones a cada indicador incluyendo, los objetivos, la demanda cognitiva o el nivel taxonómico, definiciones de precisión de ser necesarias, las especificaciones para la base del ítem y de las opciones de respuesta.

Tabla 15

Tabla resumen de habilidad de lectura con sus niveles taxonómicos por nivel de MCERL

<i>Habilidad</i>	<i>MCER</i>	<i>Nivel taxonómico*</i>
<i>Lectura</i>	<i>B1</i>	Comprender
		Recordar
	<i>B1+</i>	Analizar
		Comprender
	<i>B2</i>	Evaluar
	<i>B2+</i>	Recordar
		Analizar
		Evaluar
	<i>C1</i>	Comprender
		Analizar
Recordar		
Aplicar		
<i>C2</i>	Analizar	

*Taxonomía de Anderson y Krathwohl (2001)

Finalmente, se obtuvo la Tabla 16 de alineación de los indicadores de la habilidad de escritura con los criterios que serían evaluados en las rúbricas, además de ser clasificados de acuerdo con las categorías encontradas en la literatura. Dentro de la categorización realizada se encontró que la mayoría de indicadores se concentran en la categoría de escritura académica, dejando seis indicadores para las otras tres categorías, dos en inglés académico, dos en inglés general y dos en investigación.

Tabla 16

Tabla de alineación de indicador con criterio evaluado en las rúbricas

<i>Habilidad</i>	<i>MCER</i>	<i>Pearson</i>	<i>Indicador</i>	<i>Criterio de rúbrica</i>	<i>Categoría</i>
Escritura	<i>B1</i>	<i>50</i>	50.8 Can paraphrase information from a simple text, if guided by questions. (P)	Main idea	Academic Writing
		<i>56</i>	56.3 Can write a short, simple academic essay on a familiar topic, given a model. (P)	Sentence structure Grammar	Academic Writing
	<i>B1+</i>		58.6 Can write a detailed description of a simple process. (P)	Details, evidence, support	Academic Writing
		<i>58</i>	58.8 Can write a chronological paragraph on an academic topic, if provided with a model. (P)	Sequencing	Academic Writing
			58.9 Can write a conclusion to a simple academic essay. (P)	Lead/ conclusion	Academic Writing
		<i>59</i>	59.10 Can write an introduction to a simple academic essay. (P)	Introduction	Academic Writing
			61.7 Can signal that two ideas are similar when writing a simple academic text by using discourse markers. (P)	Connecting sentences	Academic Writing
	<i>B2</i>		61.8 Can contrast two ideas when writing a simple academic text by using discourse markers. (P)	Connecting sentences	Academic Writing
		<i>64</i>	64.3 Can write a chronological paragraph on an academic topic. (P)	Purpose/ text structure	Academic Writing
		<i>65</i>	65.10 Can explain a term in a text using synonyms, definitions or examples. (P)	Word meaning	General English
		<i>66</i>	66.6 Can use fact and opinion effectively in writing. (P)	Textual evidence	Academic English
	<i>B2+</i>	<i>67</i>	67.10 Can use descriptive language to support a main idea in written academic work. (P)	Details, evidence, support	Academic English

<i>Pearson</i>	<i>Indicador</i>	<i>Criterio de rúbrica</i>	<i>Categoría</i>
	70.4 Can start an essay with a strong thesis statement. (P)	Tone	Academic Writing
	70.5 Can use parallel structure in academic writing. (P)	Purpose/ text structure	Academic Writing
70	70.6 Can clarify a concept in a narrative text using adverbial clauses, phrases and modifiers. (P)	Details, evidence, support	Academic Writing
	70.7 Can write a critical analysis of a simple academic text. (P)	Reasoning/ thinking	Reaserch
	74.3 Can write a bibliography showing an understanding of conventions. (P)	Textual evidence	Academic Writing
74	74.7 Can identify and correct errors in a piece of academic writing. (P)	Spelling Punctuation	General English
	75.2 Can write an effective and informative abstract. (P)	Pacing	Academic Writing
	75.3 Can use information from academic texts to support a thesis. (P)	Reasoning/ thinking	Academic Writing
75	75.4 Can synthesize information from two or more academic texts. (P)	Main idea	Academic Writing
	75.5 Can refute a counter argument in written work. (P)	Reasoning/ thinking	Academic Writing
	76.3 Can use complex numerical values in an academic text and explain their significance to the reader. (P)	Results	Academic Writing
76	76.4 Can use citations effectively and appropriately in an academic paper. (P)	Textual evidence	Academic Writing
C1			
	78.2 Can use persuasive techniques appropriately in academic writing (P)	Method Tone	Academic Writing
78			
C2	87	87.2 Can write an academic article for publication in their field. (CA)	Results Reaserch

Etapa 4. Elaboración de ítems

Con respecto al resultado de la etapa anterior, se añadieron al Anexo 6 la base del ítem junto con sus opciones de respuesta para cada indicador, esto conformó la primera versión de la prueba. La prueba está conformada por una portada, una introducción, presentación, el consentimiento informado y las instrucciones, un cuestionario de contexto conformado por 13 preguntas, la sección de lectura con 22 preguntas de opción múltiple y una pregunta abierta para la escritura, que es en total 23 ítems. En cuanto a las instrucciones se muestran en la Figura 10, las cuales fueron parte de la presentación en Anexo 6.

Figura 10

Instrucciones para contestar la prueba

Esta prueba consta de dos secciones y un cuestionario de contexto. Pero antes de comenzar a responder la prueba es necesario que leas un artículo, te pedimos que lo leas con mucha atención, ya que las preguntas que estás por contestar se relacionan con dicho documento. Una vez que termines de leer el artículo, responderás un cuestionario de contexto, para luego pasar a la primera sección; la cual consta de 22 ítems que exploran el nivel de dominio que tienes para leer y comprender textos científicos en inglés. En esta sección todos los ítems son de opción múltiple y solo tiene tres opciones de respuesta. Solo en algunos casos, se te pregunta por la mejor respuesta, y en otros tienes que seleccionar la respuesta correcta. Para hacer más fluido el ejercicio, en algunos casos se te presenta la parte del texto que corresponde a lo que se te solicita, lo cual evita que regreses a leer el artículo de nueva cuenta.

En la segunda sección consta de una sola pregunta y debes escribir en inglés lo que se te pide. Debes de tomar en cuenta que se te pide un mínimo de palabras y un máximo, te pedimos que la producción escrita se ajuste a lo solicitado y consideres tus conocimientos de metodología de la investigación para dar estructura al texto.

Para la evaluación de la prueba, se generaron tres productos. Primeramente, una cadena de respuestas para preguntas de opción múltiple, además se adaptaron dos rúbricas para la evaluación de la escritura. Una rúbrica fue adaptada de Education Northwest (2021) que evalúa las convenciones de la lengua (Anexo 9), y la otra evalúa el contenido, esto es las convenciones de un texto científico, y se tomó de Office of Science de la U.S. Department of Energy (s.f.) (Anexo 10).

Basado en los ítems obtenidos, se logró alinear con las categorías encontradas en la literatura como se muestra en la Tabla 17.

Tabla 17*Categorización de ítems*

<i>Categoría</i>	<i>Ítems</i>
<i>General English Skills</i>	1, 2, 3, 4, 7, 12, 13, 23
<i>Academic Reading and Writing skills</i>	5, 8, 9, 11, 14, 17, 18, 23
<i>Academic English Skills</i>	6, 21, 23
<i>Research skills</i>	10, 15, 16, 19, 20, 22, 23

Se identificó a través de esta categorización que, en cuanto a ítems, en la prueba existe un predominio en las categorías de habilidades de inglés general, y en las habilidades de lectura y escritura académica, con ocho ítems cada una. Por otra parte, la categoría de habilidades de investigación constituye el segundo más representado con siete ítems y finalmente la categoría de habilidades de inglés académico con tres. Estos resultados eran los esperados, ya que, en la operacionalización de las categorías, la categoría con menos competencias fue la de inglés académico, por lo tanto, obtuvo menos representación.

Etapa 5. Validación de contenido de la prueba

Como se mencionó en el capítulo del método la validación de contenido se realizó gracias a la participación de tres expertos. Sus respuestas fueron concentradas en un documento Excel (Anexo 11). Los criterios considerados para la evaluación de los expertos fueron nueve: correspondencia de ítem con su respectiva especificación, correspondencia entre la especificación y el contenido esencial que evalúa, ausencia o presencia de errores conceptuales, penalización injusta, ofensa o estereotipo, apego a las convenciones de redacción técnica, demanda lingüística apropiada, demanda cognitiva apropiada y escritura. Con base en estos criterios se calcularon el coeficiente de validez de contenido de Aiken (1985) y la Razón de Validez de Contenido de Lawshe (1975) modificada por Tristán (2008) por subcriterio (Tabla 18) y por ítem (Tabla 19).

Tabla 18*Resumen de coeficiente de Aiken y Tristán por subcriterio*

<i>Criterio</i>	<i>Subcriterio</i>	<i>Aiken (V)</i>	<i>Tristán</i>
<i>Correspondencia de ítem con su respectiva especificación</i>	¿El ítem presenta todos los elementos especificados?	0.96	0.96
	¿El ítem presenta los elementos de la manera en que fueron especificados?	0.93	0.93
	¿Existe otra falla de correspondencia entre el ítem y la especificación?	0.97	0.97
<i>Correspondencia entre la especificación y el contenido esencial que evalúa</i>	¿La especificación delinea un ítem que evalúa lo importante del contenido?	0.99	0.99
<i>Corrección conceptual</i>	¿El ítem está libre de errores conceptuales?	0.97	0.97
<i>Penalización injusta+QI:AQI</i>	¿El ítem pone en ventaja a examinados de cierto nivel socioeconómico o cultural?	1	1
	¿El ítem pone en ventaja a los examinados de algún género?	1	1
	¿El ítem pone en ventaja a otra clase de examinados?	1	1
<i>Ofensa o estereotipo</i>	¿El ítem ofende a examinados de cierto nivel socioeconómico o cultural?	1	1
	¿El ítem ofende a los examinados de algún género?	1	1
	¿El ítem ofende a los examinados que provienen de otro país o ciudad?	1	1
	¿El ítem ofende a los examinados con capacidades limitadas?	1	1
	¿El ítem ofende a los examinados de zonas rurales, marginadas o indígenas?	1	1
	¿El ítem ofende a otra clase de examinados?	1	1
<i>Apego a las convenciones de redacción técnica</i>	¿El ítem presenta determinantes específicos como claves contextuales o gramaticales, distractores poco plausibles, etc.?	0.93	0.93

<i>Criterio</i>	<i>Subcriterio</i>	<i>Aiken (V)</i>	<i>Tristán</i>
<i>Demanda lingüística apropiada</i>	¿El ítem presenta una demanda lingüística excesiva o vocabulario complejo, o redacción poco familiar o términos técnicos innecesarios?	0.88	0.88
<i>Demanda cognitiva apropiada</i>	¿El ítem presenta una estructura congruente con el tipo de contenido que evalúa, ya sea un concepto o procedimiento?	0.93	0.93
	¿El ítem suscita los procesos cognitivos pretendidos, según el dominio que se pretende evaluar o el nivel cognitivo especificado?	0.97	0.97
<i>Escritura</i>	¿El ítem carece de errores ortográficos y de impresión?	0.81	0.81
	¿La redacción de ítem es clara y esencial?	0.68	0.68

Tabla 19

Resumen de coeficientes de Aiken y Tristán por ítem

<i>Ítem</i>	<i>Aiken</i>	<i>Tristán</i>
<i>IL1</i>	0.98	0.98
<i>IL2</i>	0.97	0.97
<i>IL3</i>	0.97	0.97
<i>IL4</i>	0.92	0.92
<i>IL5</i>	0.97	0.97
<i>IL6</i>	0.87	0.87
<i>IL7</i>	0.9	0.9
<i>IL8</i>	0.92	0.92
<i>IL9</i>	1	1
<i>IL10</i>	0.98	0.98
<i>IL11</i>	0.97	0.97
<i>IL12</i>	0.95	0.95
<i>IL13</i>	0.93	0.93
<i>IL14</i>	0.98	0.98
<i>IL15</i>	0.97	0.97
<i>IL16</i>	1	1

<i>ítem</i>	<i>Aiken</i>	<i>Tristán</i>
<i>IL17</i>	0.98	0.98
<i>IL18</i>	0.97	0.97
<i>IL19</i>	0.95	0.95
<i>IL20</i>	0.87	0.87
<i>IL21</i>	0.95	0.95
<i>IL22</i>	0.93	0.93
<i>IE1</i>	0.95	0.95

De acuerdo con los valores mínimos de criterio de calidad obtenidos, esto es 0.58 para el RVC' de Tristán y 0.73 para el IVC de Aiken, se observa que en general los resultados son aceptables. Sin embargo, se puede observar en la Tabla 18 que en el criterio de escritura y subcriterio “¿la redacción es clara y esencial?” la V de Aiken no alcanza el criterio de calidad, ya que solo cuenta con un 0.68. Para mayor información, consultar Anexo 12 que presenta la información sobre la IVC de Aiken y Anexo 13 con la RVC de Tristán por criterio y por ítem.

Etapa 6. Edición

Con base en las observaciones realizadas por los expertos (Tabla 20) se realizaron las modificaciones al instrumento. Esta edición dio como resultado la segunda versión de la prueba, así como la edición del documento de especificaciones. La edición consistió principalmente en cuestiones de redacción de la base del ítem, las respuestas o el texto, ortografía y congruencia con la especificación.

Tabla 20*Tipo de observaciones realizadas por los jueces a los ítems de la prueba*

<i>Habilidad</i>	<i>MCER</i>	<i>Pearson</i>	<i>Total de ítems</i>	<i>Num. de ítems observados</i>	<i>Tipo de observaciones</i>
Lectura	<i>B1</i>	45	4	4	Redacción de base del ítem
		48			Redacción del texto
		49			Ortografía
		50			Concepto
	<i>B1+</i>	52	5	4	Tamaño del texto/ especificaciones
		54			Ortografía
					Redacción del texto
		58			Base del ítem
				Formulación de distractor	
	<i>B2</i>	61	1	1	Congruencia de redacción base y respuestas
	<i>B2+</i>	68	4	4	Ortografía
					Redacción de la base del ítem
		70			Redacción de respuestas
				Unificar el vocabulario	
	<i>C1</i>	71	7	6	Redacción de base del ítem
74		Uso de vocabulario			
		Uso de lenguaje			
77		Ortografía			
				Redacción de respuestas	
	79			Congruencia de especificación- ítem	
<i>C2</i>	85	1	1	Redacción de la base del ítem	
Escritura		1	1	Especificación de la respuesta con los elementos del resumen	
				Redacción de la base del ítem	

Etapa 7. Pilotaje

Dentro de esta etapa se obtuvieron como resultado, observaciones realizadas durante el pilotaje al primer programa. Además, se realizó una breve encuesta sobre la prueba, los datos se concentraron en la Tabla 21. Las respuestas obtenidas concuerdan con los resultados obtenidos del pilotaje, ya el reactivo 23 se mencionó más veces como el más difícil y este mismo reactivo fue el menos contestado y más complejo. En cuanto a las otras dos preguntas sirvieron de referencias sobre la dificultad de la sección de lectura.

Tabla 21*Concentrado de entrevistas en pilotaje a lápiz y papel*

<i>Examinado</i>	<i>¿Cuál fue la pregunta más difícil?</i>	<i>¿Fue complicado el texto?</i>	<i>¿Existió algún término/palabra que no comprendiste? ¿Cuál?</i>
1	23	Sí	Sí, conectores de ideas
2	16	No	No
3	23	No	Sí
4	23	No	Sí
5	6	No	Sí
6	23	Sí	Sí
7	23	Sí	Sí
8	23	Sí	Sí
9	23	Sí y extenso	Sí

Los escritos del pilotaje a papel fueron evaluados tanto por humanos como por la inteligencia artificial. Debido a las instrucciones otorgadas a los jueces humanos, solo se obtuvieron puntuaciones de acuerdo con las rúbricas proporcionadas. Por lo tanto, no se contó con evaluaciones o *feedback* de su parte.

En cuanto a las evaluaciones de los escritos realizados por ChatGPT, se encontró que no solo proporcionaba una puntuación por criterio y global, sino que también proporcionaba observaciones de acuerdo con cada criterio. Para el caso de la rúbrica de usos y convenciones de la lengua, proporcionó una evaluación general, y puntualizaba en algunos errores, mayormente en aspectos de deletreo, gramática, entre otros. En la rúbrica de resumen, la inteligencia artificial, proporcionó una evaluación global, además de una descripción de las fortalezas y debilidades por cada criterio. Como se muestra de ejemplo al examinado cinco en la Tabla 22, la Inteligencia Artificial, pudo reconocer aspectos de contexto de la investigación, literatura y objetivos, además de ofrecer aspectos de mejora. Para consultar las evaluaciones a todos los examinados en cada uno de los criterios de ambas rúbricas, ir a Anexo 14.

Tabla 22

Muestra de evaluación realizada por ChatGPT al examinado 5 en criterios de Main Idea e Introduction

	<i>Usos y convenciones de la lengua: Main idea</i>	<i>Resumen de artículo científico: Introduction</i>
<i>Evaluación</i>	The main idea is clear and focused: to elaborate and validate an instrument for measuring Teachers Wellbeing (TWB) based on a theoretical framework. The complexity is present, though the idea could be presented with more depth and originality.	Assessment: The introduction effectively describes the complexity and variety of conceptualizations related to Teacher Wellbeing (TWB) and highlights the gap in existing research. It clearly states the problem and sets up the context for the research. Strengths: Provides a thorough background on TWB and identifies a gap in the literature, which is critical for understanding the research's relevance. Weaknesses: While comprehensive, it could benefit from a clearer connection to the research objectives and the significance of addressing the identified issues.
<i>Puntuación</i>	4/6	2/2

Fase III. Análisis psicométricos y descriptivos de la prueba

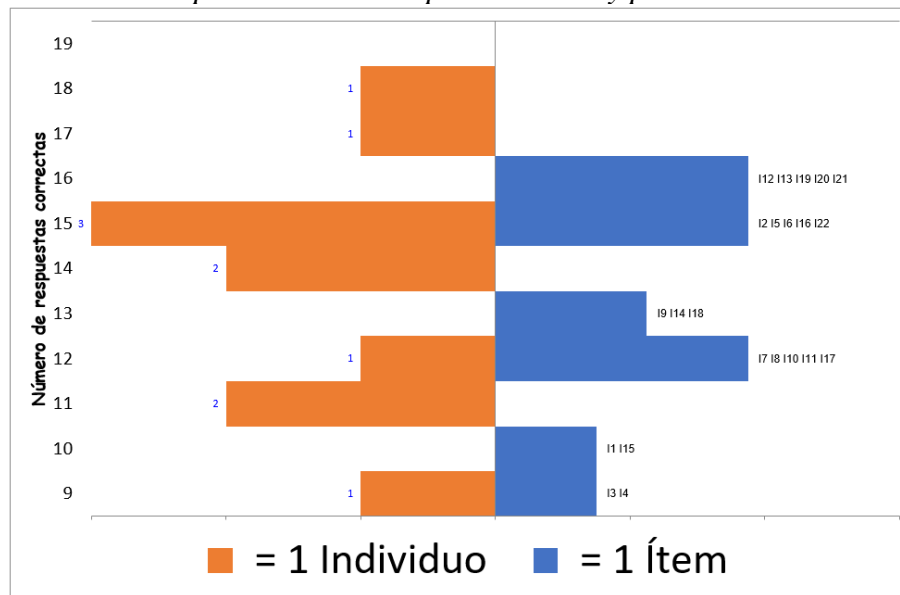
Dentro de este apartado se estarán presentado los resultados de los análisis psicométricos de la prueba. La organización de la información está dividida en dos subapartados, el primero corresponde al análisis del comportamiento de la prueba, esto es la discriminación, habilidad, distribución de aciertos y comportamiento de los distractores. El segundo apartado está enfocado en los examinados y sus resultados en la prueba. En el tratamiento de los datos obtenidos, se estandarizaron las puntuaciones con el fin de ser comparadas a otras pruebas, que es este caso es la prueba TOEFL, de manera que las puntuaciones tienen una media de 500 puntos y una desviación de 100 puntos.

Comportamiento de la prueba.

Para este apartado se utilizó la extensión Corrector 2.2 (Gaviria, 2018) a la paquetería de Excel Microsoft y se analizaron los 22 ítems correspondientes al apartado de lectura. Esta extensión permitió obtener información sobre la media de aciertos de 13.73, con un máximo de 18 y un mínimo de 9 aciertos, además de una distribución de aciertos por individuo y por ítem como se muestra en la Figura 11. En esta figura, se encuentra que los ítems 12, 13, 19, 20, 21 obtuvieron el máximo de respuestas correctas.

Figura 11

Número de respuestas correctas por individuo y por ítem



Por otra parte, el programa arrojó valores sobre la habilidad y discriminación, donde la habilidad media fue de 0.62 y la correlación biserial puntual (RBP) media de 0.27. Como se muestra en la Figura 12. Esta figura presenta en el eje de las abscisas el valor P de Facilidad y en el de las ordenadas la correlación biserial puntual. Se observó que los reactivos que alcanzaron la zona objetivo, esto es con discriminación deseable y con facilidad media, fueron solamente el I5, I18, I18, I19 e I22. Como ejemplo la Figura 13, presenta el comportamiento del ítem 16 el cual se encuentra dentro de la zona objetivo, para este pilotaje, ambos distractores funcionaron. Sin embargo, ítems como el dos no solo uno de los distractores no funcionó, sino que su nivel de habilidad fue ligeramente más alto que la media, lo cual se refleja en la Figura 14, donde más de la mitad escogió el primer distractor. Para información de los otros ítems consultar Anexo 15.

Figura 12
Correlación Biserial Puntual y Valor P de Facilidad

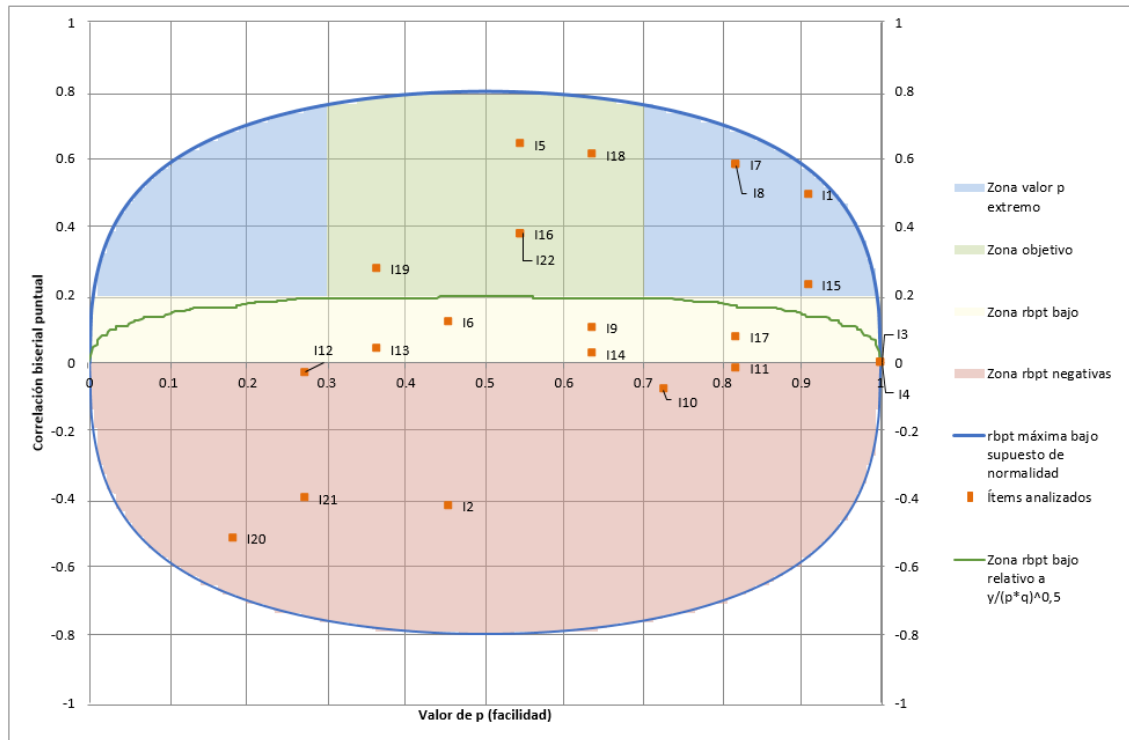


Figura 13
Porcentaje de elección de cada alternativa para ítem 16

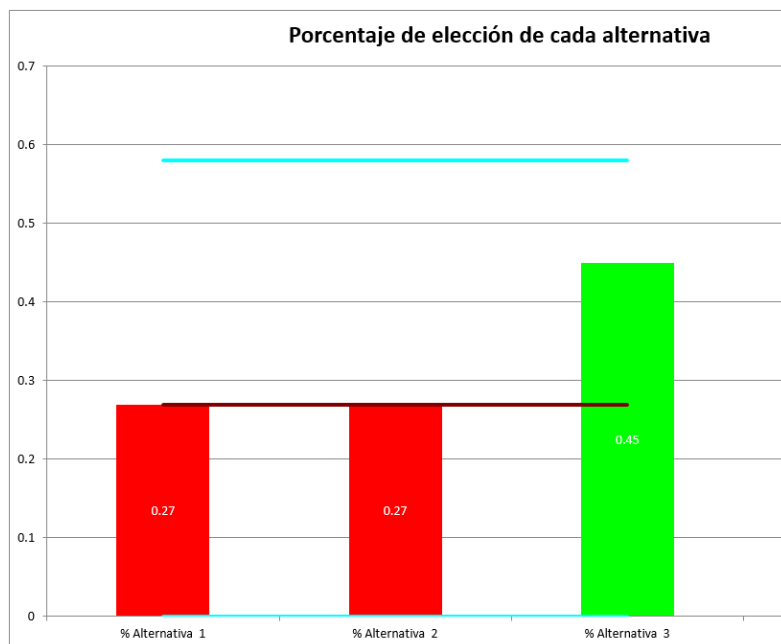
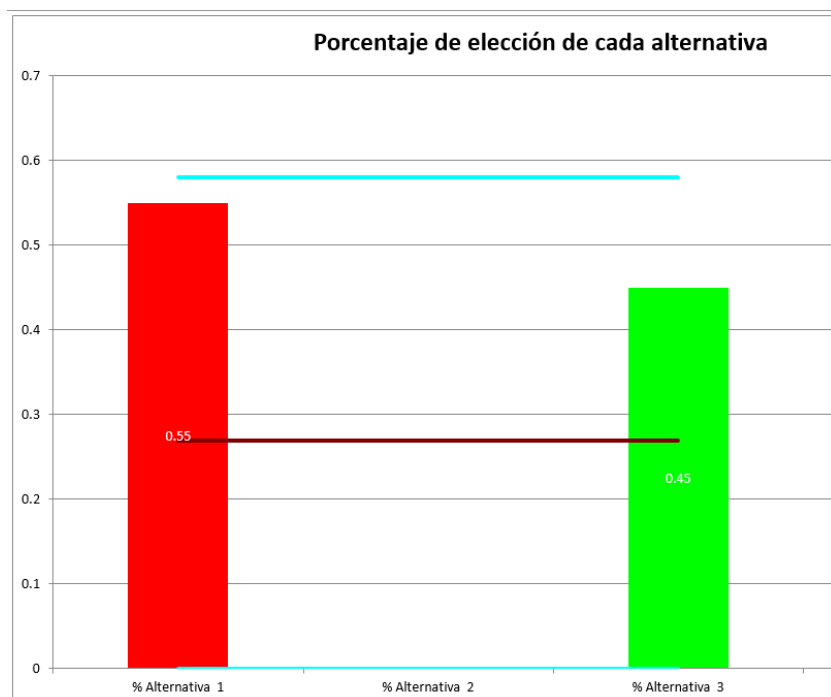


Figura 14
Porcentaje de elección de cada alternativa para ítem dos.



Por otra parte, se encontraron ítems de lectura que evaluaban criterios similares a los criterios de escritura, lo cual se esperaba por la estrecha relación lectura-escritura. En consecuencia, ítems de lectura presentaban similitud con los criterios evaluados en las rúbricas de escritura, como se muestra en la Tabla 23. Esto resulta útil sobre todo en ítems que estuvieron fuera de la zona objetivo (Figura 12), como el ítem 4, 6, 7, 8, 9, 11, 12, 13, 14 y 20, ya que ayuda a explorar o evaluar ese indicador desde una rúbrica

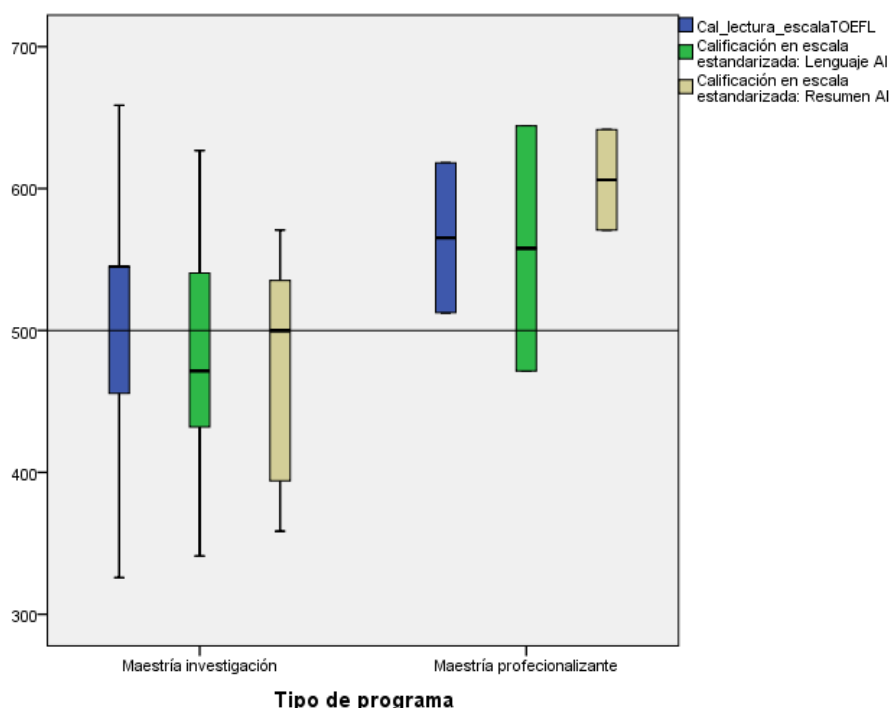
Tabla 23
Similitud entre criterios evaluados en la sección de escritura e ítems de lectura.

<i>Criterio</i>	<i>Ítems</i>
<i>Main idea</i>	4, 7, 8, 13
<i>Details, evidence, support</i>	7, 13
<i>Reasoning/ Thinking</i>	14, 18, 19, 20, 22
<i>Textual evidence</i>	6
<i>Sequencing</i>	5
<i>Grammar</i>	12
<i>Connecting sentences</i>	9, 16

Comportamiento de examinados ante la prueba.

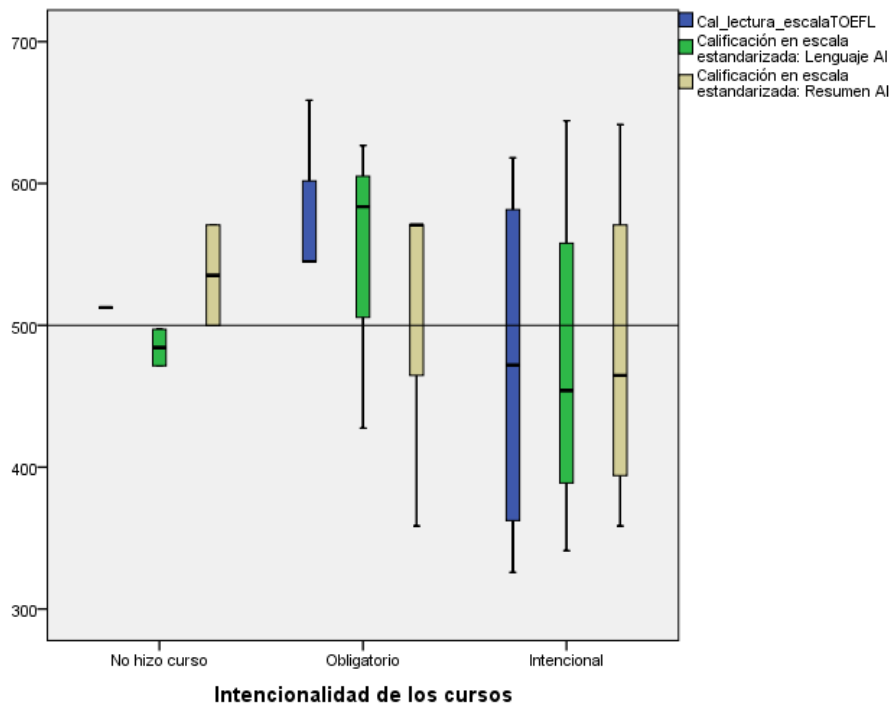
Como parte del análisis de las puntuaciones obtenidas por los examinados, se encontró que en promedio los examinados del programa de maestría profesionalizante obtuvieron mejor puntuación en lectura que los de investigación como se muestra en la Figura 15. Igualmente, se aprecia que las puntuaciones obtenidas por la inteligencia artificial en el reactivo de escritura muestran una media más alta. Para este análisis no se consideran las evaluaciones realizadas por los jueces ya que no existe evaluación para el pilotaje en el programa profesionalizante.

Figura 15
Distribución de puntajes por tipo de programa



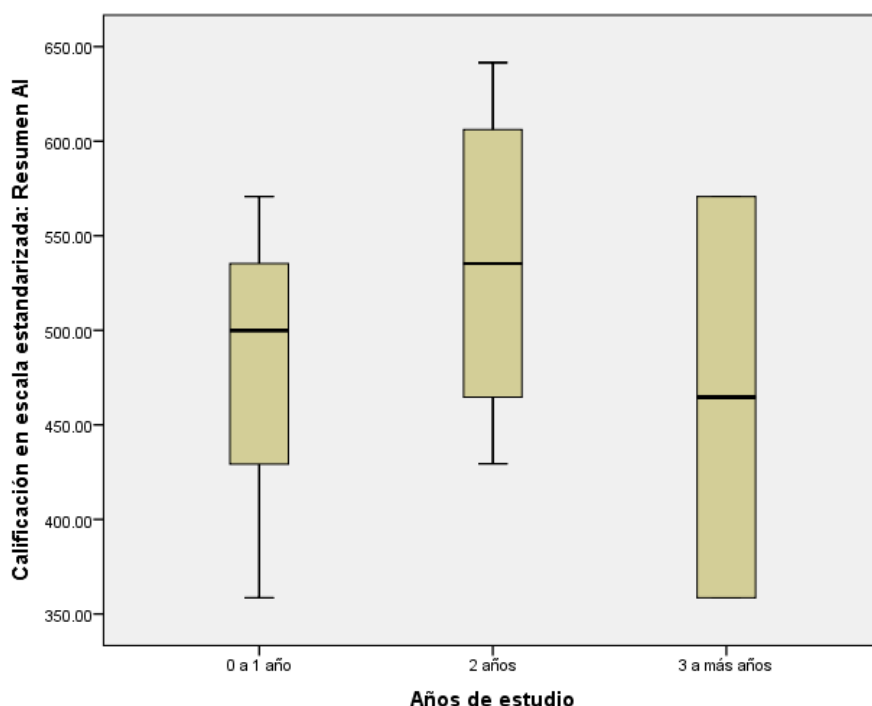
Por otra parte, el cuestionario de contexto permitió explorar si los examinados tomaron cursos de inglés con anterioridad y cuál fue la intencionalidad de estos cursos, esto es, si fueron obligatorios como parte de los programas educativos o si fueron de manera intencional o extracurriculares. En cuanto a los que solamente tomaron los cursos obligatorios, mostraron una media más alta que los que no hicieron cursos, además, aquellos que tomaron cursos extracurriculares tendieron a medias de calificaciones más bajas que incluso los que tomaron solamente cursos obligatorios, como se muestra en la Figura 16.

Figura 16
Distribución por tipo de cursos tomados



En cuanto a los años de estudios del idioma, este estudio refleja que los que realizaron tres años o más de estudio presentaron medias más bajas que los demás, como se muestra en la Figura 17. Por otra parte, los que estudiaron dos años, se encuentra que cuentan con mayores extremos en las puntuaciones obtenidas tanto en la sección de lectura como de escritura.

Figura 17
Distribución por años de estudio

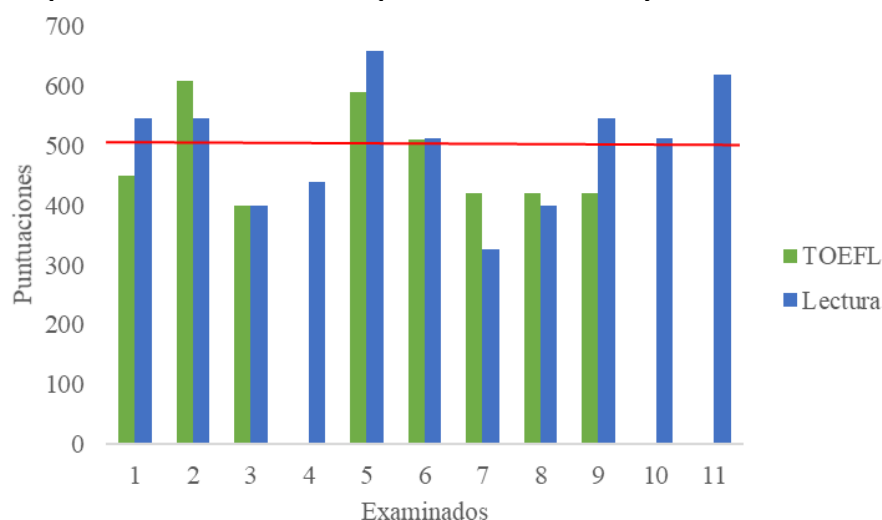


En cuanto a la edad de los examinados, se encontró que aquellos con una edad entre 20 y 29 obtuvieron en general mejores puntuaciones en la prueba. De manera contraria, aquellos que repostaron una edad de 40 o más obtuvieron calificaciones muy bajas, a excepción de la sección de lectura.

Adicionalmente, en el cuestionario de contexto se indagó sobre las puntuaciones obtenidas en el examen TOEFL, si es que se había presentado para ingresar al posgrado. Esta información fue de relevancia para comparar esa puntuación con la obtenida en la prueba. Tal es el caso de la Figura 18 que presenta una gráfica de barras que compara los resultados que reportaron de la prueba TOEFL y la puntuación obtenida en la sección de lectura. En este caso no todos los examinados reportaron haber presentado la prueba, sin embargo, los examinados 1, 5 y 9, obtuvieron puntajes mayores en la sección de lectura, sin embargo, examinados como el 2, 7 y 8 reportaron un puntaje mayor en TOEFL al que obtuvieron en la prueba en la sección de lectura.

Figura 18

Comparación de resultados reportados en TOEFL y en la sección de lectura en la prueba



Uso de la Inteligencia Artificial para la evaluación de escritos

En cuanto a la sección de escritura, se realizó una comparación entre las evaluaciones realizadas por los 3 jueces y la realizadas por la inteligencia artificial. Las comparaciones se realizaron de acuerdo con las dos áreas evaluadas: convenciones y usos del lenguaje, y resumen.

Para la rúbrica de lenguaje (Figura 19), se encontró que las evaluaciones realizadas por la IA se asemejaban más al juez más severo o intermedio. Por otra parte, se encontró que en los examinados con mejor puntuación hubo más al consenso entre expertos y la inteligencia artificial. El caso contrario ocurrió con las puntuaciones en la rúbrica de resumen (Figura 20).

Figura 19

Comparación de juicio de expertos e Inteligencia Artificial en rúbrica de lenguaje

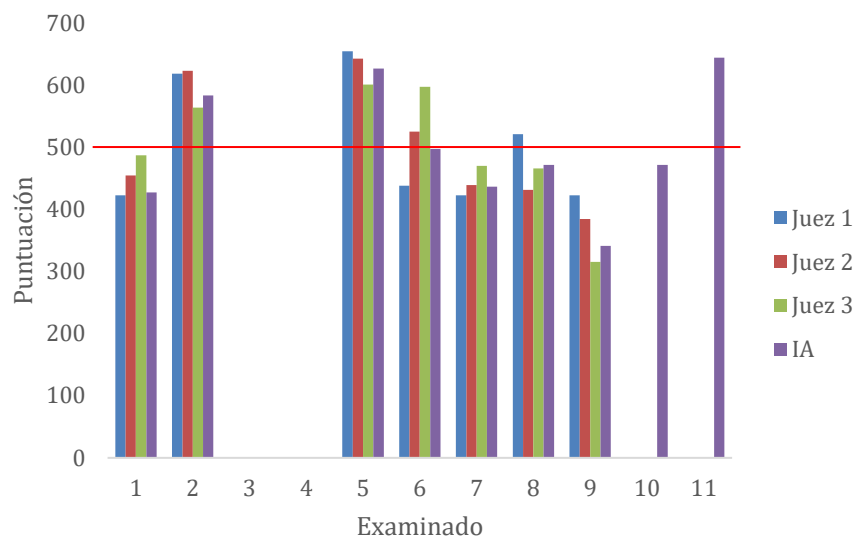
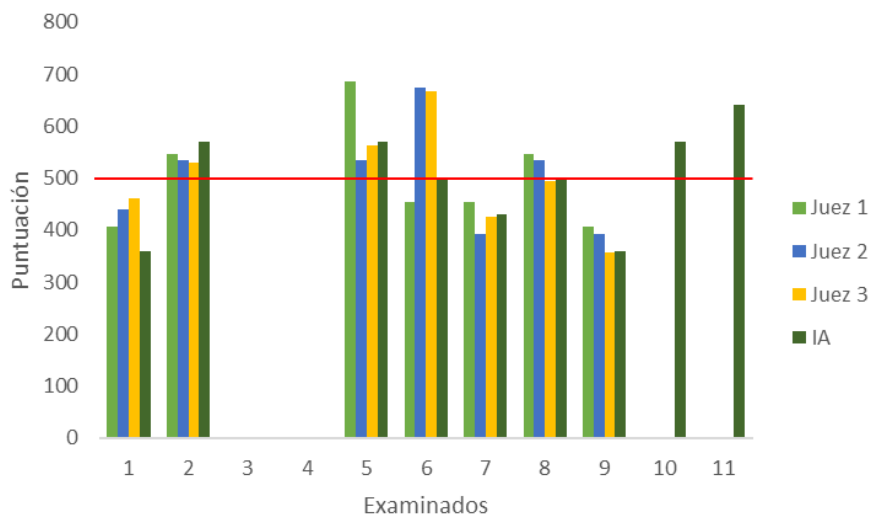


Figura 20

Comparación de juicio de expertos e Inteligencia Artificial en rúbrica de resumen



Las evaluaciones realizadas con la rúbrica de lenguaje (Figura 21), se observa que las evaluaciones realizadas por la inteligencia artificial suelen ser entre severas a medias comparadas con los juicios realizados por el humano y la puntuación en TOEFL. Para el caso del aspecto de resumen (Figura 22), la inteligencia artificial se asemejó más a las puntuaciones obtenidas en TOEFL en cuatro casos, y al humano en tres casos. Además, se encontró que las puntuaciones en la mayoría de los casos se asemejaba una evaluación severa o intermedia.

Figura 21

Comparación de resultados en sección de escritura de aspecto de lenguaje evaluado por humanos y por la inteligencia artificial

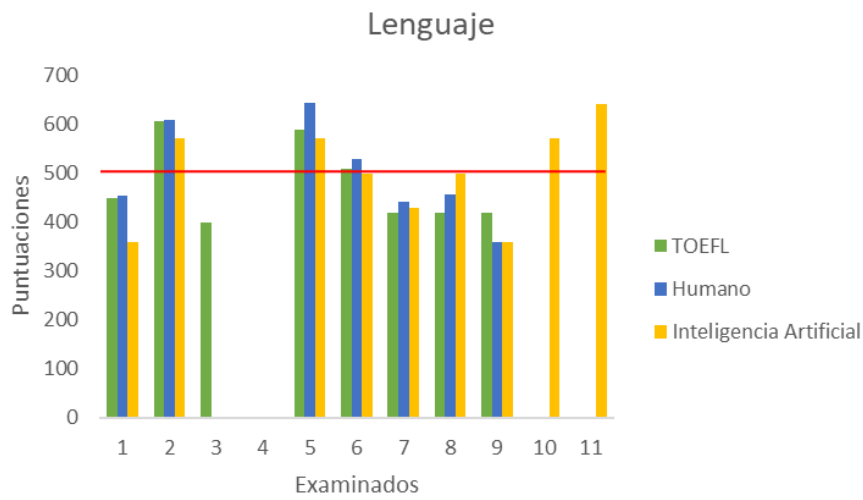
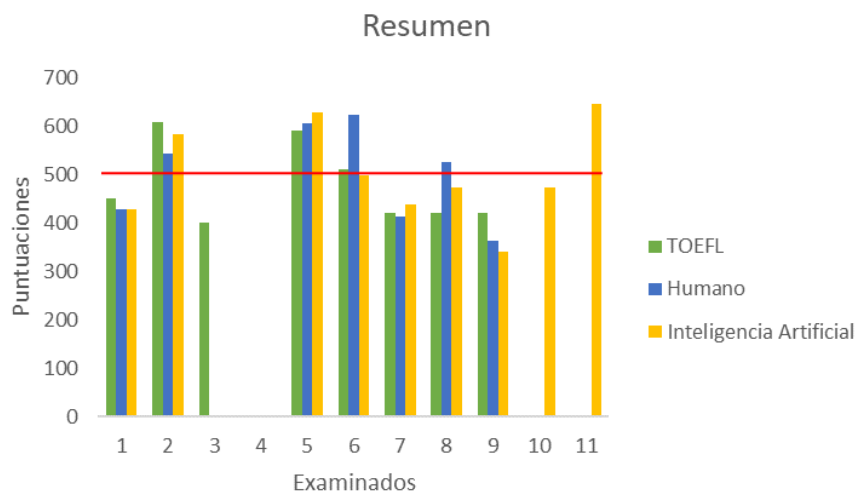


Figura 22

Comparación de resultados en sección de escritura de aspecto de resumen evaluado por humanos y por la inteligencia artificial



Capítulo V. Discusión y conclusiones

El presente capítulo es resultado de un análisis crítico y reflexivo de los principales hallazgos obtenidos a lo largo de este estudio, con el objetivo de vincularlos a las teorías, marcos conceptuales y antecedentes revisados en el campo del inglés con propósitos académicos (EAP), y específicamente en la evaluación del dominio del idioma con fines de publicación científica. Este capítulo se estructura en tres secciones principales: la discusión de resultados en relación con los objetivos planteados, las conclusiones generales que sintetizan el impacto del estudio, y las recomendaciones junto a futuras líneas de investigación derivadas del trabajo.

El subapartado de discusión analiza los resultados en función de los dos objetivos específicos. Primero, se reflexiona sobre la operacionalización de los niveles de dominio del inglés académico, destacando las competencias identificadas y su relación con habilidades clave como la lectura y la escritura. En este sentido, se confrontan los hallazgos con estudios previos y se evidencian las contribuciones específicas de este trabajo en la identificación y clasificación de habilidades, la integración de categorías relacionadas con la investigación, y las implicaciones pedagógicas para la evaluación y enseñanza del EAP. En segundo lugar, se examina el desarrollo de la prueba diagnóstica, enfatizando los niveles de complejidad taxonómica abordados, la validez de los indicadores, y la incorporación de herramientas tecnológicas como la inteligencia artificial en la evaluación de tareas de escritura. Este análisis incluye comparaciones con investigaciones previas que emplean metodologías similares y subraya las innovaciones introducidas en este trabajo. Este capítulo busca, en esencia, integrar los hallazgos de la investigación en una discusión más amplia, generando un marco conceptual y metodológico para avanzar en la evaluación del EAP que responda a las demandas crecientes de la comunicación y comunidad científica en formación.

Operacionalizar los niveles de dominio del EAP

Atendiendo al primer objetivo, que es operacionalizar los niveles de dominio EAP con enfoque en la publicación científica, se identificaron las principales habilidades a evaluar conforme a los propósitos de la prueba, estos son la habilidad de lectura y escritura. Además, de coincidir con las ideas de Prihandoko (2024) quien argumenta la relevancia de estas habilidades tanto en entornos académicos como de investigación, inclusive estableciendo relaciones entre las habilidades de lectura y escritura. Similarmente, Plakans et al., (2018) y O'Grady y

Taşkesen (2022) argumentan la importancia y la estrecha interacción de ambas habilidades tanto en la enseñanza como en la evaluación, siendo una la evaluación indirecta de la otra. Por otra parte, estudios de iniciativas y propuestas pedagógicas presentan distintos conjuntos de competencias para la lectura y la escritura por separado y en conjunto en ámbitos académico-científicos (Prapinwong y Koolsriroj, 2019; Nakamura, 2022; Song y Zhou, 2022; Grabe y Zhang, 2013; Qayyum, 2024; Zainal y Jamil, 2015; Sata y Karakaya, 2021). Estas competencias coinciden con otras pruebas analizadas donde las principales secciones o habilidades evaluadas eran lectura/comprensión lectora, escritura, vocabulario y gramática. La selección concuerda con lo encontrado por Hyland y Jiang (2021a) en su estudio bibliométrico donde analizan las principales temáticas de investigación e interés en el EAP. Dentro de este estudio se comparan dos periodos, 1980-2000 y 2001 - 2020, y las principales temáticas de investigación fueron la comprensión lectora y la redacción académica, las cuales aumentaron de manera significativa de un periodo a otro. Entre las competencias propuestas se lograron clasificar cuatro grandes categorías que ayudaron en la operacionalización de las habilidades principales en el EAP.

El desarrollo de la prueba con fines de publicación científica permitió visibilizar las competencias asociadas a la lectura y escritura de textos académicos, particularmente los de publicación científica. De acuerdo con los resultados se identificó a través del cribado que los expertos consideraron principalmente habilidades asociadas con la lectura y escritura académica, como son los aspectos de organización, redacción de ensayos, resúmenes, paráfrasis, citación, interpretación de gráficas, entre otros. Al respecto, esto coincide con los autores Li y Flowerdew (2020) que enfatizan en su revisión de iniciativas pedagógicas la importancia de las competencias lingüísticas asociadas con la escritura y lectura académica como foco principal. Esto es, con atención a la capacidad de redactar y leer textos científicos, y los aspectos como la lengua y la investigación de manera indirecta.

En cuanto a las habilidades de investigación, se asociaron con ocho indicadores. Esta habilidad se encontró que está asociada con el análisis crítico, la amplitud y profundidad de conocimientos especializados, análisis inductivo y deductivo, entre otros. Estos hallazgos van de la mano con la categoría antes mencionada, ya que como menciona Huang (2014), es de los principales problemas a los que se enfrentan los interesados en publicar que no son nativos angloparlantes. Este trabajo permitió identificar las principales categorías de habilidades que además funcionen para otras áreas y la operacionalización de conocimientos.

Desarrollar una prueba diagnóstica de inglés con propósitos de publicación científica.

Además, a partir de estas habilidades y categorías, se operacionalizaron los indicadores por medio de un cribado realizado por un grupo de expertos. De estas habilidades se rescataron 22 indicadores de lectura y 26 de escritura. Los indicadores de lectura se transformaron en focos evaluativos para el diseño de las especificaciones de los ítems. En cuanto a los indicadores de escritura se les asignaron criterios conforme a rúbricas según lo recomendado por expertos. Las rúbricas se adaptaron de otros autores, una de aspecto de convenciones y usos de la lengua y otro de formato de resumen científico. Estos indicadores se encontraron en el espectro de dificultad de intermedio – avanzado, esto es, B1 a C2, de acuerdo con el MCER. En otras palabras, más de la mitad de los indicadores se encontraban en el nivel taxonómico de aplicar o superior de acuerdo con la taxonomía de Anderson y Krathwohl. En concordancia, diversos estudios analizados en los antecedentes y en el marco teórico afirman la complejidad de cursos/ pruebas dirigidas a contextos académicos. Estos niveles de complejidad taxonómico se relacionan estrechamente con las habilidades más representadas en la prueba, ya que habilidades como el análisis, la síntesis, la argumentación, pensamiento inductivo y deductivo, conforman procesos cognitivos complejos. Además, de estar asociados al pensamiento crítico, creativo, analítico, lógico, reflexivo, reflexivo y metacognitivo, pertenecientes a las Higher Order Thinking Skills (Arif et al., 2020; Riefani y Utami, 2017). Así de esta manera, se lograron identificar los indicadores y los requerimientos a nivel taxonómico para los interesados en el inglés académico y científico.

Finalmente, se logró la implementación de Inteligencia Artificial como propuesta de retroalimentación en tareas de escritura. En concordancia con lo propuesto por Farzi y Dietz (2024), se exploró el uso de las rúbricas como recurso en el uso de las IA, esto con el fin de integrar las fortalezas del juicio humano y máquina. En el caso de este trabajo, se realizó la comparación de ambos juicios. Los resultados arrojaron que para examinados que reportaron puntuaciones TOEFL altas, obtuvieron puntuaciones similares en la prueba. Sin embargo, se encontraron inconsistencias en las de aquellos que reportaron puntuaciones bajas. Para ambos casos, se encontró que el juicio emitido por la inteligencia artificial fue de severo a medio en comparación al emitido por los expertos. Las evaluaciones realizadas por la Inteligencia Artificial demostraron en este estudio ser más severas en sus puntuaciones en comparación a la evaluación humana. Igualmente, Taskiran y Goksel (2022) utilizan la Inteligencia Artificial como propuesta de evaluaciones y retroalimentación, presentando este recurso como útil en el feedback de tareas de escritura, esto es en la detección y mejora de aspectos como de

puntuación, gramática, deletreo, estilo entre otros beneficios. Igualmente, Wang y Wang (2012) proponen en su estudio del feedback automático por medio de computadora, el cual muestra los beneficios en el uso de este recurso, sin embargo, futura exploración en temas de evaluación es recomendada. La evidencia obtenida de dicha investigación se puede utilizar para construir un argumento de evaluación empírica para el uso de esta tecnología (Cotos, 2011).

Conclusiones

A manera de conclusión, con los resultados obtenidos, se atendieron a los objetivos específicos y finalmente al objetivo general. De esta manera, este proyecto permitió el desarrollo de una prueba diagnóstica del dominio del idioma inglés con propósitos de publicación científica en estudiantes de posgrado, así como la obtención de evidencias de validez de contenido.

Este estudio aporta una prueba que atiende a las necesidades de los estudiantes de posgrado y autoridades universitarias que están interesados en el inglés académico-investigativo. Al ser una prueba diagnóstica permite que los estudiantes identifiquen sus áreas de oportunidad para mejorar su escritura científica, y a las autoridades, crear cursos, programas o estrategias de enseñanza que mejoren las habilidades de escritura científica de sus estudiantes de posgrado. Además, abre la puerta a futuras líneas de investigación y nuevos desarrollos educativos.

En suma, el estudio ofrece a nivel social una prueba que puede incidir en la difusión y transmisión del conocimiento que generan los estudiantes de posgrado, lo cual se puede ver reflejado en el volumen de artículos publicados. Por otra parte, su aportación práctica está en la operacionalización de los niveles de dominio, que se encuentra tanto en la comparación de escalas, clasificación de habilidades por competencia y el cribado de indicadores, así como la adaptación de rúbricas de evaluación de la escritura.

Adicionalmente, este trabajo propone una metodología para el diseño de pruebas de inglés, mediante un constante jueceo y cribado, esto con el fin de evitar sesgo en el contenido de la prueba. Esta aportación está principalmente dirigida a pruebas localizadas, y de inglés especializado en distintos propósitos, con el fin de responder a las necesidades particulares de cada contexto y necesidades de los alumnos.

Finalmente, se presenta la propuesta del uso de la inteligencia artificial para la evaluación de textos. Lo anterior por medio del uso de instrucciones precisas y la descripción de los criterios a evaluar. Se busca que este proyecto permita futuras líneas de investigación y

explorar las aproximaciones que tienen las evaluaciones de la inteligencia artificial a los juicios emitidos por el humano. De esta manera, reducir el recurso humano y de tiempo necesario para evaluar escritos en pruebas de tipo diagnóstico. Por otra parte, se pueda explorar sus usos en pruebas con otras funciones e incluso se incorpore como herramienta en intervenciones didácticas para la mejora de escritura.

Recomendaciones

Derivado de este proyecto, se proponen las siguientes recomendaciones:

- El desarrollo de pruebas tiene retos en las distintas etapas que comprende el proceso. En los casos en que el desarrollo presente dificultades al momento cribado, jueceo y validación, por no contar con recursos financieros para cubrir los honorarios de algunos de ellos, el uso de la Inteligencia Artificial puede ser una alternativa para los procesos de validación. Su uso es un área de investigación en sí misma, por lo que habrá que seguir explorando y evaluando. La diferencia entre el juicio humano y los juicios que emiten las aplicaciones de IA son un tema de debate, por lo tanto, se tienen que tomar con sus debidas reservas.
- Se invita a las autoridades de la Secretaría de Educación Pública para que diseñen y desarrollen iniciativas y políticas respecto a la enseñanza y evaluación de inglés en educación superior.
- Además, se recomienda la participación de diversos especialistas que ayuden en el proceso de cribado de indicadores, creación de especificaciones, validación y evaluación de la prueba. Estos expertos deben contar con un perfil que esté alineado a las necesidades de la prueba, esto es, expertos en el idioma, los propósitos específicos de la lengua (como en este caso, la publicación de artículos científicos) y psicometría. Finalmente, en la medida de lo posible, que los comités estén conformados por al menos tres expertos, esto con el fin de asegurar el desempate de los juicios y obtener evidencias de validez de contenido.
- Se invita a las instituciones de educación superior, que desarrollen pruebas de este tipo con el fin de que aporten insumos para la creación de propuestas pedagógicas que eleven el nivel de dominio del inglés académico en licenciatura y esto se vea reflejado en aspirantes a posgrado. De manera particular, los

resultados de las pruebas pueden servir como un diagnóstico para el diseño y aplicación de propuestas pedagógicas que contribuyan a mejorar el nivel de dominio del inglés para publicaciones científicas en posgrado, esto con la finalidad de aumentar la participación en revistas de alto impacto y la divulgación del conocimiento generado.

- En cuanto al uso de las GenAI, se aconseja documentar todos los procesos en los que contribuyó y reportar la versión, fechas en las que realizaron las consultas, tipo de consulta o *prompt* utilizado. En cuanto a este último punto, se recomienda realizar las consultas en un mismo día, con el fin de reducir la variación en las respuestas o productos generados por la inteligencia.

Las recomendaciones anteriores se hacen en consideración a las fortalezas y debilidades identificadas a lo largo de este trabajo. Como parte de las debilidades, se identifica cuidar el tamaño de la muestra, con el fin de presentar evidencias de validez de constructo y análisis métricos más robustos. Dentro de las fortalezas de este trabajo está la parte de diseño y desarrollo de la prueba. Esto incluye la metodología empleada, ya que se hicieron adaptaciones de diferentes metodologías para el desarrollo de pruebas alineadas a un criterio que permitieran, obtener una prueba acorde a los contextos y necesidades de los estudiantes de posgrados mexicanos. Además, los productos de la operacionalización de las categorías, indicadores y niveles de dominio de la lengua.

Futuras líneas de investigación.

Como parte del ejercicio de este trabajo, se encontraron diversas líneas de investigación que pueden dar continuidad a esta investigación. Primeramente, trabajos especializados en enriquecer el campo del inglés con propósitos de publicación científica. Además, de ampliar la muestra con el fin de abonar a los análisis realizados y extender la población objetivo, no solo a estudiantes de posgrado, sino a investigadores y estudiantes de licenciatura. Por otra parte, el diseño e implementación de diseños instruccionales y material didáctico con el fin de capacitar tanto estudiantes como investigadores. Por último, investigaciones relacionadas al uso de inteligencia artificial como evaluadores de textos, reduciendo recursos humanos y temporales en la evaluación

Referencias

- Aiken, L. R. (1985). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955-959.
- Altay, M., Curle, S., Dogan Yuksel, D. y Soruç, A. (2022). Investigating academic achievement of English medium instruction courses in Turkey. *Studies in Second Language Learning and Teaching*, 12(1), 117-141 <http://dx.doi.org/10.14746/ssllt.2022.12.1.6>.
- American Educational Research Association, American Psychological Association, Y National Council on Measurement in Education [AERA, APA, NCME]. (2018). Estándares para pruebas psicológicas y educativas. Washington, DC: American Educational Research Association.
- Anderson, L.W. y D. Krathwohl (Eds.) (2001). A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives. Longman.
- Arif, S., Basyaruddin y Yuhdi, A. (2020). Integration of high order thinking skills in research method subject in university. *Britain International of Linguistics, Arts and Education (BIoLAE) Journal*, 2(1), 378-383. doi: 10.33258/BIOLAE.V2I1.207
- Arriaga, M. (2015). El diagnóstico educativo, una importante herramienta para elevar la calidad de la educación en manos de los docentes. *Atenas*, 3(31), 63-74.
- Asociación Nacional de Universidades e Instituciones de Educación Superior [ANUIES]. (2016). Plan de desarrollo institucional. Visión 2030. http://www.anuies.mx/media/docs/avisos/pdf/PlanDesarrolloVision2030_v2.pdf
- Baird C. y Dooley P. (2017). Using images to facilitate writing for skills assessment: A Visual PELA. *Australian Journal of Indigenous Education*, 46(2), 1-13. Doi: 10.1017/jie.2016.32
- Banaeian H. y Gilanlioglu I. (2021). Influence of the NAO robot as a teaching assistant on university students' vocabulary learning and attitudes. *Australasian Journal of*

- Educational Technology*, 37(3), 71-87.
- Barrios, E. y López-Gutierrez, A. (2021). Language learning in a partially English-taught teacher education programme: language gains and student perceptions. *International Journal of Multilingualism*. DOI: 10.1080/14790718.2021.188895
- Bayacán, C. (2013). Operacionalización de la taxonomía de Anderson y Krathwohl para la docencia universitaria. *Paideia XXI*, 3(4), 109-119.
- Bell, D.E. (2022). Methodology in EAP: Why is it largely still an overlooked issue? *Journal of English for Academic Purposes*, 55, 1-11. <https://doi.org/10.1016/j.jeap.2021.101073>
- Benemérita Universidad Autónoma de Puebla [BUAP]. (s.f.). *Módulo II. Taxonomías del Aprendizaje*. https://repmdl.buap.mx/repositorio/Diplomado-Evaluacion-Aprendizajes_2020/Septiembre/html/modulo02/recursos/modulo02-tipos-taxonomias.pdf
- Blanco, L. A. (1996). La evaluación referida al criterio y la toma de decisiones en un centro educativo. *Educación*, 5(9), 5-20.
- Božović P., Piletić D. y Lakić I. (2022). Language for Specific Purposes in academia vs. Labor Market Needs: The Trends and Perspectives in Montenegro [NASTAVA STRANOG JEZIKA STRUKE NA UNIVERZITETU I POTREBE TRŽIŠTA RADA: TRENDovi I PERSPEKTIVE U CRNOJ GORI]. *Folia Linguistica et Litteraria*, 13(42), 31-43.
- Brosch, C. (2015). On the conceptual history of the term Lingua Franca. *Apples, Journal of applied Languages Studies*, 9(1), 71-85.
- Buisán, C. y Marín, M. P. (2001). ¿Cómo realizar un diagnóstico pedagógico? Alfa Omega. México.
- Canale, G. (2011). Planificación y políticas lingüísticas en la enseñanza de lenguas extranjeras: el acceso al inglés en la educación pública uruguaya. *Revista Digital de Políticas Lingüísticas*, (3), 49-76.

- Cardwell, R., Naismith, B., Burstein, J., Nydick, S., Goodwin, S. y Verardi, A. (2024). From pen to pixel: Rethinking english language proficiency admissions assessments in the digital era. *Calico Journal*, 41(2), 209-234. <https://doi.org/10.1558/cj.27104>
- Cargill, M. y Burgess, S. (2008). Introduction to the special issue: English for Research Publication Purposes. *Journal of English for Academic Purposes*, 7, 75-76.
doi:10.1016/j.jeap.2008.02.006
- Carrillo-Beltrán, J., Abad-Matos, N., Ramírez-Jiménez, A., Llamas-Rodríguez, V., Llanos-Ramírez, C., Gómez-Campos, S. y Salgado-Moreno, S. (2021) *Abanico Agroforestal*, (3), 1-17. <http://dx.doi.org/10.37114/abaagrof/2021.9>
- Castillo-Martínez, I. M. y Ramírez-Montoya, M. S. (2021). Research competencies to develop academic reading and writing: A Systematic Literature Review. *Frontiers in Education*, 5. <https://doi.org/10.3389/feduc.2020.576961>
- Çelik, S. y Baran, E. (2021). Student response system: its impact on EFL students' vocabulary achievement. *Technology, Pedagogy and Education*, 1-36.
- Cephe, P. T., y Toprak, T. E. (2014). The Common European Framework of Reference for Languages: Insights for language testing. *Journal of Language and Linguistic Studies*, 10(1), 79-88.
- Chan, H. (2023). Grading generative AI-based assignments using a 3R framework. 1-5. doi: 10.1109/tale56641.2023.10398408
- Chang, C. (2023). Enhancing EFL university students' readiness for learning academic content in english: the effectiveness of combining MOOCs with data-driven learning activities in an english language classroom. *English Teaching & Learning*.
<https://doi.org/10.1007/s42321-023-00139-8>
- Charge, N. y Taylor, L. (1997). Recent developments in IELTS. *ELT Journal*, 51(4), 374-380.
- Cizek, G. (2020). *Validity: An Integrated Approach to Test Score Meaning and Use*. Routledge.

- Consejo de Europa. (2020). Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. Volumen complementario. Servicio de publicaciones del Consejo de Europa: Estrasburgo. www.coe.int/lang-cefr.
- Consejo Mexicano de Estudios de Posgrado (COMPEPO). (2015). Diagnóstico del Posgrado en México: NACIONAL. <https://comepo.org.mx/wp-content/uploads/2023/06/Diagnostico-Posgrado-Me%CC%81xico-COMEPO-Nacional.pdf>
- Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT). (s.f.). *Capacitación para la Escritura de Manuscritos Académicos en Inglés (CEMAI)*. <https://conahcyt.mx/acceso-universal-al-conocimiento/cemai/>
- Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT). (s.f.). Programa Institucional CONACYT 2020-2024. Ciudad de México: Autor.
- Corcoran, J. y Englander, K. (2016). A proposal for critical-pragmatic pedagogical approaches to English for Research Publication Purposes. *Publications*, 4(6). doi:10.3390
- Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal*, 28(2), 420-459. 10.11139/cj.28.2.420-459
- Council of Europe. (2001). Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge: Cambridge University Press.
- Creswell, J.W. y Creswell, J.D. (2018) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage, Los Angeles.
- De Jong, H. A. L. y Benigno, V. (2017). Alignment of the Global Scale of English to other scales: the concordance between PTE Academic, IELTS, and TOEFL. Pearson
- Devos A.O., Torbenko I.O., Doroshenko T.V., Revenko V.V. y Shuhaiev A.V. (2021). The application of the simulation method in the in foreign language teaching in higher education institutions, the cognitive linguistic approach. *Journal of Educational and*

Social Research, 11(4). <https://doi.org/10.36941/jesr-2021-0072>

Diario Oficial de la Federación. (2011). ACUERDO número 592 por el que se establece la Articulación de la Educación Básica.

<https://www.sep.gob.mx/work/models/sep1/Resource/9721849d-666e-48b7-8433-0eec1247f1ab/a592.pdf>

Diario Oficial de la Federación. (2015). ACUERDO número 20/12/15 por el que se emiten las Reglas de Operación del Programa Nacional de Inglés para el ejercicio fiscal 2016.

https://dof.gob.mx/nota_detalle.php?codigo=5421425&fecha=27/12/2015#gsc.tab=0

Díaz-Castelazo, C. (2018). La importancia del idioma inglés para el desarrollo y la enseñanza de las ciencias. *Eduscientia*, 1(2), 60-68.

Diccionario del Español de México [DEM]. (2024). Lengua. En *Diccionario de Español de México*. Recuperado en 10 de junio de 2024, de <https://dem.colmex.mx/ver/lengua>

Eberhard, D. M., Simons, G. F. y Fennig C. D. (2022). *Ethnologue: Languages of the World*. Twenty-fifth edition. Dallas, Texas: SIL International.

<https://www.ethnologue.com/guides/ethnologue200>

Education Northwest. (Marzo 2021). *6+1 Trait Writing Rubrics: Grades 3–12* [Archivo PDF]. Resources Education Northwest.

<https://educationnorthwest.org/sites/default/files/resources/traits-rubrics-3-12.pdf>

EF English Proficiency Index. (2021). Índice EF de nivel de inglés Una clasificación de 112 países y regiones en función de su nivel de inglés. Autor.

Encinas, F., Sánchez, V. y Puon, Y., (2020). La escritura en Inglés como lengua extranjera: Una revisión de la visión de la literatura e implicaciones pedagógicas. *Leer, Escribir y Descubrir*, 1(6), 45-58.

Farnill D. y Hayes S.C. (1996). Screening higher-education students for english language problems: Development of the australian tertiary-english screening test. *International*

Journal of Phytoremediation, 21(1), 61-71.

<http://dx.doi.org/10.1080/0729436960150105>

- Farzi, N. y Dietz, L. (2024). Pencils Down! Automatic Rubric-based Evaluation of Retrieve/Generate Systems. ICTIR '24. <https://doi.org/10.1145/3664190.3672511>
- Ferguson, G. (2007). The global spread of English, scientific communication and ESP: questions of equity, access and domain loss. *Iberica*, (13), 7-38.
- Fierro López, L. (2015). *Desarrollo profesional docente y contexto de implementación del programa de inglés en primaria pública en la ciudad de Ensenada, Baja California. Perspectiva de los profesores con formación para la enseñanza del inglés y con formación en otros campos*. [Tesis de doctorado no publicada] Universidad Autónoma de Baja California.
http://iide.ens.uabc.mx/documentos/divulgacion/tesis/DCE/2010/Laura_Emilια_Fierro_Lopez.pdf
- Figueras, N., North, B., Takala, S., Verhelst, N., y Van Avermaet, P. (2005). Relating examinations to the Common European Framework: A manual. *Language Testing*, 22, 262-279.
- Finch, W. y French, B. (2019). *Educational and Psychological Measurement*. Routledge
- Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 1(4), 253-26.
- Gaona, E., Sierra, J. y Gonzalez, D. (2017). Economía del conocimiento: El caso de México en comparación con seis países. *CIMEXUS*, 12(2), 65-82.
- Garrido, J. (2010). Lengua y globalización: inglés global y español pluricéntrico. *Historia y Comunicación*, (15), 63-95.
- Gaviria, J. L. (31 de enero de 2018). Corrector 2.1 “Complemento de excel para corregir exámenes de opción múltiple y analizar sus Items. Departamento de Investigación y

Psicología en Educación. Facultad de Educación.

González Ramírez, C. (2015). English for specific purposes: brief history and definitions.

Revista de lenguas modernas, (23), 379-386.

Grabe, W. y Zhang, C. (2013). Reading and writing together: A critical component of english for academic purposes teaching and learning. *TESOL Journal*, 4(1), 9-24.

<https://doi.org/10.1002/tesj.65>

Gu, L., y So, Y. (2015), Voices from stakeholders: What makes an academic English test ‘international’?. *Journal of English for Academic Purposes*, 18, 9-24.

<http://dx.doi.org/10.1016/j.jeap.2014.10.002>

Habibie, P. y Starfield, S. (2023). AI-mediated English for research publication purposes. Are we there yet?. *Journal of English for Research Publication Purposes*, 4(1), 1-4.

<https://doi.org/10.1075/jerpp.00013.hab>

Haladyna, T. M., Downing, S. M., y Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15, 309–334.

Haro Sly, M.J. y Liaudat, S. (2021). ¿Qué podemos aprender de China en política científica y tecnológica?. *Ciencia Tecnología y Política*, 4(6).

<https://doi.org/10.24215/26183188e052>

Hashemi,A. y Daneshfar, S. (2018). A review of the IELTS test: Focus on validity, reliability, and washback. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 3(1), 39-52.

Henshall, A. (2018). English language policies in scientific journals: Signs of change in the field of economics. *Journal of English for Academic Purposes*, 36, 26-36.

<https://doi.org/10.1016/j.jeap.2018.08.001>

Hernández Madrigal, M., Ramírez Flores, E. y Gamboa Cerda, S. (2018). La implementación

- de una prueba estandarizada en una institución de educación superior. *Innovación educativa*, 18(76), 149-170.
- Hernández, J. (27 de febrero de 2019) ¿Cuándo tendremos una verdadera política de enseñanza del idioma inglés?. *Nexos*. <https://educacion.nexos.com.mx/cuando-tendremos-una-verdadera-politica-de-ensenanza-del-idioma-ingles/>
- Hu H., Said N.E.M. y Hashim H. (2022) Killing two birds with one stone? A study on achievement levels and affective factors in content and language integrated Learning (CLIL). *International Journal of Learning, Teaching and Educational Research*, 21(4) 150-167. <https://doi.org/10.26803/ijlter.21.4.9>
- Huang, J. (2014). Learning to write for publication in English through genre-based pedagogy: A case in Taiwan. *System*, 45, 175-186. 10.1016/j.system.2014.05.010
- Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of second language*, 31, 58-69. <http://dx.doi.org/10.1016/j.jslw.2016.01.005>
- Hyland, K. y Jiang, F. (2021a). A bibliometric study of EAP research: Who is doing what, where and when?. *Journal of English for Academic Purposes*, 49, 1-12. <https://doi.org/10.1016/j.jeap.2020.100929>
- Hyland, K. y Jiang, F. (2021b). Delivering relevance: The emergence of ESP as a discipline. *Journal of English for Academic Purposes*, 64, 13-25
- Hyland, K., y Hyland, F. (Eds.). (2019). Feedback in second language writing: Contexts and issues.
- Instituto Cervantes. (2022a). *El español en el mundo 2022. Anuario del Instituto Cervantes*. Autor.
- Instituto Cervantes. (2022b). El español: una lengua viva. Informe 2018. Instituto Cervantes. https://cvc.cervantes.es/lengua/espanol_lengua_viva/pdf/espanol_lengua_viva_2018.pdf

- Instituto de Investigación y Desarrollo Educativo [IIIDE] (2020). Plan de desarrollo del Instituto de Investigación y Desarrollo Educativo 2019-2023. Autor.
- Instituto Mexicano para la Competitividad (IMCO). (2015). Inglés es posible. Propuesta de una Agenda Nacional.
- Iwai, Y. (2011). The effects of metacognitive reading strategies: pedagogical implications for EFL/ESL teachers. *The Reading Matrix*, 11(2), 150-159.
https://www.readingmatrix.com/articles/april_2011/iwai.pdf
- Jin, Y. (2019). Testing tertiary-level english language learners. En *The college english test in China* (pp. 101-130). Routledge. 10.4324/9781351254021-5
- Jöreskog, K.G. y Sörbom, D. (1989): *Lisrel 7: A guide to the Program Applications*. Chicago: SPSS Inc
- Jornet, J.M. y Suárez, J.M. (1996). Pruebas estandarizadas y evaluación del rendimiento: usos y características métricas. *Revista de Investigación Educativa*, 14(2), 141-163
- Kaldaras, L., Akaeze, H. y Reckase, M. (2024). Developing valid assessments in the era of generative artificial intelligence. *Frontiers in Education*, doi: 10.3389/feduc.2024.1399377
- Kalnberzina, V. y Karapetjana, I. (2021). Assessment of academic staff performance in english medium instruction. *Baltic Journal of English Language, Literature and Culture*, 11, 52-64. <https://doi.org/10.22364/BJELLC.11.2021.04>
- Karimpour, P. y Mazlum, F. (2024). EAP practitioners' assessment behavior: Bringing the hidden-away to light. *Journal of English for Academic Purposes*, 65, 1-15.
<https://doi.org/10.1016/j.jeap.2023.101321>
- Kharchenko T., Semashko T., Dolynskiy I., Bepala L. y Ivanova T. (2021). Use of moodle lms-based tests for enhancing linguistic competence of students majoring in foreign language philology. *Journal of Curriculum and Teaching*, 10(4), 67-81. DOI:

10.5430/jct.v10n4p67

- Khlaif, Z., Odeh, A. y Bsharat, T. (2024). Generative AI-Powered adaptive assessment. En M. Sanmugam, D. Lim, N. Mohd Barkhaya, W. Wan Yahaya, y Z. Khlaif (Eds.), *Power of persuasive educational technologies in enhancing learning* (pp. 157-176). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-6397-3.ch007>
- Lawshe C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- Li, Y. y Flowerdew, J. (2020). Teaching English for Research Publication Purposes (ERPP): A review of language teachers' pedagogical initiatives. *English for Specific Purposes*, 59, 29-41.
- López-Leyva, S., Alvarado-Borrego, A. y Mungaray-Moctezuma, A. (2018). La difusión de la ciencia en México a través de artículos científicos. Condiciones y contextos. *Revista de la educación superior*, 47(188), 157-176
- Lousado, J. P., Costa, C. M. M., Oliveira I. P. y Roberto M. T. G. (2011). Audio folios management system for foreign language learning. *6th Iberian Conference on Information Systems and Technologies*, Chaves, Portugal, pp. 1-5.
- Ma'mura, Y., Durдона, B., Umida, N., Umida, Y. y E'zoza, A. (2020). Methodology of teaching english for specific purposes and difference between general english. *European Journal of Research and Reflection in Educational Sciences*, 8(1), 125-128.
- Malone, M. Y Montee, M. (2014). *Stakeholders' Beliefs About the TOEFL iBT® Test as a Measure of Academic Language Ability* (Reporte No. RR-14-42). TOEFL iBT®. <https://files.eric.ed.gov/fulltext/EJ1109315.pdf>
- Marsaulina, R. M. (s. f.). Web 2.0 Technology Integrated Personalized Learning in CLT for EAP to at least CEFR Level B2. *Indonesia: Institut Teknologi Del*. <https://conference.pixel-online.net/FOE/files/foe/ed0009/FP/5967-ITLM4095-FP->

FOE9.pdf

- Martínez-Hernández, A. y Bellés-Fortuño. B. (2021). Accommodating the syllabus to visually impaired students in the english language classroom: Challenges and concerns. *International Journal of English Studies*, 21(1), 75–92. doi: 10.6018/ijes.438891.
- McDowell, L. y Liardet, C. (2019). Japanese materials scientists' experiences with English for research publication purposes. *Journal of English for Academic Purposes*. 37, 141-153. <https://doi.org/10.1016/j.jeap.2018.11.011>
- Meneses, J., Barrios Cerrejón, M., Bonillo Martín, A., Meneses Naranjo, J., Cosculluela Mas, A., Turbany Oset, J., Lozano Fernández, L. M. y Valero Ventura, S. (2013). *Psicometría*. Editorial UOC, S.L.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. <https://doi.org/10.1037/0003-066X.50.9.74>
- Morales, J. A. (2020). Leer e investigar en Educación Superior. *MLS Educational Research*, 4(2), 34-53. doi:10.29314/mlser.v4i2.355
- More, A. (2024). *Generative AI. Advances in computational intelligence and robotics book series*. IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-8557-9.ch002>
- Moscoso, M.L., Vega, M.V. y Martínez, M. (2019). La evolución del inglés como lengua franca, su enseñanza y aplicación en los negocios. Casa Editora.
- Muijs, D. (2011). *Doing quantitative research in education with SPSS*. SAGE Publications Ltd, <https://doi.org/10.4135/9781446287989>
- Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuestas a los ítems. *Papeles del psicólogo*, 31(1), 57-66.
- Muñiz, J. y Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema*,

31(1), 7-16. doi: 10.7334/psicothema2018.291

Nakamura, K. (2022). Investigating the factor structure of the Test of English for Academic Purposes (TEAP) and its relation to test takers' perceived test task value. *Language Testing in Asia*, 12(35).

Nithideechaiwarachok, B., Maneekeanon, O. y Bubphad, T. (2022). Exploring english language proficiency, english language problems, and english needs among first year undergraduate students. *International Journal of Learning, Teaching and Educational Research*, 21(12), 272-289. <https://doi.org/10.26803/ijlter.21.12.15>

North, B. (2000). The development of a common framework scale of language proficiency. Peter Lang

O'Grady, S. y Taşkesen, Ö. (2022). Developing a rating scale for integrated assessment of reading-into-writing skills. *Language Learning in Higher Education*, 12(1), 159–183. doi: 10.1515/cercles-2022-2049

O'Neill, G. y McMahan, T. (2005). Student-centred learning: what does it mean for students and lecturers? En G. O'Neill, S. Moore y B. McMullin (Eds.), *Emerging Issues in the Practice of University Learning and Teaching* (pp. 27-36). AISHE

Pearson, W. (2021). A comparative study of lexical bundles in IELTS Writing Task 1 and 2 simulation essays and tertiary academic writing. *Journal of Academic Language and Learning*, 15(1), 27–52. <https://journal.aall.org.au/index.php/jall/article/view/717>

Pearson, W. (2022). Test review: LanguageCert IESOL B1 (Achiever) SELT. *Language Assessment Quarterly*, 19(55), 526-537. DOI: 10.1080/15434303.2022.2103420

Pearson. (2015). Global Scale of English Learning Objectives for Academic English. Autor.

Pérez, M. (2010). Using virtual learning environments and computer-mediated communication to enhance the lexical competence of pre-service English teachers: a quantitative and qualitative study. *Computer Assisted Language Learning*, 23(2), 129-150.

<http://dx.doi.org/10.1080/09588221003666222>.

Plakans, L., Liao, J., Wang, F. (2018). Integrated assessment research: Writing-into-reading.

Language Teaching, 51(3), 430–434. doi:10.1017/S0261444818000149

Plews, R. C. y Amos, M. L. (2020). *Evidence-based Faculty Development through the*

Scholarship of Teaching and Learning (SoTL). Hershey, PA: Information Science Reference.

Popham, W.J. (1983). *Evaluación basada en criterios*. Magisterio Español.

Prapinwong, M. y Koolsriroj, U. (2019). ICCP: An instructional model for academic English

skills development of educational researchers. *Kasetsart Journal of Social Sciences*, (40), 677–683. <https://doi.org/10.34044/j.kjss.2019.40.3.03>

Prihandoko, L.A. (2024). Exploring the impact of academic reading and research competence

on academic writing: A multi-regression analysis among EFL undergraduates. *JOLLT Journal of Languages and Language Teaching*, 12(1), 451-462.

<https://doi.org/10.33394/jollt.v%vi%i.9738>

Qayyum, R. (2024). A teacher's inquiry into diagnostic assessment in an EAP writing course.

Assessing Writing, 61, 1-16. <https://doi.org/10.1016/j.asw.2024.100848>

Real Academia Española. (s.f.). Pidgin. En *Diccionario de la lengua española*. Recuperado el 5

de noviembre de 2022, de <https://dle.rae.es/pidgin>

Reidl-Martínez, L. (2013). Confiabilidad en la medición. *Investigación en educación médica*,

2(6), 107-111. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572013000200007&lng=es&tlng=es.

Riefani, M. y Utami, N. (2017). The assesment of high order thinking skills of undergraduate

students in biology education department. *Advances in Social Science, Education and Humanities Research*. 100, 350-351. doi: 10.2991/SEADRIC-17.2017.75

Rodríguez, M. (2021). *Desarrollo y validación de una prueba para el diagnóstico cognitivo de*

los aprendizajes del inglés en educación secundaria. [Tesis de maestría no publicada]

Universidad Autónoma de Baja California.

http://iide.ens.uabc.mx/documentos/divulgacion/tesis/DCE/2016/Maite_Rodriguez_de_1_a_Vega_Elu.pdf

Said, M. y Weda, S. (2018). English language anxiety and its impacts on students' oral communication among Indonesian students: a case study at Tadulako University and Universitas Negeri Makassar. *TESOL International Journal*, 13(3), 21-30.

Salager-Meyer, F. (2008). Scientific publishing in developing countries: Challenges for the future. *Journal of English for Academic Purposes*, 7, 121-132.

<https://doi.org/10.1016/j.jeap.2008.03.009>

Şata, M. y Karakaya, I. (2021). Investigating the effect of rater training on differential rater function in assessing academic writing skills of higher education students. *Journal of Measurement and Evaluation in Education and Psychology*, 12(2), 163-181. doi: 10.21031/epod.842094

Scientific Journals Ranking. (marzo de 2024). *Scimago Journal and Country Rank*.

Recuperado el 12 de abril de 2024. <https://www.scimagojr.com/countryrank.php>

Secretaría de Economía. (2019). TLCAN: Inicia el año 26 de su entrada en vigor.

<https://www.gob.mx/se/prensa/tlcan-inicia-el-ano-26-de-su-entrada-en-vigor> Secretaría

de Educación Pública. 2015. *Diagnóstico Ampliado S270: Programa Nacional de Inglés para Alumnos en Educación Básica*.

https://www.gob.mx/cms/uploads/attachment/file/64151/Diagnostico_del_programa.pdf

Secretaría de Educación Pública [SEP] (2023). Preguntas y respuestas CENNI.

https://dgair.sep.gob.mx/storage/recursos/CENNI/cenni_faq.pdf

Secretaría de Educación Pública [SEP]. (2015). Diagnóstico ampliado: S270. Programa

nacional de inglés para alumnos en educación básica.

https://www.gob.mx/cms/uploads/attachment/file/64151/Diagno_stico_del_programa.pdf

Serrano, D. (2016). El inglés, lingua franca de la globalización y su repercusión en diversos ámbitos sociales y académicos. *Revista Fuentes Humanísticas*, (52), 177-190.

Sharpling, G.P. (2010). When BAWE meets WELT: The use of a corpus of student writing to develop items for a proficiency test in grammar and English usage. *Journal of Writing Research*, 2(2), 179-195.

Simonova, I. (2019). Autonomous and collaborative e-learning in English for Specific Purposes. En I. O. Pappas, P. Mikalef, Y.K. Dwivedi, L. Jaccheri, J. Krogstie, y M. Mäntymäki (Eds). *Digital transformation for a sustainable society in the 21st century* (pp. 609–620). Springer, Cham. https://doi.org/10.1007/978-3-030-29374-1_49

Sistema Educativo Estatal. (2010). *Programa de inglés en educación básica*.

Song, Y. y Zhou, J. (2022). Revising english language course curriculum among graduate students: An EAP needs analysis study. *SAGE Open*, 12(3).

<https://doi.org/10.1177/21582440221093040>

Steger, M. (2009). *Globalization. A very short introduction*, 2nd edition. Oxford University Press

Steger, M. B. (2019). *Globalización: una breve introducción*. ANTONI BOSCH EDITOR.

Stevens, P. (1977). Special-purposed language learning: A perspective. *Language Teaching*, 10, 145-163.

Taskiran, A. y Goksel, N. (2022). Automated feedback and teacher feedback: writing achievement in learning English as a foreign language at a distance. *Turkish Online Journal of Distance Education-TOJDE*, 23(2), 120-139.

Tobler, S. (2024). Smart grading: A generative AI-based tool for knowledge-grounded answer

- evaluation in educational assessments. *MethodsX*, 12.
<https://doi.org/10.1016/j.mex.2023.102531>
- Toledo, D.G. y García, L. (2020). La realidad lingüística en la frontera Tijuana (México)- San Diego (Estados Unidos). *Traslaciones. Revista Latinoamericana de Lectura y Escritura*, 7(13), 207-233.
- Tristán, A. (2008). Modificación al Modelo de Lawshe para el dictamen Cuantitativo de la validez de contenido de un instrumento objetivo. *Revista Avances de medición*, 6, 32-48.
- Tristán, A. y Vidal, R. (2006). *Estándares de calidad para pruebas Objetivas*. Magisterio
- Tristán, A., y Pedraza, N. Y. (2017). La Objetividad en las pruebas estandarizadas. *Revista Iberoamericana de Evaluación Educativa*, 10(1), 11-31.
<https://doi.org/10.15366/riee2017.10.1.001>
- Tuela, A., y Palar, Y. (2022). Analysis of Higher Order Thinking Skills (HOTS) based on bloom taxonomy in comprehensive examination questions. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 8(4), 957-971. <https://doi.org/10.33394/jk.v8i4.5885>
- U.S. Department of Energy (s.f.) Writing your Abstract.
<https://dl.icdst.org/pdfs/files2/c20dc88a777d07244363b20bcf3edf2f.pdf>
- Universidad Autónoma de Baja California [UABC] (2018). Modelo educativo de la UABC. Mexicali: UABC.
- Universidad Autónoma de Baja California [UABC] (2019a). Estatuto general de la universidad autónoma de Baja California. UABC.
- Universidad Autónoma de Baja California [UABC] (2019b). Plan de desarrollo institucional 2019-2023. UABC.
- Universidad Autónoma de Baja California [UABC] (2023). Plan de desarrollo institucional

2023-2027. UABC.

- Urgo, K., Arguello, J. y Capra, R. (2019). Anderson and Krathwohl's two-dimensional taxonomy applied to task creation and learning assessment. ICTIR '19, October 2–5. <https://doi.org/10.1145/3341981.3344226>
- Wang, F y Wang, S. (2012). A Comparative Study on the Influence of Automated Evaluation System and Teacher Grading on Students' English Writing. *Procedia Engineering*, 29, 993-997. 10.1016/j.proeng.2012.01.077
- Zabala-Delgado, J. y Sawicka, B. (2019). Mutual recognition of certification systems: The case of SERMO and ACLES. *Language Learning in Higher Education*, 9(2), 289–319. <https://doi.org/10.1515/cercles-2019-0017>.
- Zainal, S. A. y Jamil, A. (2015). Toward an english proficiency test for postgraduates in Malaysia. *SAGE Open*, 5(3). DOI: 10.1177/2158244015597725
- Zhang, L. (2022). College English Test-Spoken English Test (CET-SET). *STUDIES IN LANGUAGE ASSESSMENT*, 11(2), 164-180.
- Zorrilla, M., (2004). La educación secundaria en México: al filo de su reforma. REICE. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2(1).

Anexos

Anexo 1

Ejemplo de presentación de capacitación a expertos

INTRODUCCIÓN

Diseñar, desarrollar y validar una prueba diagnóstica del dominio del idioma inglés con propósitos de publicación científica en estudiantes de posgrado.

Anotar la calificación de acuerdo a su criterio: 0 irrelevante; 1 relevante

MCER	GSE	READING	Habilidad requerida	Observaciones
		GSE 10-21/ Below A1	0-1	
	10	10.1 Can recognise cardinal numbers up to 10. (P)		
	10	10.2 Can recognise the letters of the alphabet (P)		
A1-	12	12 Can read and understand simple prices. (P)		
	14	14.1 Can identify very common food and drink on a menu. (P)		
	14	14.2 Can recognise basic plural for ms of nouns (e.g cars, books). (P)		
	15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)		
		GSE 22-29/ A1		
A1	26	26 Can follow short, simple written directions (e.g to go from X to Y). (C)		
	27	27 Can understand short written notices, signs and instructions with visual support. (P)		
		GSE 30-35/ A2		
	31	31.1 Can understand simple questions in questionnaires on familiar topics. (P)		
	31	31.2 Can find specific, predictable information in everyday materials (e.g menus, timetables). (CA)		
A2	31	31.3 Can understand short, simple messages on postcards, emails and social networks. (CA)		
	34	34.1 Can understand the general meaning of short, simple informational material and descriptions if there is visual support. (CA)		
	34	34.2 Can understand simple instructions on everyday equipment (e.g cash machines). (CA)		

Ejemplo

Si lemos la habilidad 14.1 "can identify very common food and drink on a menu" reconocemos que no tiene relevancia en la redacción de textos científicos

MCER	GSE	READING	Habilidad requerida	Observaciones
		GSE 10-21/ Below A1	0-1	
	10	10.1 Can recognise cardinal numbers up to 10. (P)		
	10	10.2 Can recognise the letters of the alphabet (P)		
A1-	12	12 Can read and understand simple prices. (P)		
	14	14.1 Can identify very common food and drink on a menu. (P)		
	14	14.2 Can recognise basic plural for ms of nouns (e.g cars, books). (P)		
	15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)		
		GSE 22-29/ A1		
A1	26	26 Can follow short, simple written directions (e.g to go from X to Y). (C)		
	27	27 Can understand short written notices, signs and instructions with visual support. (P)		
		GSE 30-35/ A2		
	31	31.1 Can understand simple questions in questionnaires on familiar topics. (P)		
	31	31.2 Can find specific, predictable information in everyday materials (e.g menus, timetables). (CA)		
A2	31	31.3 Can understand short, simple messages on postcards, emails and social networks. (CA)		
	34	34.1 Can understand the general meaning of short, simple informational material and descriptions if there is visual support. (CA)		
	34	34.2 Can understand simple instructions on everyday equipment (e.g cash machines). (CA)		

Anexo 2

Tabla de evaluación de indicadores

MCER	GSE	READING	Habilidad requerida	Observaciones
A1-		GSE 10-21/ Below A1	0-1	
	10	10.1 Can recognise cardinal numbers up to 10. (P)		
		10.2 Can recognise the letters of the alphabet (P)		
	12	12 Can read and understand simple prices. (P)		
	14	14.1 Can identify very common food and drink on a menu. (P)		
	14	14.2 Can recognise basic plural for ms of nouns (e.g cars, books). (P)		
	15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)		
A1		GSE 22-29/ A1		
	26	26 Can follow short, simple written directions (e.g.to go from X to Y). (C)		
	27	27 Can understand short written notices, signs and instructions with visual support. (P)		
A2		GSE 30-35/ A2		

	31	31.1 Can understand simple questions in questionnaires on familiar topics. (P)		
		31.2 Can find specific, predictable information in everyday materials (e.g. menus, timetables) (CA)		
		31.3 Can understand short, simple messages on postcards, emails and social networks. (CA)		
	34	34.1 Can understand the general meaning of short, simple informational material and descriptions if there is visual support. (CA)		
		34.2 Can understand simple instructions on everyday equipment (e.g. cash machines). (CA)		
	A2 +	GSE 36-42/ A2(+)		
37		37.1 Can understand rules and regulations (e.g. safety) if expressed in simple language (C)		
		37.2 Can identify specific information in simple letters, brochures and short articles. (CA)		
		37.3 Can understand short, simple personal emails and letters. (CA)		
		37.4 Can understand reports of travel delays and cancellations. (P)		
38		38 Can make basic inferences from simple information in a short text (P)		
39		39.1 Can understand simple factual titles and headlines relating to common events. (P)		
	39.2 Can identify specific information in a simple factual text (P)			
	39.3 Can understand the instructions to buy tickets online. (P)			
42	42 Can understand the main information from simple diagrams (eg. graphs, bar charts) . (P)			
GSE 43-50/ B1				
B1	43	43.1 Can understand simple technical information (e.g. instructions for everyday equipment). (CA)		
		43.2 Can make basic inferences or predictions about text content from headings, titles or headlines. (P)		
	44	44 Can understand written instructions for taking medication. (P)		
	45	45.1 Can understand information in advertisements for jobs and services. (P)		
		45.2 Can find and understand information in advertisements for sporting or cultural events. (P)		
		45.3 Can skim a simple text to identify key concepts. (P)		
	46	46.1 Can derive the probable meaning of simple unknown words from short, familiar contexts. (CA)		
		46.2 Can understand clearly written, straightforward instructions on how to use a piece of equipment. (CA)		
		46.3 Can understand basic types of standard letters and emails on familiar topics (e.g. enquiries, complaints). (CA)		
		46.4 Can generally understand straightforward factual texts on familiar topics. (CA)		

	47	47 Can identify the topic sentence of a paragraph. (P)		
	48	48.1 Can extract relevant details in everyday letters, brochures and short official documents. (CA)		
		48.2 Can recognise the writer's point of view in a simple academic text. if guided by questions. (P)		
		48.3 Can predict the content of a simple academic text, using headings, images. and captions. (P)		
		48.4 Can take basic notes on a text about a familiar topic in their field of study. (P)		
		48.5 Can identify key information in a simple academic text. if guided by questions. (P)		
	49	49.1 Can identify the main topic and related ideas in a structured text. (P)		
		49.2 Can scan a simple academic text to find specific information. (P)		
	50	50.1 Can understand the relationship between a main point and an example in a structured text. (P)		
		50.2 Can understand the main idea of a passage using textual clues. (P)		
		GSE 51-58/ B1(+)		
B1 +	51	51.1 Can recognise the direct repetition of ideas as a simple cohesive device. (P)		
		51.2 Can distinguish between fact and opinion in relation to common topics. (P)		
		51.3 Can make simple inferences based on information given in a short article. (P)		
		51.4 Can generally understand details of events, feelings and wishes in letters, emails and online postings (CA)		
	52	52.1 Can follow chronological sequence in a formal structured text. (P)		
		52.2 Can recognise the organisational structure of a paragraph in a simple academic text. (P)		
		52.3 Can understand cause and effect relationships in a simple academic text, if clearly signalled. (P)		
	53	53.1 Can identify whether an author is quoting or paraphrasing another person. (P)		
		53.2 Can recognise common discourse markers that convey emphasis in a simple text (P)		
		53.3 Can identify the sources of information in a simple academic text. (P)		
		53.4 Can understand the writer's purpose in a simple academic text, if guided by questions. (P)		
	54	54.1 Can recognise examples and their relation to the idea they support. (P)		
		54.2 Can infer meaning in a simple academic text. in order to answer specific questions (P)		
		54.3 Can distinguish between fact and opinion in a simple academic text. (P)		
		54.4 Can distinguish between the main idea and related ideas in a simple academic text in order to answer specific questions. (P)		

		54.5 Can understand numerical values in graphs and charts in a simple academic text (PI)		
		54.6 Can identify the key points presented in graphs and charts in a simple academic text, if guided by questions. (P)		
	55	55.1 Can understand written advice and instructions for resolving a problem with a product or piece of equipment (P)		
		55.2 Can guess the meaning of an unfamiliar word from context. (P)		
		55.3 Can distinguish between different viewpoints in a simple academic text (P)		
	56	56.1 Can recognise significant points and arguments in straightforward newspaper articles on familiar topics (CA)		
		56.2 Can follow the chronological sequence of events in an academic text using numbers, times and dates. (P)		
		56.3 Can scan an interview transcript for key information. (P)		
		56.4 Can recognise that ideas are parallel in a simple academic text (P)		
	57	57.1 Can understand cause and effect relationships in a structured text (P)		
		57.2 Can recognise the general line of a written argument though not necessarily all the details. (CA)		
		57.3 Can identify different types of supporting details in a simple academic text, in order to answer specific questions. (P)		
		57.4 Can distinguish between active and passive voice in an academic text (P)		
	58	58.1 Can understand most correspondence relating to their field of interest (CA)		
		58.2 Can recognise the writer's point of view in a structured text. (P)		
		58.3 Can understand problem and solution relationships in a structured text (P)		
		58.4 Can synthesise information from two or more basic texts. if guided by questions. (P)		
		58.5 Can understand relationships between ideas in a simple academic text. if guided by questions. (P)		
		GSE 59-66/ B2: Reading		
B2	59	59.1 Can use a monolingual dictionary to check the meaning of words without needing to refer to a bilingual dictionary (P)		
		59.2 Can understand instructions for making financial transactions online. (P)		
		59.3 Can identify the main conclusions in a text that presents and contrasts arguments in a clearly signalled way. (CA)		
	60	60.1 Can make inferences or predictions about the content of newspaper and magazine articles from headings, titles or headlines. (P)		
		60.2 Can identify the use of paraphrasing in a simple academic text (P)		
		60.3 Can understand simple metaphors in an academic text (P)		

		60.4 Can recognise inferred meaning in a structured text, if guided by questions. (P)		
	61	61.1 Can follow the exchanges on the discussion board of a website. (P)		
		61.2 Can distinguish supporting details from the main points in a text (P)		
		61.3 Can critically evaluate the quality of sources used in a simple text (P)		
		61.4 Can critically evaluate the effectiveness of a simple descriptive essay. (P)		
		61.5 Can critically evaluate the effectiveness of a simple problem-solution essay. (P)		
		62	62.1 Can interpret the main message from complex diagrams and visual information. (P)	
	62.2 Can recognise the author's use of irony in a simple text. if guided by questions. (P)			
	62.3 Can understand the author's purpose and intended audience. (P)			
	63	63.1 Can use reference materials to check factual information, when guidance on finding relevant sources of information is provided. (P)		
		63.2 Can scan a long text or a set of related texts in order to find specific information. (CA)		
		63.3 Can critically evaluate the effectiveness of a simple discursive essay. (P)		
		63.4 Can understand the use of quotes in an academic text. (P)		
		63.5 Can identify the main line of argument in an academic text. (P)		
	65	65.1 Can recognise the tone and intended audience of a structured text. (P)		
		65.2 Can identify examples in an academic text to support an argument. (P)		
	66	66.1 Can summarise, comment on and discuss a wide range of factual and imaginative texts. (CA)		
		66.2 Can critically evaluate the effectiveness of a simple argumentative essay. (P)		
		GSE 67-75/ B2(+)		
B2+	67	67 Can recognise the repetition of ideas expressed by substitution, paraphrasing, etc. (P)		
	68	68.1 Can recognise contrasting arguments in structured, discursive text. (P)		
		68.2 Can understand complex technical information such as operating instructions, specifications for familiar products and services. (C)		
		68.3 Can infer what will come next in an unstructured text by using contextual, grammatical and lexical cues. (P)		
		68.4 Can understand the details of long complex instructions in their field, rereading as necessary. (CA)		
		68.5 Can identify common features of an academic abstract. (P)		

	68.6 Can identify different types of supporting details in an academic text. (P)		
	68.7 Can recognise common discourse markers that convey emphasis in a linguistically complex text. (P)		
	68.8 Can distinguish between active and passive voice in a linguistically complex academic text. (P)		
69	69 Can synthesise information from different sources in order to give a written or oral summary (P)		
	69.2 Can quickly scan long, complex texts for key information. (P)		
	69.3 Can use a variety of reference materials to check factual information quickly and efficiently. (P)		
	69.4 Can recognise contrasting ideas in a linguistically complex academic text when signalled by discourse markers. (P)		
	69.5 Can use a synopsis to identify where specific information can be located in a long text. (P)		
70	70.1 Can understand inferred meaning in formal structured text. (P)		
	70.2 Can understand complex, detailed correspondence, with occasional support from a dictionary. (CA)		
	70.3 Can distinguish between the main idea and related ideas in a linguistically complex academic text. (P)		
	70.4 Can evaluate information in an academic text using specific criteria. (P)		
71	71.1 Can distinguish between fact and opinion in complex formal contexts. (P)		
	71.2 Can understand complex questions in questionnaires designed to elicit opinions. (P)		
	71.3 Can predict the content of a linguistically complex academic text by reading introductory and summary statements. (P)		
72	72.1 Can guess the meaning of an unfamiliar word from context in a linguistically complex academic text (P)		
	72.2 Can recognise poetic devices such as rhythm, alliteration or repetition. (P)		
	72.3 Can scan a long and linguistically complex interview transcript for key information. (P)		
	72.4 Can use relevant material in academic textbooks and articles to support or challenge an argument. (CA)		
73	73.1 Can identify specific information in a linguistically complex factual text (P)		
	73.2 Can recognise the author's use of irony in a text . (P)		
	73.3 Can scan a linguistically complex academic text to find specific information. (P)		
	73.4 Can identify the sources of information in a linguistically complex academic text. (P)		
74	74.1 Can understand the intended double meaning of a word in a written text. (P)		
	74.2 Can understand cause and effect relationships in a linguistically complex academic text. (P)		

		74.3 Can take effective notes on a complex and unfamiliar text. (P)		
		74.4 Can recognise organisational patterns within a complex academic text. (P)		
		74.5 Can recognise the writer's point of view in a linguistically complex academic text. (P)		
		74.6 Can understand the writer's purpose in a linguistically complex academic text (P)		
		74.7 Can distinguish between different viewpoints in a linguistically complex academic text (P)		
	75	75 Can identify the main line of argument in a linguistically complex academic text (P)		
		GSE 76-84/ C1		
C1	76	76.1 Can understand correspondence containing idiomatic or non-standard language (CA)		
		76.2 Can understand the use of paraphrasing in a linguistically complex academic text. (P)		
		76.3 Can infer the author's attitude in a linguistically complex academic text. (P)		
		76.4 Can recognise that ideas are parallel in a linguistically complex academic text (P)		
	77	77.1 Can identify logical fallacies in an argument in an academic paper. (P)		
		77.2 Can distinguish between literal and allegorical meaning in a literary text. (P)		
		77.3 Can follow abstract argumentation, for example the balancing of alternatives and the drawing of a conclusion. (CA)		
	78	78.1 Can extract information, ideas and opinions from highly specialised sources within their field. (CA)		
		78.2 Can infer meaning in a linguistically complex academic text. (P)		
	79	79.1 Can understand complex arguments in newspaper articles. (P)		
		79.2 Can research a topic by reading linguistically complex academic texts. (P)		
		79.3 Can critically evaluate the effectiveness of a linguistically complex problem-solution essay. (P)		
		79.4 Can critically evaluate the effectiveness of a linguistically complex descriptive essay. (P)		
		79.5 Can compare the presentation of a key concept in different texts by different authors using different styles of writing. (P)		
	80	80.1 Can understand the details of long complex texts in their field without needing to reread. (CA)		
		80.2 Can critically evaluate the quality of sources used in a linguistically complex text. (P)		
		80.3 Can infer the interviewee's opinion on a subject from a long and complex interview transcript. (P)		
81	81.1 Can recognise multiple purposes in a linguistically complex academic text. (P)			

		81.2 Can critically evaluate the effectiveness of a linguistically complex argumentative essay. (P)		
		81.3 Can understand a critique of a linguistically complex academic text. (P)		
	82	82.1 Can critically evaluate the effectiveness of a linguistically complex discursive essay. (P)		
		82.2 Can understand a linguistically complex poem. (P)		
	84	84 Can understand complex or extended metaphors in an academic text. (P)		
		GSE 85-90/ C2		
C2	85	85.1 Can understand linguistically complex academic texts in specialised fields. (CA)		
		85.2 Can understand complex arguments in technical or academic journals. (CA)		
	87	87.1 Can understand highly colloquial language in unstructured texts that use complex structures. (P)		
		87.2 Can recognise subtle distinctions of style in linguistically complex academic texts. (CA)		
		WRITING		
		GSE 10-21/ Below A1		
A1 -	10	10.1 Can write the letters of the alphabet in upper and lower case. (P)		
		10.2 Can write their name, address and nationality. (CA)		
	18	18 Can copy familiar words and short phrases about everyday objects and set phrases. (CA)		
	20	20 Can write consistently with joined-up letters. (P)		
		GSE 22-29/ A1		
A1	23	23 Can complete simple forms with basic personal details. (CA)		
	24	24 Can copy short sentences on everyday subjects (e.g. directions how to get somewhere). (C)		
	25	25 Can write simple sentences about things that they and other people have. (P)		
	26	26 Can use basic punctuation (e.g. commas, full stops, question marks). (P)		
	27	27.1 Can spell a range of common names. (P)		
		27.2 Can write simple sentences about their family and where they live. (CA)		
		27.3 Can write simple sentences about personal interests. (P)		
		27.4 Can complete a simple form requiring travel information (e.g. landing card, customs declaration). (P)		
	28	28.1 Can write short, simple notes, emails and postings to friends. (N2000A)		
		28.2 Can spell a range of common greetings. (P)		
28.3 Can spell a range of common jobs. (P)				

		28.4 Can write dates using both digits and words. (P)		
		28.5 Can write simple sentences about someone's life and routines. (P)		
	29	29 Can write times using both digits and words. (P)		
		GSE 30-35/ A2		
A2	30	30 Can write simple sentences about what they and other people do. (CA)		
	31	31.1 Can write simple sentences about someone's work and duties. (P)		
		31.2 Can use very basic connectors like <i>and</i> , <i>but</i> , <i>so</i> and <i>then</i> . (CA)		
		31.3 Can write a simple description of a room, house or apartment. (P)		
		31.4 Can give personal details in written form in a limited way. (CA)		
	32	32 Can write very short, basic directions. (P)		
	33	33 Can write simple sentences about personal skills. (P)		
	34	34 Can give an example of something in a very simple text using like or for example. (P)		
		GSE 36-42/ A2(+)		
A2+	36	36.1 Can make simple comparisons between people, places or things. (P)		
		36.2 Can ask for personal details in written form in a limited way (CA)		
	38	38.1 Can write short, simple notes, emails and messages relating to everyday matters. (CA)		
		38.2 Can write short, basic descriptions of places, people or things. (P)		
		38.3 Can write a description of a simple everyday process (e.g. a recipe). (P)		
		38.4 Can write very simple personal emails/letters expressing thanks and apology. (CA)		
	39	39.1 Can signal contrast in simple text with a limited range of language. (P)		
		39.2 Can write basic instructions with a simple list of points. (P)		
		39.3 Can take a short, simple message if they are able to ask the speaker to repeat and reformulate it. (CA)		
		39.4 Can write short basic descriptions of past events and activities. (P)		
	40	40.1 Can use appropriate openings and endings in simple informal emails. (P)		
		40.2 Can write a brief summary of their own personal details, e.g. for a business card. (P)		
	41	41.1 Can write a basic informal email/letter of invitation with simple, key details. (P)		
		41.2 Can write a basic email/letter accepting or declining an invitation. (P)		

		41.3 Can write about everyday things (e.g people. places, job. study) in linked sentences. (CA)		
		41.4 Can write simple sentences about their educational background and present or past job. (CA)		
		GSE 43-50/ B1		
B1	43	43 Can use simple headings to organise written work. (P)		
	44	44.1 Can clearly signal the end of a simple narrative or description. (P)		
		44.2 Can write a basic description of experiences, feelings and reactions, given a model. (P)		
	45	45.1 Can make simple. logical paragraph breaks in a longer text. (P)		
		45.2 Can write a short, simple description of a familiar device or product. (P)		
		45.3 Can write an email/letter congratulating someone on something. (P)		
		45.4 Can write a basic paragraph containing a topic sentence and related details. if provided with a model. (P)		
	46	46.1 Can show a basic direct relationship between a simple problem and a solution. (P)		
		46.2 Can write a basic formal email/letter requesting information. (P)		
		46.3 Can write a description of a future event or activity (P)		
		46.4 Can write the concluding sentence or sentences of a basic paragraph, given a model. (P)		
		46.5 Can write simple instructions on how to use a device or product. given a model. (P)		
		46.6 Can write simple informal emails/letters and online postings giving news or opinions. (P)		
		46.7 Can write short, simple essays with basic structure on familiar topics. (P)		
	47	47.1 Can write about personal interests in some detail. (CA)		
		47.2 Can write descriptions of past events, activities, or personal experiences. (P)		
		47.3 Can write an everyday connected text using a set of short elements or facts and building them into a sequence. (CA)		
		47.4 Can describe the sequence in a process when writing a simple text, using common discourse markers. (P)		
	48	48.1 Can write personal emails/letters giving some details of events. experiences and feelings. (CA)		
		48.2 Can write a basic letter of application with limited supporting details. (P)		
		48.3 Can complete a detailed form requiring travel information (e.g. visa application). (P)		
		48.4 Can use common discourse markers to show order of importance (P)		
		48.5 Can prepare a simple questionnaire in order to gather data. (P)		

49	49.1 Can use simple cohesive devices such as repetition in a structured text. (P)			
	49.2 Can express a personal opinion in a simple academic text. (P)			
50	50.1 Can write fractions using both digits and words. (P)			
	50.2 Can write about experiences, feelings and reactions in a simple connected text. (CA)			
	50.3 Can write a basic summary of a simple text using the original wording and paragraph order. (CA)			
	50.4 Can write personal emails/letters giving advice. (P)			
	50.5 Can take notes on a simple presentation or lecture aimed at a general audience. (P)			
	50.6 Can take notes on a simple academic text. (P)			
	50.7 Can write an introduction to a simple essay, if provided with a model. (P)			
	50.8 Can paraphrase information from a simple text, if guided by questions. (P)			
	GSE 51-58/ B1(+)			
B1+	51	51.1 Can use limited discourse devices to link sentences smoothly into connected discourse. (CA)		
		51.2 Can write a basic email/letter of complaint requesting action. (P)		
		51.3 Can use common connectors to tell a story or describe an event in writing. (CA)		
		51.4 Can write short, simple biographies about real or imaginary people (CA)		
		51.5 Can complete a form requiring educational information (e.g. applying for a course of study). (P)		
		51.6 Can write an email/letter sending a message of sympathy. (P)		
	52	52.1 Can summarise the main message from simple diagrams (e.g. graphs, bar charts). (P)		
		52.2 Can clearly signal chronological sequence in narrative text. (P)		
		52.3 Can write a simple review of a film, book or TV programme using a limited range of language. (P)		
		52.4 Can write a review of a simple text, using appropriate conventions, if provided with a model. (P)		
		52.5 Can write a conclusion to a simple essay, if provided with a model. (P)		
	53	53.1 Can take messages, communicate enquiries and explain problems. (CA)		
		53.2 Can write emails/letters exchanging information, emphasising the most important point. (CA)		
		53.3 Can give someone clear, detailed directions on how to get somewhere in a letter, email, or online posting. (P)		
		53.4 Can write a description of a real or imagined event (e.g. a recent trip). (C)		
53.5 Can take notes while researching a familiar topic. (P)				

	53.6 Can write a simple descriptive essay, if provided with a model. (P)		
54	54.1 Can write a formal email/letter requesting information. (P)		
	54.2 Can complete a form requiring health information. (P)		
	54.3 Can respond to and comment on other people's personal updates on a social media website. (P)		
	54.4 Can summarise information from a simple presentation or lecture aimed at a general audience. (P)		
	54.5 Can write a simple essay in response to a specific question (P)		
	54.6 Can write a transcript of a simple interview. (P)		
55	55.1 Can write a formal email/letter accepting or declining an invitation. (P)		
	55.2 Can take notes of key points during a talk on a familiar topic, if delivered clearly. (CA)		
	55.3 Can complete a form requiring financial information (e.g. application for a bank account or credit agreement). (P)		
	55.4 Can write personal updates on a social media website using an appropriate style (P)		
	55.5 Can post comments on the discussion board of a website. (P)		
	55.6 Can vary the formality of greetings in emails/letters based on intended recipients. (P)		
	55.7 Can edit and improve a simple text (P)		
	55.8 Can write a simple discursive essay, if provided with a model. (P)		
	55.9 Can summarise simple research findings in an academic text, if provided with a model summary. (P)		
	55.10 Can write bullet points to summarise key points in a structured text (P)		
56	56.1 Can introduce a counter-argument in a simple discursive text using however. (P)		
	56.2 Can write instructions on how to use a device or product (P)		
	56.3 Can write a short, simple academic essay on a familiar topic, given a model. (P)		
57	57.1 Can summarise factual information within their field of interest (CA)		
	57.2 Can write a description of items for sale on a trading website. (P)		
	57.3 Can write emails/letters of complaint with supporting details. (P)		
	57.4 Can write a bibliography, if provided with a model. (P)		
	57.5 Can support a main idea with examples and reasons. (P)		
	57.6 Can write an essay in response to a specific question, if provided with a model. (P)		
	57.7 Can use simple graphs and charts to convey information in academic written work. (P)		

58	58.1 Can write instructions on how to look after an object, device or product. (P)			
	58.2 Can write emails/letters responding to personal news and views in detail. (CA)			
	58.3 Can collate short pieces of information and summarise them for somebody else. (CA)			
	58.4 Can write a description of a problem with a product or piece of equipment (P)			
	58.5 Can write comments and complaints about products and services. (P)			
	58.6 Can write a detailed description of a simple process. (P)			
	58.7 Can use appropriate outlines to organise ideas. (P)			
	58.8 Can write a chronological paragraph on an academic topic. if provided with a model. (P)			
	58.9 Can write a conclusion to a simple academic essay. (P)			
	GSE 59-66/ B2			
B2	59	59.1 Can write detailed descriptions of real or imaginary people (P)		
		59.2 Can demonstrate understanding of formality and conventions in standard letters. (P)		
		59.3 Can comment on factual information within their field of interest. (CA)		
		59.4 Can write a formal email/letter of thanks or apology with appropriate conventions. (P)		
		59.5 Can support ideas with relevant examples. (P)		
		59.6 Can write a detailed description of an object. device or product. (P)		
		59.7 Can write a strong topic sentence within a clear paragraph. (P)		
		59.8 Can write a simple discursive essay. (P)		
		59.9 Can write a review of a simple text. using appropriate conventions. (P)		
		59.10 Can write an introduction to a simple academic essay. (P)		
60	60.1 Can write a formal email/letter of invitation with appropriate register and conventions. (P)			
	60.2 Can clearly signal cause and effect relationships in structured text. (P)			
	60.3 Can write personal emails/letters giving and commenting on news in detail. (CA)			
61	61.1 Can reformulate an idea in different words to emphasise or explain a point. (P)			
	61.2 Can end a discursive argument with a clear conclusion and opinion. (P)			
	61.3 Can write instructions on how to repair an object. device or product. (P)			
	61.4 Can clearly signal the difference between fact and opinion in structured text. (P)			

	61.5 Can write a brief standard report conveying factual information, stating reasons for actions. (CA)		
	61.6 Can take notes on a presentation or lecture in their field of specialisation. (P)		
	61.7 Can signal that two ideas are similar when writing a simple academic text by using discourse markers. (P)		
	61.8 Can contrast two ideas when writing a simple academic text by using discourse markers. (P)		
62	62.1 Can signal additional information in a formal structured text with a range of language. (P)		
	62.2 Can write a structured text clearly signalling main points and supporting details. (P)		
	62.3 Can clearly signal problem and solution relationships in structured text. (P)		
	62.4 Can write informal emails or letters using a range of idiomatic phrases. (P)		
	62.5 Can use statistical data, fractions, and percentages in an academic text. (P)		
	62.6 Can modify a sentence using adverbial clauses. (P)		
63	63.1 Can write detailed descriptions of real or imaginary places. (P)		
	63.2 Can employ simple time-saving strategies when taking notes (leaving out words, abbreviations etc. (P)		
64	64.1 Can write a concise summary of the main ideas of a longer structured text. (P)		
	64.2 Can write personal emails/letters about abstract or cultural topics (e.g music, films) (CA)		
	64.3 Can write a chronological paragraph on an academic topic. (P)		
65	65.1 Can systematically evaluate the advantages and disadvantages of various options. (P)		
	65.2 Can structure longer texts in clear, logical paragraphs. (P)		
	65.3 Can give a structured written explanation of a problem (CA)		
	65.4 Can write a letter of complaint with appropriate register, structure and conventions. (P)		
	65.5 Can develop a clear written description or narrative with relevant supporting detail and examples. (CA)		
	65.6 Can write a letter of application with appropriate register, conventions and supporting detail. (P)		
	65.7 Can express news and views effectively in writing and relate to those of others. (C)		
	65.8 Can summarise information from a presentation or lecture in their field of specialisation (P)		
	65.9 Can use correct formatting in an academic essay. (P)		
	65.10 Can explain a term in a text using synonyms, definitions or examples. (P)		
	65.11 Can signal cause and effect relationships when writing an academic text by using discourse markers. (P)		

66	66.1 Can show the relationship between an opinion and a counter argument in a discursive text. (P)			
	66.2 Can compare and evaluate ideas in a structured and logical text. (P)			
	66.3 Can adopt a level of formality appropriate to the circumstances. (N2000)			
	66.4 Can write an accurate summary of an essay or article on a familiar topic. (P)			
	66.5 Can write engaging headlines or titles to capture a reader's attention. (P)			
	66.6 Can use fact and opinion effectively in writing. (P)			
	66.7 Can correct errors in a piece of academic writing if helped to identify them first. (P)			
	66.8 Can write bullet points to summarise key points in an academic text. (P)			
	GSE 67-75/ B2(+)			
B2+	67	67.1 Can write relevant subheadings to structure longer more complex texts. (P)		
		67.2 Can systematically develop an argument giving the reasons for or against a point of view. (P)		
		67.3 Can write about feelings and the personal significance of experiences in detail. (CA)		
		67.4 Can write clear, detailed descriptions on a variety of subjects related to their field of interest. (C)		
		67.5 Can write a forceful but polite letter of complaint, including supporting details and a statement of the desired outcome. (P)		
		67.6 Can synthesise and evaluate familiar information and arguments from a number of sources. (CA)		
		67.7 Can write effective and appropriate paragraphs in a range of genres. (P)		
		67.8 Can write an effective and informative summary. (P)		
		67.9 Can write a response to an academic text giving opinions. (P)		
		67.10 Can use descriptive language to support a main idea in written academic work. (P)		
68	68	68.1 Can seek confirmation of information and ideas on abstract and concrete topics. (CA)		
		68.2 Can respond in writing to other people's arguments in an appropriate style (N2007A)		
		68.3 Can write clear and precise emails intended to create rapport and put the addressee at ease. (CA)		
		68.4 Can take notes while researching an unfamiliar topic. (P)		
		68.5 Can summarise sources and data appropriately. (P)		
		68.6 Can support a line of argument in an academic text using direct quotes. (P)		
		68.7 Can embed quotations and paraphrases in written academic work, if provided with a model. (P)		

69	69 Can write a structured review of a film, book or play with some references and examples. (CA)		
70	70.1 Can express the same idea with a different level of formality appropriate to the audience. (P)		
	70.2 Can convey information and ideas on abstract and concrete topics. (CA)		
	70.3 Can describe the sequence in a process when writing a linguistically complex text, using common discourse markers. (P)		
	70.4 Can start an essay with a strong thesis statement. (P)		
	70.5 Can use parallel structure in academic writing. (P)		
	70.6 Can clarify a concept in a narrative text using adverbial clauses, phrases and modifiers. (P)		
	70.7 Can write a critical analysis of a simple academic text. (P)		
71	71.1 Can demonstrate understanding of structure and conventions of different written genres. (CA)		
	71.2 Can introduce and reference sources in written academic work (P)		
	71.3 Can support a line of argument in an academic text using indirect quotes. (P)		
72	72.1 Can check and correct spelling, punctuation and grammar mistakes in long written texts. (P)		
	72.2 Can write a detailed, reasoned argument for or against a case. (N2007A)		
	72.3 Can structure an essay to incorporate counterarguments. (P)		
74	74.1 Can structure longer complex texts using a range of cohesive devices. (P)		
	74.2 Can write a detailed description of a complex process. (P)		
	74.3 Can write a bibliography showing an understanding of conventions. (P)		
	74.4 Can take notes on a linguistically complex presentation or lecture in their field of specialisation. (P)		
	74.5 Can take notes on a linguistically complex academic text. (P)		
	74.6 Can prepare a linguistically complex questionnaire in order to gather data. (P)		
	74.7 Can identify and correct errors in a piece of academic writing. (P)		
75	75.1 Can develop an argument with appropriate highlighting of significant points and relevant supporting detail (CA)		
	75.2 Can write an effective and informative abstract. (P)		
	75.3 Can use information from academic texts to support a thesis. (P)		
	75.4 Can synthesise information from two or more academic texts. (P)		
	75.5 Can refute a counterargument in written work. (P)		

		75.6 Can write a linguistically complex essay in response to a specific question. (P)		
		75.7 Can use appropriate tone and register when writing academic texts. (P)		
		75.8 Can use a range of idiomatic phrases as part of a structured text. (P)		
		75.9 Can support a thesis in an academic text using statistical examples. (P)		
		GSE 76-84/ C1		
C1	76	76.1 Can write a clear summary of a complex factual text, maintaining its original tone and message. (P)		
		76.2 Can make jokes in writing using words with similar spelling but different meanings. (P)		
		76.3 Can use complex numerical values in an academic text and explain their significance to the reader. (P)		
		76.4 Can use citations effectively and appropriately in an academic paper. (P)		
		76.5 Can proofread their own academic work and make corrections. (P)		
		76.6 Can clarify a concept in an academic text using adverbial clauses, phrases or modifiers. (P)		
		76.7 Can paraphrase the main idea of a linguistically complex quotation to support a line of argument. (P)		
	77	77.1 Can write an accurate summary of a complex, discursive text. (P)		
		77.2 Can write a detailed critical review of cultural events (e.g plays, films, concerts) or literary works (P)		
		77.3 Can express themselves fluently in writing, adapting the level of formality to the context. (P)		
		77.4 Can summarise information from a linguistically complex presentation or lecture. (P)		
	78	78.1 Can write effective and appropriate essays in a range of genres (P)		
		78.2 Can use persuasive techniques appropriately in academic writing (P)		
		78.3 Can write effective conclusions to a range of academic essays. (P)		
		78.4 Can embed quotations and paraphrases appropriately in written work. (P)		
	79	79.1 Can write the introduction to a thesis or dissertation. (P)		
		79.2 Can summarise research methodology and findings appropriately in a report. (P)		
		79.3 Can support a thesis in an academic text using research findings. (P)		
	80	80 Can write a review of a linguistically complex text, using appropriate conventions. (P)		
	81	81.1 Can confidently argue a case in writing, specifying needs and objectives precisely and justifying them as necessary. (N2007A)		
81.2 Can write a linguistically complex discursive essay (P)				

		81.3 Can write the conclusion to a thesis or dissertation. (P)		
	82	82.1 Can smoothly switch between a range of writing styles to address specific audiences and topics in a personal way (P)		
		82.2 Can use linguistically complex vocabulary and structures to enhance the impact in academic reports and essays. (CA)		
	83	83.1 Can write an effective and detailed research proposal. (P)		
		83.2 Can write a detailed account of an experiment, including theoretical background, findings, and conclusions. (CA)		
	84	84 Can write a transcript of a linguistically complex interview. (P)		
		GSE 85-90/ C2		
C2	85	85.1 Can create well-structured complex texts with underlying inferred meaning (CA)		
		85.2 Can edit and improve a linguistically complex text. (P)		
		85.3 Can write an effective and correctly formatted research paper (P)		
	86	86 Can write a critical analysis of a linguistically complex academic text. (CA)		
	87	87.1 Can enhance a text using figurative language such as onomatopoeia, alliteration or hyperbole. (P)		
		87.2 Can write an academic article for publication in their field. (CA)		
89	89 Can write a research report including detailed analysis and evaluation of own and others' work on the topic of investigation. (CA)			
		ORIGEN		
	(C)	Common European Framework descriptor, verbatim, © Council of Europe		
	(CA)	Common European Framework descriptor, adapted or edited, © Council of Europe		
	(N2000)	North (2000) descriptor, verbatim		
	(N2000A)	North (2000) descriptor, adapted or edited		
	(N2007A)	North (2007) expanded set of C1 and C2 descriptors, adapted or edited		
	(P)	New Pearson descriptor		

Anexo 3

Ejemplo de llenado de tabla de indicadores

MCER	GSE	READING	Habilidad	Observaciones
		GSE 10-21/ Below A1	0-1	I consider most of the skills irrelevant only because they are basic; it would be impossible to have any of the other relevant skills without having mastered these ones.
A1-	10	10.1 Can recognise cardinal numbers up to 10. (P)	0	
		10.2 Can recognise the letters of the alphabet (P)	0	
	12	12 Can read and understand simple prices. (P)	0	
	14	14.1 Can identify very common food and drink on a menu. (P)	0	
		14.2 Can recognise basic plural for ms of nouns (e.g cars, books). (P)	0	
15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)	0		
		GSE 22-29/ A1		
A1	26	26 Can follow short. simple written directions (e.g.to go from X to Y). (C)	0	
	27	27 Can understand short written notices. signs and instructions with visual support. (P)	0	
		GSE 30-35/ A2		
A2	31	31.1 Can understand simple questions in questionnaires on familiar topics. (P)	0	
		31.2 Can find specific. predictable information in everyday materials (e.g. menus, timetables) (CA)	0	
		31.3 Can understand short. simple messages on postcards, emails and social networks. (CA)	0	

Anexo 4

Ejemplo de cribado de indicadores entre expertos

GSE	READING	Código	Habilidad 2	Habilidad 3	Pearson AE	Total
	GSE 10-21/ Below A1		0-1	0-1	0-1	
	10.1 Can recognise cardinal numbers up to 10. (P)	0	1	0	1	0 1
10	10.2 Can recognise the letters of the alphabet (P)	0	1	0	1	0 1
12	12 Can read and understand simple prices. (P)	0	1	0	1	0 1
	14.1 Can identify very common food and drink on a menu. (P)	0	1	0	1	0 1
14	14.2 Can recognise basic plural for ms of nouns (e.g cars, books). (P)	0	1	0	1	0 1
15	15 Can recognise familiar names, words and very basic phrases on simple notices. (CA)	0	1	0	1	0 1
	GSE 22-29/ A1					0 0
26	26 Can follow short. simple written directions (e.g.to go from X to Y). (C)	0	1	0	1	0 1
27	27 Can understand short written notices. signs and instructions with visual support. (P)	0	1	0	1	0 1
	GSE 30-35/ A2					0 0
	31.1 Can understand simple questions in questionnaires on familiar topics. (P)	0	1	0	1	0 1
	31.2 Can find specific. predictable information in everyday materials (e.g. menus, timetables) (CA)	0	1	0	1	0 1
31	31.3 Can understand short. simple messages on postcards, emails and social networks. (CA)	0	1	0	0	0 0

Anexo 5

Tabla de ítems y taxonomía

<i>Habilidad</i>	<i>MCE R</i>	<i>Pearson</i>	<i>Descriptor</i>	<i>Nivel taxonómico</i> *	<i>Nº ítems</i>	<i>Tipo ítem</i>
<i>Lectura</i>	<i>B1</i>	45	45.3 Can skim a simple text to identify key concepts. (P)	Recordar		
		48	48.5 Can identify key information in a simple academic text, if guided by questions. (P)	Recordar		
		49	49.2 Can scan a simple academic text to find specific information. (P)	Recordar		
		50	50.2 Can understand the main idea of a passage using textual clues. (P)	Comprender		
	<i>B1+</i>	52	52.2 Can recognise the organizational structure of a paragraph in a simple academic text. (P)	Recordar		
		54	54.3 Can distinguish between fact and opinion in a simple academic text. (P)	Analizar		
			54.4 Can distinguish between the main idea and related ideas in a simple academic text in order to answer specific questions. (P)	Analizar		
		58	58.4 Can synthesize information from two or more basic texts, if guided by questions. (P)	Comprender		
	58.5 Can understand relationships between ideas in a simple academic text, if guided by questions. (P)		Comprender			
	<i>B2</i>	<i>61</i>	61.5 Can critically evaluate the effectiveness of a simple problem-solution essay. (P)	Evaluar		
	<i>B2+</i>	68	68.5 Can identify common features of an academic abstract. (P)	Recordar		
			68.8 Can distinguish between active and passive voice in a linguistically complex academic text. (P)	Analizar		
		70	70.3 Can distinguish between the main idea and related ideas in a linguistically complex academic text. (P)	Analizar		
			70.4 Can evaluate information in an academic text using specific criteria. (P)	Evaluar		
<i>C1</i>	71	71.3 Can predict the content of a linguistically complex academic text by reading introductory and summary statements. (P)	Comprender			
	74	74.2 Can understand cause and effect relationships in a linguistically complex academic text. (P)	Analizar			

			74.7 Can distinguish between different viewpoints in a linguistically complex academic text (P)	Analizar
		77	77.1 Can identify logical flaws in an argument in an academic paper. (P)	Recordar
			77.3 Can follow abstract argumentation, for example the balancing of alternatives and the drawing of a conclusion. (CA)	Aplicar
		79	79.2 Can research a topic by reading linguistically complex academic texts. (P)	Aplicar
			79.5 Can compare the presentation of a key concept in different texts by different authors using different styles of writing. (P)	Analizar
			85.2 Can understand complex arguments in technical or academic journals. (CA)	Comprender
<i>Escritura</i>	<i>B1</i>	48	48.5 Can prepare a simple questionnaire in order to gather data. (P)	Crear
		50	50.8 Can paraphrase information from a simple text, if guided by questions. (P)	Comprender
	<i>B1+</i>	55	55.9 Can summarize simple research findings in an academic text, if provided with a model summary. (P)	Comprender
		56	56.3 Can write a short, simple academic essay on a familiar topic, given a model. (P)	Crear
		58	58.6 Can write a detailed description of a simple process. (P)	Crear
			58.8 Can write a chronological paragraph on an academic topic, if provided with a model. (P)	Crear
			58.9 Can write a conclusion to a simple academic essay. (P)	Crear
	<i>B2</i>	59	59.10 Can write an introduction to a simple academic essay. (P)	Crear
		61	61.7 Can signal that two ideas are similar when writing a simple academic text by using discourse markers. (P)	Analizar
			61.8 Can contrast two ideas when writing a simple academic text by using discourse markers. (P)	Analizar
		64	64.3 Can write a chronological paragraph on an academic topic. (P)	Crear
		65	65.10 Can explain a term in a text using synonyms, definitions or examples. (P)	Comprender
		66	66.6 Can use fact and opinion effectively in writing. (P)	Aplicar
			66.7 Can correct errors in a piece of academic writing if helped to identify them first. (P)	Evaluar
	<i>B2+</i>	67	67.10 Can use descriptive language to support a main idea in written academic	Aplicar

		work. (P)	
	70	70.4 Can start an essay with a strong thesis statement. (P)	Crear
		70.5 Can use parallel structure in academic writing. (P)	Aplicar
		70.6 Can clarify a concept in a narrative text using adverbial clauses, phrases and modifiers. (P)	Comprender
		70.7 Can write a critical analysis of a simple academic text. (P)	Crear
	74	74.3 Can write a bibliography showing an understanding of conventions. (P)	Crear
		74.6 Can prepare a linguistically complex questionnaire in order to gather data. (P)	Analizar
		74.7 Can identify and correct errors in a piece of academic writing. (P)	Evaluar
	75	75.2 Can write an effective and informative abstract. (P)	Crear
		75.3 Can use information from academic texts to support a thesis. (P)	Aplicar
		75.4 Can synthesize information from two or more academic texts. (P)	Comprender
		75.5 Can refute a counter argument in written work. (P)	Analizar
<i>CI</i>	76	76.3 Can use complex numerical values in an academic text and explain their significance to the reader. (P)	Aplicar
		76.4 Can use citations effectively and appropriately in an academic paper. (P)	Aplicar
		76.5 Can proofread their own academic work and make corrections. (P)	Analizar
	78	78.2 Can use persuasive techniques appropriately in academic writing (P)	Evaluar
	79	79.1 Can write the introduction to a thesis or dissertation. (P)	Crear
		79.2 Can summarize research methodology and findings appropriately in a report. (P)	Comprender
		79.3 Can support a thesis in an academic text using research findings. (P)	Analizar
	81	81.3 Can write the conclusion to a thesis or dissertation. (P)	Crear
	83	83.1 Can write an effective and detailed research proposal. (P)	Crear
		83.2 Can write a detailed account of an experiment, including theoretical background, findings, and conclusions. (CA)	Crear

C2	85	85.3 Can write an effective and correctly formatted research paper (P)	Crear
	86	86 Can write a critical analysis of a linguistically complex academic text. (CA)	Crear
	87	87.2 Can write an academic article for publication in their field. (CA)	Crear
	89	89 Can write a research report including detailed analysis and evaluation of own and others' work on the topic of investigation. (CA)	Crear

*Taxonomía de Anderson y Krathwohl (2001)

Origen del descriptor

(CA) Common European Framework descriptor, adapted or edited, © Council of Europe

(P) New Pearson descriptor

Anexo 6

Formato para la especificación de ítems e ítems muestra del examen de inglés con propósito de publicación científica en posgrado.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se evaluará en qué medida el estudiante puede inferir el contenido de un texto académico complejo lingüísticamente, por lo tanto, se elaborará una especificación para producir un ítem	
	Definición de precisión:	Demanda cognitiva
	Se considerará un texto académico complejo lingüísticamente aquel que contenga las siguientes características:	Comprender
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	71.3 Can predict the content of a linguistically complex academic text by reading introductory and summary statements	<ul style="list-style-type: none"> Se presentará un <i>abstract</i> de no más de 10 renglones (Arial,10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:	Se le presentará los términos temática o contenido general, y se le solicitará que los identifique. O en su caso se le presentará la temática o contenido general de un texto y se le solicitará que identifique al <i>abstract</i> que corresponde de entre diferentes textos.	
• Especificación de la respuesta correcta a.	Será la temática y contenido general del texto que corresponde al <i>abstract</i> . O en su defecto será el <i>abstract</i> que corresponde a la temática presentada.	
• Especificación del distractor b.	Será una temática o contenido del texto que no corresponde al tema principal pero que esté relacionado, o bien será un <i>abstract</i> que no corresponde a la temática y contenido general del texto.	
• Especificación del distractor c.	Será una temática similar a la temática o contenido principal del texto, o bien un <i>abstract</i> sobre un tema similar a la temática o contenido principal.	
• Ítem muestra (ítem 1, versión 1):	<p>Plagiarism is a major problem for research. There are, however, divergent views on how to define plagiarism and on what makes plagiarism reprehensible. In this paper we explicate the concept of “plagiarism” and discuss plagiarism normatively in relation to research. We suggest that plagiarism should be understood as “someone using someone else’s intellectual product (such as texts, ideas, or results), thereby implying that it is their own” and argue that this is an adequate and fruitful definition. We discuss a number of circumstances that make plagiarism more or less grave and the plagiariser more or less blameworthy. As a result of our normative analysis, we suggest that what makes plagiarism reprehensible as such is that it distorts scientific credit. In addition, intentional plagiarism involves dishonesty. There are, furthermore, a number of potentially negative consequences of plagiarism.</p> <p>What is the abstract about?</p> <p>a. Plagiarism in research</p> <p>b. Scientific credit</p>	
	c.	Plagiarism as a research topic

Información contextual								
Habilidad	Intención evaluativa:							
Lectura	Se evaluará en qué medida el estudiante puede identificar conceptos clave por medio de la lectura rápida de un texto, por lo tanto, se elaborará una especificación para producir un ítem							
	Definición de precisión:	Demanda cognitiva						
	Se considera al concepto de <i>skim</i> a la acción de hojear o de hacer lectura rápida.	Comprender						
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):						
Ítem 1	45.3 Can skim a simple text to identify key concepts. (P)	<ul style="list-style-type: none"> Se presentará un párrafo introductorio sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que haga una lectura rápida del texto y responda lo que se le pregunta. 						
• Especificación de la base del ítem:		Se le presentará el término de Key concepts, y se le solicitará que determine aquellos que están relacionados con el texto (de no más de 6 renglones, Arial 10). O en su caso se le presentará los conceptos clave y se le solicitará que identifique el párrafo a los que corresponde.						
• Especificación de la respuesta correcta a.		Serán los conceptos clave que corresponden al párrafo. O en su defecto será el párrafo que corresponde a los conceptos clave.						
• Especificación del distractor b.		Serán conceptos clave que no correspondan al párrafo, o bien será un párrafo que no corresponda a los conceptos clave.						
• Especificación del distractor c.		Será conceptos clave relativamente asociados al párrafo, o bien un párrafo relativamente asociado a los conceptos clave.						
• Ítem muestra (ítem 1, versión 1):		<p>Plagiarism is a well-known and growing issue in the academic world. It is estimated to make up a substantial part of the total number of serious deviations from good research practice (Titus et al. 2008; Vitse and Poland 2012). For some journals it is indeed a serious problem, with up to a third of the published papers containing plagiarism (Zhang 2010; Bazđaric et al. 2012; Butler 2010). Given that plagiarism is perceived as a considerable problem for the research community, spelling out in some detail what is to count as plagiarism becomes a matter of pressing concern.</p> <p>Which of the following key concepts are related to the text?</p> <table border="1"> <tr> <td>a.</td> <td>Plagiarism, definition, journals</td> </tr> <tr> <td>b.</td> <td>Originality, software, writing issues</td> </tr> <tr> <td>c.</td> <td>Scientific credit, academic world, scientific research</td> </tr> </table>	a.	Plagiarism, definition, journals	b.	Originality, software, writing issues	c.	Scientific credit, academic world, scientific research
a.	Plagiarism, definition, journals							
b.	Originality, software, writing issues							
c.	Scientific credit, academic world, scientific research							

Información contextual								
Habilidad	Intención evaluativa:							
Lectura	Se identificará en qué medida el estudiante puede identificar la información clave en un texto académico simple, si es guiado por preguntas.							
	Definición de precisión:	Demanda cognitiva						
		Recordar						
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):						
	48.5 Can identify key information in a simple academic text, if guided by questions. (P)	•Se presentará un párrafo introductorio de no más de 10 renglones (Arial, 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.						
• Especificación de la base del ítem:	Se le presentará el término <i>key information</i> , y se le solicitará que la identifique. O en su caso se le presentará la información clave y se le solicitará que identifique el párrafo al que corresponde.							
• Especificación de la respuesta correcta a .	Serán la información clave que corresponda al párrafo. O en su defecto será el párrafo que corresponde a la información clave.							
• Especificación del distractor b .	Será información que no corresponda al párrafo. O en su defecto será un párrafo que no corresponde a la información clave.							
• Especificación del distractor c .	Será la información clave relativamente asociada al párrafo. O en su defecto será un párrafo relativamente asociado a la información clave.							
• Ítem muestra (ítem 1, versión 1):	<p>In this paper we will mainly do two things. First, we will explicate the concept of “plagiarism”, i.e. present an analysis of the concept aimed at further clarifying it. This means that we will look at previous uses of the term and through critical analysis come up with what we take to be an improved definition. The point of the definition that we present is not to identify the essence or ‘real nature’ of plagiarism (we doubt that there is such a thing), but rather to extract one that is useful for the purpose of clarifying normative issues related to plagiarism, while being true to common uses of the term. Second, we will discuss plagiarism normatively, by taking a closer look at different aspects of it. We restrict our analysis to the context of research, since plagiarism in the arts, for instance, raise a partly different set of issues, and includes partly different normative intuitions, which would require a separate analysis.</p> <p>Which is the key information of the paragraph?</p> <table border="1"> <tr> <td>a.</td> <td>The concept of plagiarism and its normative implications.</td> </tr> <tr> <td>b.</td> <td>The concept of plagiarism, appropriation and not giving proper credit.</td> </tr> <tr> <td>c.</td> <td>The concept of plagiarism and its ethic dilemmas scientific research.</td> </tr> </table>		a.	The concept of plagiarism and its normative implications.	b.	The concept of plagiarism, appropriation and not giving proper credit.	c.	The concept of plagiarism and its ethic dilemmas scientific research.
a.	The concept of plagiarism and its normative implications.							
b.	The concept of plagiarism, appropriation and not giving proper credit.							
c.	The concept of plagiarism and its ethic dilemmas scientific research.							

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para encontrar información específica dentro de un texto académico.	
	Definición de precisión:	Demanda cognitiva
		Recordar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	49.2 Can scan a simple academic text to find specific information. (P)	•Se presentará un párrafo introductorio de no más de 6 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:		Se le presentará una información específica, y se le solicitará identificar a qué se refieren los autores al respecto. O en su caso se le presentarán las ideas que refieren los autores y se le solicitará identificar la información específica que corresponde.
• Especificación de la respuesta correcta a.		Serán la información específica que corresponde a lo mencionado por los autores en el texto.
• Especificación del distractor b.		Será información presentada en el texto, pero no se refiere a la información específica solicitada.
• Especificación del distractor c.		Será información relacionada con la información específica.
• Ítem muestra (ítem 1, versión 1):		The conditions of adequacy should identify relevant restrictions on any suggested definition for the definition to be reasonably adequate for the intended purpose in the intended context. Partly different criteria may become relevant depending on the intended use of the definition. We suggest, inspired by Bruñide and Tengland (2003), that the following criteria for adequacy are relevant to a definition of "plagiarism" for our intended use: fitting language use, precision, Reliability, theoretical fruitfulness, relevance for normative purposes, simplicity.
		Which criteria do the authors suggest for defining plagiarism?
		a. Fitting language use, precision, reliability, theoretical fruitfulness, relevance for normative purposes, simplicity.
		b. Stealing, copying, appropriation of material, inappropriate use, pretending, not giving credit
		c. Adequacy, definition, intellectual property, normative, ideas and work, self- plagiarism, originality,

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para entender la idea principal de un pasaje utilizando pistas textuales.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	50.2 Can understand the main idea of a passage using textual clues. (P)	•Se presentará un párrafo de no más de 6 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:		Se le presentará el término de idea principal y se le solicitará que la encuentre a partir de un texto dado. O en su caso se le presentará la idea principal y se le solicitará que identifique el párrafo al que corresponde.
• Especificación de la respuesta correcta a.		Será la idea principal que corresponde al texto. O en su defecto será el párrafo que corresponde a la idea principal.
• Especificación del distractor b.		Será una idea secundaria. O en su defecto será el texto que no corresponde a la idea principal.
• Especificación del distractor c.		Será una idea relacionada con la principal. O en su defecto será un párrafo relativamente asociado a la idea principal.
• Ítem muestra (ítem 1, versión 1):		Since it is important to determine what constitutes misconduct in scientific writing, and “plagiarism” is a much used concept in discussions of scientific misconduct, one could perhaps expect agreement and a fairly high level of precision regarding what constitutes plagiarism. However, while there is agreement about paradigmatic cases of plagiarism, there is less agreement regarding how plagiarism should be defined. In fact, the issue is rarely discussed in detail.
		What is the main idea of the paragraph?
a.		The concept of plagiarism is used to refer to misconduct in the scientific community, however there isn't agreement in the definition.
b.		The technology helps to identify plagiarism in the research community
c.		Plagiarism is a much used concept than misconduct in scientific writing.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para reconocer la estructura organizacional de un párrafo en un texto académico simple	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	52.2 Can recognise the organizational structure of a paragraph in a simple academic text. (P)	<ul style="list-style-type: none"> • Se presentará un párrafo de no más de 6 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:	Se le presentará el término de estructura organizacional, y se le solicitará que reconozca la que corresponde a la parte del texto que se le presenta. O bien, se le presentará la estructura organizacional de un texto, y se le solicitará que reconozca a qué párrafo corresponde.	
• Especificación de la respuesta correcta a.	Será la estructura que corresponde al texto. O en su defecto, será el texto que corresponde con la estructura organizacional dada.	
• Especificación del distractor b.	Será una estructura organizacional que corresponde a otro texto. O en su defecto, será un texto que guarda una estructura similar.	
• Especificación del distractor c.	Será una estructura organizacional similar. O en su defecto, será un texto que no corresponde con la estructura dada.	
• Ítem muestra (ítem 1, versión 1):	Let us first ask what it means to appropriate someone else's work. In some definitions, plagiarism is characterised as stealing. However, if plagiarism by definition concerns stealing, then it is not theft in the traditional sense of taking a thing, where if person A takes it from person B, then B will no longer have it. What is appropriated in such instances of plagiarism is intellectual property, as when people download copyright-protected films or music from the Internet. Thus, to the extent that plagiarism is theft, it is stealing someone else's intellectual work by copying. Related to research papers, it is about copying another's text, tables, graphs, or pictures into one's own paper without having permission to do so (and with certain pretense, a point we shall be returning to presently).	
	What's the organizational structure of the paragraph?	
	a.	Topic sentence, supporting detail (fact), supporting detail (example), supporting detail (fact), concluding sentence.
	b.	Topic sentence, supporting detail (example), supporting detail (fact), supporting detail (example), supporting detail (fact), concluding sentence.
c.	Topic sentence, supporting detail (fact), supporting detail (definition), supporting detail (fact), concluding sentence.	

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para distinguir entre un hecho y una opinión en un texto académico simple.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	54.3 Can distinguish between fact and opinion in a simple academic text. (P)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 6 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:		Se le presentará el término <i>opinión</i> , y se le solicitará que seleccione la que corresponde al texto dado. O en su caso, se le presentará el término <i>hecho</i> , y se le solicitará que lo distinga de un texto cuyo contenido tiene varias opiniones.
• Especificación de la respuesta correcta a.		Será la <i>opinión</i> que corresponda al texto presentado. O en su caso, el <i>hecho</i> que corresponda al texto presentado.
• Especificación del distractor b.		Será un <i>hecho</i> que corresponde al texto. O en su caso será una <i>opinión</i> que corresponde al texto.
• Especificación del distractor c.		Será una <i>opinión</i> que no corresponde al texto presentado. O en su caso será un <i>hecho</i> que no corresponde al texto.
• Ítem muestra (ítem 1, versión 1):		<p>We are, however, disinclined to include stealing in our definition. Although one may steal intellectual as well as non-intellectual property, and even talk about “theft of the recognition due to the original contributor” (Rathod 2012), talking about plagiarism as stealing is nevertheless misguided, at least as part of a definition. This is so because using someone else’s text, say, and passing it off as one’s own can be done regardless of whether one steals the text or not. One can do it by finding the text in a journal or book or by using an unpublished paper.</p> <p>What is the author's opinion about plagiarism?</p> <p>a. Stealing doesn’t belong to the concept of plagiarism.</p> <p>b. Plagiarism occurs even if you imply that you create it.</p> <p>c. Making statements about the results of research should be considered plagiarism.</p>

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para distinguir entre la idea principal y las secundarias, en un texto académico simple con el fin de responder a preguntas específicas.	
	Definición de precisión:	Demanda cognitiva
		Analizar
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	54.4 Can distinguish between the main idea and related ideas in a simple academic text in order to answer specific questions. (P)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 6 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presenten. -
• Especificación de la base del ítem:	Se le presentará el término de idea principal, y se le solicitará que la extraiga de un texto dado. O en su caso, se le presentará la idea principal de un texto, y se le solicitará que reconozca el texto que le corresponda.	
• Especificación de la respuesta correcta a.	Será la idea principal del texto. O en su caso, será el texto que corresponde a la idea principal.	
• Especificación del distractor b.	Será una idea secundaria. O en su caso, un texto que contiene una idea secundaria.	
• Especificación del distractor c.	Será una idea que no esté relacionada con el texto. O en su caso, un texto relacionado con el tema.	
• Ítem muestra (ítem 1, versión 1):	<p>It may still be argued that there is a sense of “stealing” that concerns appropriating someone else’s intellectual work and passing it off as one’s own. In this sense you may steal someone’s song if you play it and claim to have composed it yourself. This seems to mean that there is a sense of “stealing” that is equivalent to “plagiarizing”. If so, this second understanding of “stealing”, which is distinct from the one discussed above, cannot contribute anything to a definition of “plagiarism”. The conclusion remains: stealing, or theft, cannot be used as part of the definition of “plagiarism”.</p> <p>What is the main idea of the paragraph?</p> <p>a. Stealing should not be part of the definition of plagiarism.</p> <p>b. One can steal the ideas and pass them on to our own.</p> <p>c. “Steal”, “appropriate” and “use” are words in the plagiarism definition.</p>	

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para sintetizar la información de un texto académico.	
	Definición de precisión:	Demanda cognitiva
		Comprender
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	58.4 Can synthesize information from the parts of academic texts, if guided by questions. (P)	<ul style="list-style-type: none"> Se presentará un texto académico de no más de 7 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:	Se le presentará el término de síntesis, y se le solicitará que determine la mejor. O en su caso, se le presentará una síntesis, y se le solicitará que determine el término al que se refiere.	
• Especificación de la respuesta correcta a.	Será la mejor síntesis del texto que se le presenta. O en su caso, ser el término síntesis.	
• Especificación del distractor b.	Serán una síntesis de una parte del extracto de texto académico. O en su caso, será un término relacionado síntesis.	
• Especificación del distractor c.	Será una síntesis que no corresponde al extracto del texto académico. O en su caso, un término subordinado al de síntesis.	
• Ítem muestra (ítem 1, versión 1):	It is no accident that plagiarism is discussed in relation to research, although it is also clearly relevant in relation to music, literature, art, and design, since it relates to using the product of someone else's intellectual work while passing it off as one's own. Note that there is no reason to restrict the use of the term to published work, since you may use someone else's work while passing it off as your own even if it is not published. For instance, you may do it by first stealing the manuscript from the author, by using passages from an unpublished manuscript circulated at a seminar, or by using ideas communicated at a lecture.	
	Which is the best synthesis of the text?	
	a.	Plagiarism is also relevant in relation to music, literature, art, and design and there is no reason to restrict the use of the term to published work.
	b.	Plagiarism can be relevant to other fields like music, literature and art. Plagiarism can happen when you try to take an unpublished manuscript and you use it.
c.	Plagiarism is not restricted to published works and can be with or without the intention of do it.	

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para entender las relaciones entre ideas en un texto académico simple.	
	Definición de precisión:	Demanda cognitiva
		Comprender
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	58.5 Can understand relationships between ideas in a simple academic text, if guided by questions. (P)	•Se presentará un par de ideas sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:	Se le presentarán dos ideas extraídas de un texto académico, y se le solicitará que encuentre la relación entre ambas. O en su caso, se le presentará un texto con varias ideas, y se solicitará que explique la relación entre ellas.	
• Especificación de la respuesta correcta a.	Será la relación que guardan las ideas presentadas. O en su caso, una explicación de la relación de ideas.	
• Especificación del distractor b.	Serán una relación opuesta a las ideas presentadas. O en su caso, una explicación de una relación similar a las ideas presentadas.	
• Especificación del distractor c.	Será una relación diferente a la que guardan las ideas presentadas. O en su caso, una explicación que no guarde relación con las ideas presentadas.	
• Ítem muestra (ítem 1, versión 1):	"If someone uses another's idea and implies that it is an idea of their own, that someone is plagiarizing" and "Ideas are not always documented, but might be presented at conferences or in personal conversation, etc."	
	What's the relationship between the ideas?	
	a.	It may often be much more difficult to verify that an idea has been plagiarized compared to research results or text.
	b.	Research results are easier to plagiarize than ideas.
	c.	There is no smart software to identify if an idea has been plagiarized.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para evaluar críticamente la efectividad de un problema-solución simple.	
	Definición de precisión:	Demanda cognitiva
		Evaluar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	61.5 Can critically evaluate the effectiveness of a simple problem-solution	<ul style="list-style-type: none"> Se presentará un párrafo sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
	• Especificación de la base del ítem:	Se le presentará un problema, y se le solicitará que identifique la mejor solución. O bien, se le presentará una solución, y se le solicitará que identifique el problema.
	• Especificación de la respuesta correcta a.	Será la solución al problema planteado. O en su caso, será el problema la solución dada.
	• Especificación del distractor b.	Será una solución similar planteada por el autor. O en su caso, será un problema similar a la solución dada.
	• Especificación del distractor c.	Será una solución opuesta a la propuesta por el autor. O en su caso, será un problema que no guarde correspondencia con la solución dada.
	• Ítem muestra (ítem 1, versión 1):	According to the author of the article, what is the best solution for scientists to publish in English?
		a. They wrote their papers in their native language and had them translated by professional translators.
		b. They paid for private English classes to improve their writing and reading skills.
		c. They just published their papers in their native language.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para identificar las características comunes de un resumen académico.	
	Definición de precisión:	Demanda cognitiva
		Recordar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	68.5 Can identify common features of an academic abstract. (P)	<ul style="list-style-type: none"> Se presentará un texto sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:		Se le presentará el término de <i>abstract</i> , y se le solicitará que identifique las características más comunes que lo componen. O en su caso, presentará las características más comunes de un <i>abstract</i> , y solicitará que identifique el término.
• Especificación de la respuesta correcta a.		Serán las características más comunes de un <i>abstract</i> . O en su caso, será el término <i>abstract</i> .
• Especificación del distractor b.		Serán algunas características de un artículo científico. O en su caso, será un término referido a un documento de difusión.
• Especificación del distractor c.		Serán características que no estén relacionadas con el resumen. O en su caso, será un término referido a un documento de divulgación.
• Ítem muestra (ítem 1, versión 1):		Which are the common features of an abstract?
a.		Context, objective, method, results, conclusions.
b.		Introduction, theoretical framework, method, results, discussion.
c.		Introduction, problem statement, theory, method, Discussion

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para distinguir entre la voz pasiva y activa en un texto académico lingüísticamente complejo.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	68.8 Can distinguish between active and passive voice in a linguistically complex academic text. (P)	<ul style="list-style-type: none"> •Se presentará un párrafo sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:	Se le presentará el término de enunciado, y se le solicitará que identifique el que se encuentre la voz pasiva. O en su caso, se le presentará el término de enunciado, y se le solicitará que identifique la voz activa.	
• Especificación de la respuesta correcta a.	Será la oración que esté en voz pasiva. O en su caso, la oración que esté en voz activa.	
• Especificación del distractor b.	Será una oración que esté en voz activa. O en su caso, una oración que esté en voz pasiva.	
• Especificación del distractor c.	Será una oración que esté en voz reflexiva. O en su caso, una oración que esté en voz reflexiva.	
• Ítem muestra (ítem 1, versión 1):		
	Which of the following sentences are in passive voice?	
	a.	...researchers with insufficient skills in the English language have been encouraged by their supervisors or colleagues...
	b.	...to use another paper as a template and change data in order to include their own results instead of those in the template, with a considerable text overlap as a consequence.
c.	Regardless of what the underlying explanations are, they have no bearing on whether or not a certain act is an act of plagiarism.	

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para distinguir entre la idea principal y las secundarias, en un texto académico complejo.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	70.3 Can distinguish between the main idea and related ideas in a linguistically complex academic text. (P)	<ul style="list-style-type: none"> • Se presentará un párrafo de no más de 7 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
	• Especificación de la base del ítem:	Se le presentará el término de idea principal, y se le solicitará que la identifique. O en su caso, se le presentará el término de idea secundaria, y se le solicitará que la identifique.
	• Especificación de la respuesta correcta a .	Será la idea principal del párrafo. O en su caso, será la idea secundaria.
	• Especificación del distractor b .	Será la idea secundaria. O en su caso, la idea principal.
	• Especificación del distractor c .	Será un argumento. O en su caso, un argumento secundario.
	• Ítem muestra (ítem 1, versión 1):	<p>If something is to be considered plagiarism or not is also independent of who is plagiarized. For instance, it makes no difference if the person plagiarized is considered insignificant. It is also irrelevant to the evaluation of whether or not there is plagiarism if the plagiarized and plagiarising texts are used for different purposes, have different intended readers, or are of very different dimensions. It is still plagiarism if someone copies something from a short research paper and includes it in an extensive book. Whether or not the plagiarized text was published in an indexed, peer-reviewed journal is equally irrelevant.</p> <p>What is the main idea of the paragraph?</p> <p>a. If something is to be considered plagiarism or not is also independent of who is plagiarized</p> <p>b. It is still plagiarism if someone copies something from a short research paper and includes it in a long book.</p> <p>c. To evaluate the existence of plagiarism, it is irrelevant to consider the purposes or recipients of the plagiarized text.</p>

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para evaluar información en un texto académico usando un criterio específico.	
	Definición de precisión:	Demanda cognitiva
		Evaluar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	70.4 Can evaluate information in an academic text using specific criteria. (P).	<ul style="list-style-type: none"> Se presentará un párrafo sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
	• Especificación de la base del ítem:	Se le presentará un problema, y se le solicitará que evalúe una posible solución a partir de uno o más criterios. O bien, presentará una solución, y solicitará que identifique los criterios que se utilizaron.
	• Especificación de la respuesta correcta a.	Será la solución que corresponda a los criterios dados. O en su caso, será el criterio empleado para llegar a la solución presentada.
	• Especificación del distractor b.	Será una solución que no corresponde a los criterios dados. O en su caso, será un criterio que no corresponde a la solución presentada.
	• Especificación del distractor c.	Será una solución que aplique incorrectamente el criterio dado. O en su caso, será un criterio diferente para llegar a la solución.
	• Ítem muestra (ítem 1, versión 1):	Plagiarism = def. an instance of someone using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own.
		Based on the definition that appears in the article, what would you answer to a person who thinks that ideas should be free and not be someone's intellectual property?
		a. Ideas are an intellectual product and, therefore, recognition must be given to the person.
		b. Ideas are not an intellectual product of a single person, therefore, they must be free.
		c. Ideas are a collective product, therefore, they must be recognized by society.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para entender las relaciones de causa y efecto en textos académicos lingüísticamente complejos.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	74.2 Can understand cause and effect relationships in a linguistically complex academic text. (P)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 3 renglones (Arial, 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:		Se le presentará el término de causa-efecto, y se le solicitará que lo identifique entre varios enunciados. O bien, presentará un enunciado, y solicitará que identifique el tipo de relación que guardan los elementos presentados.
• Especificación de la respuesta correcta a.		Será el enunciado que guarda una relación de causa y efecto. O en su caso, será el término caus-efecto.
• Especificación del distractor b.		Será una idea que no guarda una relación de causa y efecto
• Especificación del distractor c.		Será una idea que no guarda una relación de causa y efecto
• Ítem muestra (ítem 1, versión 1):	Someone who is plagiarizing is always blameworthy, to the extent that readers thereby are misled about who deserves credit for the work. Still, those who mislead intentionally are more blameworthy.	
	From the previous text, what type of relationship exists between plagiarism and blaming?	
	a.	Cause-effect relationship.
	b.	Sequence relationship.
	c.	Contrast relationship.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para distinguir entre diferentes puntos de vista en un texto académico lingüísticamente complejo.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	74.7 Can distinguish between different viewpoints in a linguistically complex academic text (P)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 7 renglones (Arial, 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:	Se le presentará el término de punto de vista del autor, y se le solicitará que lo identifique. O en su caso, se presentará un punto de vista, y se le solicitará que señale el texto en el que aparece.	
• Especificación de la respuesta correcta a.	Será el punto de vista que corresponde al autor. O en su caso, será el texto que corresponde al punto de vista dado.	
• Especificación del distractor b.	Será un punto de vista que no corresponde al autor. O en su caso, será un texto diferente al que aparece el punto de vista dado.	
• Especificación del distractor c.	Será una afirmación relacionada con el punto de vista que corresponde al autor. O en su caso, será un texto que tiene un punto de vista similar al punto de vista dado.	
• Ítem muestra (ítem 1, versión 1):	<p>Even though a case of plagiarism is judged differently depending on whether the offender did it on purpose or not, one may also be held responsible for one's ignorance. Good research practice involves knowing where to draw the line between acceptable and unacceptable behaviors relating to research, such as the unacceptability of fabricating or tampering with data or results. That it is unacceptable to cut and paste from other people's published work, without stating the source and showing exactly what passages are quoted, is part of that required knowledge of good research practice. If one has not been taught this, part of the blame for plagiarism must fall on one's teachers and supervisors.</p> <p>What is the authors' point of view on the importance that teachers have on plagiarism?</p>	
	a.	Teachers are responsible for teaching good research practices to their students to avoid plagiarism.
	b.	Teachers are responsible for plagiarism committed by their students.
	c.	Teachers are responsible for teaching their students the risks of plagiarism.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para identificar errores lógicos en un argumento en un texto académico.	
	Definición de precisión:	Demanda cognitiva
		Recordar
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	77.1 Can identify logical flaws in an argument in an academic paper. (P)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 5 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:		Se le presentará un texto que contenga un tipo de falla lógica en su argumentación, y se le solicitará que identifique el tipo de falla. O en su caso, se presentará el tipo de falla lógica en la argumentación, y se solicitará que identifique el texto que la contiene.
• Especificación de la respuesta correcta a .		Será el tipo de falla lógica que se presenta en el argumento dado. O en su caso, será el texto que corresponde al tipo de falla lógica dada.
• Especificación del distractor b .		Será un tipo de falla lógica que no está presente en el argumento dado. O en su caso, será un texto que no corresponde al tipo de falla lógica dada.
• Especificación del distractor c .		Será un tipo de falla lógica diferente al presentado en la opción b. O en su caso, será un texto que contenga un tipo de falla lógica diferente a la opción b.
• Ítem muestra (ítem 1, versión 1):		Plagiarism can be more or less extensive, ranging from whole chapters of books, or entire academic papers, to shorter passages. We will argue that to the extent that quantity words matter, it depends on whether quantity has an impact on quality. This is to say that quantity matters only indirectly, while quality matters directly.
		What type of logical flaw does the argument of the following text present?
		a. Lack of adequate evidence.
		b. Lack of coherence.
		c. Lack of logical structure.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para seguir una argumentación abstracta, por ejemplo, el equilibrio de alternativas y la extracción de conclusiones.	
	Definición de precisión:	Demanda cognitiva
		Aplicar
	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	77.3 Can follow abstract argumentation, for example the balancing of alternatives and the drawing of a conclusion. (CA)	<ul style="list-style-type: none"> • Se presentará un párrafo de no más de 5 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:	Se le presentará un párrafo con una serie de argumentos, y se le solicitará que extraiga una conclusión. O en su caso, se le presentará una conclusión, y se le solicitará que identifique el texto donde se extrajo.	
• Especificación de la respuesta correcta a.	Será la conclusión que corresponde a los argumentos presentados en el texto. O en su caso, será el texto que contiene los argumentos para llegar a la conclusión dada.	
• Especificación del distractor b.	Será una conclusión que no corresponde a los argumentos presentados en el texto. O en su caso, será un texto que no contiene los argumentos para llegar a la conclusión dada.	
• Especificación del distractor c.	Será una conclusión que cumple parcialmente con los argumentos presentados en el texto. O en su caso, será un texto que contiene argumentos que llevan a una conclusión diferente.	
• Ítem muestra (ítem 1, versión 1):	We claim that plagiarism (in principle) can consist in as little as one word, while there are many standard sentences describing research methods that will not be plagiarism even if, in fact, copied from someone else. This is to say that the unmarked reuse of some very short passages might be plagiarism, even though the reuse of other equally short passages would not.	
	Following the author's argumentation, which is the conclusion of his ideas?	
	a.	The plagiarism has to do with quality rather than quantity.
	b.	The plagiarism has to do with intention rather than quantity.
c.	The plagiarism has to do with quantity rather than quality.	

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para investigar un tema mediante la lectura de textos académicos lingüísticamente complejos.	
	Definición de precisión:	Demanda cognitiva
		Aplicar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	79.2 Can research a topic by reading linguistically complex academic texts.	<ul style="list-style-type: none"> •Se presentará un párrafo sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:	Se le presentará el texto completo, y a partir de la lectura se le solicitará que seleccione el mejor tema de investigación que dé seguimiento a la lectura. O en su caso, se le presentará un tema de investigación que da continuidad a la lectura, y se le solicitará que identifique el área del conocimiento que corresponde la investigación.	
• Especificación de la respuesta correcta a.	Será el tema que da mayor continuidad a la lectura. O en su caso, será el área de conocimiento al que corresponde el tema de investigación propuesto.	
• Especificación del distractor b.	Será un tema similar al tratado en la lectura pero no relacionado cercanamente. O en su caso, será un área de conocimiento que no corresponde al tema de investigación propuesto.	
• Especificación del distractor c.	Será un tema que no esté relacionado. O en su caso, será un área de conocimiento que no corresponde el tema de investigación propuesto, pero diferente de la opción b.	
• Ítem muestra (ítem 1, versión 1):	Which of the following research topics is the best to continue the reading?	
	a.	Plagiarism in social sciences.
	b.	Ethics in the research community.
	c.	Scientific credit.

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para comparar la presentación de un concepto clave en diferentes textos por diferentes autores usando diferentes estilos de escritura.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
Ítem 1	79.5 Can compare the presentation of a key concept in different texts by different authors using different styles of writing. (P)	<ul style="list-style-type: none"> Se presentará un párrafo sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan.
• Especificación de la base del ítem:	Se le presentará un párrafo de no más de 15 renglones, y se le solicitará que identifique las palabras clave que diferencian a las definiciones.	
• Especificación de la respuesta correcta a.	Serán las palabras diferenciadoras entre definiciones.	
• Especificación del distractor b.	Serán algunas palabras diferenciadoras y algunas similares.	
• Especificación del distractor c.	Serán palabras que no sean diferenciadoras	
• Ítem muestra (ítem 1, versión 1):	<p>When the concept is explained in a recent newsletter from the US Office of Research Integrity, it looks deceptively simple: "It involves stealing someone else's work and lying about it afterward" (Sox 2012). Others prefer to speak of "copying" part of someone else's published work and using it without showing that it is borrowed from someone else. In the Longman Contemporary English Advanced Learner's Dictionary, the act of plagiarism is defined as "when someone uses another person's words, ideas, or work and pretends they are their own".</p> <p>In the scholarly definitions, the more technical notions of "appropriation" and "credit" re central: "Plagiarism is the appropriation of other people's material without giving proper credit" (The European Code of Conduct for Research Integrity); "Plagiarism is the appropriation of another person's ideas, processes, results, or words without giving appropriate credit" (US Federal Policy on Research Misconduct). So the basic ideas seem to be that someone deliberately takes someone else's work, whether in the form of an idea, a method, data, results, or text, and presents it as their own instead of giving credit to the person whose ideas, results, or words it is. This is mirrored in the definition given by Merriam-Webster: "to steal and pass off (the ideas or words of another) as one's own: use (another's production) without crediting the source".</p>	
	Which of the following key words make a difference between definitions?	
	a.	Stealing, Pass, Use, Appropriate, Credit, Pretend
	b.	Stealing, Work, Credit, Ideas, Appropriate
	c.	Use, Work, Ideas, Product, Plagiarism

Información contextual		
Habilidad	Intención evaluativa:	
Lectura	Se identificará la habilidad del estudiante para entender argumentos complejos en revistas técnicas o académicas.	
	Definición de precisión:	Demanda cognitiva
		Analizar
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	85.2 Can understand complex arguments in technical or academic journals. (CA)	<ul style="list-style-type: none"> Se presentará un párrafo de no más de 7 renglones (Arial 10) sobre un tema de interés general para los estudiantes de los diferentes posgrados; y se le solicitará que lea detenidamente el texto y responda las preguntas que se le presentan. -
• Especificación de la base del ítem:		Se le presentará un argumento, y se le solicitará que seleccione la oración que contiene el significado principal. O bien, se le presentará el significado principal, y se le solicitará que identifique el argumento.
• Especificación de la respuesta correcta a.		Será oración que contenga el argumento principal. O en su caso, será el argumento que corresponde al significado principal que se presenta.
• Especificación del distractor b.		Será una idea similar a la presentada por el autor. O en su caso, será un argumento que no corresponde al significado principal que se presenta.
• Especificación del distractor c.		Será una idea contraria a la presentada por el autor. O en su caso, será un argumento que no corresponde al significado principal que se presenta, pero diferente de la opción b.
• Ítem muestra (ítem 1, versión 1):		<p>While ordinary language use of “plagiarism” certainly allows for the act of plagiarizing being intentional, it seems clear that plagiarism does not necessarily involve an intention to deceive. We therefore would like to suggest a definition that does not require intention. The notion of “passing something off” also seems to imply intention, and therefore ought preferably to be avoided. A slightly modified definition, where we define “plagiarism” rather than “to plagiarize”, would therefore read: Plagiarism = an instance of someone using someone else’s work, thereby implying that it is their own.</p> <p>What is the main meaning of the following argument?</p> <p>a. According to the authors, plagiarism can be intentional or unintentional.</p> <p>b. According to the author, not all authors have the intention to deceive.</p> <p>c. According to the author, intention matters when it comes to plagiarize.</p>

Información contextual		
Habilidad	Intención evaluativa:	
Escritura	Se identificará la habilidad del estudiante para redactar el resumen de un artículo científico	
	Definición de precisión:	Demanda cognitiva
		Crear
Ítem 1	Foco evaluativo del ítem:	Criterios sobre la información a emplear y las indicaciones para responder este ítem (cuando es el caso):
	Can write an abstract of a research topic	-
• Especificación de la base del ítem:	Se le solicitará que redacte el resumen sobre su proyecto o intervención de acuerdo al esquema de un artículo científico	
• Especificación de la respuesta correcta a.	Coherencia, cohesión, Gramática, rango de vocabulario.	
• Ítem muestra (ítem 1, versión 1):	Write an abstract about your research project or intervention (150 a 250 words). Remember to include all the parts of an abstract: Context, objective, method, and possible results.	

Anexo 7

Primera versión de la prueba

Prueba de diagnóstico de inglés para estudiantes de posgrado con propósitos de publicación científica



2024

Presentación

El Examen diagnóstico de inglés con propósitos de publicación científica en posgrado fue creado con la finalidad de identificar las áreas de oportunidad en el idioma extranjero, y a partir de ellas crear programas, estrategias y acciones de mejora institucionales. A partir de los resultados de la prueba se pretende aportar a la creación de intervenciones didácticas y políticas educativas, que impulsen a los estudiantes de posgrado a publicar artículos científicos en revistas de alto impacto en su disciplina. Esta prueba es de carácter diagnóstico, por lo tanto, no tiene repercusiones en sobre la calificación en alguna materia ni se usará como elemento para tomar decisiones que afecten la continuidad de tus estudios dentro del posgrado.

Esta prueba consta de dos secciones y un cuestionario de contexto. Pero antes de comenzar a responder la prueba es necesario que leas un artículo, te pedimos que lo leas con mucha atención, ya que las preguntas que estás por contestar se relacionan con dicho documento. Una vez que termines de leer el artículo, responderás un cuestionario de

contexto, para luego pasar a la primera sección; la cual consta de 22 ítems que exploran el nivel de dominio que tienes para leer y comprender textos científicos en inglés. En esta sección todos los ítems son de opción múltiple y solo tiene tres opciones de respuesta. Solo en algunos casos, se te pregunta por la mejor respuesta, y en otros tienes que seleccionar la respuesta correcta. Para hacer más fluido el ejercicio, en algunos casos se te presenta la parte del texto que corresponde a lo que se te solicita, lo cual evita que regreses a leer el artículo de nueva cuenta.

En la segunda sección consta de una sola pregunta y debes escribir en inglés lo que se te pide. Debes de tomar en cuenta que se te pide un mínimo de palabras y un máximo, te pedimos que la producción escrita se ajuste a lo solicitado y consideres tus conocimientos de metodología de la investigación para dar estructura al texto.

Si llegas a tener alguna duda acerca de la prueba o sobre cómo responder, por favor comunícate con la Lic. Liliana Parra González al correo liliana.parra@uabc.edu.mx.

Muchas gracias por tu participación.

SECTION I

1. What is the abstract about?

Plagiarism is a major problem for research. There are, however, divergent views on how to define plagiarism and on what makes plagiarism reprehensible. In this paper we explicate the concept of “plagiarism” and discuss plagiarism normatively in relation to research. We suggest that plagiarism should be understood as “someone using someone else’s intellectual product (such as texts, ideas, or results), thereby implying that it is their own” and argue that this is an adequate and fruitful definition. We discuss a number of circumstances that make plagiarism more or less grave and the plagiariser more or less blameworthy. As a result of our normative analysis, we suggest that what makes plagiarism reprehensible as such is that it distorts scientific credit. In addition, intentional plagiarism involves dishonesty. There are, furthermore, a number of potentially negative consequences of plagiarism.

a. Plagiarism in research

- b. Scientific credit
- c. Plagiarism as a research topic

2. Which of the following key concepts are related to the text?

Plagiarism is a well-known and growing issue in the academic world. It is estimated to make up a substantial part of the total number of serious deviations from good research practice (Titus et al. 2008; Vitse and Poland 2012). For some journals it is indeed a serious problem, with up to a third of the published papers containing plagiarism (Zhang 2010; Baždaric et al. 2012; Butler 2010). Given that plagiarism is perceived as a considerable problem for the research community, spelling out in some detail what is to count as plagiarism becomes a matter of pressing concern.

- a. Plagiarism, definition, journals
- b. Originality, software, writing issues
- c. Scientific credit, academic world, scientific research

3. Which is the key information of the paragraph?

In this paper we will mainly do two things. First, we will explicate the concept of “plagiarism”, i.e. present an analysis of the concept aimed at further clarifying it. This means that we will look at previous uses of the term and through critical analysis come up with what we take to be an improved definition. The point of the definition that we present is not to identify the essence or ‘real nature’ of plagiarism (we doubt that there is such a thing), but rather to extract one that is useful for the purpose of clarifying normative issues related to plagiarism, while being true to common uses of the term. Second, we will discuss plagiarism normatively, by taking a closer look at different aspects of it. We restrict our analysis to the context of research, since plagiarism in the arts, for instance, raise a partly different set of issues, and includes partly different normative intuitions, which would require a separate analysis.

- a. The concept of plagiarism and its normative implications.
- b. The concept of plagiarism, appropriation and not giving proper credit.
- c. The concept of plagiarism and its ethic dilemmas scientific research.

4. Which criteria do the authors suggest for defining plagiarism?

The conditions of adequacy should identify relevant restrictions on any suggested definition for the definition to be reasonably adequate for the intended purpose in the intended context. Partly different criteria may become relevant depending on the intended use of the definition. We suggest, inspired by Brülde and Tengland (2003), that the following criteria for adequacy are relevant to a definition of “plagiarism” for our intended use: fitting language

use, precision, Reliability, theoretical fruitfulness, relevance for normative purposes, simplicity.

- a. Fitting language use, precision, reliability, theoretical fruitfulness, relevance for normative purposes, simplicity.
- b. Stealing, copying, appropriation of material, inappropriate use, pretending, not giving credit
- c. Adequacy, definition, intellectual property, normative, ideas and work, self- plagiarism, originality

5. What is the main idea of the paragraph?

Since it is important to determine what constitutes misconduct in scientific writing, and “plagiarism” is a much used concept in discussions of scientific misconduct, one could perhaps expect agreement and a fairly high level of precision regarding what constitutes plagiarism. However, while there is agreement about paradigmatic cases of plagiarism, there is less agreement regarding how plagiarism should be defined. In fact, the issue is rarely discussed in detail.

- a. The concept of plagiarism is used to refer to misconduct in the scientific community, however there isn't agreement in the definition.
- b. The technology helps to identify plagiarism in the research community
- c. Plagiarism is a much used concept than misconduct in scientific writing.

6. What's the organizational structure of the paragraph?

Let us first ask what it means to appropriate someone else's work. In some definitions, plagiarism is characterised as stealing. However, if plagiarism by definition concerns stealing, then it is not theft in the traditional sense of taking a thing, where if person A takes it from person B, then B will no longer have it. What is appropriated in such instances of plagiarism is intellectual property, as when people download copyright-protected films or music from the Internet. Thus, to the extent that plagiarism is theft, it is stealing someone else's intellectual work by copying. Related to research papers, it is about copying another's text, tables, graphs, or pictures into one's own paper without having permission to do so (and with certain pretense, a point we shall be returning to presently).

- a. Topic sentence, supporting detail (fact), supporting detail (example), supporting detail (fact), concluding sentence.
- b. Topic sentence, supporting detail (example), supporting detail (fact), supporting detail (example), supporting detail (fact), concluding sentence.
- c. Topic sentence, supporting detail (fact), supporting detail (definition), supporting detail (fact), concluding sentence.

7. What is the author's opinion about plagiarism?

We are, however, disinclined to include stealing in our definition. Although one may steal intellectual as well as non-intellectual property, and even talk about “theft of the recognition due to the original contributor” (Rathod 2012), talking about plagiarism as stealing is nevertheless misguided, at least as part of a definition. This is so because using someone else’s text, say, and passing it off as one’s own can be done regardless of whether one steals the text or not. One can do it by finding the text in a journal or book or by using an unpublished paper.

- a. Stealing doesn’t belong to the concept of plagiarism.
- b. Plagiarism occurs even if you imply that you create it.
- c. Making statements about the results of research should be considered plagiarism.

8. What is the main idea of the paragraph?

It may still be argued that there is a sense of “stealing” that concerns appropriating someone else’s intellectual work and passing it off as one’s own. In this sense you may steal someone’s song if you play it and claim to have composed it yourself. This seems to mean that there is a sense of “stealing” that is equivalent to “plagiarising”. If so, this second understanding of “stealing”, which is distinct from the one discussed above, cannot contribute anything to a definition of “plagiarism”. The conclusion remains: stealing, or theft, cannot be used as part of the definition of “plagiarism”.

- a. Stealing should not be part of the definition of plagiarism.
- b. One can steal the ideas and pass them on to our own.
- c. “Steal”, “appropriate” and “use” are words in the plagiarism definition.

9. Which is the best synthesis of the text?

It is no accident that plagiarism is discussed in relation to research, although it is also clearly relevant in relation to music, literature, art, and design, since it relates to using the product of someone else’s intellectual work while passing it off as one’s own. Note that there is no reason to restrict the use of the term to published work, since you may use someone else’s work while passing it off as your own even if it is not published. For instance, you may do it by first stealing the manuscript from the author, by using passages from an unpublished manuscript circulated at a seminar, or by using ideas communicated at a lecture.

- a. Plagiarism is also relevant in relation to music, literature, art, and design and there is no reason to restrict the use of the term to published work.

- b. Plagiarism can be relevant to other fields like music, literature and art. Plagiarism can happen when you try to take an unpublished manuscript and you use it.
- c. Plagiarism is not restricted to published works and can be with or without the intention of doing it.

10. What's the relationship between the ideas?

"If someone uses another's idea and implies that it is an idea of their own, that someone is plagiarizing" and "Ideas are not always documented, but might be presented at conferences or in personal conversation, etc."

- a. It may often be much more difficult to verify that an idea has been plagiarized compared to research results or text.
- b. Research results are easier to plagiarize than ideas.
- c. There is no smart software to identify if an idea has been plagiarized.

11. According to the author of the article, what is the best solution for scientists to publish in English?

- a. They wrote their papers in their native language and had them translated by professional translators.
- b. They paid for private English classes to improve their writing and reading skills.
- c. They just published their papers in their native language.

12. Which are the common features of an abstract?

- a. Context, objective, method, results, conclusions.
- b. Introduction, theoretical framework, method, results, discussion.
- c. Introduction, problem statement, theory, method, Discussion

13. Which of the following sentences are in passive voice?

- a. ...researchers with insufficient skills in the English language have been encouraged by their supervisors or colleagues...
- b. ...to use another paper as a template and change data in order to include their own results instead of those in the template, with a considerable text overlap as a consequence.

c. Regardless of what the underlying explanations are, they have no bearing on whether or not a certain act is an act of plagiarism.

14. What is the main idea of the paragraph?

If something is to be considered plagiarism or not is also independent of who is plagiarized. For instance, it makes no difference if the person plagiarized is considered insignificant. It is also irrelevant to the evaluation of whether or not there is plagiarism if the plagiarized and plagiarizing texts are used for different purposes, have different intended readers, or are of very different dimensions. It is still plagiarism if someone copies something from a short research paper and includes it in an extensive book. Whether or not the plagiarized text was published in an indexed, peer-reviewed journal is equally irrelevant.

- a. If something is to be considered plagiarism or not is also independent of who is plagiarized
- b. It is still plagiarism if someone copies something from a short research paper and includes it in a long book.
- c. To evaluate the existence of plagiarism, it is irrelevant to consider the purposes or recipients of the plagiarized text.

15. Based on the definition that appears in the article, what would you answer to a person who thinks that ideas should be free and not be someone's intellectual property?

Plagiarism = def. an instance of someone using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own.

- a. Ideas are an intellectual product and, therefore, recognition must be given to the person.
- b. Ideas are not an intellectual product of a single person, therefore, they must be free.
- c. Ideas are a collective product, therefore, they must be recognized by society.

16. From the previous text, what type of relationship exists between plagiarism and blaming?

Someone who is plagiarizing is always blameworthy, to the extent that readers thereby are misled about who deserves credit for the work. Still, those who mislead intentionally are more blameworthy.

- a. Cause-effect relationship.
- b. Sequence relationship.
- c. Contrast relationship.

17. What is the authors' point of view on the importance that teachers have on plagiarism?

Even though a case of plagiarism is judged differently depending on whether the offender did it on purpose or not, one may also be held responsible for one's ignorance. Good research practice involves knowing where to draw the line between acceptable and unacceptable behaviors relating to research, such as the unacceptability of fabricating or tampering with data or results. That it is unacceptable to cut and paste from other people's published work, without stating the source and showing exactly what passages are quoted, is part of that required knowledge of good research practice. If one has not been taught this, part of the blame for plagiarism must fall on one's teachers and supervisors.

- a. Teachers are responsible for teaching good research practices to their students to avoid plagiarism.
- b. Teachers are responsible for plagiarism committed by their students.
- c. Teachers are responsible for teaching their students the risks of plagiarism.

18. What type of logical flaw does the argument of the following text present?

Plagiarism can be more or less extensive, ranging from whole chapters of books, or entire academic papers, to shorter passages. We will argue that to the extent that quantity words matter, it depends on whether quantity has an impact on quality. This is to say that quantity matters only indirectly, while quality matters directly.

- a. Lack of adequate evidence.
- b. Lack of coherence.
- c. Lack of logical structure.

19. Following the author's argumentation, which is the conclusion of his ideas?

We claim that plagiarism (in principle) can consist in as little as one word, while there are many standard sentences describing research methods that will not be plagiarism even if, in fact, copied from someone else. This is to say that the unmarked reuse of some very short passages might be plagiarism, even though the reuse of other equally short passages would not.

- a. The plagiarism has to do with quality rather than quantity.
- b. The plagiarism has to do with intention rather than quantity.
- c. The plagiarism has to do with quantity rather than quality.

20. Which of the following research topics is the best to continue the reading?

- a. Plagiarism in social sciences.
- b. Ethics in the research community.
- c. Scientific credit.

21. Which of the following key words make a difference between definitions?

When the concept is explained in a recent newsletter from the US Office of Research Integrity, it looks deceptively simple: “It involves stealing someone else’s work and lying about it afterward” (Sox 2012). Others prefer to speak of “copying” part of someone else’s published work and using it without showing that it is borrowed from someone else. In the Longman Contemporary English Advanced Learner’s Dictionary, the act of plagiarism is defined as “when someone uses another person’s words, ideas, or work and pretends they are their own”.

In the scholarly definitions, the more technical notions of “appropriation” and “credit” re central: “Plagiarism is the appropriation of other people’s material without giving proper credit” (The European Code of Conduct for Research Integrity); “Plagiarism is the appropriation of another person’s ideas, processes, results, or words without giving appropriate credit” (US Federal Policy on Research Misconduct). So the basic ideas seem to be that someone deliberately takes someone else’s work, whether in the form of an idea, a method, data, results, or text, and presents it as their own instead of giving credit to the person whose ideas, results, or words it is. This is mirrored in the definition given by Merriam-Webster: “to steal and pass off (the ideas or words of another) as one’s own: use (another’s production) without crediting the source”.

- a. Stealing, Pass, Use, Appropriate, Credit, Pretend
- b. Stealing, Work, Credit, Ideas, Appropriate
- c. Use, Work, Ideas, Product, Plagiarism

22. What is the main meaning of the following argument?

While ordinary language use of “plagiarism” certainly allows for the act of plagiarising being intentional, it seems as clear that plagiarism does not necessarily involve an intention to deceive. We therefore would like to suggest a definition that does not require intention. The notion of “passing something off” also seems to imply intention, and therefore ought preferably to be avoided. A slightly modified definition, where we define “plagiarism” rather than “to plagiarise”, would therefore read: Plagiarism = an instance of someone using someone else’s work, thereby implying that it is their own.

- a. According to the authors, plagiarism can be intentional or unintentional.
- b. According to the author, not all authors have the intention to deceive.

c. According to the author, intention matters when it comes to plagiarize.

Section II

1. Write an abstract about your research project or intervention (150 a 250 words). Remember to include all the parts of an abstract: Context, objective, method, and possible results.

Anexo 8

Cuestionario de contexto

CUESTIONARIO DE CONTEXTO

1. ¿Cuántos años tienes de edad?
2. ¿En qué país y ciudad naciste?
3. ¿De ser extranjero, cuántos años tienes viviendo en México?
4. ¿Nombre del posgrado que cursas?
5. ¿En qué Unidad Académica se encuentra el posgrado?

6. ¿Has tomado un curso de inglés?
7. En caso de que hayas tomado un curso de inglés, ¿cómo se llama la institución dónde tomaste el curso?
8. ¿Cuántos cursos de inglés has tomado?
9. ¿Has presentado el TOEFL?
10. Si presentaste el TOEFL, ¿cuál fue el puntaje que obtuviste?
11. ¿Has presentado algún otro examen de certificación del dominio del inglés?
12. En caso de haber presentado un examen diferente al TOEFL, ¿cuál fue?
13. ¿Has redactado algún texto académico/artículo en inglés?

Anexo 9

Rúbrica de evaluación de uso y convenciones de la lengua.

	Not proficient			Proficient		
	1 Beginning	2 Emerging	3 Developing	4 Capable	5 Experienced	6 Exceptional
Main idea	Does not have a main idea or topic	Focuses on a general topic, but main idea is unclear	Suggests a main idea, but the direction of the piece is confusing	Presents an adequately clear, focused main idea	Develops a clear, focused, and somewhat complex and/or original main idea	Conveys a clear, focused, complex, and original main idea that drives the piece
Details, evidence, support	Does not support a main idea	Attempts to support the main idea with some information and/or details, but these are unfocused, unclear, and/or unrelated	Provides incidental support of the main idea with information and/or details that lacks specificity, relevance, and/or accuracy	Supports the main idea with generally accurate, specific, and relevant information and/or details	Supports the main idea fully, with specific, credible, relevant, information and/or vivid details	Supports the main idea convincingly, with highly specific, credible information and/or striking details that go beyond the obvious
Reasoning/ thinking	Does not show writer's thinking in developing or connecting ideas	Generalizes about ideas without analyzing or connecting them logically	Attempts to identify relationships among ideas, but lacks clarity or logical connection	Helps the reader understand the writer's reasoning or perspective Connects ideas	Presents fresh reasoning or perspective using higher-order thinking and connects Clarifies ideas	Conveys unique, useful insights using higher-order thinking and convincing reasoning Clarifies and enhances ideas.
Textual evidence	Contains no textual evidence	Uses very weakly integrated textual evidence	Contains weakly integrated textual evidence	Provides adequately integrated textual	Includes well-integrated textual evidence	Incorporates very smoothly integrated textual

				evidence		evidence
Lead/ conclusion	Uses no lead and no conclusion	Includes either a lead or a conclusion If lead is present, fails to establish purpose; if conclusion is present, fails to provide closure	Includes both a lead and a conclusion Lead fails to establish purpose, and/or conclusion fails to provide closure	Contains a lead that establishes purpose, though may be formulaic or obvious Contains a conclusion that provides closure, though may be formulaic or obvious	Features a lead that creates anticipation and establishes clear purpose Includes a conclusion that ties up loose ends, providing a satisfying sense of closure	Creates an inviting lead that establishes clear purpose, draws reader in, and creates a strong sense of anticipation Develops a satisfying conclusion that conveys a powerful and thoughtful sense of closure
Sequencing	Uses no sequencing of ideas	Uses very limited sequencing that fails to show how ideas fit together	Uses sequencing that fails to showcase ideas or becomes formulaic	Provides logical sequencing of ideas	Employs sequencing that builds connections to create a unified whole	Utilizes highly effective sequencing, making best choices for progression and enrichment of reader's understanding
Pacing	Uses no evident pacing	Uses uneven pacing, slowing down or speeding up inappropriately or awkwardly	Appropriately controls pacing in some sections but not in others	Evenly controls pacing in most places	Exhibits well-controlled pacing throughout	Skillfully uses pacing to compel the reader through the text and enhance its impact

Purpose/ text structure	Uses no discernible text structure or purpose	Uses loose text structure that leaves reader unclear or confused about purpose	Uses text structure inconsistently, affecting the reader's ability to identify purpose	Uses text structure consistently to reflect purpose, moving the reader through the text logically, with minimal confusion	Employs text structure that clarifies and supports purpose throughout	Utilizes text structure that enhances understanding of purpose and flows very smoothly
Tone	Uses no evident tone Uses inappropriate level of formality	Uses a tone that does not support the purpose Often uses inappropriate level of formality	Uses a flat, disinterested tone or one that does not always support the purpose Sometimes uses inappropriate level of formality	Uses a tone that supports purpose and reflects audience Nearly always uses appropriate level of formality	Uses a tone that is frequently appropriate for purpose and audience Consistently uses appropriate level of formality	Employs a tone that is consistently appropriate for purpose and audience, clarifying and validating the message Consistently uses appropriate level of formality
Word meaning	Consistently uses vocabulary that is incorrect in meaning	Uses flawed and/or simplistic vocabulary that impairs understanding	Uses vocabulary that is sometimes incorrect or limited, or uses inaccurate domain-specific vocabulary	Uses mostly correct, clear vocabulary, including domain-specific vocabulary if appropriate	Frequently uses precise vocabulary, including accurate domain-specific vocabulary if appropriate; vocabulary may be creative	Consistently features precise, creative, and powerful vocabulary, including exact domain-specific vocabulary if appropriate

Grammar	Misuses parts of speech frequently, clouding meaning	Uses limited variety in parts of speech; may use jargon or clichés that detract from meaning	Uses passive or boring verbs and/or overuses nouns and/or modifiers that limit meaning; may sound mechanical	Uses accurate and functional parts of speech that start to shape meaning, including some active verbs and specific nouns and modifiers	Chooses precise and varied parts of speech (e.g., active verbs, specific nouns and modifiers) to effectively communicate meaning	Chooses powerful parts of speech, including expressive, active verbs and polished nouns/modifiers, to best convey exact meaning
Sentence structure	Uses choppy sentence structure that is incomplete, run-on, or rambling	Uses sentence structure that is simplistic and/or rarely correct	Uses sentence structure that may be technically correct yet isn't smooth, or overuses complicated sentences	Uses smooth, correct sentence structure that may be somewhat mechanical in places	Utilizes well-developed sentence structure throughout text	Controls strong sentence structure for maximum impact throughout text
Connecting sentences	Uses weak or no transitional words or phrases to join sentences or sentence parts and/or has choppy, disconnected language and/or run-on sentences	Uses basic transitions that do little to join sentences and sentence parts within paragraphs; transitional words or phrases may be missing or randomly applied	Joins sentences and sentence parts with few transitional words or phrases; some transitional words or phrases may be overused	Joins sentences and sentence parts with varied transitional words or phrases	Joins sentences and sentence parts fluidly, with thoughtful and varied transitional words or phrases	Joins sentences and sentence parts throughout paragraphs with creative and varied transitional words or phrases that highlight relationships between sentences
Spelling	Makes frequent	Uses phonetic spelling, with many errors	Frequently spells simple words incorrectly,	Uses correct spelling for most common grade-	Correctly spells common grade-level words and	Exhibits correct spelling, including

	spelling errors, even with common words		though reader can still understand the meaning	level words; may be inaccurate with more difficult words	often more difficult words	more difficult words
Punctuation	Does not use or misuses punctuation nearly all the time	Uses punctuation randomly and/or incorrectly	Uses simple end punctuation that is usually correct; internal punctuation (e.g., comma, apostrophe, semicolon) contains errors	Uses mostly correct end and internal punctuation at grade level	Correctly uses both end and internal punctuation	Employs correct end and internal punctuation; may use creative punctuation that enhances readability

Anexo 10

Rúbrica de evaluación de elaboración de un resumen de artículo científico.

Research Abstract Components	0 points A component in this category:	1 point A component in this category:	2 points A component in this category:
<p>INTRODUCTION</p> <p>The introduction succinctly describes and appropriately connects the subject and context/background to the purpose of the investigation.</p>	<ul style="list-style-type: none"> • Describes and connects subject and context/background to the purpose of the investigation, but lacks organization, lacks specific details or is overly wordy, • Or fails to make connections, • Or contains no introduction. 	<ul style="list-style-type: none"> • Describes and connects the subject and context/background to the purpose of the investigation, but is somewhat disorganized, somewhat lacking in specifics or somewhat wordy, • Or contains only implicit or superficial connections. 	<ul style="list-style-type: none"> • Describes and connects the subject and context/background to the purpose of the investigation in an organized, specific and concise manner.
<p>METHODS</p> <p>The methods section succinctly identifies the methods used to study the subject of the investigation OR succinctly identifies the existing procedures, models or programs used.</p>	<ul style="list-style-type: none"> • Discusses the methods or resources used in the investigation, but lacks organization, lacks specific details or is overly wordy, • Or provides no discussion of the methods or resources used. 	<ul style="list-style-type: none"> • Discusses the methods or resources used to study the subject of the investigation, but is somewhat disorganized, somewhat lacking in specifics or somewhat wordy. 	<ul style="list-style-type: none"> • Discusses the methods or resources used to study the subject of the investigation in an organized, specific and concise manner
<p>RESULTS</p> <p>The results section provides a succinct and specific explanation of what was discovered, accomplished, collected or produced.</p>	<ul style="list-style-type: none"> • Provides no results, • Or does not explain why results were not possible, • Or states only “No results.” 	<ul style="list-style-type: none"> • States results including quantitative data, but is somewhat disorganized, somewhat lacking in specifics or somewhat wordy, • Or states results only qualitatively, • Or provides only a superficial explanation of why no data or results were achieved. 	<ul style="list-style-type: none"> • States results, including quantitative data, in an organized, specific and concise manner, • Or provides a thorough, organized and specific explanation of why no data or results were achieved.*

*This exception is meant to cover instances where the investigation was not completed during the internship.

Anexo 11

Concentrado de juicios emitidos por expertos para la validación de contenido

Ítem	Correspondencia de ítem con su respectiva especificación									Correspondencia entre la especificación y el contenido esencial que evalúa			Corrección conceptual			Penalización injusta+Q1:AQ1								
	¿El ítem presenta todos los elementos especificados?			¿El ítem presenta los elementos de la manera en que fueron especificados?			¿Existe otra falla de correspondencia entre el ítem y la especificación?			¿La especificación delinea un ítem que evalúa lo importante del contenido?			¿El ítem está libre de errores conceptuales?			¿El ítem pone en ventaja a examinados de cierto nivel socioeconómico o cultural?			¿El ítem pone en ventaja a los examinados de algún género?			¿El ítem pone en ventaja a otra clase de examinados?		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Expertos	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
What is the abstract about?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Which of the following key concepts are related to the text?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Which is the key information of the paragraph?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Which criteria do the authors suggest for defining plagiarism?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
What is the main idea of the paragraph?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Anexo 12

Promedio de ítem por criterio en el coeficiente de Aiken.

Promedio ítem por criterio en el coeficiente de Aiken									
Ítems	Ítem con su especificación	Especificación y el contenido que evalúa	Corrección conceptual	Penalización injusta	Ofensa o estereotipo	Apego a las convenciones de redacción técnica	Demanda lingüística apropiada	Demanda cognitiva apropiada	Escritura
IL1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL2	1,00	1,00	1,00	1,00	1,00	0,67	0,67	1,00	1,00
IL3	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,67
IL4	1,00	1,00	0,67	1,00	1,00	0,67	1,00	1,00	0,50
IL5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,67
IL6	0,89	1,00	1,00	1,00	1,00	0,67	0,67	1,00	0,17
IL7	1,00	1,00	0,67	1,00	1,00	0,67	1,00	1,00	0,33
IL8	1,00	1,00	1,00	1,00	1,00	1,00	0,67	0,67	0,67
IL9	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL10	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL11	0,89	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL12	0,78	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL13	0,78	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,67
IL14	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL15	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,67
IL16	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL17	1,00	1,00	1,00	1,00	1,00	0,67	1,00	1,00	1,00
IL18	0,89	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL19	1,00	1,00	1,00	1,00	1,00	1,00	0,67	1,00	0,67
IL20	0,78	0,67	1,00	1,00	1,00	1,00	0,67	0,67	0,67
IL21	1,00	1,00	1,00	1,00	1,00	1,00	0,67	0,83	0,83
IL22	1,00	1,00	1,00	1,00	1,00	1,00	0,67	0,67	0,83
IE1	0,89	1,00	1,00	1,00	1,00	1,00	0,67	1,00	0,83

Anexo 13

Promedio de ítem por criterio en el coeficiente de Tristán.

Promedio ítem por criterio en el coeficiente de Tristán									
Ítems	Ítem con su especificación	Especificación y el contenido que evalúa	Corrección conceptual	Penalización injusta	Ofensa o estereotipo	Apego a las convenciones de redacción técnica	Demanda lingüística apropiada	Demanda cognitiva apropiada	Escritura
IL1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83
IL2	1,00	1,00	1,00	1,00	1,00	0,67	0,67	1,00	1,00
IL3	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,67
IL4	1,00	1,00	1,00	0,89	1,00	1,00	1,00	0,83	1,00
IL5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	1,00
IL6	1,00	0,67	1,00	1,00	1,00	1,00	1,00	0,67	1,00
IL7	1,00	1,00	1,00	0,89	1,00	1,00	1,00	0,83	1,00
IL8	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	0,67
IL9	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL10	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL11	0,67	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL12	0,67	0,67	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL13	0,67	0,67	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL14	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL15	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL16	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL17	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	1,00
IL18	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
IL19	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	1,00
IL20	1,00	0,67	1,00	0,89	1,00	1,00	1,00	0,83	0,67
IL21	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	0,83
IL22	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,83	0,67
IE1	1,00	0,67	1,00	1,00	1,00	1,00	1,00	0,83	1,00

Anexo 14

Tabla de resumen de observaciones realizados por Chat GPT

<i>Criterio/ examen ado</i>	<i>1</i>	<i>2</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>
<i>Main idea</i>	The main idea is somewhat clear but lacks complexity and originality. It identifies the subject of the research but does not articulate a nuanced or focused central argument.	The main idea is clear and relevant, focusing on applying diagnostic cognitive models to admission tests. However, it lacks complexity and depth in presenting the significance and implications of this approach.	The main idea is clear and focused: to elaborate and validate an instrument for measuring Teachers Wellbeing (TWB) based on a theoretical framework. The complexity is present, though the idea could be presented with more depth and originality.	The main idea is clear: identifying factors influencing student reprobation at UABC. However, the idea is not highly complex or original, and its presentation lacks depth.	The main idea is somewhat clear: a comparison of two schools in Ensenada and the strategies used. However, it is not well-developed or original, and it lacks a clear focus on what the study aims to reveal about these strategies.	The main idea is present: examining difficulties in access and growth in higher education (HE) at public state universities. However, it is not fully developed or complex. The idea is somewhat clear but lacks depth and originality.	The main idea is unclear and not well articulated. It seems to introduce a topic related to the construction of an analytical program for a secondary school but lacks detail or focus.	The main idea is to analyze teachers' perceptions, experiences, and needs regarding project methodologies in primary education. While the topic is relevant, the idea is not fully developed or complex.	The main idea is clearly conveyed: learning communities are seen as transformation projects aimed at identifying and solving common problems. The idea is focused and relevant, but it could benefit from more complexity and depth.
<i>Details, evidence, support</i>	The text lacks specific details and supporting evidence. It provides only a general outline without any in-depth information or examples.	The text briefly mentions diagnostic cognitive models and their application but lacks specific details or examples to	The text outlines the objective of the dissertation and the need for a new instrument. However, it lacks specific details about	The text mentions some details about the study but lacks specific, convincing evidence or in-depth support. It identifies some programs	The text provides minimal details and lacks specific evidence. It briefly describes the schools but does not offer	The text provides limited details and lacks specific evidence. It mentions some difficulties but does not provide	The text does not provide any details, evidence, or support. It is too brief and vague to support any main idea effectively.	The text provides some basic details but lacks specific evidence or comprehensive support. For instance, there is no detailed	The text provides a basic overview of the research objectives and methods but lacks specific details or evidence. It mentions the

		convincingly support the main idea.	the conceptualizations of TWB and the nature of the theoretical framework being used.	and methods but does not provide detailed evidence or examples.	in-depth analysis or supporting information about the rules and strategies.	detailed examples or data to support the claims about the challenges faced by universities.	information on the results or how the analysis was conducted.	use of mixed methods but does not elaborate on how these methods will be applied or what specific problems will be addressed.	
<i>Reasoning/ thinking</i>	The text shows minimal reasoning. It mentions concepts like "metaevaluation" and "formative evaluation" but does not explore them or their implications in depth.	The text demonstrates some reasoning by connecting diagnostic cognitive models to formative assessments. However, the insights are not deeply explored or particularly unique.	The reasoning is logical, pointing out the gap in current research and the need for a new measurement tool. However, more unique insights or deeper exploration of how the instrument will address existing gaps could enhance this.	The reasoning is basic. The text outlines the study's goal and methods but does not delve into the implications or provide unique insights into the factors influencing reprobation.	The text lacks detailed reasoning or insight. It mentions that the rules and strategies are different but does not explain the significance or provide a deeper analysis of these differences.	The reasoning is basic. The text outlines some difficulties but does not delve deeply into the reasons behind these issues or provide unique insights into the broader implications.	There is no reasoning or thinking conveyed. The text does not provide any insights or analysis, making it difficult to understand the purpose or implications of the topic.	The text lacks depth in reasoning. It outlines objectives and methods but does not provide insight into the significance of the findings or deeper analysis.	The reasoning is straightforward but not deeply developed. The text explains the purpose and methodology but does not provide much insight into the potential impact or unique aspects of the research.
<i>Textual evidence</i>	There is no incorporation of textual evidence or examples. The text is very general and does not include	The text lacks integrated textual evidence. While it references concepts like diagnostic cognitive	The text lacks specific textual evidence or examples from the literature to support the claims. While it references the gap in	There is minimal integration of textual evidence or examples. The text describes the study's approach but	There is no integration of textual evidence or examples. The text is too brief and lacks detailed examples to	The text lacks textual evidence or examples. It mentions interviewing actors and policy makers but does not	There is no textual evidence or examples included. The text is too sparse to incorporate or integrate any	There is no integration of textual evidence. The text does not incorporate any findings or examples from the interviews	There is minimal textual evidence. The text lacks integration of specific examples or detailed

	supporting material.	models and formative assessments, it does not provide supporting data or specific examples.	research, it does not provide detailed evidence or examples.	does not provide any detailed evidence from the study itself or relevant literature.	support its claims.	integrate any specific findings or examples from these interviews.	evidence.	or group discussions.	descriptions that would support the claims or objectives more convincingly.
<i>Lead, conclusion</i>	The lead is not inviting or engaging. The text does not have a clear conclusion or a satisfying sense of closure.	The lead introduces the topic effectively but could be more engaging. The conclusion is implicit and could be more clearly articulated to provide a strong sense of closure.	The lead introduces the topic effectively but could be more engaging. The conclusion about the dissertation's aim is stated but could be more compelling and clearer about the expected impact.	The lead introduces the problem but is not very engaging or detailed. The conclusion is vague and does not clearly convey a strong sense of closure or impact.	The lead is not engaging and does not clearly introduce the purpose of the text. The conclusion is also vague and does not provide a clear sense of closure or impact.	The lead introduces the topic but is not very engaging or detailed. The conclusion is also vague and does not provide a strong sense of closure or clearly convey the impact of the research.	The lead is incomplete and does not establish a clear purpose or engage the reader. There is no conclusion or sense of closure provided.	The lead introduces the topic but is not particularly engaging or clear. The conclusion is also vague and does not effectively summarize the research outcomes.	The lead introduces the topic but is somewhat abrupt. The conclusion is implied rather than explicitly stated. The text could benefit from a clearer, more engaging conclusion that summarizes the expected outcomes.
<i>Sequencing</i>	The sequencing is rudimentary. The ideas are presented in a somewhat logical order but lack smooth transitions and coherent progression.	The sequencing is logical but basic. The text moves from describing admission tests to introducing diagnostic cognitive models and their application, though transitions	The sequencing is logical, moving from identifying the issue to stating the objective of the dissertation. Improvements could be made in the flow and connections between ideas to enhance	The sequencing is logical but basic. The text moves from identifying the problem to describing the study's methodology and potential outcomes, but the transitions could be smoother.	The sequencing is basic. The text mentions the schools, objectives, method, and results but does so in a disjointed manner. Improved transitions and structure are needed for	The sequencing is basic and somewhat disjointed. The text jumps between discussing objectives, methods, and findings without smooth transitions or clear organization.	The text lacks effective sequencing. It is fragmented and does not present information in a coherent or organized manner.	The text is somewhat disorganized. The sequencing of information is unclear, making it difficult to follow the progression from context to objectives and methods.	The text is logically sequenced, moving from the introduction of learning communities to the research objectives and methods. However, the flow could be improved with more detailed

		could be smoother.	understanding.		better flow.				transitions between sections.
<i>Pacing</i>	The text does not effectively manage pacing. It reads as a series of disjointed statements rather than a cohesive narrative or argument.	The pacing is generally appropriate but could benefit from more dynamic shifts to enhance reader engagement. Some sections feel a bit rushed or underdeveloped.	The pacing is generally appropriate but could benefit from a more dynamic approach. Some sections feel a bit dense, and a more engaging pace could help maintain reader interest.	The pacing is uneven. Some parts feel rushed, particularly where the methodology is described. The text could benefit from a more balanced approach to maintain reader interest.	The pacing is uneven. The text is brief and lacks sufficient detail, making it difficult to follow or engage with the content effectively.	The pacing is uneven. The text is brief and lacks sufficient detail, making it difficult to follow or engage with the content fully.	Pacing is not applicable as the text is too brief and does not provide sufficient content to evaluate pacing.	Pacing is uneven. The text is brief and lacks sufficient detail, affecting the flow and reader engagement.	Pacing is adequate but could be improved. The text moves through the information relatively quickly without delving deeply into any particular aspect, which affects the overall impact.
<i>Purpose/ text structure</i>	The purpose is somewhat clear but not well-developed. The text structure is basic and does not enhance understanding or flow smoothly.	The text structure is straightforward but could be more refined. The purpose is clear, but the flow between ideas could be improved for better coherence.	The text structure is clear and supports the purpose of outlining the dissertation's aim. However, the structure could be refined to improve the flow and coherence of the argument.	The text structure supports the purpose but is somewhat disjointed. The flow of ideas could be improved for better clarity and coherence.	The text structure is rudimentary. While it has sections for context, objectives, method, and results, the organization is not clear or well-developed. The purpose is not fully realized or conveyed.	The text structure is rudimentary. While it has sections for objectives, methods, and findings, the organization is not clear or well-developed. The purpose is not fully realized or conveyed.	The text structure is incomplete and does not clearly articulate the purpose. It lacks sections or organization to enhance understanding.	The text structure is rudimentary. It has sections for objectives and methods, but the organization is not well-developed, and the purpose is not clearly articulated.	The text structure is generally effective, with clear sections for objectives and methodology. However, the text could be more smoothly organized with better integration of ideas.

<i>Tone</i>	The tone is appropriate for an academic context but lacks consistency and sophistication. The formality is not well-maintained throughout.	The tone is appropriate for an academic context, maintaining formality and clarity. However, it could be more engaging or varied to better captivate the audience.	The tone is formal and appropriate for an academic context. It maintains a professional level of formality but could be more engaging or varied to better captivate the audience.	The tone is appropriate for an academic context but lacks consistency and sophistication. It is formal but could be more engaging and polished.	The tone is informal and inconsistent. It does not maintain the level of formality expected in an academic or analytical text, and the language is simplistic.	The tone is formal but inconsistent. It is appropriate for an academic context but lacks sophistication and polish. The language could be more engaging and precise.	The tone is not established due to the lack of content. It is impossible to assess the appropriateness or formality of the tone with such a brief excerpt.	The tone is formal but inconsistent. The language is appropriate for an academic context but lacks polish and coherence.	The tone is appropriate for an academic context, formal and objective. However, it could be more engaging and nuanced to better connect with the reader.
<i>Word meaning</i>	Vocabulary is basic and lacks precision. Domain-specific terminology is used but not explained or contextualized.	Vocabulary is adequate but not particularly precise or creative. Terms like "diagnostic cognitive models" are used correctly but not elaborated upon in a meaningful way.	Vocabulary is generally appropriate but lacks precision in some areas (e.g., "myrlad" should be "myriad," "intrepretated" should be "interpreted," "necesary" should be "necessary"). Some domain-specific terms are used but could be explained more clearly.	Vocabulary is often imprecise or incorrect. Terms like "reprobation," "myxt," and "cualitative" are used incorrectly or are misspelled. Domain-specific terms are not clearly defined.	Vocabulary is basic and sometimes incorrect. Words like "tallo" and "localite" are either misspelled or used incorrectly. More precise and accurate terminology is needed.	Vocabulary is often imprecise or incorrect. Words like "scolanty" (should be "enrollment"), "pretends" (should be "aims"), and "planification" (should be "planning") are used incorrectly or are misspelled. Domain-specific terms are not clearly defined.	The text includes some errors in word usage and spelling, such as "Analitic" (should be "Analytical") and "secondary" (should be "secondary"). The vocabulary is not precise or effective.	There are multiple issues with word usage, such as "Teching" (should be "Teaching"), "primaria" (should be "primary"), "analiyze" (should be "analyze"), and "capation" (should be "training"). The vocabulary is often imprecise.	The vocabulary is appropriate but not particularly precise or powerful. Terms like "problematic" and "optimazation" (should be "optimization") are used, but there is room for more precise and creative vocabulary.
<i>Grammar</i>	The text has noticeable grammatical errors, including	Grammar is mostly correct but includes a few errors, such as missing	Grammar is mostly correct but includes errors such as "myrlad"	Grammar issues are present, such as incorrect verb forms	Grammar issues are prevalent. Errors include incorrect verb	Grammar issues are present, such as incorrect verb forms	Grammar issues are present, such as the incorrect use of articles	The text contains numerous grammatical errors,	The text contains some grammatical issues, such as incorrect verb

	incorrect usage of articles and verb forms (e.g., "it's" instead of "it is" or "its," "objectiv" instead of "objective," "perspectie" instead of "perspective").	articles ("the topics the tests measure" should be "the topics that the tests measure") and inconsistent verb forms ("assessment" should be "assessment").	(should be "myriad"), "theoretical" (should be "theoretical"), and incorrect verb forms. These errors slightly detract from the clarity.	("pretends" should be "aims"), incorrect word forms ("mixt" should be "mixed"), and awkward phrasing. These errors detract from clarity.	forms ("have" should be "has"), incorrect word choices ("diferent" should be "different"), and awkward phrasing. These detract from clarity.	("pretends" should be "aims"), awkward phrasing, and incorrect use of articles and prepositions. These errors affect clarity.	and unclear phrasing. The text is not grammatically coherent.	including incorrect verb forms ("was with six teacher" should be "was conducted with six teachers") and awkward phrasing.	forms and awkward phrasing. For example, "it is inserted in" could be revised for clarity.
<i>Sentence structure</i>	Sentence structure is basic and sometimes awkward. The text lacks variety and does not fully utilize complex sentence constructions.	Sentence structure is generally correct but could be more varied and sophisticated. Some sentences are straightforward but could be combined or restructured for better impact.	Sentence structure is generally correct but could be more varied. Some sentences are complex and could be simplified or restructured for better readability.	Sentence structure is basic and sometimes awkward. Many sentences are convoluted or poorly constructed, affecting readability.	Sentence structure is often awkward and simplistic. Many sentences are short and lack variety, which affects readability and impact.	Sentence structure is basic and often awkward. Many sentences are short or poorly constructed, which affects readability and impact.	Sentence structure is lacking. The text consists of fragments rather than complete sentences, making it difficult to assess the effectiveness of sentence structure.	Sentence structure is basic and often awkward. Many sentences are incomplete or poorly constructed, which affects clarity.	Sentence structure is generally correct but sometimes awkward. Some sentences could be rephrased for greater impact and clarity.
<i>Connecting sentences</i>	There is little to no use of transitional phrases. The connections between sentences are weak and contribute to the disjointed nature of the text.	Transitional phrases are minimal, leading to somewhat abrupt shifts between sentences and ideas. More creative transitions would help highlight	Transitions between sentences are somewhat basic. Improved use of transitional phrases could better highlight relationships between ideas and enhance overall flow.	Transitions between sentences and ideas are weak. The text could benefit from better use of transitional phrases to improve coherence and highlight relationships.	There is minimal use of transitional phrases or connectors. The text feels disjointed, and the relationships between sentences are not clearly highlighted.	There is minimal use of transitional phrases or connectors. The text feels disjointed, and the relationships between sentences are not clearly highlighted.	There are no connecting sentences or transitional phrases. The text is fragmented and does not establish connections between ideas.	There are minimal connecting sentences or transitions. The text is fragmented, making it difficult to follow the relationships between ideas.	Transitions between sentences are present but could be more effective. The connections between ideas are sometimes unclear or abrupt.

		relationships and improve flow.							
<i>Spelling</i>	Spelling errors are present, such as "objectiv," "perspectie," and "analys."	Spelling is generally correct but includes some errors such as "strenghts" (should be "strengths") and "assesment" (should be "assessment").	Spelling errors are present, including "myrlad" (should be "myriad"), "intrepretated" (should be "interpreted"), and "necesary" (should be "necessary"). These errors affect readability.	Spelling errors include "myrlad" (should be "myriad"), "speciffically" (should be "specifically"), and "entervIEWS" (should be "interviews"). These errors affect readability.	Spelling errors are present, such as "fist" (should be "first"), "tallo" (unclear; should be "tallo" or "tallo"), and "diferent" (should be "different").	Spelling errors include "scolanty" (should be "enrollment"), "diferent" (should be "different"), and "acces" (should be "access"). These errors detract from readability.	Spelling errors include "Analitic" (should be "Analytical") and "secundary" (should be "secondary"). The spelling issues detract from readability.	Spelling errors include "Teching" (should be "Teaching"), "primaria" (should be "primary"), "analiyze" (should be "analyze"), and "capation" (should be "training"). These errors impact readability.	Spelling is mostly correct, but there is a notable error with "optimazation" (should be "optimization"). This impacts readability.
<i>Punctuation</i>	Punctuation is mostly incorrect or missing, contributing to the lack of clarity and readability.	Punctuation is mostly correct but could be improved. For example, the text could benefit from commas to separate clauses and enhance readability.	Punctuation is mostly correct but could be improved. For example, additional commas could help clarify complex sentences and improve readability.	Punctuation is generally incorrect or missing. For instance, commas are needed to separate clauses, and the overall punctuation could be improved to enhance readability.	Punctuation is generally incorrect or missing. There are issues with commas and periods that affect the readability and flow of the text.	Punctuation is generally incorrect or missing. There are issues with commas and periods that affect the readability and flow of the text.	Punctuation is minimal and incorrect. The text does not use punctuation effectively, affecting readability and coherence.	Punctuation is inconsistent. There are missing commas and periods that affect readability and coherence.	Punctuation is generally correct, but there are areas where it could be improved for clarity. The text could benefit from more nuanced use of punctuation to enhance readability.

<i>Introducción</i>	<p>Assessment: The introduction gives a brief idea of the research topic (art evaluation in higher education) and mentions the purpose (metaevaluation of an examination of a formative evaluation perspective). However, it lacks clarity and depth in describing the context or background. The introduction is somewhat fragmented and contains errors that obscure the intended message. Strengths: It identifies the general research topic and purpose. Weaknesses: The introduction is</p>	<p>Assessment: The introduction effectively outlines the general context of college admission tests and their typical use in summative assessments. It also introduces the concept of using diagnostic cognitive models, setting up the research focus. Strengths: Clearly states the subject (college admission tests), provides context (summative vs. formative/diagnostic assessments), and introduces the purpose of the research. Weaknesses: The introduction could be clearer in</p>	<p>Assessment: The introduction effectively describes the complexity and variety of conceptualizations related to Teacher Wellbeing (TWB) and highlights the gap in existing research. It clearly states the problem and sets up the context for the research. Strengths: Provides a thorough background on TWB and identifies a gap in the literature, which is critical for understanding the research's relevance. Weaknesses: While comprehensive, it could benefit from a clearer</p>	<p>Assessment: The introduction outlines the problem of reprobation (failure) in educational systems, identifies specific programs with high reprobation rates, and sets up the purpose of the study. Strengths: Clearly identifies the context (educational problems related to reprobation) and specifies the focus on programs at UABC. Weaknesses: Contains grammatical errors and lacks detail on why this problem is significant or how the study will address it.</p>	<p>Assessment: The introduction briefly mentions visiting two schools in Ensenada and provides some context about the schools. However, it lacks clarity and depth. Strengths: Identifies the schools and some basic characteristics. Weaknesses: The introduction is disjointed and does not clearly explain the purpose of the investigation or how it relates to the context. There are numerous spelling and grammatical errors that hinder understanding.</p>	<p>Assessment: The introduction highlights the issue of access to higher education and the goal of the federal administration, but it lacks coherence and clarity. Strengths: Identifies the main problem and the objective related to increasing access to higher education. Weaknesses: The text contains grammatical errors and awkward phrasing, such as "In the actual Federal administration pretends to arrive al 50% scolanty," which is unclear. It does not clearly</p>	<p>Assessment: The introduction is incomplete and does not provide a clear description or connection between the subject and the purpose of the investigation. Strengths: Mentions the construction of "Analytic Programs of study plan 2022," which indicates a focus on secondary education. Weaknesses: The text is fragmented and lacks a full introduction. It does not clearly describe the subject, background, or the purpose of the investigation. The context and objective sections are left</p>	<p>Assessment: The introduction provides some context but lacks clarity and precision. It identifies the location and scope of the study but does so in a fragmented manner. Strengths: It mentions the location (Ensenada, Baja California), the subject (teaching practice in project methodologies), and the target group (six teachers). Weaknesses: The description is not well-organized, contains grammatical errors, and lacks a clear connection between the background</p>	<p>Assessment: The introduction provides a general overview of learning communities and their purpose. It establishes the context and the problem to be addressed but does so in a somewhat broad manner. Strengths: It introduces the concept of learning communities as transformation projects and outlines the objective to identify problems and develop strategies. Weaknesses: The introduction lacks specific details about the context and background of the learning community</p>
---------------------	---	--	--	--	--	---	---	--	--

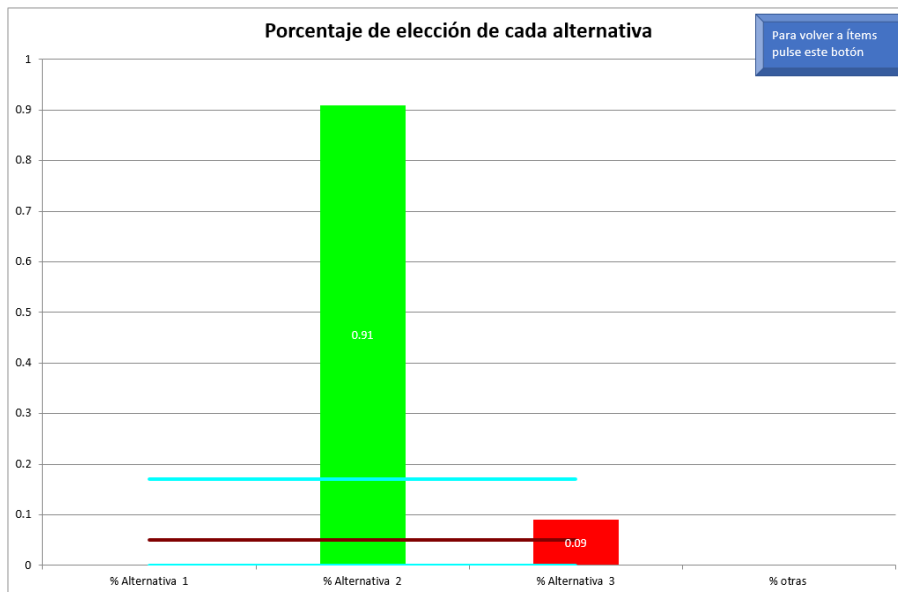
	incomplete, with grammatical issues and a lack of detailed background or context.	connecting the background information to the specific purpose of the research, but it does provide a solid foundation.	connection to the research objectives and the significance of addressing the identified issues.	The introduction could benefit from clearer language and more specific background information.		connect the context to the research purpose or explain why this issue is significant.	blank or incomplete.	and the purpose of the study. For example, "teching practice" should be "Teaching practice," and "primaria" should be "primary."	being studied. The explanation is somewhat vague, and the connection between the context and the purpose could be clearer.
<i>Método</i>	Assessment: The methods are vaguely mentioned as involving "metaevaluatoin," which involves verifying the design, process, and analysis of evaluations. However, specific methods or procedures used to conduct the research are not described. Strengths: Mentions a broad method (metaevaluatoin) and aspects it will cover (design,	Assessment: The methods section briefly describes the intention to construct and adjust cognitive models but lacks specific details on how this will be achieved. The text suggests the use of diagnostic cognitive models but does not describe the procedures, tools, or steps involved. Strengths: Mentions the general method (constructing	Assessment: The methods section outlines the goal to develop and validate an instrument based on a theoretical framework. However, it lacks specific details on how this will be accomplished, such as the exact procedures or steps involved. Strengths: Mentions the development of a model and validation of the instrument according to established	Assessment: The methods section describes a mixed methodology approach, including online tests and interviews. It lists two types of interviews conducted (collective and semi-structured) and the participants involved. Strengths: Provides some details on the methods and the types of data collected. Weaknesses: There are several	Assessment: The methods section is vague and does not provide sufficient detail on how the study was conducted. Strengths: Mentions that an intervention was involved. Weaknesses: Lacks specific details about the methods used to study the schools. The term "This intervention is quality" is unclear and does not describe the methodology.	Assessment: The methods section is mentioned but is not clearly described. Strengths: States that policy documents will be analyzed and that interviews will be conducted with actors and policy makers. Weaknesses: The description of methods is vague and lacks detail. Terms like "pretends to analysis" and "interviewed actors involve" are	Assessment: The methods section is missing. Strengths: No methods are outlined, so there is no information to evaluate. Weaknesses: The methods section should detail how the study will be conducted or what existing procedures or models will be used. Since this information is not provided, it is not possible to assess the methods.	Assessment: The methods section is vague and unclear. It attempts to describe the methodology but lacks detail and coherence. Strengths: It mentions that interviews and focus groups (grupo focal) are used. Weaknesses: The methods are described imprecisely with several grammatical errors ("analysis it result rhe diagnostic was throught the	Assessment: The methods section mentions the use of a mixed-method approach, combining both quantitative and qualitative strategies. However, it lacks detailed information on how these methods will be applied. Strengths: The use of a mixed-method approach is identified, which indicates a comprehensive approach to

	<p>process, analysis). Weaknesses: Lacks detailed explanation of how the methods will be applied or any specific models, procedures, or tools used in the investigation.</p>	<p>and adjusting cognitive models). Weaknesses: Lacks detail on how the models will be constructed and adjusted or how the research will be conducted. More specific information on the methodology would improve clarity.</p>	<p>standards. Weaknesses: Does not provide detailed descriptions of the methods used to develop the model, how the instrument will be tested, or specific procedures for validation.</p>	<p>grammatical errors, and the explanation of the methodology is somewhat unclear. The description of the "mixed methodology" could be more precise, and the text lacks specific details about the procedures for the online tests and interviews.</p>		<p>grammatically incorrect and unclear. The methods section should provide a more precise and detailed account of how the study will be conducted.</p>	<p>interviews a her teachers" should be revised). The text does not clearly explain how the data was collected or analyzed.</p>	<p>research. Weaknesses: The description of the methods is not detailed. There is no information about specific procedures, data collection techniques, or analysis methods. The mention of "mix method" and "optimazation" (likely meant to be "optimization") are not sufficiently elaborated upon.</p>	
<i>Resultados</i>	<p>Assessment: The text does not provide any results or findings of the research. There is no information on what was discovered, accomplished, or produced through the study. Strengths:</p>	<p>Assessment: The text does not provide any results or findings of the research. There is no information on what was discovered, accomplished, or produced through the study. Strengths:</p>	<p>Assessment: The text does not include any results or findings from the research. It only describes the intended approach and objectives. Strengths: None provided in the text. Weaknesses: There is no</p>	<p>Assessment: The results section mentions that the study aims to provide descriptive and comparative information but does not present specific findings or results. Strengths: States the type</p>	<p>Assessment: The results section states that the rules and strategies in the two schools are different but does not provide detailed findings or specific information on what was</p>	<p>Assessment: The results section provides some insights into the difficulties faced but lacks a clear, specific summary of findings. Strengths: Mentions issues like insufficient budget and</p>	<p>Assessment: The results section is not present. Strengths: No results are mentioned, so there is nothing to critique here. Weaknesses: The text does not include any findings or explanations of</p>	<p>Assessment: The results section is very brief and lacks specificity. Strengths: It identifies a general finding that there is a need for training (capation) in project methodologies. Weaknesses:</p>	<p>Assessment: There is no results section in the text. The text is still in the diagnostic stage and does not provide any specific results or findings. Strengths: None in this section, as results are not provided.</p>

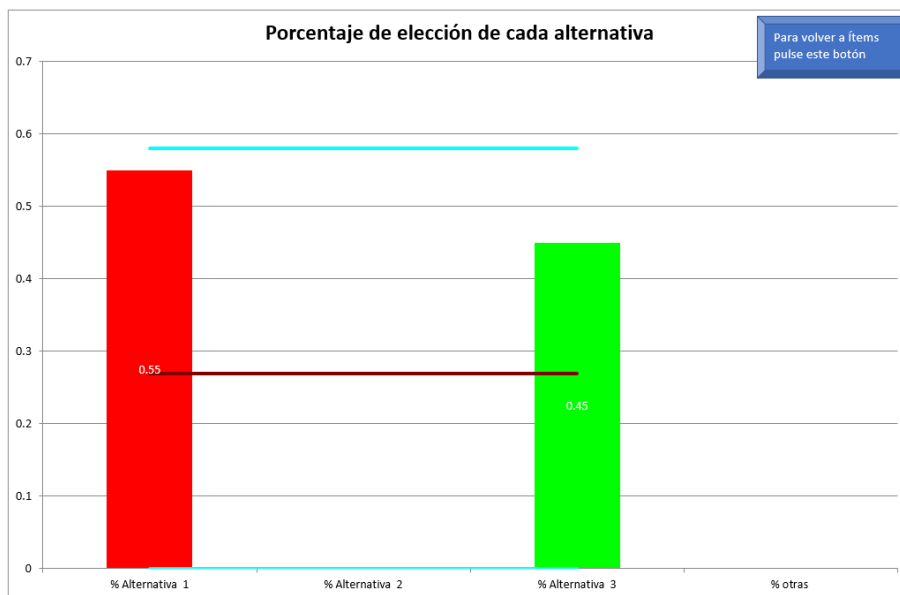
<p>None provided in the text. Weaknesses: No results or outcomes are mentioned, making it impossible to evaluate what was achieved or found.</p>	<p>None provided in the text. Weaknesses: No results or outcomes are mentioned.</p>	<p>discussion of what was discovered, accomplished, or produced from the research.</p>	<p>of results expected (descriptive and comparative). Weaknesses: No actual results or findings are provided. The section is more about what the study aims to achieve rather than what was discovered or accomplished.</p>	<p>discovered. Strengths: Provides a general observation about the difference in rules and strategies. Weaknesses: The results are not sufficiently detailed or specific. The statement "Because the familys organize the rules" is vague and does not explain how it was determined or its significance.</p>	<p>operational inequalities. Weaknesses: Results are not specifically detailed or organized. The text lacks clarity in presenting what was discovered or accomplished. The explanation is broad and does not provide specific results or data.</p>	<p>what was discovered, accomplished, or produced.</p>	<p>The results are not detailed or specific. It does not provide a clear explanation of what was discovered or accomplished from the study. The term "capation" seems to be a typographical error for "capacitation" or "training," which impacts clarity.</p>	<p>Weaknesses: The text does not offer any information about what has been discovered, accomplished, or collected so far.</p>
--	---	--	---	---	--	--	--	---

Anexo 15

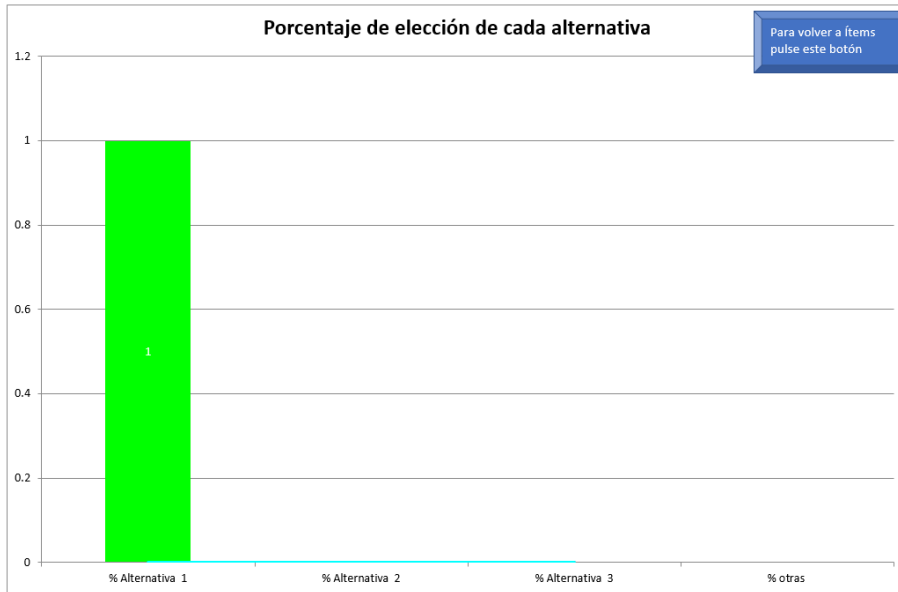
Información de ítem



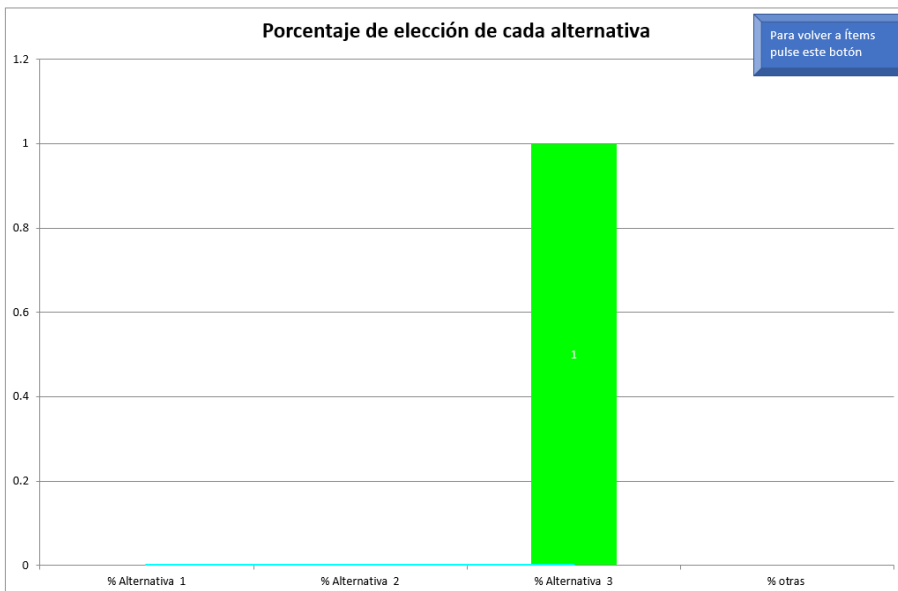
Nota. Ítem 1



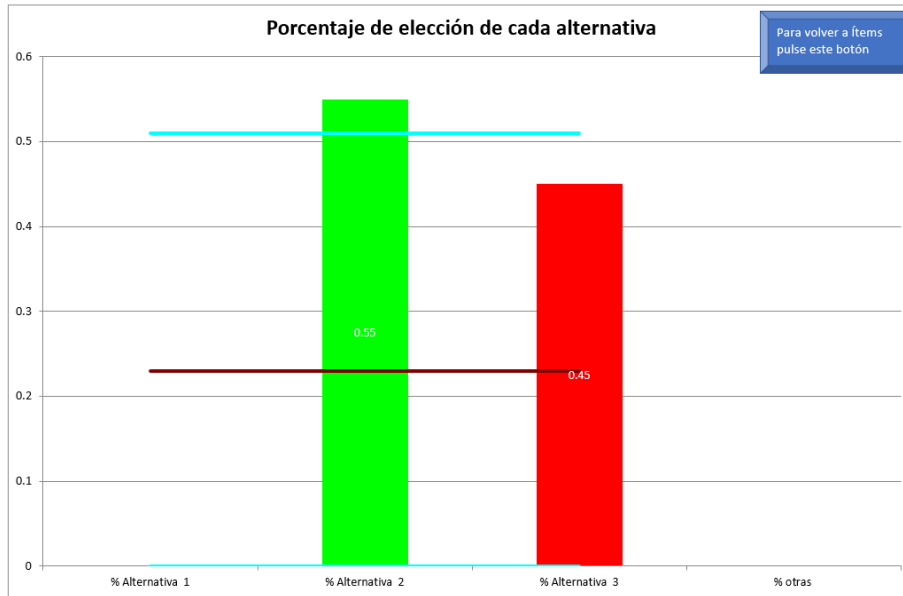
Nota. Ítem 2



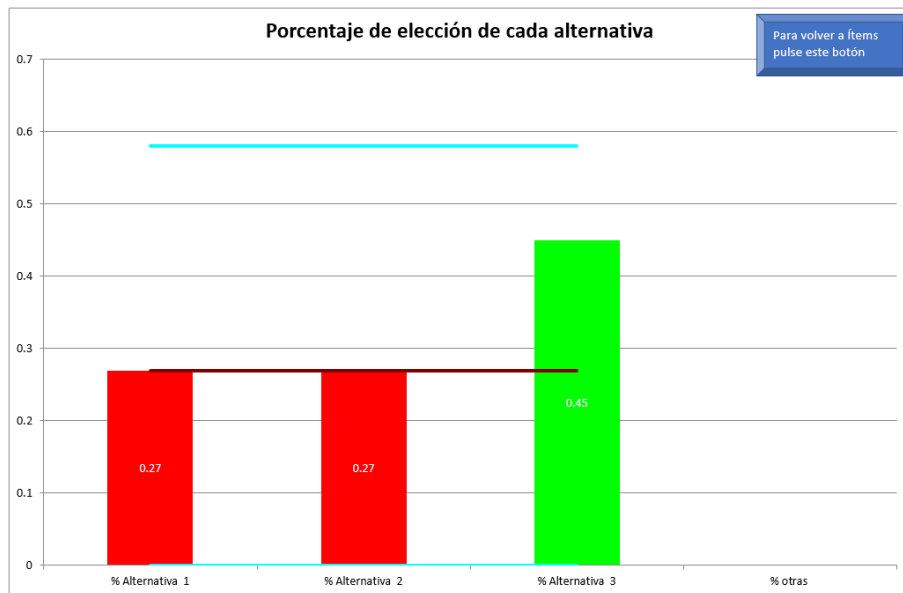
Nota. Ítem 3



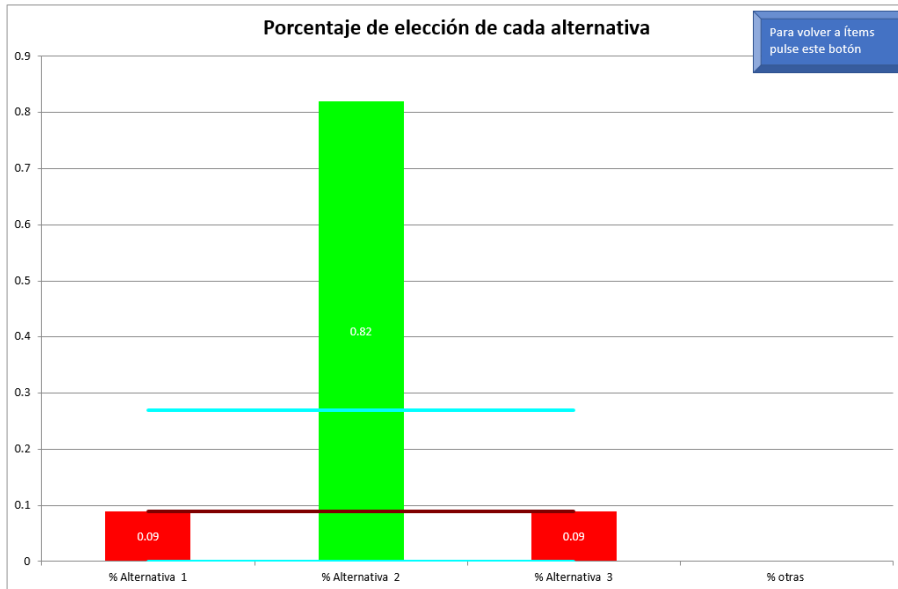
Nota. Ítem 4



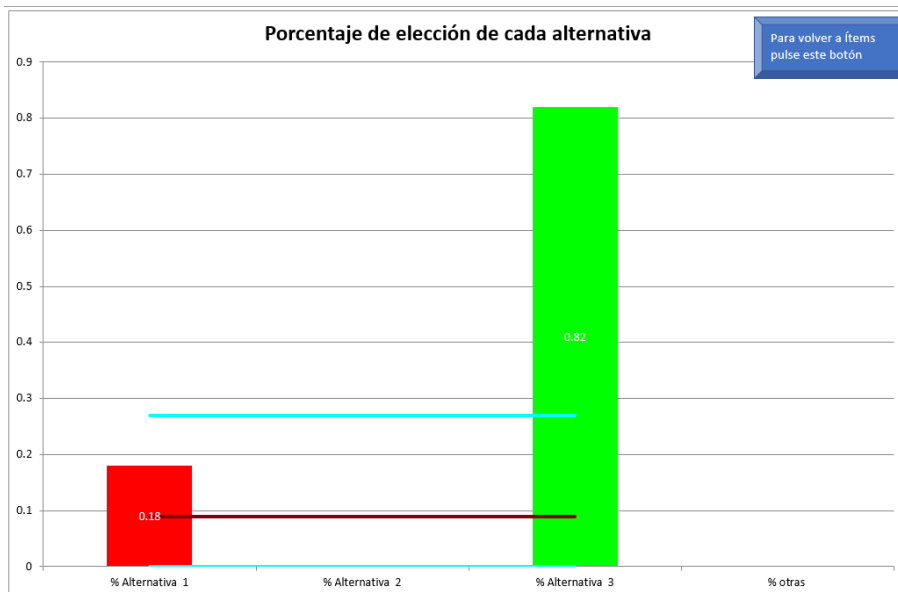
Nota. Ítem 5



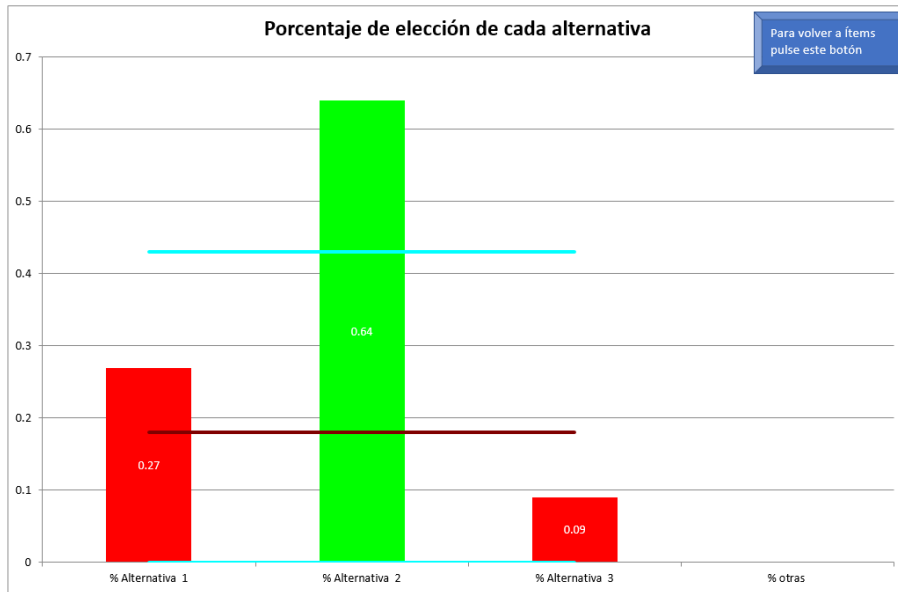
Nota. Ítem 6



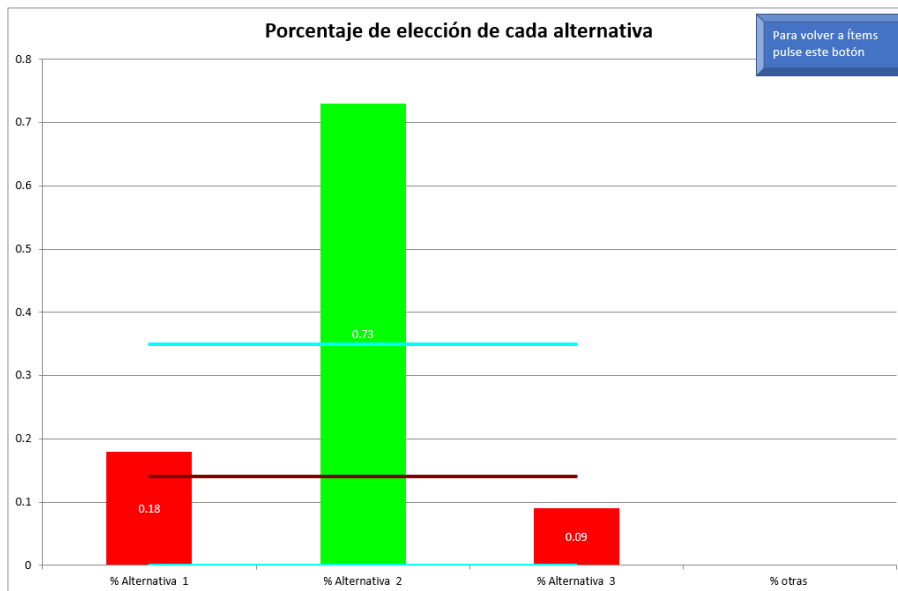
Nota. Ítem 7



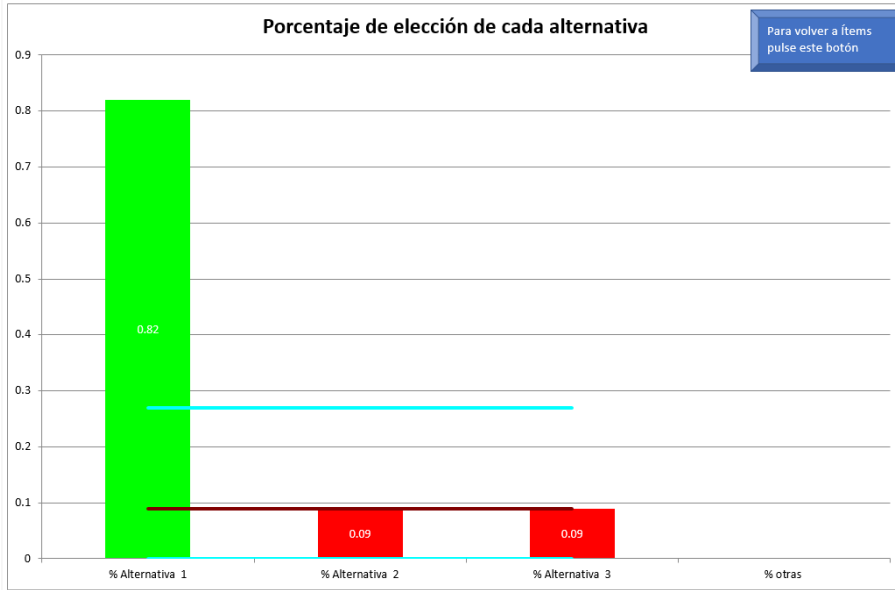
Nota. Ítem 8



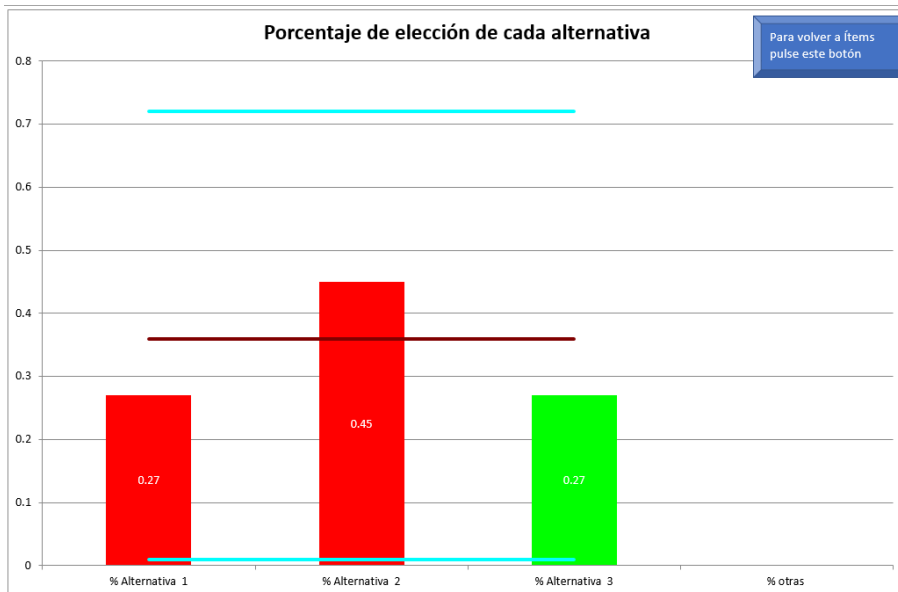
Nota. Ítem 9



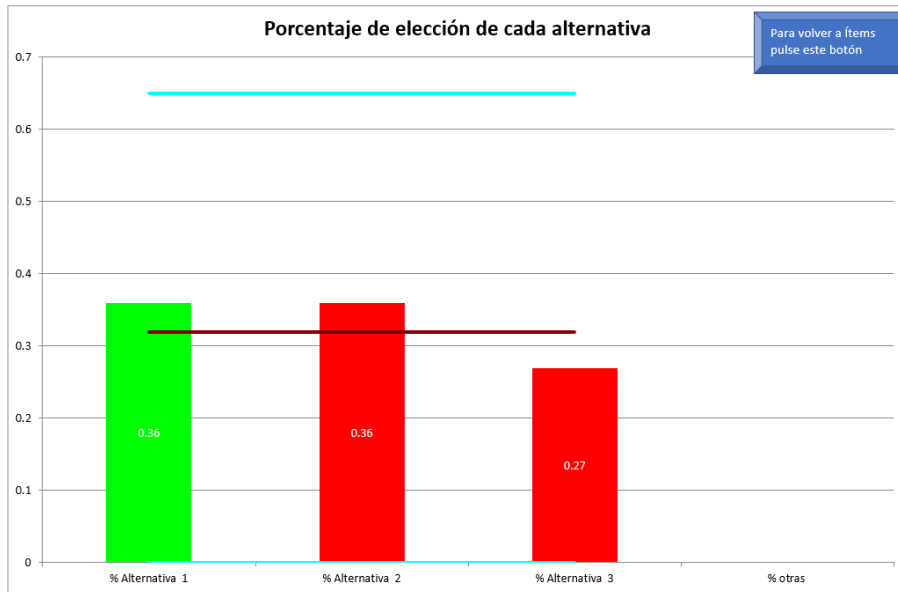
Nota. Ítem 10



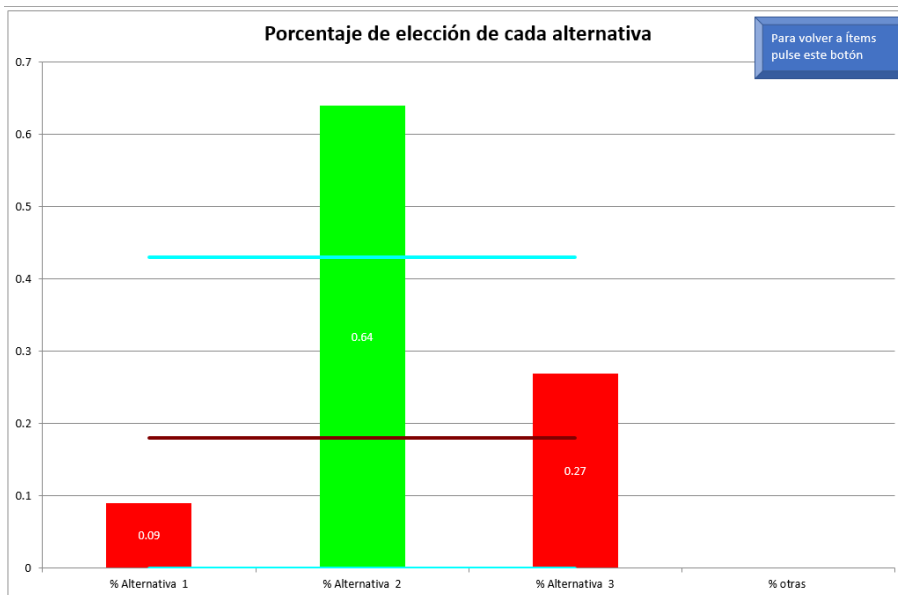
Nota. Ítem 11



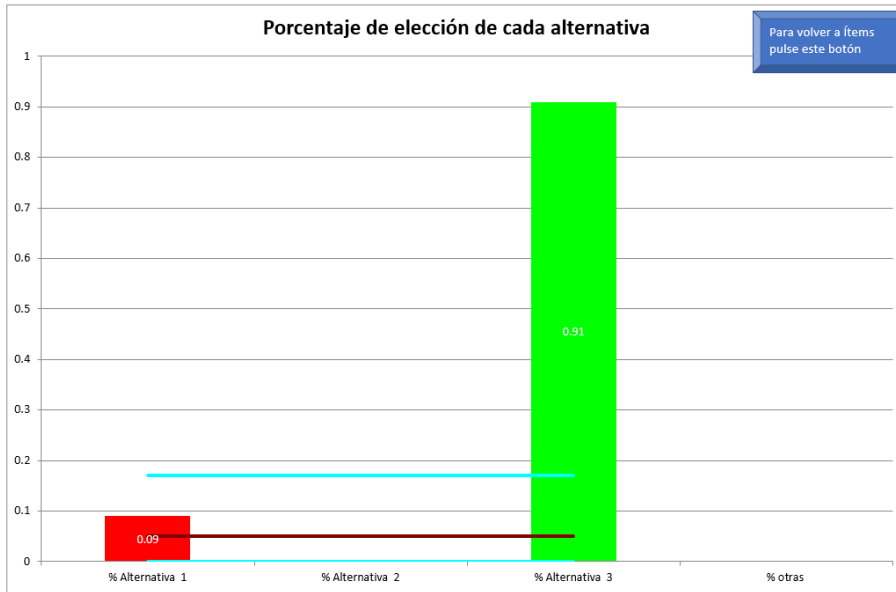
Nota. Ítem 12



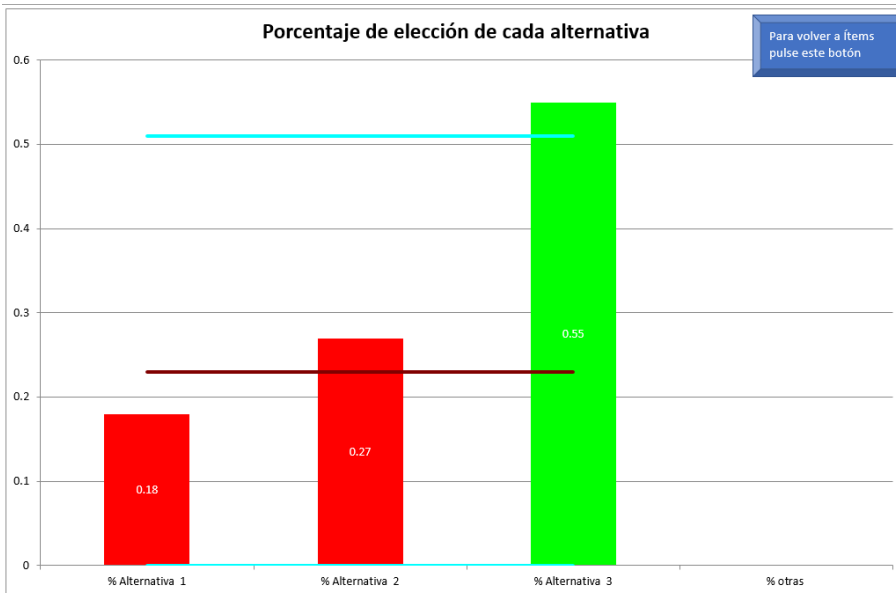
Nota. Ítem 13



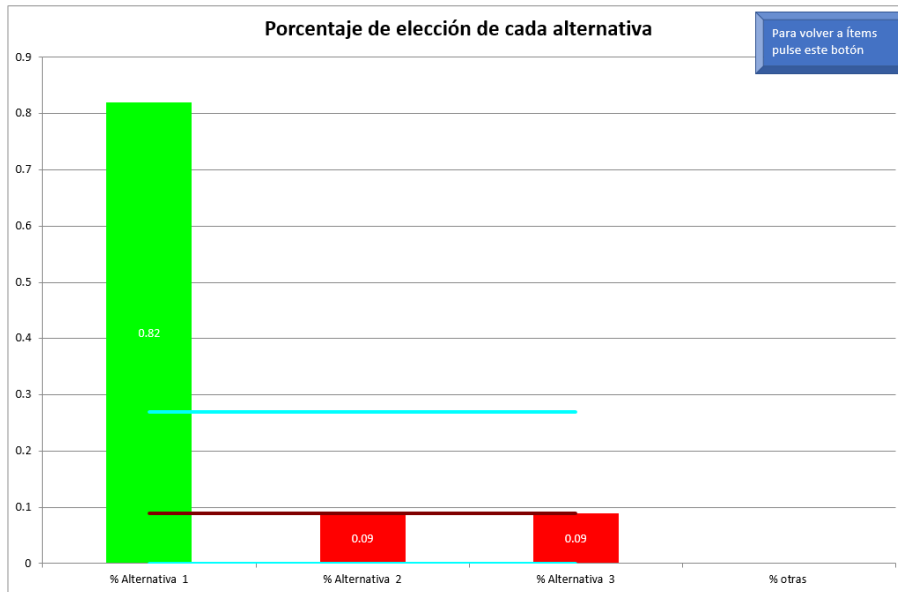
Nota. Ítem 14



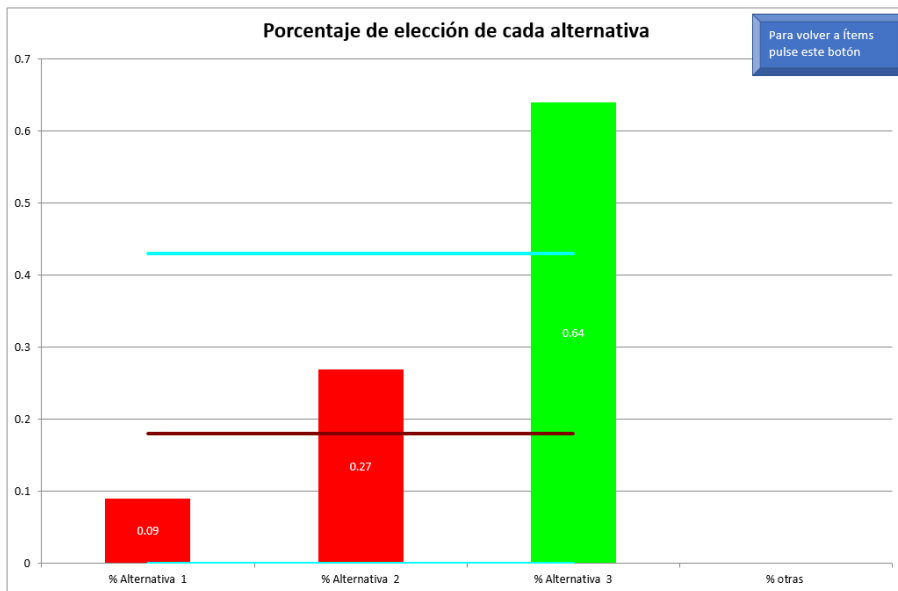
Nota. Ítem 15



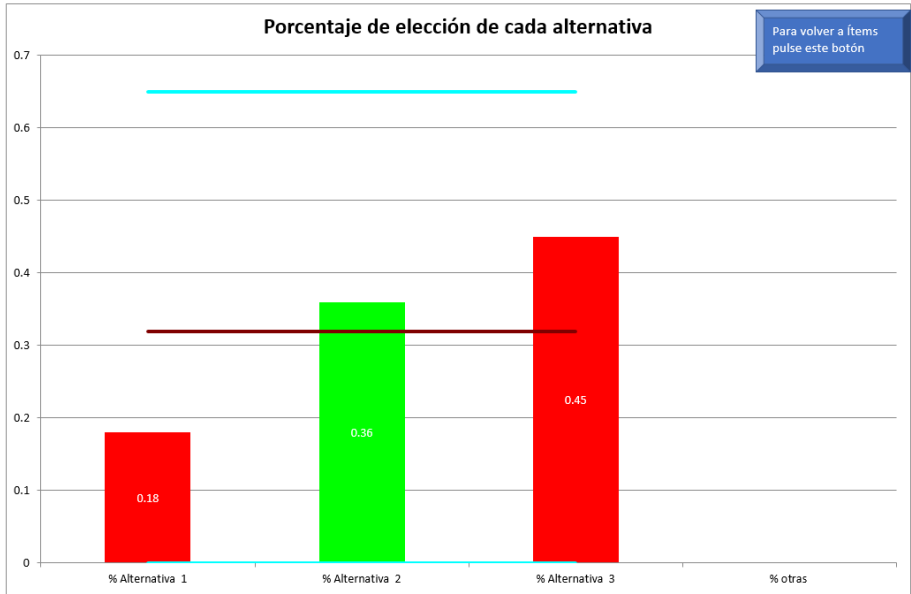
Nota. Ítem 16



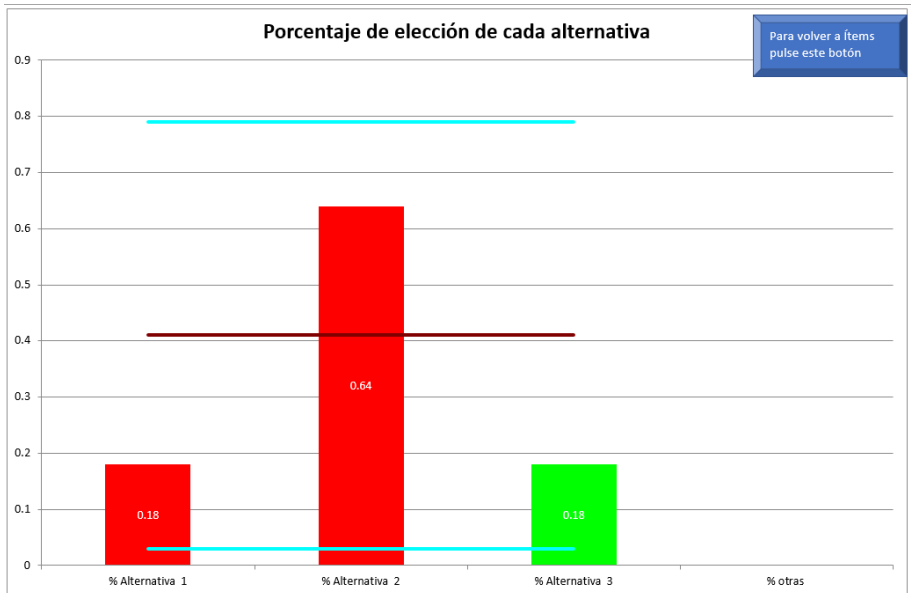
Nota. Ítem 17



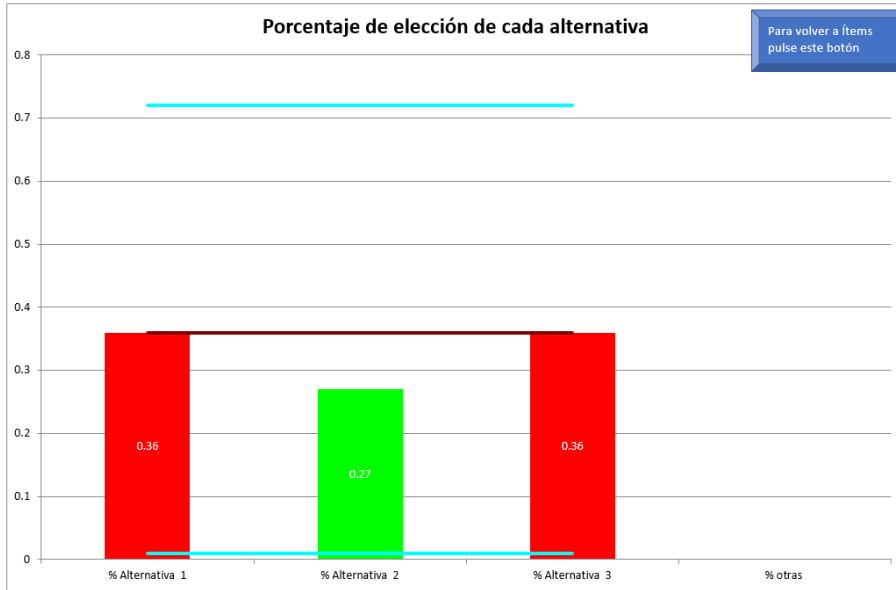
Nota. Ítem 18



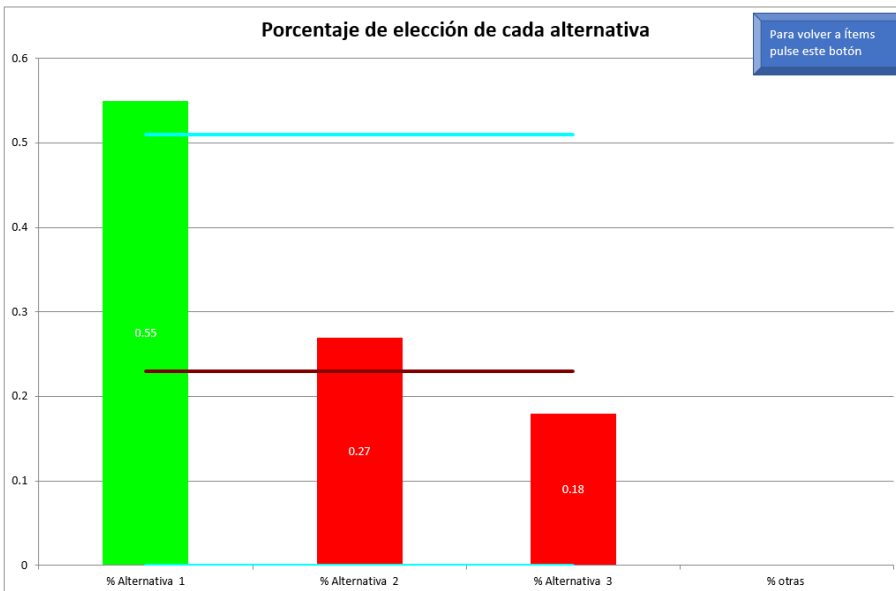
Nota. Ítem 19



Nota. Ítem 20



Nota. Ítem 21



Nota. Ítem 2