

DEGT-UNAH

UNIVERSIDAD NACIONAL AUTONOMA DE HONDURAS
FACULTAD DE CIENCIAS SOCIALES
ESCUELA DE CIENCIAS PSICOLOGICAS
MAESTRÍA EN PSICOMETRÍA Y EVALUACIÓN EDUCATIVA



Tesis

Prueba de Aptitudes de Pensamiento Básico de la Universidad Pedagógica Nacional Francisco Morazán: Análisis de sus características psicométricas y relación con rendimiento académico

Presentada por:

Julio Fernando Cerrato Mejía

Para optar al grado de Máster en Psicometría y Evaluación Educativa

Asesor temático y metodológico:

German Edgardo Moncada Godoy, PhD

Tegucigalpa, M.D.C., Honduras

Mayo, 2025

Cesión de derechos

Tegucigalpa, M.D.C., 25 de mayo del 2025

Señores

Universidad Nacional Autónoma de Honduras

Coordinadores Maestría en Psicometría y Evaluación Educativa

Estimados señores:

Por este medio autorizo a la UNAH para que publique sin fines de lucro en su revista digital, web, cd y cualquier otro medio de difusión de productos de conocimiento para el uso de beneficios sociales y académicos de la sociedad hondureña mi trabajo de tesis titulado: **“Prueba de Aptitudes de Pensamiento Básico de la Universidad Pedagógica Nacional Francisco Morazán: Análisis de sus características psicométricas y relación con rendimiento académico”**, el cual he presentado como parte de los requisitos para optar por la Maestría en Psicometría y Evaluación Educativa.

Atentamente,

Julio Fernando Cerrato Mejía
Cuenta No. PSE100307

AUTORIDADES UNIVERSITARIAS

Doctor Odir Aarón Fernández Flores
Rector

Doctora Lourdes Rosario Murcia Carbajal
Vice-Rectora Académica

Abogado José Alexander Ávila
Secretario General

Máster Oscar Arquímedes Zelaya Villafranca
Director Sistema de Estudios de Postgrado

Máster Carmen Julia Fajardo
Decana Facultad de Ciencias Sociales

Doctora María Victoria Ponce Mendoza
Coordinadora General Postgrados Facultad de Ciencias Sociales

Doctor German Edgardo Moncada Godoy
Coordinador Académico Maestría en Psicometría y Evaluación Educativa

Doctor Marcio Alexander Castillo Díaz
Coordinador de Investigación y Vinculación

Dedicatoria

A la memoria de mi padre, quien, con orgullo y entusiasmo, siempre encontró en mis proyectos y logros motivo de conversación y aliento. Su partida coincidió con el inicio de este camino académico, dejándome con el silencio de sus palabras, pero con la fuerza de su ejemplo y su amor incondicional. Este trabajo es también suyo, porque su presencia sigue viva en cada decisión, en cada esfuerzo y en cada paso dado hacia esta meta.

Agradecimientos

Durante el proceso de elaboración de esta tesis, muchas personas e instituciones me brindaron su apoyo y acompañamiento, haciendo posible la culminación de este proyecto académico. Dedico este espacio a expresar mi sincero agradecimiento por su colaboración, confianza y generosidad.

En primer lugar, agradezco mi asesor académico doctor German Moncada, por su orientación constante, la claridad de sus observaciones y su disposición durante todo el proceso de investigación. Asimismo, al doctor Marcio Castillo, que, al haber impartido los seminarios de investigación científica y el taller de tesis, los cuales fueron esenciales para establecer el enfoque metodológico de este estudio. También expreso un agradecimiento especial a maestra Lina Mendoza, por su acompañamiento cercano a lo largo de toda la maestría, demostrando siempre disposición y compromiso con nuestro proceso formativo.

Extiendo mi reconocimiento a la Universidad Pedagógica Nacional Francisco Morazán (UPNFM), por ser una institución abierta al pensamiento crítico, al análisis riguroso y al fortalecimiento de la educación. A las autoridades universitarias, agradezco su confianza, así como su apertura y cercanía para atender y facilitar los requerimientos necesarios para llevar a cabo esta investigación.

En el ámbito personal, manifiesto mi profunda gratitud a mi familia, por su apoyo incondicional, comprensión y paciencia durante los momentos en que debí priorizar el estudio por encima del tiempo compartido. También reconozco a mis compañeros de trabajo, en particular a Osman Borjas y Ricardo López, por su constante respaldo técnico, profesional y moral, el cual resultó invaluable en diversas etapas del proceso.

A mis compañeros y compañeras de la maestría, les agradezco por haber construido un vínculo de amistad y solidaridad, incluso en medio de la distancia física. El compañerismo y la empatía demostrada por todos fueron factores motivadores que enriquecieron mi experiencia académica.

Finalmente, agradezco a todas las personas que, de forma directa o indirecta, contribuyeron a la realización de esta tesis. Cada gesto de colaboración, consejo y palabra de aliento fue esencial para llegar a buen término con este importante capítulo académico.

Resumen

El presente estudio tuvo como propósito evaluar las propiedades psicométricas de la prueba de Aptitudes de Pensamiento Básico utilizada en la Universidad Pedagógica Nacional Francisco Morazán (UPNFM), así como su relación con el rendimiento académico de los estudiantes de primer ingreso. La investigación se desarrolló bajo un enfoque cuantitativo, de tipo no experimental y alcance correlacional, con la participación de 1,531 estudiantes de diferentes carreras académicas. La prueba evaluada está compuesta por cinco subescalas: Discernimiento Verbal, Habilidad Matemática, Pensamiento Lógico, Apreciación Espacial y Elocuencia Verbal. Se aplicaron análisis psicométricos clásicos para estimar la dificultad, discriminación y confiabilidad de cada ítem, así como análisis factoriales exploratorios y confirmatorios para validar la estructura interna de la prueba. Los resultados mostraron una estructura factorial consistente en cuatro de las cinco subescalas, con niveles aceptables de confiabilidad y validez estructural. La subescala de Elocuencia Verbal, sin embargo, presentó una ausencia total de consistencia interna y dificultad extrema, por lo que fue excluida de los modelos factoriales confirmatorios. Adicionalmente, se identificaron relaciones estadísticamente significativas entre los puntajes obtenidos en la prueba y el rendimiento académico posterior, aportando evidencia a favor de la validez predictiva del instrumento. No obstante, el modelo de segundo orden propuesto mostró un ajuste limitado, lo que sugiere la independencia relativa entre las subescalas. La prueba de Aptitudes de Pensamiento Básico presenta propiedades psicométricas satisfactorias en la mayoría de sus componentes y puede constituirse en una herramienta útil para procesos de orientación académica y toma de decisiones institucionales, siempre que se continúe con su refinamiento y validación continua.

Palabras clave: pruebas psicométricas, validez, confiabilidad, rendimiento académico, pensamiento básico, análisis factorial.

Índice de Contenido

RESUMEN	7
DEDICATORIA.....	5
AGRADECIMIENTOS.....	6
INTRODUCCIÓN	1
<i>1. Planteamiento del Problema</i>	<i>3</i>
1.1. SITUACIÓN PROBLEMÁTICA	3
1.2. PREGUNTA PROBLEMA.....	6
1.3. OBJETIVOS DE LA INVESTIGACIÓN	6
<i>1.3.1. Objetivo General.....</i>	<i>6</i>
<i>1.3.2. Objetivos Específicos</i>	<i>6</i>
1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN	7
2.1. ORIENTACIÓN VOCACIONAL	8
2.2. PRUEBAS DE ORIENTACIÓN VOCACIONAL.....	10
2.3. HABILIDADES MENTALES PRIMARIAS	10
<i>2.3.1. Comprensión Verbal.....</i>	<i>11</i>
<i>2.3.2. Destreza de Cálculo.....</i>	<i>11</i>
<i>2.3.3. Pensamiento Lógico.....</i>	<i>12</i>
<i>2.3.4. Percepción Espacial</i>	<i>12</i>
<i>2.3.5. Fluidez Verbal.....</i>	<i>13</i>
2.4. RENDIMIENTO ACADÉMICO.....	13
2.5. PSICOMETRÍA	15
<i>2.5.1. Validez</i>	<i>16</i>
<i>2.5.2. Confiabilidad</i>	<i>18</i>
<i>2.5.3. Teoría Clásica de los Test</i>	<i>19</i>
<i>2.5.4. Teoría de Respuesta al Ítem.....</i>	<i>20</i>
3. MARCO CONTEXTUAL.....	21
3.1. PROCESOS DE ORIENTACIÓN VOCACIONAL	21

3.2. LA UNIVERSIDAD NACIONAL PEDAGÓGICA FRANCISCO MORAZÁN (UNPFM)	23
3.2.1 Vicerrectoría de Vida Estudiantil	23
3.3. PROYECTO DE FORTALECIMIENTO DE LA GESTIÓN ACADÉMICO	
ADMINISTRATIVO DE LA UPNFM MEDIANTE EL APROVECHAMIENTO DE LA ENERGÍA	
FOTOVOLTAICA	24
4.1. TIPO DE INVESTIGACIÓN	25
4.3. POBLACIÓN Y MUESTRA	28
4.3.1. Población	28
4.3.2. Muestra	28
4.3.3. Justificación del Tamaño de la Muestra	29
4.3.4. Representatividad de la Muestra	29
4.4.1. Estructura de la Prueba	30
4.4.2. Puntuación e Interpretación	30
4.4.3. Proceso de Validación del Instrumento	31
4.4.4. Limitaciones en la Documentación Técnica	31
4.5 . PLAN DE ANÁLISIS DE DATOS	31
4.5.1 Confiabilidad por Consistencia Interna	32
4.5.2 Validez	32
4.5.3. Análisis de Rendimiento Académico	33
4.5.6. Validación Cruzada	34
4.6. PROCEDIMIENTOS	34
4.6.1. Recolección de Datos	34
4.6.2. Transferencia de Datos	35
4.6.3. Herramientas y Tecnologías Utilizadas	35
4.6.4. Protocolo de Manejo de Datos	35
4.6.5. Verificación de la Integridad de los Datos	35
4.6.6. Documentación de Procedimientos	36
4.6.7. Revisión y Validación de los Procesos	36
5. RESULTADOS DE LA INVESTIGACIÓN	37
5.1. ANÁLISIS DESCRIPTIVO DE LA MUESTRA Y DE LAS PUNTUACIONES	38

5.2. ANÁLISIS FACTORIAL EXPLORATORIO (AFE)	40
5.2.1. Subescala de Discernimiento Verbal	40
5.2.2. Subescala de Destreza de Cálculo	42
5.2.3. Subescala de Pensamiento Lógico	43
5.2.4. Subescala de Apreciación Espacial	44
5.2.5. Subescala de Elocuencia Verbal	46
5.3. ANÁLISIS FACTORIAL CONFIRMATORIO (AFC)	46
5.3.1. Subescala de Discernimiento Verbal	46
5.3.2. Destreza de Cálculo	49
5.3.3. Subescala Apreciación Espacial	51
5.3.4. Subescala Pensamiento Lógico	54
5.3.5. Subescala de Elocuencia Verbal	56
5.8. ANÁLISIS DE LOS ÍTEMS	66
5.8.1. Subescala de Discernimiento Verbal	66
5.8.2. Análisis Psicométrico de los Ítems de la Subescala Habilidad Matemática	68
5.8.3. Análisis Psicométrico de los Ítems de la Subescala de Apreciación Espacial	70
5.8.4. Análisis Psicométrico de la Subescala de Pensamiento Lógico	72
5.9. VALIDEZ CONVERGENTE Y DISCRIMINANTE	74
5.9.1. Validez Convergente	74
5.9.2. Validez Discriminante	75
6. DISCUSIÓN DE LOS RESULTADOS	76
6.1. ANÁLISIS DE LAS CARACTERÍSTICAS PSICOMÉTRICAS DE LOS ÍTEMS	76
6.2. VALIDEZ Y CONSISTENCIA INTERNA DE LAS SUBESCALAS	76
6.3. VALIDEZ PREDICTIVA Y RELACIÓN CON EL RENDIMIENTO ACADÉMICO	77
6.4. MODELO GENERAL DE APTITUD COGNITIVA	77
LIMITACIONES DEL ESTUDIO	78
CONCLUSIONES	78
RECOMENDACIONES	79

REFERENCIAS.....	80
-------------------------	-----------

DEGT-UNAH

Índice de Tablas

Tabla 1:Operacionalización de variables en la cual se detallan las dos variables latentes, sus cinco dimensiones con el respectivo número de ítems.	26
Tabla 2 Cargas estandarizadas	52
Tabla 3 Fiabilidad por consistencia interna de la prueba de Aptitudes de Pensamiento Básico	58
Tabla 4 Estadísticos descriptivos por subescalas de la prueba de Aptitudes de Pensamiento Básico	60
Tabla 5 Coeficientes de correlación	64
Tabla 6 Varianza media extraída por subescalas de la prueba de Aptitudes de Pensamiento Básico	74
Tabla 7 Análisis de la validez discriminante entre subescalas	75

Índice de Figuras

Figura 1. <i>Distribución de la edad de los sujetos incluidos en la muestra</i>	38
Figura 2. <i>Distribución de la edad de los participantes incluidos en la muestra en función del género</i>	39
Figura 3. <i>Gráfico de sedimentación para la subescala de Discernimiento Verbal</i> ..	41
Figura 4. <i>Gráfico de sedimentación para la subescala de Destreza de Cálculo</i>	43
Figura 5. <i>Gráfico de sedimentación para la subescala de Pensamiento Lógico</i>	44
Figura 6. <i>Gráfico de sedimentación para la subescala de apreciación Espacial</i>	45
Figura 7. <i>Cargas factoriales estandarizadas para la subescala de discernimiento verbal</i>	47
Figura 8. <i>Modelo gráfico del análisis omega de McDonald para la subescala de Discernimiento Verbal</i>	48
Figura 9. <i>Cargas factoriales estandarizadas por ítems para la subescala Destreza de Cálculo</i>	50
Figura 10. <i>Cargas factoriales por ítem y factor de la subescala Apreciación Espacial</i>	52
Figura 11 <i>Cargas estandarizadas por ítem para la subescala de Pensamiento Lógico</i>	55
Figura 12 <i>Representación gráfica de los factores del modelo</i>	56
Figura 13 <i>Distribución de puntajes por subescalas de la prueba de Aptitudes de Pensamiento Básico</i>	59
Figura 14. <i>Matriz de correlaciones entre subescalas de la prueba de Aptitudes de Pensamiento Básico</i>	61
Figura 15. <i>Comparación del índice académico antes y después de la aplicación de la prueba de Aptitudes de Pensamiento Básico</i>	65
Figura 16. <i>Distribución de los índices de dificultad de la subescala de discernimiento verbal</i>	66
Figura 17. <i>Distribución de los coeficientes de discriminación de la subescala de discernimiento verbal</i>	67
Figura 18. <i>Frecuencia de los ítems por coeficiente de dificultad de la subescala de habilidad matemática</i>	68

Figura 19. Frecuencia de los ítems por coeficiente de discriminación para la subescala de habilidad matemática	69
Figura 20. Distribución de la dificultad de la subescala de apreciación espacial...	70
Figura 21. Distribución de la discriminación de la subescala de Apreciación Espacial	71
Figura 22. Discriminación de ítems según su dificultad para la subescala de Pensamiento Lógico.....	72
Figura 23. Discriminación de los ítems para la subescala de Pensamiento Lógico	73

Introducción

La educación superior contemporánea se enfrenta a múltiples retos que van desde el acceso equitativo hasta la permanencia y culminación oportuna de los estudios. Uno de los problemas más persistentes en este nivel educativo es el rezago académico, el cual se manifiesta en la prolongación del tiempo que un estudiante necesita para completar su formación, respecto al tiempo establecido por el plan de estudios. Esta situación afecta no solo a los estudiantes, quienes pueden verse desmotivados o abandonar sus estudios, sino también a las instituciones educativas, al tener que destinar recursos adicionales para atender cohortes extendidas y reorganizar continuamente la planificación académica.

Numerosos estudios han identificado una multiplicidad de factores que influyen en el rendimiento y permanencia académica, entre los que se destacan las condiciones socioeconómicas, los niveles de motivación, el acompañamiento institucional, la adecuación de los programas educativos y, particularmente, la correspondencia entre las aptitudes cognitivas de los estudiantes y las exigencias de la carrera elegida. En este sentido, la orientación vocacional y el uso de pruebas psicométricas adquieren una función estratégica en la educación superior, ya que permiten valorar objetivamente las capacidades de los estudiantes y brindarles herramientas para tomar decisiones informadas respecto a su futuro académico.

La Universidad Pedagógica Nacional Francisco Morazán (UPNFM), consciente de la importancia de estos elementos, ha desarrollado y aplicado desde el año 2018 la Prueba de Aptitudes de Pensamiento Básico, con el propósito de apoyar procesos de cambio de carrera, asignación de becas y orientación vocacional. Esta prueba evalúa cinco dimensiones fundamentales del pensamiento: discernimiento verbal, destreza de cálculo, pensamiento lógico, apreciación espacial y elocuencia verbal. Las áreas que mide se vinculan directamente con las competencias requeridas para el desempeño académico en distintas carreras universitarias, especialmente aquellas orientadas a la formación docente.

Sin embargo, a pesar del uso continuo de esta herramienta, hasta la fecha no se contaba con un análisis psicométrico riguroso que permitiera establecer con claridad sus niveles de confiabilidad y validez, ni se había documentado de manera formal su capacidad predictiva en relación con el rendimiento académico de los estudiantes que se someten a ella.

Esta ausencia de evidencia pone en cuestión la precisión y utilidad de la prueba para cumplir con los fines con los que fue concebida, haciendo necesario un estudio sistemático que examine su funcionamiento desde una perspectiva técnica y científica.

En ese contexto, el presente estudio tiene como objetivo principal evaluar las propiedades psicométricas de la Prueba de Aptitudes de Pensamiento Básico de la UPNFM, mediante el enfoque de la Teoría Clásica de los Tests (TCT), analizando indicadores como la dificultad, la discriminación, la consistencia interna y la validez predictiva de sus ítems. Asimismo, se pretende explorar la relación entre los resultados obtenidos por los estudiantes en la prueba y su rendimiento académico posterior, con el fin de determinar si existe una asociación significativa que permita justificar su uso como herramienta predictiva.

Este trabajo adopta un enfoque cuantitativo, de tipo correlacional y diseño transversal, empleando como muestra a los estudiantes que han realizado cambios de carrera utilizando esta prueba como requisito. A partir de los datos extraídos de los sistemas académicos institucionales y de la aplicación del instrumento psicométrico, se efectuaron análisis estadísticos que permitieron establecer la calidad técnica de la prueba y su utilidad para fines institucionales.

La relevancia de este estudio es doble. Por un lado, ofrece evidencia empírica para fortalecer el proceso de toma de decisiones dentro de la UPNFM, en lo que respecta a la mejora y adecuación de sus procesos de evaluación y orientación. Por otro, aporta al campo de la psicometría aplicada en contextos educativos latinoamericanos, mostrando la importancia de validar los instrumentos de evaluación en función de las características específicas del entorno cultural, académico y tecnológico.

Se espera que los hallazgos derivados de esta investigación no solo contribuyan al perfeccionamiento del instrumento, sino que también abran la posibilidad de replicar estudios similares en otros contextos institucionales. En última instancia, se aspira a que la validación de esta prueba se traduzca en una herramienta más robusta y confiable que, alineada con las políticas de equidad y calidad educativa, contribuya al logro de una educación superior más justa, eficiente y orientada al éxito académico del estudiante.

1. Planteamiento del Problema

1.1. Situación Problemática

En la educación superior un tema que ha sido estudiado ampliamente es la situación del rezago estudiantil, tiempo que tarda un estudiante en culminar sus estudios en comparación con lo que estipulan los planes de estudio, problema que puede ser atribuido a distintas causas internas y externas a la universidad; en las externas se pueden atribuir a situaciones socio-económicas que afronta el estudiante; y las internas debido a poca integración del estudiante en la universidad (Del Valle & Cumsille, 2019). En Costa Rica existen estudios que manifiestan que en la carrera de educación primaria el rezago puede ser atribuido a motivos familiares y socioeconómicos (Montero-Méndez et al., 2021).

De forma similar en la Universidad Autónoma de Baja California (UABC), México concluyen que para ellos el factor más importante al que se puede atribuir el fracaso estudiantil que es como se suele definir el abandono y al rezago, se deben a la falta de preparación previa y a un apoyo institucional deficiente y es que el rezago es un problema que afecta a la educación superior, en el caso de México la tasa de graduación fue del 24% y un abandono de 244,210 estudiantes y que el 68.91% están en condición de rezago (Murillo-García & Luna-Serrano, 2021).

Constante-Amores et al. (2021) también hablan de los factores asociados al abandono universitario e indican que el factor académico dentro de las cuales están el rendimiento previo y la motivación del estudiante respecto a la carrera elegida son determinantes en el abandono o no de los estudiantes. Por lo que es necesario que las universidades creen estrategias de acompañamiento y apoyo para promover la permanencia y el éxito académico estrategias que pueden ser orientar y capacitar a los estudiantes previo a la elección de una carrera universitaria y con mecanismos de motivación y orientación.

La orientación vocacional es uno de los mecanismos que pueden ayudar en la lograr mejorar los índices de abandono y de rezago debido a que ayuda a los estudiantes que ingresan a la educación universitaria a tomar decisiones informadas que afectarían su probabilidad de éxito académico. Según Erazo Guerra & Rosero Morales (2021) una correcta orientación vocacional permite que los estudiantes conozcan sus habilidades y como se

pueden alinear con sus intereses para la elección de una carrera, para que puedan mantener una adecuada motivación y así lograr un mejor desempeño.

Por su parte, Yalanda Tombé, (2022) indica que la orientación vocacional es un componente básico para el desarrollo de un plan de vida de los estudiantes, porque a través del proceso de orientación los estudiantes pueden reflexionar sobre sus habilidades y limitaciones para poder tomar decisiones adecuadas en lo que concierne a su formación en la universidad. En consonancia con lo mencionado anteriormente otros autores mencionan sobre que, la orientación vocacional ayuda a tomar decisiones que han sido reflexionadas sobre los intereses y que contribuye no solo al éxito en los estudios universitarios si no que contribuye a la satisfacción personal y profesional (Ramos Monsiváis & González, 2020).

De acuerdo con lo anterior si un estudiante no está en una carrera que desea estudiar o que no tiene las aptitudes necesarias para el desarrollo de estas puede influir en el rendimiento dentro de la carrera acabando en rezago o abandono en los estudios. Para una institución académica es una situación problemática, ya que el tener repitencia y abandonos altos conlleva en un gasto de recursos y dificulta la planificación académica porque hay acumulación de matrícula, y dentro del aula también se presentan dificultades cuando hay presencia de estudiantes rezagados debido a que la planificación se hace en función de los estudiantes no tienen conocimiento de la asignatura (Izaguirre Bordelois et al., 2019).

Para poder ayudar al estudiante a seleccionar una carrera donde pueda desarrollar de la mejor forma sus habilidades y aptitudes de forma que su aprovechamiento académico sea optimo es necesario contar con una prueba que sea capaz de medir y predecir de forma correcta cómo será el rendimiento del estudiante en una carrera (Izaguirre Bordelois et al., 2019).

Actualmente en Honduras se pueden encontrar en el mercado varias pruebas de orientación vocacional y de aptitudes, pero la mayoría son trabajos realizados para otras culturas y contextos, y que en algunos casos ofrecen baremos hondureños en las revisiones Online, pero que resultan costosas para aplicaciones masivas. En Honduras en la década de 1980 de manera inicial y posteriormente en 2014 ha habido trabajos que buscan la baremación de pruebas para tener datos normativos que respondan al contexto nacional (Mendoza Recarte, 2017).

Es así como la UPNFM en 2018 encarga la creación de la Prueba de Aptitudes de Pensamiento Básico como parte de un conjunto de pruebas para poder brindar orientación vocacional para la selección de carreras y el otorgamiento de becas académicas. Pero después de varios años de uso uno de los problemas que enfrenta la UPNFM es la falta de evidencia sólida respecto a la validez predictiva de su Prueba de Aptitudes de Pensamiento Básico. Ya que no se ha revisado si los resultados que esta prueba brinda son buenos predictores del rendimiento académico que los estudiantes tendrán en su vida estudiantil en la universidad.

Por lo descrito en el epígrafe anterior se considera de suma importancia eliminar toda incertidumbre en el funcionamiento de la prueba en cuestión, ya que la efectividad de esta prueba de orientación está asociada a la capacidad para predecir con precisión el rendimiento académico.

La validez predictiva de una prueba es fundamental en el ámbito educativo, ya que se refiere a la capacidad de un instrumento para predecir el desempeño futuro de los individuos en actividades académicas. Es por ello que en la UPNFM, es esencial que la prueba no solo mida las habilidades de los estudiantes en el momento de su aplicación, sino que ofrezca una predicción confiable de su rendimiento futuro en la universidad pues, sin esta capacidad predictiva, la utilidad de la prueba es limitada ya que se corre el riesgo de que la prueba no prediga el rendimiento académico y que los estudiantes sean aceptados en carreras para las que no tienen las aptitudes necesarias y aumente el rezago y abandono.

Así mismo para la universidad es un problema ya que no puede darse una correcta planificación y asignación de recursos cuando existe un alto nivel de rezago y abandono. En las últimas décadas ha aumentado el uso de técnicas matemáticas y estadísticas que permiten valorar la confiabilidad y la validez de las pruebas educativas; incluyendo los reactivos que son utilizados para su construcción, todo esto para mejorar los procesos de evaluación. En este sentido es importante el poder evaluar como es el funcionamiento de dicha prueba su confianza y validez para poder predecir el rendimiento y aprovechamiento educativo de los estudiantes que cambian de una carrera (Menacho, 2020).

Por todo ello, este estudio contribuirá en el campo de la psicometría aplicada en la educación, proporcionando una evaluación exhaustiva de la validez predictiva de la Prueba de Aptitudes de Pensamiento Básico de la UPNFM. Los resultados permitirán a la

universidad tomar decisiones informadas sobre la utilidad de esta prueba en la orientación vocacional y en los procesos de cambio de carrera y otorgamiento de becas; ofreciendo oportunidades de mejora para la prueba haciéndola más efectiva y pudiendo ampliar su uso en otras áreas de interés.

1.2. Pregunta Problema

¿Cuáles son las propiedades psicométricas de la prueba de aptitudes de pensamiento básico en la UPNFM y su relación con el rendimiento académico?

1.3. Objetivos de la Investigación

1.3.1. Objetivo General

Evaluar las propiedades psicométricas de la prueba Aptitudes de pensamiento básico de la UPNFM y su relación con el rendimiento académico de los estudiantes que se han sometido a la prueba.

1.3.2. Objetivos Específicos

- Medir las características psicométricas mediante la Teoría Clásica de los Test de la prueba Aptitudes de Pensamiento Básico de la UPNFM.
- Determinar la confiabilidad de la prueba de Aptitudes de Pensamiento Básico de la UPNFM por medio del Alpha de Cronbach y el Omega de McDonald.
- Evaluar la validez predictiva de la prueba psicométrica mediante la comparación de los resultados de la prueba de Aptitudes de Pensamiento Básico de la UPNFM, con el rendimiento académico de los estudiantes.

1.4. Justificación de la Investigación

La validación de la Prueba de Aptitudes de Pensamiento Básico de la Universidad Pedagógica Nacional Francisco Morazán (UPNFM) es esencial para garantizar una adecuada orientación vocacional y una asignación justa de becas. Y dará nuevos elementos a la teoría educativa y la práctica psicométrica. Este estudio es relevante porque sus resultados beneficiarán tanto a la universidad como a sus estudiantes. Además, podrían ser útiles para otras instituciones educativas y desarrolladores de pruebas que busquen mejorar sus procesos de orientación en contextos similares (Erazo Guerra & Rosero Morales, 2021).

La importancia de este estudio está en que validar la prueba asegura resultados confiables y precisos que ayudara a mejorar las decisiones relacionadas con la orientación vocacional y la asignación de las becas. Una prueba efectiva de aptitudes puede desempeñar un papel crucial en la reducción de la tasa de abandono escolar y en la mejora del rendimiento académico. Al garantizar que los estudiantes se matriculen en carreras más acorde a sus habilidades e intereses, para alcanzar el éxito académico (Yalanda Tombé, 2022).

Diversos estudios han demostrado que una adecuada correspondencia entre las habilidades del estudiante y la elección de carrera es un predictor significativo del éxito académico, lo que subraya la importancia de contar con herramientas de evaluación válidas y precisas. Este estudio puede contribuir en la creación de políticas educativas relacionadas con la orientación académica y la asignación eficaz de las becas. Al validar esta prueba, se fortalece la capacidad de la universidad para tomar decisiones informadas sobre el otorgamiento de becas, promoviendo que las becas sean asignadas a los estudiantes con mejores aptitudes académicas (Erazo Guerra & Rosero Morales, 2021).

Aprovechar al máximo las tecnologías empleadas en la creación de la prueba también es clave. Por ejemplo, el uso de energía fotovoltaica proporciona energía limpia a la universidad. Además, la aplicación digital de la prueba reduce el uso de papel y optimiza el tiempo al calificar automáticamente. Estos factores no solo contribuyen a la sostenibilidad del proceso, sino que también permiten ampliar su alcance en el futuro.

Es fundamental actualizar y validar regularmente las herramientas psicométricas. Esto asegura que respondan a los cambios en las características y necesidades de la población estudiantil, adaptándose a los avances tecnológicos y a los cambios sociales. En contextos

educativos y laborales en constante cambio, las herramientas de evaluación deben ser flexibles y ajustables a nuevas realidades. Validar la Prueba de Aptitudes de Pensamiento Básico no solo cubre la necesidad de una herramienta confiable en la UPNFM, sino que ofrece beneficios a largo plazo para el sistema educativo. Con esta validación, se contribuye a reducir el abandono y el rezago escolar (Mendoza Recarte, 2017).

2. Marco Teórico

En este capítulo se abordarán los elementos teóricos que permitan comprender las bases conceptuales en que se basa la investigación, dando un recorrido por el desarrollo de la orientación vocacional, las pruebas que se han usado para ofrecerla y como estas buscan medir habilidades puntuales del pensamiento que sirven para poder predecir el rendimiento que puede tener un estudiante en sus estudios y en sus actividades laborales. También se fundamentará las teorías de análisis y medición que se utilizan para la creación y validación de las pruebas psicométricas.

En el año 2018 la UPNFM en el marco del proyecto de “Fortalecimiento de la Gestión Educativa de la UPNFM mediante el Uso y Aprovechamiento de la Energía Fotovoltaica” desarrollo un sistema para la aplicación de pruebas que sirven como insumo para determinar si un estudiante puede cambiar de carrera o se le puede otorgar una beca, como una vía para favorecer y facilitar la formación docente a nivel superior y fortalecer el rendimiento a nivel académico.

La prueba que se desarrolló evalúa tres áreas: personalidad, habilidades e intereses, que son elementos que pueden determinar el desempeño académico de un estudiante y como mencionan Carrasco et al. (2021). En los años 90 en Estados Unidos surge la Teoría socio cognitiva de carrera, que se centra en el estudio de cómo se van formando los intereses profesionales y académicos, realizan sus elecciones y el éxito que pueden alcanzar, evaluando el agrado o no que se tiene por las actividades que se deben de realizar.

2.1. Orientación Vocacional

La orientación vocacional es la herramienta que se utiliza para ayudar a los jóvenes a poder seleccionar de forma adecuada la carrera o vocación que esté acorde con sus

capacidades y deseos, en esta orientación lo que se desea es mostrar las capacidades que tiene cada persona en cada área para poder explotar al máximo estas habilidades y ahí la necesidad que tienen los jóvenes de que alguien los oriente en la decisión para elegir una carrera universitaria donde puedan desarrollar sus habilidades, (De León Mendoza & Rodríguez Martínez, 2007).

Los procesos de orientación vocacional se vienen observando desde principios del siglo XIX con publicaciones como “*Choosing a Vocation*” en 1909 por Frank Parson, que muchos lo llaman el fundador de la orientación vocacional, luego de esta publicación se produjo el primer congreso de orientación en la ciudad de Boston y un año después el primer curso de orientación vocacional (Pérez Escoda & Bisquerra Alzina, 2009).

En esa época la orientación se basaba en los rasgos de las personas en comparación con los requisitos que tenía una vocación, pero ya en una segunda etapa se habla que la selección de una ocupación está influenciada por elementos de sociales, culturales y económicos que se dan a lo largo de la vida y que la elección viene determinada por las necesidades y capacidades que se dan en un contexto único para cada persona (Howell et al., 1977).

Dentro de este mismo pensamiento surge la teoría de Holland donde integra la motivación, personalidad y el ambiente y plantea que en la sociedad existe un número finito de ambientes laborales: motrices, intelectuales, de apoyo, de conformidad, de persuasión y estéticos, aportando en esta etapa John D. Krumboltz acorde con la teoría de Bandura que menciona que las personas cuentan con creencias fundamentales como la autoeficacia, que es la representación que tienen los seres humanos sobre sus capacidades para alcanzar ciertos niveles de rendimiento (Rascovan, 2013).

Tomando estas teorías es que surgen los mecanismos para evaluar de forma objetiva las capacidades e intereses de los individuos para orientarlos en la elección de una carrera que cumpla con los elementos mencionados de interés, aspiraciones, necesidades y capacidades.

2.2. Pruebas de Orientación Vocacional

Para brindar una orientación vocacional objetiva y confiable, los centros educativos utilizan herramientas psicométricas que evalúan capacidades e intereses de los estudiantes, como inventarios de intereses y aptitudes (Pereyra Quiñones, 2020).

Las pruebas más representativas a nivel mundial dentro de los inventarios de intereses han sido el IIS Inventario de intereses de Strong (1927); el Registro de preferencias vocacionales de Kuder (1939); y la Investigación auto dirigida de Holland (1978). Dentro de los inventarios de aptitudes se puede encontrar a PMA-R Aptitudes mentales primarias revisado o el HMP Habilidades mentales primarias.

2.3. Habilidades Mentales Primarias

En diferentes estudios se ha logrado identificar que la inteligencia no está constituida por un solo rasgo, sino que por diferentes habilidades que son las capacidades que los seres humanos pueden tener respecto a distintas situaciones cotidianas. Las habilidades mentales primarias son características y capacidades propias del ser humano que ayudan a realizar actividades de forma ágil y correcta. El tener presente el desarrollo de cada una de estas habilidades le permite al estudiante caracterizarse en qué áreas del aprendizaje posee mejores condiciones para alcanzar el éxito académico (Martínez de la Rosa et al., 2012).

En el campo de la psicometría Thurstone, (1938) mediante un análisis factorial logro extraer ocho factores: comprensión verbal, concepción espacial, razonamiento, cálculo, fluidez verbal, memoria, coordinación motora y rapidez perceptiva. Estudios que se pueden ver en sus publicaciones Análisis Multifactorial y Habilidades Mentales Primarias (PMA) creando una prueba que medía estos factores (Seclen Gamarra, 2019).

Esta batería de pruebas ha dado origen a diversas adaptaciones orientadas a medir factores de inteligencia y predecir el rendimiento del estudiado en ámbitos laborales o académicos y la PMA ha sido una herramienta útil para evaluar la inteligencia como rasgo de personalidad en el psicodiagnóstico (Seclen Gamarra, 2019).

La inteligencia pese a no ser el único elemento que determina cómo será el rendimiento de una persona en el área de interés si es un factor muy importante y que puede servir como base para poder tomar una decisión sobre las oportunidades que se le pueden dar

al individuo o los requerimientos que esté pueda necesitar para mejorar su rendimiento (Seclen Gamarra, 2019).

La UPNFM para la medición de las habilidades que puede tener un estudiante durante los procesos de orientación vocacional ha utilizado la prueba de habilidades mentales primarias en su edición de Manual Moderno para evaluar estas habilidades mentales en los estudiantes y así poder ofrecer una orientación vocacional adecuada a los estudiantes que desean realizar cambios de carrera. Ya que cada carrera necesita que los estudiantes posean habilidades diferentes según el área del saber, ya que las carreras científicas o técnicas necesitan un mayor desarrollo en las áreas del cálculo y lógica mientras que las carreras humanísticas necesitan un desarrollo en las áreas de la comprensión y fluidez verbal.

2.3.1. Comprensión Verbal

La comprensión verbal constituye una habilidad esencial en el desarrollo cognitivo, ya que permite interpretar ideas expresadas mediante el lenguaje, sirviendo como base para procesos superiores como el análisis, la evaluación y la creación. El discernimiento verbal, en este contexto, se refiere a la capacidad de diferenciar entre conceptos o expresiones similares, lo cual es fundamental en el ámbito académico y profesional. De acuerdo con el manual del PMA-R, esta habilidad se define como “la capacidad para comprender ideas expresadas en palabras. Se necesita en actividades en las cuales haya que captar los problemas por medio de la palabra escrita o hablada. Es una aptitud muy conveniente para el aprendizaje de lenguas, taquigrafía, historia, ciencia y, en general, para ser un buen estudiante” (Departamento de I+D+i, 2002, p. 8).

2.3.2. Destreza de Cálculo

La destreza de cálculo es una habilidad fundamental para el manejo eficiente de cantidades numéricas y la resolución rápida y precisa de problemas cuantitativos, siendo indispensable tanto en la vida cotidiana como en el ámbito académico, especialmente en asignaturas del área fisicomatemática (Aguirre & Pérez, 1995). Esta aptitud, además de ser observable con facilidad, implica agilidad y precisión en el procesamiento numérico. En el manual del PMA-R se define esta habilidad como “la capacidad de manejar números, de resolver rápidamente y con acierto problemas simplemente cuantitativos. Calculadores, cajeros, estadísticos, contables, banqueros, dependientes de comercio, jefes de almacén,

etcétera, destacan generalmente por este factor N. La capacidad de cálculo es útil para el éxito escolar en aritmética, estadística y toda clase de disciplinas en que entre como componente principal el cálculo matemático” (Departamento de I+D+i, 2002, p. 9). En el sistema del programa de Becas está definida como “la facilidad que pueda tener el estudiante para resolver rápida y correctamente, sencillos problemas numéricos. Es una actividad fácil de observarse ya que implica ante todo rapidez y asertividad en la respuesta dada.” (Barranco, 2019).

2.3.3. Pensamiento Lógico

El pensamiento lógico se refiere a la capacidad de aplicar razonamientos coherentes para resolver problemas, prever consecuencias y estructurar planes de acción. Según Godoy (2022), como se cita en Andonegui, esta habilidad se sustenta en tres formas lógicas del pensamiento: el concepto, el juicio y el razonamiento, elementos fundamentales en la construcción del pensamiento científico y académico.

En el contexto de la prueba PMA, esta aptitud se concibe como “la capacidad para resolver problemas lógicos, prever y planear. Contribuye extraordinariamente a la formación de inventores, médicos, maestros y educadores, magistrados, estadistas, científicos y gestores de empresa. Cuanto más asciende un estudiante en los estudios, mayor necesidad tiene del factor R para el éxito. La comprensión de la ciencia y de las matemáticas requiere una buena dosis del factor R. Diversas investigaciones muestran que el razonamiento implica dos capacidades diferentes: una, inductiva, la aptitud para inferir de los casos particulares la norma general, y otra, deductiva, la capacidad para extraer de las premisas la conclusión lógica” (Departamento de I+D+i, 2002, p. 9).

2.3.4. Percepción Espacial

La capacidad de percibir el tamaño, forma, color y adicionalmente determinar su posición con respecto a otros objetos. Está definida por la capacidad de poder visualizar objetos de dos o tres dimensiones, mostrando la destreza para poder imaginarse una figura o un objeto al cambiar de posición y para percibir las relaciones de los objetos que hay en un espacio al colocarlos de diferente manera (Braun, 1997).

2.3.5. *Fluidez Verbal*

Thurstone (1975) define la fluidez verbal como la habilidad de hablar o escribir con facilidad. Siendo diferente con la comprensión verbal en el sentido que está se centra en la velocidad y facilidad para encontrar las palabras más que en su comprensión. La prueba de aptitudes de pensamiento básico la definen como la facilidad que tiene la persona para hablar o escribir, la diferencia con la comprensión verbal es que está se relaciona con la velocidad y la facilidad para encontrar palabras que con el grado de comprensión de las ideas expresadas verbalmente.

2.4. Rendimiento Académico

El rendimiento académico es el resultado de múltiples elementos que interactúan como ser el nivel socioeconómico, nivel educativo de los padres, ambiente académico, motivación, intereses y las estrategias que pueda tener el docente (Palacio et al, 2006).

El rendimiento que presenta un estudiante en los diferentes niveles educativos está fuertemente influenciado por elementos externos e internos de la persona lo que pueden influenciar el nivel de aprovechamiento que puede lograr esté a lo largo de su vida estudiantil. El nivel universitario requiere habilidades sociales y cognitivas que permitan al estudiante lograr los objetivos planteados de manera óptima logrando el mejor desempeño sin necesidad de una exigencia mayor. Es por ello por lo que resulta importante conocer con certeza el nivel posee en estas áreas y así poder aprovechar al máximo las habilidades que ya tenga desarrolladas y mejorar las que no (Kohler Herrera, 2013).

Pero siempre existen algunas habilidades que, aunque se traten de reforzar no alcanzaran el nivel necesario para poder tener un desempeño aceptable en algunas áreas que necesitan habilidades específicas para poder ser aprovechadas de una forma adecuada.

Como indican Sierra et al. (2020) el rendimiento académico está correlacionado con el nivel de habilidades por lo que los estudiantes que tenían un alto nivel de desarrollo en sus habilidades mentales primarias que identificó Thurstone, Comprensión verbal (V), Comprensión Espacial (E), Razonamiento (R), Números (N), Fluidez Verbal (FV), por lo que el conocer el nivel que posee cada estudiante en cada una de estas áreas es útil para poder asesorar y orientar a un estudiante en la elección de una carrera universitaria.

Ya que éstas pueden predecir el desempeño que pueden llegar a tener las habilidades mentales son predictoras del rendimiento académico, por lo cual se considera importante analizar la relación entre las habilidades mentales de comprensión verbal y las otras como comprensión espacial, raciocinio, manejo de número, fluidez verbal en estudiantes universitarios (Sierra et al., 2020).

Las pruebas de aptitud académica abarcan tres áreas fundamentales: personalidad, habilidades e intereses, las cuales pueden influir significativamente en el desempeño académico del estudiante. Esta relación resulta especialmente relevante, ya que cuando las habilidades no se alinean con las exigencias de la carrera elegida, es común que se presenten bajos niveles de rendimiento. Esta discrepancia, a su vez, puede generar frustración en el estudiante y, en consecuencia, llevarlo a considerar el abandono de sus estudios ante los resultados insatisfactorios obtenidos en las asignaturas cursadas, como lo mencionan algunos autores (Carrasco et al., 2021).

Cada una de las fortalezas y debilidades en las cinco habilidades mentales presentadas hacen que cada joven universitario tenga las posibilidades de tener éxito en sus estudios, o, por el contrario, fracasar en el intento, pues todo dependerá de cómo se direccionan sus aptitudes, es decir, cada fortaleza y debilidad hacen que el estudiante pueda ejercer alguna profesión en específico. El problema reside cuando los jóvenes se encuentran estudiando o deciden estudiar alguna profesión en dónde las fortalezas y debilidades que se requieren son todas las contrarias a las que el alumno tiene, pues será muy probable que en su camino se encuentre con problemas de aprendizaje hasta llegar a una deserción (Espíndola et al., 2018).

En el caso de las instituciones educativas, esta también representa una situación problemática, ya que los altos índices de repitencia y abandono implican un gasto considerable de recursos y dificultan la planificación académica, especialmente al generar acumulación de matrícula en determinadas asignaturas, lo que merma las capacidades organizativas. Esta problemática se agrava en contextos donde la educación no está al alcance de toda la población. Además, factores externos, como una pandemia, pueden intensificar los niveles de deserción. Un ejemplo claro de ello fue el año 2020, cuando las instituciones de educación superior en Honduras registraron un abandono del 20 % al 30 % de su población estudiantil (Tiempo.hn, 2020).

Por lo que es necesario el poder conocer las habilidades mentales primarias que poseen los estudiantes realizándoles pruebas de estas para poder determinar las áreas donde poseen mejores aptitudes (Cueva & Quizocala, 2017). Ramos-Estrada et al., (2020) también manifiestan la importancia de estudios que exploren las propiedades de los instrumentos para tener sus cualidades de confiabilidad y validez y así poder valorar si estos son adecuados para ser utilizados para generar conocimiento con bases sólidas.

Dentro de la UPNFM el rendimiento académico del estudiante se define en los Art. 49,50,51,52,53,54,55 del reglamento del régimen académico. Indicando la definición del rendimiento académico del estudiante, el grado de asimilación de conocimientos durante su proceso de aprendizaje en relación con los objetivos, logros y fines de la educación propuestos en el plan de estudios (UPNFM, 2005).

La evaluación del rendimiento académico de los alumnos de pregrado en cada asignatura se evalúa basándose en una escala de 01 a 100, siendo 65% (sesenta y cinco por ciento) la calificación mínima para la aprobación de cada asignatura. Para calcular el rendimiento académico se deberán hacer las siguientes consideraciones: por período; es el promedio que obtiene un estudiante durante un período académico determinado; acumulado, es el promedio obtenido del estudiante en todo el proceso en la universidad (UPNFM, 2005).

2.5. Psicometría

Muniz define la psicometría como un conjunto de métodos, técnicas y teoría usadas para la medición de variables psicológicas. Pero estas mediciones pueden abarcar un espectro muy amplio de elementos que se pueden medir como que se pueden dividir en: teoría de la medición, teoría de las pruebas, escalamiento psicológico, escalamiento psicofísico (Muñiz, 2018).

Históricamente se pueden encontrar evidencias de sistemas de evaluación de individuos en función de su habilidad en el imperio chino o en el siglo XIII en Europa. Pero en un ambiente más actual trabajos como los de Galton, Cattell y Binet a finales del siglo XIX e inicios del XX (Meneses, 2014).

González Llana, (2007) destaca los trabajos de Binet quien deja de estudiar funciones sensoriales y motoras y se centra en la evaluación de funciones psíquicas superiores y propone ciertas características que deben de poseer las pruebas como son:

- Sencillez
- Cortas
- Independientes del examinador
- Contrastables

Creando con esto una serie de técnicas psicométricas que permiten evaluar rasgos internos y presentes en los individuos y estos rasgos a la vez pueden ser predictores del rendimiento académico (González Llana, 2007). Estos elementos y experiencias que buscan tener mediciones claras y objetivas han confluído a ciertos elementos para garantizar la calidad de las medidas que son la validez y fiabilidad de las pruebas (Martínez Arias et al., 2014).

En la psicometría, para evaluar la calidad de las pruebas psicológicas es el análisis de sus propiedades psicométricas. En este sentido, se entiende por propiedades psicométricas al conjunto de características de una prueba que permiten evaluar su calidad como instrumento de medición, y que están relacionadas con la validez, fiabilidad y precisión de las medidas que se obtienen. Para Muñiz et al. (2013) las propiedades psicométricas son las características que permiten evaluar la calidad de las medidas que proporciona una prueba, como la consistencia, la precisión, la validez y la fidelidad.

2.5.1. Validez

Ventura-León, (2016) presenta una breve descripción de la evolución del concepto de validez en las pruebas psicológicas a lo largo del tiempo donde se puede ver que la validez en la psicología ya tiene un largo recorrido ya que surge a finales del siglo XIX con el coeficiente de correlación de Pearson que mide la relación entre dos variables continuas. En 1937 Garret ya define lo que es la validez como la fidelidad con que mide lo que se pretende medir (QuestionPro, 2019). Ávila et al. 2019 como se citó en Kerlinger & Lee (2020) indican que se requiere evidencia robusta de la relación existente entre lo que realmente se está midiendo y el atributo que se quiere medir.

En 1952 la Asociación Americana de Psicología APA lanza una publicación con recomendaciones técnicas para los test psicológicos y técnicas diagnósticas, y una segunda versión donde aparece el concepto de validez predictiva, validez de constructo e introduce un nuevo concepto de validez, como “el grado en el que la prueba es capaz de lograr ciertos

objetivos” en 1955 Cronbach y Meehl enfatizan en la validez del constructo y asocian la validez con la técnica de análisis factorial dando recomendaciones para su uso.

Es hasta 1999 en los estándares para pruebas educativas y psicológicas que se indica que la validez tiene cinco fuentes de evidencia: Basado en el contenido, procesos de respuesta, estructura interna, relación con otras variables, consecuencias de la prueba. Y en el 2014 en la última publicación ya se incluyen elementos como equidad.

Bakieva, (2019) afirma que no es el instrumento en si el que se valida, si no que información que ofrecen sea adecuada y correspondiente a los alcances y limitaciones para los que fueron creados. Dando a entender que la validez de una prueba está en los resultados que puede entregar y que estos resultados sean de utilidad para un propósito designado.

Pero toda medida que se pueda realizar ya trae incorporada en ella un error que aleja la medición del valor real, lo que dificulta lo que es la comparación de diferentes test o diferentes aplicaciones que se puedan haber desarrollado lo que podemos ver en la Teoría Clásica del Test por lo que surge una propuesta que busca evitar estos errores en las mediciones surgiendo así la Teoría de respuesta al Ítem que busca mediciones que no sean susceptibles a cambio debido al instrumento y sean invariables respecto al sujeto que se evalúa (Valdés Veloz, 2014).

Otro elemento para considerar es la precisión en la estimación de los resultados en relación con el constructo medido por ambas metodologías. El modelo de Rasch, perteneciente a la Teoría de Respuesta al Ítem (TRI), permite identificar qué sujetos no están siendo evaluados adecuadamente por los ítems, o si el instrumento carece de sensibilidad. Esto evidencia que la TRI posibilita analizar el funcionamiento de los reactivos, verificar si discriminan correctamente entre los individuos y detectar posibles fallos en su formulación o interpretación. En consecuencia, la TRI aporta evidencias más sólidas de confiabilidad y validez (Acosta, 2011).

Tal como se expuso en el párrafo anterior, puede afirmarse que la Teoría de Respuesta al Ítem (TRI) es más robusta que la Teoría Clásica de los Tests (TCT), ya que esta última no alcanza los mismos niveles de calidad y precisión. Por tanto, la TRI representa una alternativa metodológica más adecuada para la construcción y validación de instrumentos de medición psicológica (Aune, 2019).

Cuando se realiza la validación de un instrumento se deben de realizar varias pruebas para determinar que todos los criterios de validez como son la estructural, de criterio, de constructo y la fiabilidad para poder tener un instrumento que esté adaptado al contexto, en su estructura factorial, semántica y sintáctica para poder demostrar que mide los rasgos y habilidades esperadas (Menacho Chiok & Alva Trinidad, 2020; Sánchez-Palacio, 2021).

2.5.2. Confiabilidad

Meneses (2014) señala que, desde el enfoque psicométrico, la fiabilidad de un instrumento se refiere a su capacidad para ofrecer mediciones consistentes y precisas. Es decir, se espera que, al aplicar una prueba en diferentes momentos o contextos, los resultados obtenidos se mantengan estables y reflejen con exactitud el constructo evaluado. En este sentido, una prueba será más confiable en la medida en que presente menos errores y los puntajes obtenidos se aproximen al valor real del evaluado. No obstante, asumir que un instrumento está completamente libre de error es un supuesto inalcanzable. Lo que sí es posible es identificar y controlar las fuentes de error, con el fin de minimizar su impacto sobre las mediciones realizadas.

Y es que cualquier herramienta de medición que se utilice tiene implícita en ella un error asociado a está, sea física o psicológica y en la psicología este error de medida conlleva ciertos problemas al momento de querer controlarlos porque existen factores físicos, mentales, ambientales y sociales que pueden afectar el desempeño de un sujeto durante la evaluación; dentro de este tipo de evaluaciones encontramos los errores sistemáticos (Martínez Arias et al., 2014).

Luego de la construcción y aplicación de la prueba se debe de buscar si existe una relación entre las respuestas y la variable que se busca medir. Existen diferentes formas de encontrar la relación entre los resultados y lo que se quiere medir y están agrupados de la siguiente forma:

- Los que consideran la puntuación final para emitir un juicio como la Teoría Clásica de los Test
- Los que consideran los ítems de forma individual como es la Teoría de respuesta al ítem.

2.5.3. Teoría Clásica de los Test

Spearman sienta las bases de la teoría clásica de las pruebas donde establece un modelo lineal para la fundamentar las puntuaciones de las pruebas, en su modelo establece que la puntuación empírica que una persona obtiene en una prueba (X) está conformada por dos componentes, la puntuación verdadera (V) y el error (e) que es algo que es inevitable al efectuar una medición (Muñiz J., 2018).

$$X = V + e$$

Donde se asume que V sería la puntuación que una persona sacaría en infinitas aplicaciones y que los errores e no están relacionadas con las puntuaciones verdaderas y estos errores tampoco se relacionan entre ellos.

2.5.3.1. Dificultad. En su desarrollo la TCT en el contexto de las pruebas de rendimiento donde lo que se evalúa es si se contestó bien o mal la pregunta recibiendo así puntuaciones dicotómicas. Por lo que el índice de dificultad se expresaría de la siguiente forma:

$$p_i = \frac{\sum_{s=1}^n X_{si}}{N} = \frac{A_i}{N}$$

Donde lo que se mide es la razón de los aciertos (puntuaciones 1) entre el número de evaluados que intentan responder el ítem. Aunque en ciertas condiciones el número de sujetos está determinado por la cantidad total de los que fueron expuestos al ítem sin importar si lo trato de responder o no y se les califica con la puntuación de 0. Este cálculo lo que devolverá es un valor cercano a 1 si el ítem es fácil y cercano a 0 si es difícil, es así como algunos lo consideran como índice de facilidad (Muñiz J., 2018).

2.5.3.2. Consistencia. La consistencia de una prueba se ve afectada por diversos factores, como la calidad de los ítems, la homogeneidad del constructo medido, el tamaño de la muestra y el grado de dificultad de los ítems. Una prueba con alta consistencia será aquel que muestre una alta correlación entre las puntuaciones obtenidas en diferentes administraciones de la prueba o entre diferentes subconjuntos de ítems dentro de la misma prueba (Muñiz J., 2018).

Uno de los métodos más utilizados para medir la consistencia interna de un instrumento es el alfa de Cronbach. Sin embargo, este coeficiente puede verse influido por factores como la cantidad de reactivos, el número de alternativas de respuesta y la varianza total de la prueba. Ante estas limitaciones, se propone como alternativa el coeficiente omega de McDonald, el cual se basa en las cargas factoriales y proporciona una estimación más robusta del índice de consistencia interna. Según Ventura-León y Caycho-Rodríguez (2017), se consideran valores aceptables de consistencia interna aquellos que se encuentran entre .70 y .90.

2.5.3.3. Discriminación. Se puede definir la discriminación de un ítem para separar de forma efectiva los sujetos que tienen o no la característica que se busca medir, y siendo a su vez un buen indicador de la consistencia interna de una prueba (Ventura-León & Caycho-Rodríguez, 2017).

2.5.3.4. Validez de criterio. Dentro de las principales formas de validez se encuentran la validez de contenido, la validez de constructo y la validez de criterio. Esta última se refiere a la medida en que los resultados de un instrumento se relacionan con un criterio externo relevante. En particular, la validez predictiva, que es una subcategoría de la validez de criterio, permite anticipar el desempeño futuro del evaluado. En el presente caso, se utilizará para predecir el rendimiento académico de los estudiantes a partir de los resultados obtenidos en la prueba (Argibay, 2006).

2.5.4. Teoría de Respuesta al Ítem

Esta teoría viene a complementar algunos elementos que la TCT no respondía, esta teoría nace de un enfoque en el funcionamiento individual de los ítems y no tanto del rendimiento general de la prueba. Su objetivo central está encaminado a que las mediciones no sean afectadas por el instrumento y obtener puntuaciones que no dependan de a quienes midan (Muñiz J., 2018).

3. Marco Contextual

Este capítulo brinda un contexto en el cual se desarrolla la investigación como son los procesos de orientación vocacional, la institución en que se desarrolló y los elementos que sirvieron como base para el surgimiento de la prueba de Habilidades Básicas que son el centro del estudio.

3.1. Procesos de Orientación Vocacional

Los procesos de orientación vocacional constituyen herramientas fundamentales para guiar al estudiante en la elección de una carrera que sea coherente con sus intereses y habilidades, permitiendo que esta decisión sea informada y consciente de sus propias características y de los requerimientos de la profesión. Esta relevancia ha llevado a que el tema sea ampliamente estudiado por la psicología vocacional. Para brindar una orientación vocacional adecuada, es necesario identificar previamente aquellos aspectos de la personalidad que influyen en la elección de una carrera que favorezca tanto el desarrollo personal como la satisfacción en el ámbito laboral (Carrasco et al., 2021). Esto se relaciona directamente con el rendimiento académico y el nivel de satisfacción profesional que el individuo puede alcanzar (Crespo et al., 2019).

Esta orientación es importante que se de en los años previos a ingresar a la universidad o en los primeros años universitarios ya que el primer año puede ser el más complicado y de ahí que se presenten abandonos o solicitudes de cambio de carrera donde algunos medios indican que en algunas instituciones puede llegar a ser del 20% de sus estudiantes de primer ingreso cambian de carrera, en la UNAH que es la universidad con mayor población estudiantil de Honduras el cambio de carreras que registran es de 5,973 en el año 2018 (Fernández, 2020; UNAH, 2018).

La orientación vocacional tiene sus orígenes en Estados Unidos y fue de gran aceptación por que ayudaba a la generación de fuerza laboral motivada. Es un proceso continuo de acercamiento y acompañamiento para lograr el desarrollo integral de la persona considerando sus capacidades cognitivas y motivaciones que ayudan a tomar una elección consciente e informada que ayudaba a los estudiantes a elegir una profesión no teniendo en base únicamente las aspiraciones económicas que pueden ofrecer algunos rubros y considera

que este proceso de orientación es algo continuo que se debe de dar un seguimiento para potenciar y mejorar las habilidades del sujeto (Santana Sardi & Vigueras Moreno, 2019).

La orientación vocacional es un instrumento que de alguna manera puede ayudar a cumplir objetivos y metas como los que se plantean en los objetivos de desarrollo sostenible del Programa de las Naciones Unidas para el Desarrollo donde el cuarto objetivo es lograr una educación de calidad, teniendo al educando como un elemento activo y probado para garantizar el desarrollo sostenible. Con este fin, se busca asegurar que todos los niños completen su educación primaria y secundaria gratuita para 2030. También busca proporcionar acceso igualitario a formación técnica asequible y eliminar las disparidades de género e ingresos, además de lograr el acceso universal a educación superior (PNUD, 2022)

Para ingresar a las universidades, es necesario cumplir con ciertos requisitos que garanticen que los aspirantes poseen los conocimientos y aptitudes necesarios para alcanzar un desempeño académico adecuado. En el caso de las universidades hondureñas, uno de los principales requisitos de admisión es la aplicación de pruebas de conocimiento y de aptitud académica. Estas evaluaciones permiten determinar si el estudiante cuenta con las competencias requeridas para la carrera a la que desea ingresar. En caso de no alcanzar los resultados esperados, el aspirante deberá repetir la prueba posteriormente, ya sea tras haber fortalecido sus capacidades, o bien optar por una carrera en la que ya posea las aptitudes necesarias (UPNFM, 2022).

En Honduras existen 20 instituciones de educación superior, de las cuales seis son públicas y catorce privadas. En conjunto, atienden a 266,908 estudiantes a nivel nacional, lo que representa aproximadamente el 3 % de la población del país. Dentro del sistema público, las dos universidades con mayor cobertura —que concentran el 52 % de la población estudiantil universitaria— son la Universidad Nacional Autónoma de Honduras (UNAH), en primer lugar, y la Universidad Pedagógica Nacional Francisco Morazán (UPNFM), en segundo (DES, 2020; INE, 2020).

Según Castro (2022), en el contexto nacional, la orientación vocacional no ha sido una prioridad para las autoridades y tomadores de decisiones. Esta ha sido promovida principalmente por profesionales de la psicología y la orientación educativa, de manera aislada y sin una estructura formal establecida desde el gobierno.

En este sentido, la Universidad Pedagógica Nacional Francisco Morazán inició en 2015 un proceso de sistematización de las actividades de orientación vocacional mediante el uso de medios electrónicos. Este esfuerzo tuvo como objetivo ofrecer asesoría vocacional para el cambio de carrera y facilitar el otorgamiento de becas a estudiantes (BCIE, 2015).

3.2. La Universidad Nacional Pedagógica Francisco Morazán (UNPFM)

La Universidad Pedagógica Nacional Francisco Morazán (UPNFM) es el único centro universitario en Centroamérica cuya finalidad principal es la formación de docentes para los niveles de educación prebásica, básica y media. Actualmente, ofrece una oferta académica compuesta por 23 licenciaturas, 16 maestrías, 2 doctorados y 5 carreras a nivel técnico universitario, impartidas en sus 10 centros regionales distribuidos a nivel nacional. Asimismo, la institución cuenta con unidades especializadas en promover acciones orientadas al desarrollo de competencias estudiantiles. Entre estas acciones destacan la aplicación de pruebas de habilidades para el otorgamiento de becas y las evaluaciones destinadas al cambio de carrera (UPNFM, 2015, 2022; UPNFM, s.f.).

La universidad pedagógica tiene presencia a nivel nacional con 10 centros de estudios en las principales ciudades del país, Tegucigalpa, San Pedro Sula, La Ceiba, Santa Bárbara, Santa Rosa de Copán, Gracias, La Esperanza, Choluteca, Nacaome, Danlí y Juticalpa; atendiendo a una población promedio de 20,000 estudiantes aproximadamente.

3.2.1 Vicerrectoría de Vida Estudiantil

La estructura administrativa de la UPNFM está compuesta por las direcciones que son unidades que tienen funciones especiales técnico, académico, administrativo y de vinculación social. Dentro de las vicerrectorías se encuentra la de Vida Estudiantil VIVES, quien es la encargada de dar cumplimiento a las políticas de atención a los estudiantes a través de diferentes programas y vela por el desarrollo integral de los estudiantes (UPNFM, 2005).

Dentro de sus funciones está el proveer a los alumnos de servicios de orientación personal, académica y profesional; participar en servicios de selección de becarios; mantener actualizados los instrumentos psicométricos, vocacionales, actitudinales y de personalidad

que se apliquen a los estudiantes para identificar las potencialidades de la población estudiantil y a otras personas que lo soliciten (UPNFM, 2005).

3.3. Proyecto de Fortalecimiento de la Gestión Académico Administrativo de la UPNFM Mediante el Aprovechamiento de la Energía Fotovoltaica

Con el propósito de fortalecer la gestión educativa ampliando y mejorando el Programa de Becas la UPNFM firma un convenio con BCIE en el marco del programa de conversión de la deuda de Honduras frente a España, se genera un proyecto de generación de energía eléctrica por medio del aprovechamiento de la energía fotovoltaica donde se genera 1.4 Mw de energía, lo cual constituye un ahorro para la universidad, ahorro que será trasladado al otorgamiento de becas a los estudiantes de la casa de estudios.

Dentro de las actividades de mejora que se obtienen con el proyecto se adquiere un sistema informático para la evaluación de las aptitudes e intereses profesionales de los estudiantes, sistema que servirá como proveedor de insumos para la elección de los estudiantes que se les otorgará una beca, pero también servirá para realizar evaluaciones para el cambio de carrera de los estudiantes.

En el año 2018 la UPNFM implemento un sistema informático para la aplicación de pruebas de cambio de carrera y la otorgación de becas que realiza la Dirección de Servicios Estudiantiles (VIVES), a partir de ese año los estudiantes se someten a dicha prueba con lo que se mide si tienen las habilidades necesarias para poder cambiar a otra carrera.

Para lo que cuenta con un sistema de valoración psicométrico con el siguiente objetivo “Evaluar los parámetros de personalidad, asociados con el desempeño y el desarrollo de habilidades y destrezas de los aspirantes a las becas ofertadas por la UPNFM, a través de un proceso de evaluación psicotécnico integral compuesta por tres pruebas psicométricas.” este sistema evalúa tres áreas de medición, Personalidad, Habilidades e Interés (Gamma Solutions, 2017).

4. Metodología

4.1. Tipo de investigación

El presente estudio se desarrolló bajo un enfoque cuantitativo, con un diseño no experimental, de tipo transversal y de alcance correlacional. Su propósito fue evaluar las propiedades psicométricas de la prueba de Aptitudes de Pensamiento Básico aplicada en la Universidad Pedagógica Nacional Francisco Morazán (UPNFM) y analizar su relación con el rendimiento académico de los estudiantes (Sampieri, Collado y Lucio 2014).

La naturaleza cuantitativa del estudio se evidencia en la recolección y análisis de datos numéricos obtenidos a través de las respuestas de los estudiantes a la prueba, así como de sus índices académicos registrados antes y después de la aplicación. Se emplearon técnicas estadísticas descriptivas, análisis psicométricos clásicos, análisis factorial exploratorio (AFE), confirmatorio (AFC), estimaciones de confiabilidad mediante el coeficiente omega y validación de estructuras mediante modelos de segundo orden (Sampieri, Collado y Lucio 2014).

Además, se evaluó la validez predictiva del instrumento, comparando el desempeño en la prueba con el rendimiento académico posterior de los participantes. Todo este proceso permitió determinar el grado de consistencia interna, validez estructural y utilidad práctica del instrumento como predictor del rendimiento académico, alineándose con los principios metodológicos (Sampieri, Collado y Lucio 2014).

Este estudio es de tipo correlacional donde se contrastarán los resultados de la prueba y el rendimiento presentado en la carrera a la que los estudiantes realizaron se sometieron a la evaluación. Se empleará un enfoque cuantitativo para analizar los datos almacenados en los sistemas informáticos de la universidad. El estudio es transversal, dado que, aunque se considera el rendimiento académico a lo largo del tiempo, la recolección de datos se realizará en un único momento (Sampieri, Collado y Lucio 2014).

4.2. Variables

Las variables por estudiar son las propiedades psicométricas de la prueba de pensamiento básico de la UPNFM y como los resultados de esta prueba se relacionan con las calificaciones de los estudiantes que cambian de carrera.

Tabla 1:Operacionalización de variables en la cual se detallan las dos variables latentes, sus cinco dimensiones con el respectivo número de ítems.

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicador (Ítems)
Aptitudes de pensamiento básico.	Habilidades necesarias para el correcto desempeño del trabajo de cada una de las titulaciones ofertadas por la UPNFM	Instrumento de Medición: Prueba de comprensión de lectura o evaluación de razonamiento verbal. Métrica: Puntuación obtenida en una serie de preguntas diseñadas para evaluar la capacidad de entender, interpretar y analizar texto escrito.	1. Discernimiento Verbal	Fuente de Datos: Resultados de las pruebas estandarizadas administradas a los estudiantes. (50 ITEMS).
		Instrumento de Medición: Prueba de habilidades matemáticas básicas, incluyendo operaciones aritméticas y problemas de razonamiento cuantitativo. Métrica: Puntuación obtenida en una serie de preguntas que miden la habilidad para realizar cálculos matemáticos rápidos y precisos.	2. Destreza de Cálculo	Fuente de Datos: Resultados de las pruebas estandarizadas administradas a los estudiantes (70 ITEMS).
		Instrumento de Medición: Evaluaciones que miden el razonamiento lógico, incluyendo pruebas de secuencias lógicas, patrones y analogías. Métrica: Puntuación basada en la correcta identificación de reglas, formación de juicios lógicos y solución de problemas.	3. Pensamiento Lógico	Fuente de Datos: Resultados de las pruebas estandarizadas administradas a los estudiantes (30 ITEMS).
		Instrumento de Medición: Pruebas de habilidad espacial, que pueden incluir preguntas sobre rotación de objetos, visualización espacial y percepción de relaciones espaciales.	4. Apreciación Espacial	Fuente de Datos: Resultados de las pruebas estandarizadas administradas a los estudiantes (20 ITEMS).

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicador (Ítems)
		Métrica: Puntuación obtenida en ejercicios que evalúan la capacidad para manipular y entender relaciones espaciales Instrumento de Medición: Evaluación de habilidades de expresión verbal, que puede incluir tareas de fluidez verbal, como la generación de ideas bajo restricciones de tiempo Métrica: Puntuación basada en la cantidad y calidad de las respuestas dadas en ejercicios de fluidez verbal.	5. Elocuencia Verbal	Fuente de Datos: Resultados de las pruebas estandarizadas administradas a los estudiantes.
Rendimiento Académico	El Rendimiento Académico de los alumnos de pregrado en cada asignatura, será evaluado basándose en una escala de 01 a 100, siendo 65% (sesenta y cinco por ciento) la calificación mínima para la aprobación de cada asignatura.” (UPNFM, 2005).	Instrumento de Medición: Registros académicos oficiales de la UPNFM. Métrica: Calificaciones finales de los estudiantes en cada asignatura, registradas en una escala de 01 a 100. Las calificaciones se analizarán tanto individualmente por asignatura como acumulativamente para obtener el promedio general de cada estudiante. Se considerarán aprobados aquellos estudiantes cuyas calificaciones sean iguales o superiores al 65% en cada asignatura.	Rendimiento Académico	Fuente de Datos: Base de datos académica de la UPNFM donde se almacenan las calificaciones finales de cada estudiante por asignatura.

Fuente. (Gamma Solution, 2018).

4.3. Población y Muestra

4.3.1. Población

La población de estudio comprende a todos los estudiantes que han realizado un cambio de carrera a través del sistema de evaluación de la Vicerrectoría de Vida Estudiantil (VIVES) hasta la fecha de recopilación de los datos. Esta delimitación asegura que se consideran únicamente aquellos estudiantes que han pasado por el proceso de cambio de carrera mediante el sistema evaluativo específico de la DISE.

4.3.2. Muestra

La muestra consiste en los datos de 1587 estudiantes evaluados mediante el sistema de la DISE. Este conjunto de datos representa la totalidad de los estudiantes que han realizado un cambio de carrera hasta el momento de la recopilación. La selección de esta muestra es intencional, ya que se han considerado a todos los casos disponibles dentro del sistema sin realizar una selección aleatoria o de conveniencia, asegurando así que la muestra incluye a todos los estudiantes que cumplan con el criterio de cambio de carrera en el marco de evaluación establecido.

4.3.2.1. Descripción Detallada de la Muestra. Para proporcionar mayor claridad y contexto sobre la muestra de estudio, se destacan las siguientes características relevantes de los 1587 estudiantes evaluados:

4.3.2.1.1. Carreras Representadas. La muestra incluye estudiantes de 25 carreras diferentes. Las carreras con mayor representación en la muestra son:

- Enseñanza del Inglés
- Enseñanza del Español
- Educación Básica (I y II ciclo)
- Ciencias Sociales
- Ciencias Naturales

Estas cinco carreras constituyen aproximadamente el 50% del total de estudiantes evaluados. Este desglose permite observar cómo se distribuye la muestra entre diferentes áreas académicas de la universidad.

4.3.2.1.2. Género. El 69% de los estudiantes en la muestra son de sexo femenino, lo que refleja una mayor participación de mujeres en los cambios de carrera dentro del sistema de evaluación de la VIVES.

4.3.2.1.3. Modalidad de Estudios. El 75% de los estudiantes pertenece a la modalidad presencial, lo cual podría influir en los resultados al reflejar las características de aquellos en esta modalidad educativa.

4.3.3. Justificación del Tamaño de la Muestra

El tamaño de la muestra de 1587 estudiantes se considera adecuado en relación con el número de ítems (171) de la prueba a validar, ya que se logra una proporción de 1:9 (participantes: ítems). Esta proporción es recomendada en estudios psicométricos para asegurar que cada ítem tenga suficiente variabilidad de respuesta y garantizar así la estabilidad de los parámetros en modelos de análisis como la Teoría de Respuesta al Ítem (TRI) y Teoría Clásica de los Test (TCT). La proporción de 1:9 ha sido respaldada por estudios que sugieren su idoneidad para análisis en instrumentos con un elevado número de ítems, lo cual contribuye a la confiabilidad de los resultados obtenidos (Gunawan, 2021).

4.3.4. Representatividad de la Muestra

Dado que ciertas carreras están sobrerrepresentadas en la muestra (como las áreas de Enseñanza y Educación Básica), los resultados psicométricos podrían estar sesgados hacia las características de estos grupos. Para abordar este posible sesgo, se considerarán análisis de subgrupos o se aplicarán métodos de ponderación para equilibrar las diferencias en la representación de cada carrera. Esto garantizará que los resultados sean interpretables y aplicables a una población más amplia, además de reducir el impacto de la sobrerrepresentación de ciertos grupos.

4.4. Instrumentos

El instrumento principal utilizado en este estudio es la Prueba de Aptitudes de Pensamiento Básico Académico desarrollada por Gamma Solutions en 2018, bajo la supervisión de la Dirección de servicios estudiantiles (DISE) ahora vicerrectoría de Vida Estudiantil (VIVES) de la Universidad Pedagógica Nacional Francisco Morazán (UPNFM). Esta prueba se emplea específicamente para evaluar las habilidades cognitivas fundamentales

de los estudiantes, en el contexto de solicitudes de becas y cambios de carrera dentro de la UPNFM, permitiendo valorar su aptitud en competencias críticas para el desempeño exitoso en diversas titulaciones académicas (Gamma Solutions, 2017).

4.4.1. Estructura de la Prueba

- La Prueba de Aptitudes de Pensamiento Básico Académico consta de un total de 171 ítems que se agrupan en cinco dimensiones clave:
 - Discernimiento Verbal: Capacidad para comprender y evaluar ideas expresadas en palabras.
 - Destreza de Cálculo: Habilidad para realizar operaciones matemáticas fundamentales.
 - Pensamiento Lógico: Aptitud para aplicar razonamiento lógico en la resolución de problemas.
 - apreciación Espacial: Competencia en la visualización y manipulación de objetos en el espacio.
 - Elocuencia Verbal: Habilidad en la comunicación oral y escrita efectiva.
- Los ítems presentan distintos tipos de formatos de respuesta, tales como:
 - Selección múltiple: Con 4 o 5 opciones de respuesta por ítem.
 - Preguntas de completación: Requieren una respuesta específica en un rango de 0 a 1 puntos, según la precisión de la respuesta.
 - Verdadero/Falso: Ofrecen dos opciones de respuesta.

4.4.2. Puntuación e Interpretación

Cada respuesta correcta en los ítems de selección múltiple y verdadero/falso otorga 1 punto. Los ítems de completación tienen una valoración variable de 0 a 1 punto, basada en la exactitud de la respuesta proporcionada. La puntuación total en la prueba oscila entre 0 y 270 puntos, donde un puntaje más alto refleja un mayor nivel de competencia en las áreas evaluadas. Para la construcción de la prueba se dieron 5 pasos, que van de la conceptualización de la prueba mediante a un estudio piloto y entrevistas estructuradas para determinar cómo se iba a medir el constructo objetivo.

4.4.3. Proceso de Validación del Instrumento

La prueba fue validada siguiendo un riguroso proceso metodológico que incluyó (Gamma Solutions, 2017):

1. Conceptualización y Diseño: Se inició con entrevistas estructuradas y un análisis de constructo para definir claramente las dimensiones que la prueba debía medir.
2. Desarrollo de Ítems: Basándose en la conceptualización, se elaboraron los ítems preliminares, definiendo las escalas y formatos de respuesta adecuados para cada dimensión.
3. Estudio Piloto y Revisión: Se aplicó la prueba a una muestra piloto de estudiantes para evaluar el rendimiento de los ítems y su alineación con el propósito de la prueba.
4. Validación de Contenido: Expertos en psicometría revisaron los ítems para asegurar que midieran adecuadamente los constructos definidos, aunque la documentación técnica no detalla los métodos estadísticos específicos empleados en las pruebas de validez y confiabilidad.
5. Análisis Psicométrico: Tras la prueba piloto, se realizaron análisis de validez y confiabilidad, junto con pruebas específicas de índice de dificultad, análisis de distractores, y normas de ejecución máxima.
6. Validación Cruzada: Finalmente, se llevó a cabo una validación cruzada en una segunda muestra de la misma población, permitiendo verificar la estabilidad de los resultados obtenidos.

4.4.4. Limitaciones en la Documentación Técnica

Aunque el proceso de validación fue exhaustivo, la documentación técnica de la prueba no incluye detalles completos sobre los métodos estadísticos específicos aplicados en los análisis de validez y confiabilidad, lo cual representa una limitación en la transparencia de la metodología de desarrollo.

4.5. Plan de Análisis de Datos

Se emplearon diversos análisis estadísticos y psicométricos para evaluar la validez y confiabilidad del instrumento de prueba. Para este estudio, se utilizó R (versión 4.1.2), junto

con sus librerías especializadas, como psych para análisis psicométricos, lavaan para modelos factoriales, y ltm para análisis bajo la Teoría de Respuesta al Ítem (TRI). Esta selección de herramientas asegura la reproducibilidad de los análisis y facilita la manipulación de los datos para obtener resultados confiables.

4.5.1 Confiabilidad por Consistencia Interna

4.5.1.1. Alpha de Cronbach. Se calculará el coeficiente Alpha de Cronbach para determinar la consistencia interna del instrumento, interpretando los valores de acuerdo con los estándares: valores entre 0.7 y 0.8 indican una consistencia aceptable, mientras que valores superiores a 0.8 reflejan alta confiabilidad. Este análisis se realizará usando la función alpha() de la librería psych.

4.5.1.2. Omega de McDonald. Para fortalecer la medición de la confiabilidad, se calculará el Omega de McDonald, que proporciona una evaluación adicional de la consistencia interna en estructuras con posibles factores multidimensionales. Los valores se interpretarán bajo el mismo umbral que el Alpha de Cronbach, utilizando la función omega() en psych.

4.5.2 Validez

4.5.2.1 Validez de Contenido. La validez de contenido se garantizó mediante revisiones de expertos en las dimensiones evaluadas durante la fase de construcción del instrumento. Se incluyó una validación cruzada utilizando una muestra adicional de la misma población para verificar la representatividad y adecuación de los ítems a las dimensiones. Esto asegura que el contenido de los ítems cubre el dominio teórico de las competencias evaluadas, reduciendo el riesgo de sesgos y omisiones.

4.5.2.2 Validez de Constructo. Para determinar que el instrumento evalúa adecuadamente los constructos teóricos, se realizará un Análisis Factorial Exploratorio (AFE) al 50% de la muestra. El método de extracción se basará en la técnica de Mínimos Cuadrados No Ponderados (ULS) o Máxima Verosimilitud (ML), en función de los resultados preliminares de los supuestos estadísticos.

4.5.2.2.1. Pruebas de Supuestos

Matriz de Correlación. Se evaluará que al menos el 30% de las correlaciones sean mayores a 0.30.

Prueba de Esfericidad de Bartlett. Un valor de $p < 0.05$ indicará que la matriz de correlaciones no es una matriz identidad, y que los datos son aptos para el análisis factorial.

Índice KMO. Un valor mayor a 0.5 confirmará la adecuación muestral para el análisis factorial.

Identificación de Factores. Se emplearán métodos como Análisis Paralelo (PA) y Minimum Average Partial (MAP) para determinar el número de factores. Posteriormente, los factores seleccionados serán rotados ortogonalmente (Varimax) u oblicuamente (Promax) según la estructura de los datos.

Análisis Factorial Confirmatorio (AFC). Con el 50% restante de los datos, se realizará un AFC para validar la estructura obtenida en el AFE, aplicando modelos de TRI y TCT para afinar la selección de ítems en cada dimensión.

4.5.3. Análisis de Rendimiento Académico

4.5.3.1. Consulta y Análisis de Notas. Se accedirá a las notas de los estudiantes en una escala de 1 a 100 para evaluar cambios en el rendimiento académico tras el cambio de carrera. Se utilizarán modelos de regresión lineal para evaluar si existe una relación significativa entre el cambio de carrera y el rendimiento académico, controlando por variables como género y edad, que podrían influir en el desempeño. Este análisis permitirá observar tendencias en el rendimiento a través de distintos programas.

4.5.4. Análisis de Correlación

Selección de la Prueba de Correlación: Dependiendo de la distribución de los datos, se aplicarán:

4.5.4.1. Correlación de Pearson. Para datos normalmente distribuidos.

4.5.4.2. Correlación de Spearman. para datos no normales o en caso de relaciones no lineales.

4.5.4.3. Interpretación. Los coeficientes de correlación serán interpretados bajo criterios estándar: valores cercanos a ± 0.3 se consideran de baja magnitud, ± 0.5 moderados, y ± 0.7 o superiores reflejan correlaciones fuertes.

4.5.5. Análisis de las Características de los Ítems

4.5.5.1. Dificultad del Ítem. La dificultad se evaluará mediante la proporción de respuestas correctas en cada ítem; una proporción cercana a 0.5 indica una dificultad adecuada, mientras que valores extremos pueden sugerir ítems muy fáciles o difíciles.

4.5.5.2. Discriminación del Ítem. Se calculará el coeficiente de correlación ítem-total corregido para cada ítem, con valores cercanos a 0.3 o superiores considerados aceptables. Este análisis se realizará con la función `item.analysis()` en `ltm` para identificar ítems que discriminan efectivamente entre niveles de habilidad.

4.5.6. Validación Cruzada

4.5.6.1. Procedimiento de Validación Cruzada. Para corroborar la consistencia del modelo de validez de constructo, se aplicará una validación cruzada utilizando una muestra adicional de la población original. Esto implicará replicar el AFC y los análisis de correlación en la muestra cruzada, asegurando así que los resultados observados no son específicos de una muestra única. Este paso ayuda a garantizar la generalización de los resultados y fortalecer la confiabilidad de las conclusiones.

4.6. Procedimientos

Para este estudio, se implementará un proceso estructurado para la recolección, transferencia, manejo y verificación de datos. A continuación, se describe cada fase en detalle.

4.6.1. Recolección de Datos

Se accedió a las bases de datos relacionales de la Universidad Pedagógica Nacional Francisco Morazán (UPNFM) mediante el uso de credenciales autorizadas, de acuerdo con las políticas de seguridad de la institución. Para extraer los datos específicos de la Prueba de Aptitudes de Pensamiento Básico y los registros académicos relevantes de los estudiantes en el Sistema de Gestión Académica de la UPNFM, se diseñaron scripts en SQL. Estos scripts

estaban orientados a recuperar únicamente la información necesaria para el análisis, preservando así la eficiencia y minimizando el riesgo de exposición de datos.

4.6.2. Transferencia de Datos

Los datos extraídos de las bases de datos relacionales se exportarán en formatos compatibles con el entorno de análisis, como archivos CSV. Este paso es crucial, porque facilita la transferencia a software de análisis, preservando toda la información sin alteraciones. Una vez exportados, los datos se importaron en RStudio, versión 4.1.2. que es donde se realizó la mayor parte del análisis estadístico y psicométrico, aprovechando las características robustas de esta herramienta.

4.6.3. Herramientas y Tecnologías Utilizadas

Para el análisis estadístico, se utilizó RStudio como principal plataforma de trabajo. Se emplearon librerías como dplyr y tidyverse para manipular datos. También se usaron psych y ltm para análisis psicométricos específicos. Además, se desarrollaron herramientas o scripts adicionales que automatizaron la extracción de datos, logrando mayor eficiencia y precisión en el proceso. Esto optimiza tanto la repetibilidad como la gestión de los datos.

4.6.4. Protocolo de Manejo de Datos

Todos los datos transferidos se almacenaron en un servidor seguro. Solo el personal autorizado tiene acceso, conforme a las políticas de privacidad de la UPNFM. Esto asegura la confidencialidad de la información de los estudiantes en cada etapa del proceso.

4.6.5. Verificación de la Integridad de los Datos

Antes de proceder con los análisis, se realizó una verificación exhaustiva de la integridad y precisión de los datos. Este proceso incluyó:

- Detección de valores atípicos y datos faltantes.
- Verificación de la consistencia en los formatos de datos y de las categorías utilizadas en cada variable.
- Se utilizarán las funciones summary () y describe () de R para detectar posibles discrepancias o inconsistencias en los datos extraídos.

4.6.6. Documentación de Procedimientos

Cada paso del proceso de extracción y análisis de datos se documentó detalladamente en un archivo de protocolo. Este archivo incluirá:

- Los scripts y consultas SQL utilizados.
- Cualquier problema detectado y los ajustes realizados.
- Observaciones sobre el comportamiento de los datos durante el análisis.

Esta documentación asegura la replicabilidad del estudio y facilita la revisión del proceso completo.

4.6.7. Revisión y Validación de los Procesos

Durante la fase inicial de extracción de datos, se realizaron revisiones periódicas de los datos extraídos para garantizar que las sintaxis funcionen correctamente y que los datos obtenidos coincidan con las especificaciones del estudio. Se efectuaron revisiones de los registros de extracción y todos los ajustes necesarios se implementaron de inmediato para asegurar la precisión y confiabilidad de los datos.

5. Resultados de la Investigación

El presente capítulo tiene como propósito exponer y analizar los resultados obtenidos durante el proceso de evaluación psicométrica de la prueba de aptitudes, así como su relación con el rendimiento académico de los participantes. Esta sección constituye un elemento clave para validar empíricamente el instrumento aplicado, permitiendo valorar su adecuación en función de las propiedades psicométricas y su capacidad predictiva respecto al desempeño académico.

Desde un enfoque cuantitativo, se emplearon técnicas estadísticas acordes con los objetivos del estudio y la naturaleza de los datos. Inicialmente, se procedió a la caracterización de la muestra mediante un análisis descriptivo, con el propósito de examinar la distribución de las puntuaciones obtenidas tanto en la prueba global como en cada una de sus subpruebas.

Posteriormente, se aplicó un Análisis Factorial Exploratorio (AFE) con el fin de identificar la estructura subyacente del instrumento y explorar la agrupación de los ítems en factores latentes. Esta estructura fue posteriormente evaluada mediante un Análisis Factorial Confirmatorio (AFC), utilizando indicadores de ajuste tales como el índice de ajuste comparativo (CFI), el índice Tucker-Lewis (TLI), el error cuadrático medio de aproximación (RMSEA) y la raíz del cuadrado medio del residuo estandarizado (SRMR), lo que permitió confirmar la validez estructural del modelo propuesto.

De igual forma, se llevó a cabo un análisis de los ítems utilizando la Teoría de Respuesta al Ítem (TRI) y el modelo de Rasch, con el objetivo de estimar los parámetros de dificultad y discriminación de cada reactivo, proporcionando así una visión más precisa del funcionamiento individual de los ítems dentro del constructo evaluado. Finalmente, se calcularon coeficientes de correlación entre las puntuaciones obtenidas en la prueba y el rendimiento académico de los participantes, con el objetivo de valorar la validez predictiva del instrumento.

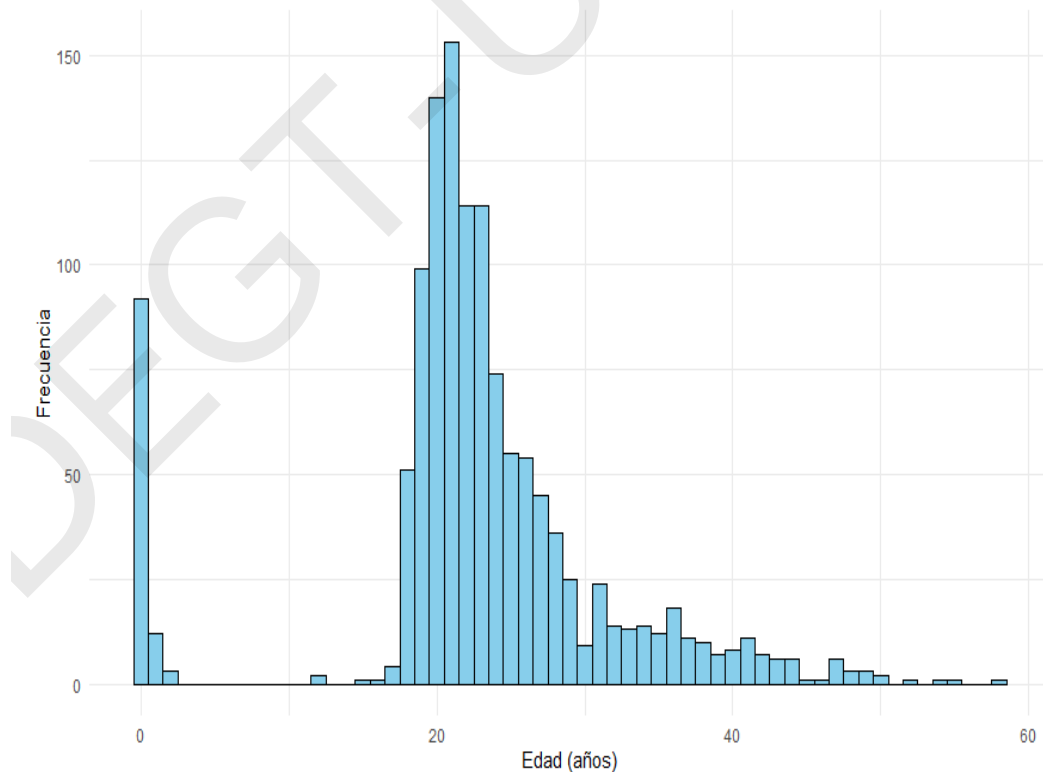
Los resultados derivados de estos análisis se presentan a continuación, organizados en subsecciones temáticas que facilitan su interpretación y discusión posterior.

5.1. Análisis Descriptivo de la Muestra y de las Puntuaciones

Con el propósito de caracterizar la muestra y describir el comportamiento general de las puntuaciones obtenidas en la prueba de aptitudes, se realizó un análisis descriptivo preliminar. Esta etapa resultó esencial para contextualizar los análisis psicométricos posteriores y verificar la calidad de los datos.

En primer lugar, se calcularon las edades de los participantes al momento de la aplicación de la prueba, derivadas de la diferencia entre la fecha de nacimiento y la fecha de presentación del instrumento. Se excluyeron del análisis los casos con edades no válidas (menores de cero años o superiores a 80 años), considerando que dichos registros podían deberse a errores en la digitación de fechas.

Figura 1. *Distribución de la edad de los sujetos incluidos en la muestra*

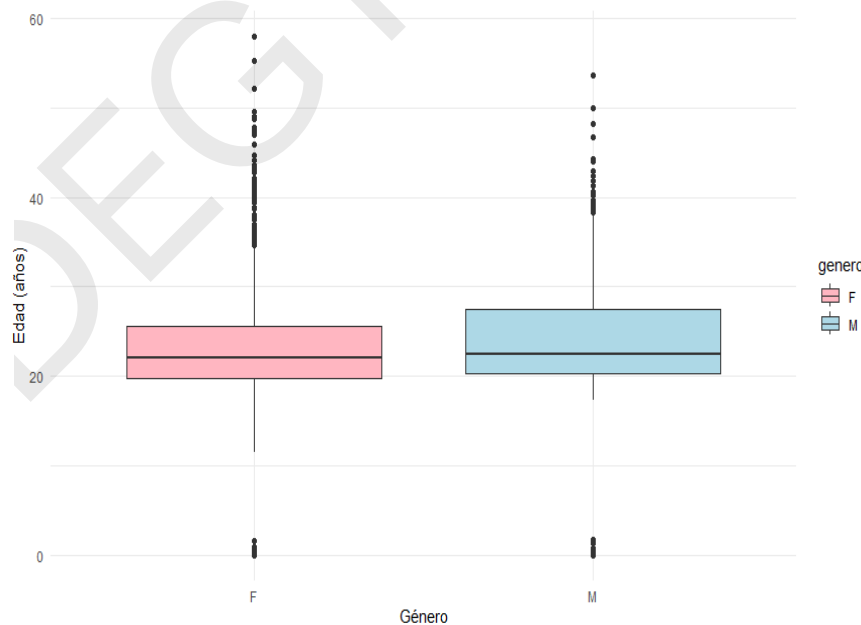


Nota. *Edades de los participantes al momento de la aplicación de la prueba, derivadas de la diferencia entre la fecha de nacimiento y la fecha de presentación del instrumento.*

La muestra final estuvo compuesta por 1,487 estudiantes, con una mayor proporción de mujeres (71.6%) respecto a hombres (28.4%). La edad promedio de los participantes fue de 22.08 años ($DE = 9.56$), con un rango que osciló entre los 0.87 y 59.92 años. Esta distribución permitió verificar la diversidad etaria de la población evaluada, aunque se observó un pequeño número de casos con edades anómalas que fueron tratados según criterios de inclusión previamente definidos.

En cuanto a las puntuaciones de la prueba, se calcularon medidas de tendencia central (media y mediana), dispersión (desviación estándar y rango intercuartílico), así como los valores mínimo y máximo para cada una de las cinco áreas evaluadas. Los resultados mostraron una variabilidad considerable entre áreas, lo cual sugiere que los componentes evaluados presentan diferentes niveles de dificultad o discriminación. Estos hallazgos preliminares se profundizan en los análisis factoriales y de validez presentados en secciones posteriores.

Figura 2. *Distribución de la edad de los participantes incluidos en la muestra en función del género*



Nota. Elaboración propia

5.2. Análisis Factorial Exploratorio (AFE)

Con el objetivo de identificar la estructura subyacente de la prueba de aptitudes y establecer una base empírica para su validación, se realizó un Análisis Factorial Exploratorio (AFE) sobre los datos obtenidos. Esta técnica permitió examinar la agrupación natural de los ítems en dimensiones latentes, considerando las correlaciones existentes entre ellos.

Previo a la extracción de los factores, se evaluó la adecuación de los datos para el análisis factorial mediante el índice de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett. Un valor alto del índice KMO indica que los patrones de correlación entre los ítems son suficientemente compactos para que el análisis factorial proporcione resultados confiables. Por su parte, una prueba de Bartlett significativa ($p < .05$) confirma que las correlaciones entre las variables son adecuadas para proceder con el AFE.

Para la extracción de los factores se utilizó el método de máxima verosimilitud, debido a que este permite realizar inferencias estadísticas y obtener estimaciones robustas bajo el supuesto de normalidad multivariada. Asimismo, se aplicó una rotación ortogonal tipo varimax con el propósito de facilitar la interpretación de los factores, maximizando la varianza de las cargas factoriales entre ellos. El número de factores a retener se determinó utilizando el criterio de eigenvalores mayores a uno (Kaiser), el análisis visual del gráfico de sedimentación (scree plot) y, de manera complementaria, la proporción acumulada de varianza explicada.

Los resultados indicaron la presencia de [N] factores que explican conjuntamente el [X]% de la varianza total. Las cargas factoriales obtenidas evidencian una estructura clara, en la cual la mayoría de los ítems se agrupan de forma coherente según su contenido. Sin embargo, se observaron algunos ítems con cargas cruzadas o valores inferiores al umbral mínimo de .30, los cuales fueron considerados para una posible revisión en fases posteriores de ajuste del instrumento.

En general, el análisis permitió establecer una estructura preliminar que será contrastada mediante un Análisis Factorial Confirmatorio (AFC) en el siguiente apartado, con el fin de evaluar el ajuste del modelo factorial propuesto y consolidar la validez estructural del instrumento.

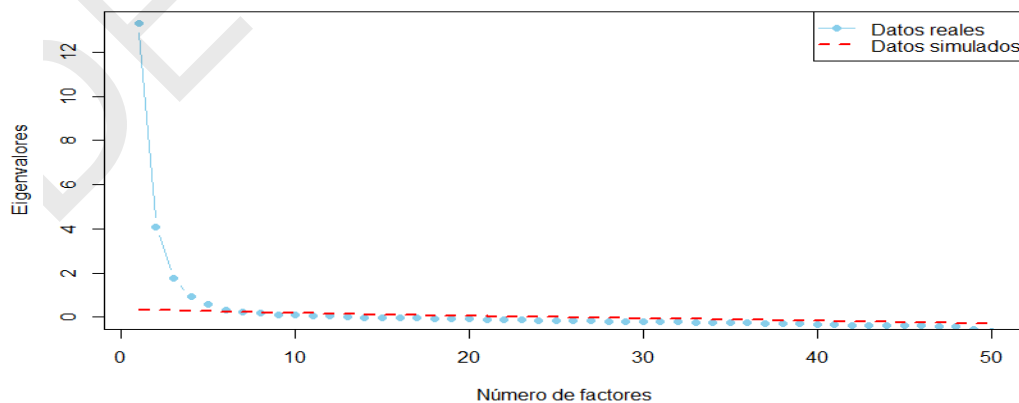
5.2.1. Subescala de Discernimiento Verbal

La subescala de Discernimiento Verbal está integrada por 50 ítems, donde el participante debe identificar sinónimos entre un conjunto de opciones. Esta tarea evalúa el dominio léxico, la comprensión semántica y la rapidez en la toma de decisiones verbales. El tiempo asignado para la prueba es de cuatro minutos.

Tras comprobar una excelente adecuación muestral ($KMO = 0.96$) y una prueba de esfericidad de Bartlett altamente significativa ($\chi^2 = 2260.68$; $gl = 1225$; $p < .001$), se realizó un Análisis Factorial Exploratorio utilizando el método de máxima verosimilitud y rotación varimax. El análisis paralelo sugirió la retención de **seis factores**, los cuales fueron estimados en el modelo factorial.

Las cargas factoriales mostraron agrupamientos interpretables. Por ejemplo, los ítems DV37 a DV52 presentan cargas elevadas en el **primer factor (ML1)**, lo cual sugiere un posible componente de vocabulario contextual o sinónimos de uso frecuente. A su vez, los ítems DV28 a DV36 se agrupan en el **tercer factor (ML3)**, lo que podría reflejar una dimensión relacionada con sinónimos formales o abstractos. Por su parte, el **factor ML6** reúne ítems como DV6, DV9, DV11, y DV18 a DV24, posiblemente vinculados a vocabulario más básico o literal. Otros factores como ML2 y ML4 agrupan ítems con características particulares en términos de complejidad o ambigüedad semántica.

Figura 3. Gráfico de sedimentación para la subescala de Discernimiento Verbal



Nota. Elaboración propia

El modelo explicó un **44.5% de la varianza total**, lo cual es adecuado dada la naturaleza compleja del contenido verbal. Estos resultados respaldan la validez estructural del instrumento y evidencian la existencia de múltiples dimensiones subyacentes que intervienen en la comprensión de sinónimos, facilitando su análisis confirmatorio posterior y la mejora de su estructura interna.

5.2.2. Subescala de Destreza de Cálculo

La subescala de Destreza de Cálculo evalúa la capacidad de los participantes para verificar la corrección de operaciones aritméticas básicas, como sumas y restas, mediante la identificación de resultados ciertos o erróneos. Esta tarea se ejecuta en un entorno controlado por tiempo, con un máximo de ocho minutos para responder 70 ítems, lo que demanda rapidez en el procesamiento numérico y toma de decisiones bajo presión.

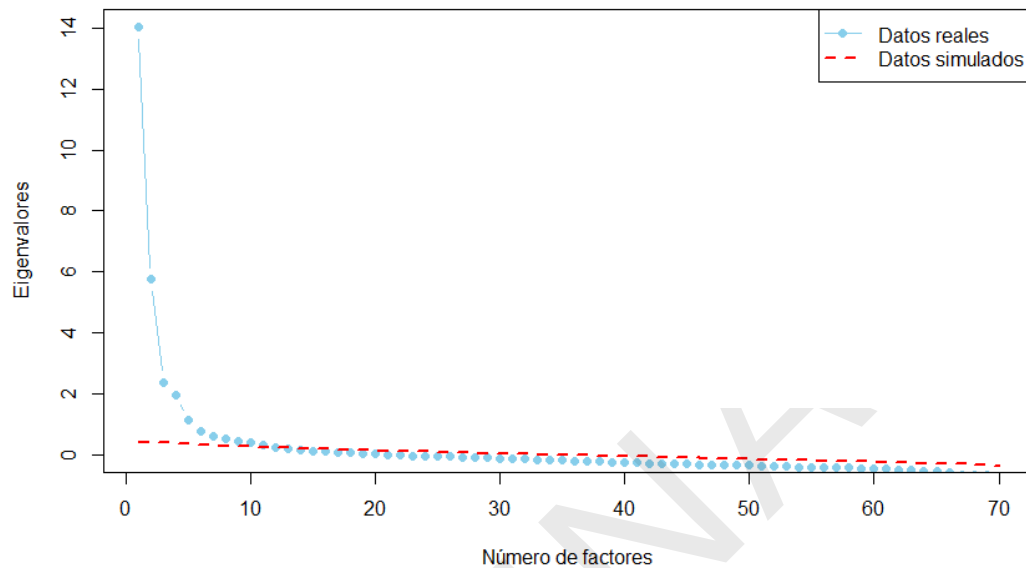
Tras verificar la adecuación muestral ($KMO = 0.94$) y una prueba de esfericidad de Bartlett significativa ($\chi^2 = 49198.11$; $gl = 2415$; $p < .001$), se aplicó un Análisis Factorial Exploratorio utilizando el método de máxima verosimilitud con rotación ortogonal tipo varimax. El análisis paralelo indicó la retención de 11 factores, los cuales fueron estimados en el modelo final.

Las cargas factoriales evidenciaron agrupamientos temáticamente coherentes entre los ítems. Por ejemplo, varios ítems relacionados con operaciones de similar estructura se agruparon en factores específicos como el ML7 y ML3, mientras que otros factores como ML1 y ML2 mostraron fuertes cargas asociadas a ítems que podrían representar niveles de dificultad, diferencias en magnitud numérica o precisión en la validación de resultados correctos versus incorrectos.

En conjunto, los once factores explicaron el 44.4% de la varianza total, lo cual es consistente con la diversidad de habilidades implicadas en esta subescala. Esta estructura multidimensional puede reflejar procesos como la estimación numérica rápida, el reconocimiento de errores sistemáticos, el manejo de cargas cognitivas simultáneas y el control del tiempo durante la ejecución.

Estos resultados apoyan la validez estructural de la prueba y servirán de base para la modelación confirmatoria y la depuración de ítems que presenten cargas cruzadas o baja saturación.

Figura 4. Gráfico de sedimentación para la subescala de Destreza de Cálculo



Nota. Elaboración propia

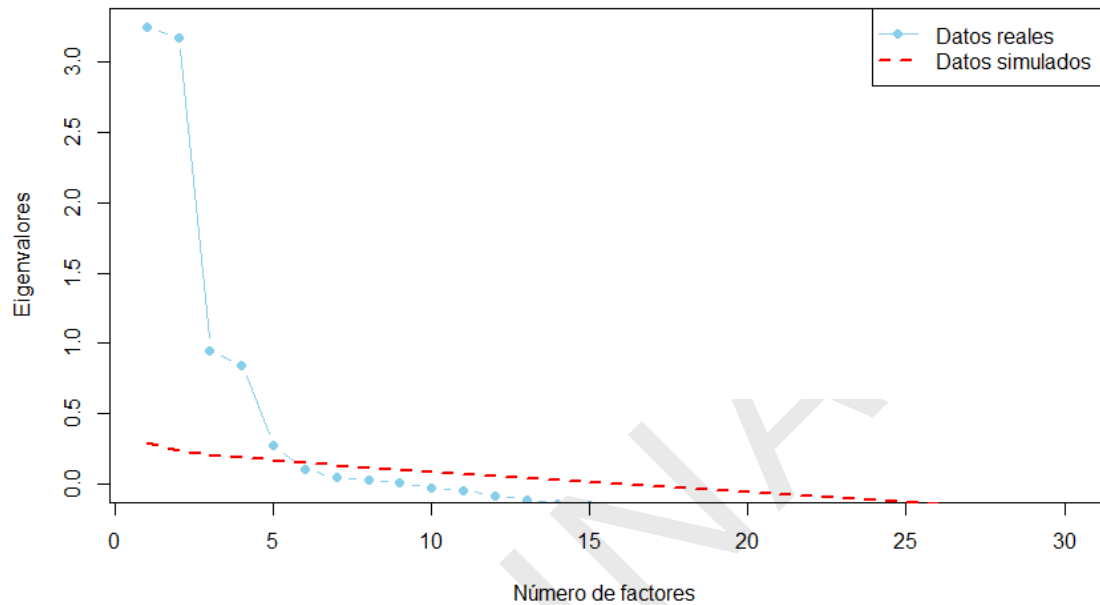
5.2.3. Subescala de Pensamiento Lógico

La subescala de Pensamiento Lógico está compuesta por ítems que presentan secuencias de letras y números, exigiendo que el participante determine cuál es el elemento que sigue de forma lógica. Esta tarea implica identificar patrones secuenciales, aplicar razonamiento inductivo y mantener temporalmente la información en la memoria de trabajo. La resolución se efectúa en un tiempo límite de seis minutos.

Luego de comprobar la adecuación muestral mediante el índice de Kaiser-Meyer-Olkin ($KMO = 0.85$) y la prueba de esfericidad de Bartlett ($\chi^2 = 9723.67$; $gl = 435$; $p < .001$), se procedió a la estimación del modelo factorial. El análisis paralelo sugirió la retención de cinco factores, los cuales fueron confirmados mediante un Análisis Factorial Exploratorio (AFE), utilizando el método de máxima verosimilitud con rotación ortogonal tipo varimax.

La matriz de cargas factoriales reveló una organización interpretable de los ítems. Por ejemplo, los ítems PL156 a PL160 se agruparon de forma clara bajo el primer factor (ML1), mientras que los ítems PL138, PL139, PL142 y PL144 mostraron cargas significativas en el

Figura 5. Gráfico de sedimentación para la subescala de Pensamiento Lógico



Nota. Elaboración propia.

tercer factor (ML3). Otros ítems se asociaron a factores distintos, probablemente en función del tipo de secuencia (alfabética, numérica o mixta) o del grado de dificultad que implican.

En conjunto, los cinco factores explicaron el 32.2% de la varianza total, lo cual es coherente con la diversidad de patrones lógicos implicados en los ítems. Esta estructura factorial sugiere la presencia de múltiples dimensiones dentro del razonamiento lógico, como la inferencia simbólica, el reconocimiento de regularidades abstractas o la manipulación de secuencias estructuradas.

Estos resultados aportan evidencia preliminar sobre la validez estructural de la subescala y constituyen un punto de partida para futuros análisis confirmatorios.

5.2.4. Subescala de Apreciación Espacial

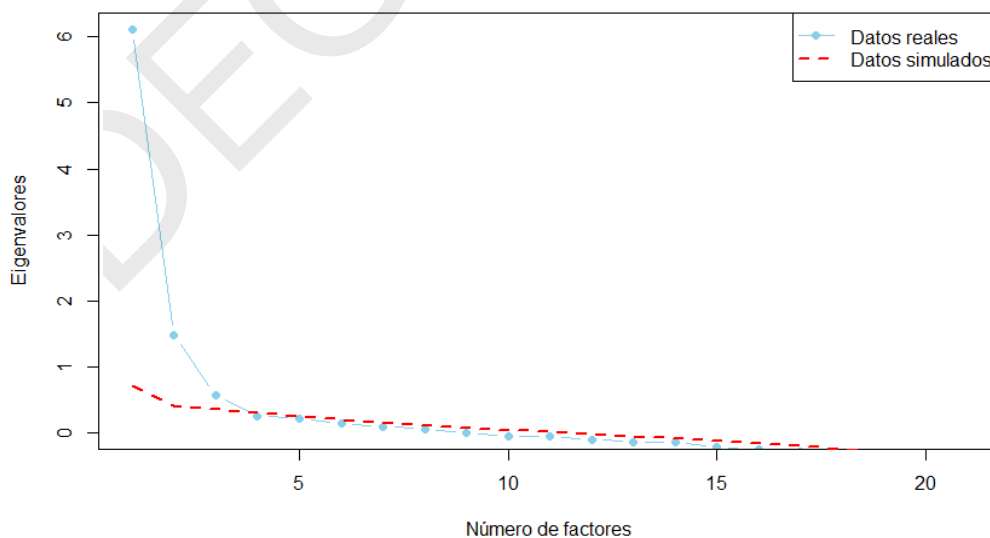
La subescala de Apreciación Espacial evalúa la capacidad del participante para identificar diferencias e inversiones en figuras geométricas. Cada ítem presenta una figura

original acompañada de varias opciones, entre las cuales el evaluado debe seleccionar aquellas que sean diferentes o estén invertidas. Debido a que algunos ítems pueden tener más de una opción correcta, la calificación se realizó considerando como correcta únicamente la selección exacta de todas las respuestas válidas y la omisión de aquellas incorrectas.

El análisis factorial comenzó con la evaluación de la adecuación muestral. El índice de Kaiser-Meyer-Olkin (KMO) general fue de 0.90, lo que indica una excelente adecuación para el análisis. La mayoría de los ítems reportaron medidas individuales de adecuación superiores a 0.85. Por su parte, la prueba de esfericidad de Bartlett fue significativa ($\chi^2 = 2472.17$; $gl = 210$; $p < .001$), lo cual indica que las correlaciones entre los ítems son suficientemente altas para justificar la aplicación del análisis factorial.

El análisis paralelo sugirió la retención de tres factores, los cuales fueron estimados utilizando el método de máxima verosimilitud con rotación ortogonal tipo varimax. Las cargas factoriales revelaron agrupamientos coherentes entre los ítems. Por ejemplo, los ítems X174 a X180 mostraron cargas consistentes sobre el primer factor (ML1), mientras que los ítems X167 a X172 y algunos adicionales se agruparon principalmente en el segundo factor (ML3). El tercer factor (ML2) capturó ítems como X173, X181, X183 y X184, que podrían estar asociados con formas específicas de inversión o rotación compleja.

Figura 6. Gráfico de sedimentación para la subescala de *Apreciación Espacial*



Nota. Elaboración propia

En conjunto, los tres factores explicaron el 40.9% de la varianza total, un valor adecuado considerando la complejidad visual y perceptiva de los estímulos. Estos resultados apoyan la existencia de múltiples dimensiones subyacentes en el desempeño espacial, tales como el reconocimiento de inversión, la discriminación de diferencias y la orientación mental. Este modelo provee una base sólida para la posterior validación estructural de la subescala y su eventual integración en modelos confirmatorios más amplios.

5.2.5. Subescala de Elocuencia Verbal

La subescala de Elocuencia Verbal evalúa la capacidad del participante para identificar la letra correcta que completa palabras presentadas dentro de un texto. Cada ítem corresponde a una posición específica dentro de una palabra incompleta, y el participante debe seleccionar la letra adecuada para formar la palabra correcta.

A diferencia de las demás subescalas, en este caso no se realizó un análisis factorial exploratorio, debido a que los resultados preliminares mostraron una ausencia de consistencia interna entre los ítems. El índice de Kaiser-Meyer-Olkin (KMO) general fue de 0.5, lo cual representa el límite inferior aceptable para análisis factorial, y todos los ítems reportaron valores mínimos en la medida de adecuación muestral ($MSA = 0.5$). Adicionalmente, el alfa de Cronbach fue de 0, lo cual indica que no existe correlación sistemática entre los ítems, y por tanto, no se puede asumir que miden una misma dimensión subyacente.

Este patrón sugiere que cada ítem actúa de forma independiente, posiblemente evaluando micro habilidades ortográficas específicas en lugar de una capacidad verbal general integrada. Por lo tanto, se optó por analizar esta subescala mediante estadísticos descriptivos, confiabilidad individual por ítem y puntuación total acumulada, sin aplicar técnicas de reducción dimensional como el análisis factorial.

5.3. Análisis Factorial Confirmatorio (AFC)

5.3.1. Subescala de Discernimiento Verbal

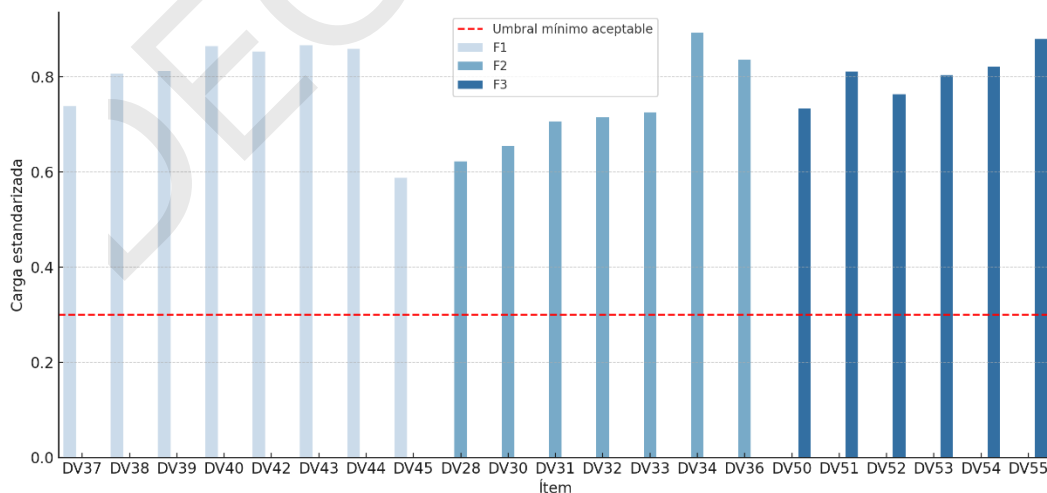
Con base en los resultados del análisis factorial exploratorio, se propuso un modelo reducido de tres factores latentes para la subescala de Discernimiento Verbal. Este modelo fue evaluado mediante análisis factorial confirmatorio (AFC), utilizando el estimador de máxima verosimilitud robusta (MLR). La muestra analizada estuvo compuesta por 1,528

observaciones válidas. El modelo incluyó tres factores que agruparon los ítems según sus cargas factoriales: el primer factor integrado por los ítems DV37, DV38, DV39, DV40, DV42, DV43, DV44 y DV45; el segundo factor por los ítems DV28, DV30, DV31, DV32, DV33, DV34 y DV36; y el tercer factor DV50, DV51, DV52, DV53, DV54 y DV55.

Los resultados indicaron un ajuste aceptable del modelo a los datos. El índice de ajuste comparativo (CFI) fue de 0.872 y el índice de Tucker-Lewis (TLI) fue de 0.856, valores cercanos al umbral recomendado de 0.90. El error cuadrático medio de aproximación (RMSEA) fue de 0.084, situándose en el límite superior del rango aceptable (< 0.08), mientras que el residuo cuadrático medio estandarizado (SRMR) fue de 0.069, considerado adecuado (< 0.08).

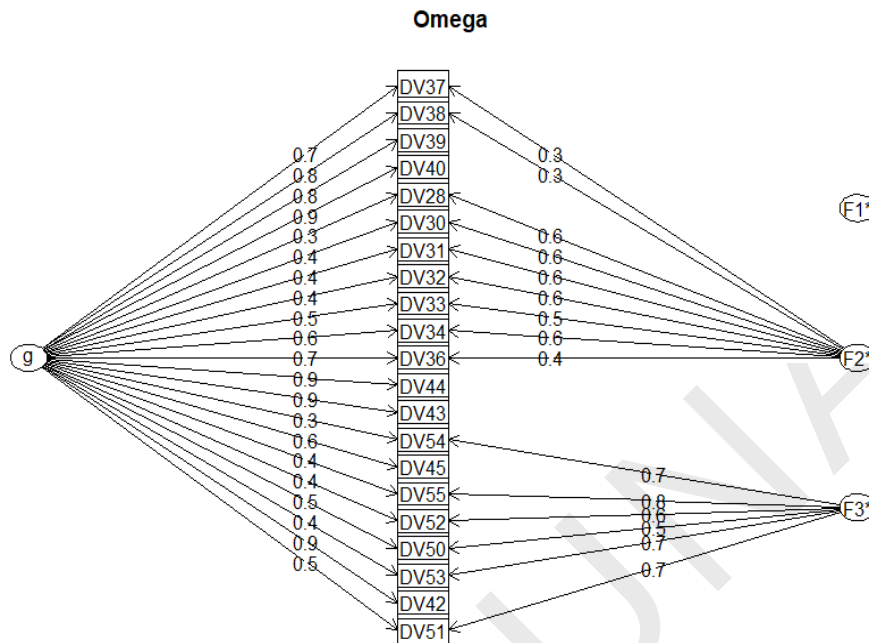
Todas las cargas factoriales estandarizadas fueron estadísticamente significativas ($p < .001$) y superiores a 0.58, lo cual indica una adecuada saturación de los ítems en sus respectivos factores. Las covarianzas entre los factores también fueron significativas, lo que sugiere relaciones consistentes entre las dimensiones latentes del modelo.

Figura 7. Cargas factoriales estandarizadas para la subescala de discernimiento verbal



Nota. Elaboración propia

Figura 8. Modelo gráfico del análisis omega de McDonald para la subescala de Discernimiento Verbal



Nota. Elaboración propia

Estos hallazgos respaldan la estructura tridimensional propuesta para esta subescala y ofrecen evidencia preliminar sobre la validez estructural del instrumento en lo que respecta a la medición del discernimiento verbal. Con el objetivo de complementar el análisis factorial confirmatorio y evaluar de forma más robusta la consistencia interna de la subescala de Discernimiento Verbal, se estimó el coeficiente omega de McDonald. Esta medida se considera superior al alfa de Cronbach en contextos donde se identifican múltiples factores y no se cumple el supuesto de tau-equivalencia, ya que toma en cuenta las cargas factoriales específicas de cada ítem (McDonald, 1999; Revelle & Zinbarg, 2009).

El análisis, realizado sobre el modelo reducido de tres factores previamente confirmado, mostró una omega total de 0.96, lo cual indica una excelente fiabilidad para la escala en su conjunto. Asimismo, el valor de omega jerárquico fue de 0.75, lo que sugiere que una proporción sustancial de la varianza puede atribuirse a un factor general subyacente. Estos resultados respaldan la homogeneidad relativa de los ítems, a pesar de la existencia de factores específicos.

Además, se calcularon las estimaciones de omega específicas para cada factor: $\omega(F2) = 0.93$ y $\omega(F3) = 0.92$, mientras que el factor principal mostró un comportamiento más general ($\omega(F1)$) sin valor específico por su saturación en el componente g). Estos resultados refuerzan la validez interna de las agrupaciones propuestas en el AFC reducido.

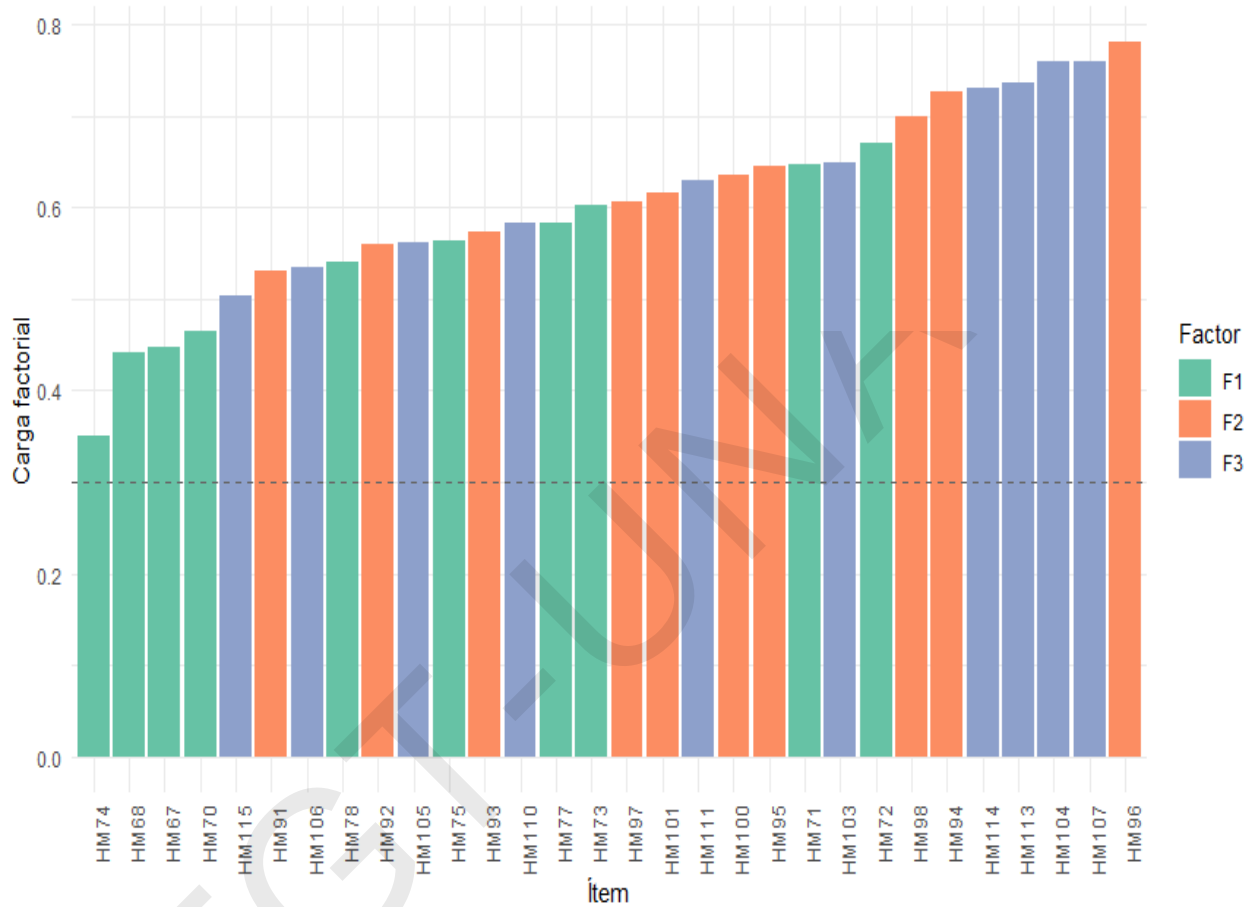
5.3.2. Destreza de Cálculo

Con el propósito de validar la estructura interna de la subescala de Destreza de Cálculo, se llevó a cabo un Análisis Factorial Confirmatorio (AFC) bajo el supuesto de un modelo estructurado en tres factores latentes. Esta decisión metodológica se fundamenta tanto en criterios teóricos como empíricos. Desde el plano conceptual, se considera que la destreza matemática en contextos evaluativos tiende a organizarse en tres dominios interrelacionados: el cálculo básico, la manipulación de expresiones numéricas, y el razonamiento cuantitativo aplicado. Esta estructura permite mantener la coherencia con las subescalas anteriores y garantiza un modelo ajustado y eficiente, facilitando su interpretación psicométrica.

El modelo AFC propuesto incluyó un total de 30 ítems, agrupados en los tres factores mencionados. La estimación del modelo se realizó utilizando el estimador robusto MLR (Máxima Verosimilitud con Corrección de Yuan-Bentler), adecuado para variables categóricas con distribución no normal.

Los indicadores globales de ajuste del modelo fueron los siguientes: CFI = 0.850, TLI = 0.838, RMSEA = 0.064 [IC90%: 0.060 – 0.068], y SRMR = 0.061. Si bien el valor del RMSEA se sitúa en el límite superior aceptable, los demás índices indican un ajuste razonable, considerando la complejidad del instrumento y el tamaño de la muestra ($n = 1580$). Las cargas estandarizadas de los ítems fueron en su mayoría superiores a 0.50, lo que evidencia una adecuada contribución de los ítems a sus respectivos factores. En el siguiente gráfico se visualizan las cargas factoriales estandarizadas por ítem:

Figura 9. Cargas factoriales estandarizadas por ítems para la subescala Destreza de Cálculo



Nota. Elaboración propia

Asimismo, se estimó el coeficiente omega de McDonald para evaluar la consistencia interna del modelo, tanto de manera global como por factor:

- **Omega Total:** 0.92
- **Omega General:** 0.70
- **Omega Hierarchical:** 0.70
- **Omega por factores:**

- Factor 1 (Cálculo básico): $\omega = 0.80$
- Factor 2 (Procesamiento numérico): $\omega = 0.54$
- Factor 3 (Razonamiento cuantitativo): $\omega = 0.07$

Estos resultados sugieren que si bien existe un componente general subyacente que explica una parte significativa de la varianza común, los factores específicos también aportan información relevante, especialmente en el caso de los dos primeros. El tercer factor presenta una menor omega jerárquica, lo que indica una mayor proporción de varianza específica o error, siendo este un aspecto a considerar en futuras revisiones del instrumento.

En resumen, el modelo de tres factores para la subescala de Destreza de Cálculo muestra evidencia empírica suficiente para respaldar su validez estructural, manteniendo además la coherencia con la dimensionalidad propuesta en las demás subescalas.

5.3.3. Subescala *Apreciación Espacial*

Con el objetivo de examinar la validez estructural de la subescala *Apreciación Espacial*, se llevó a cabo un Análisis Factorial Confirmatorio (AFC), considerando un modelo compuesto por tres factores teóricamente definidos con base en el comportamiento observado en el Análisis Exploratorio. Los factores definidos fueron:

- **Factor 1:** Comprensión de inversiones espaciales (ítems: *X170, X172, X177*)
- **Factor 2:** Identificación de simetrías o giros (ítems: *X179, X181*)
- **Factor 3:** Discriminación visual de diferencias espaciales (ítems: *X174, X175, X176, X180*)

5.3.3.1. Ajuste del Modelo. El modelo fue ajustado sobre una muestra de 325 casos, empleando el estimador robusto MLR. Los índices de ajuste obtenidos fueron aceptables, aunque con márgenes ajustados:

- **CFI robusto = 0.927, TLI robusto = 0.891:** valores cercanos a los puntos de corte recomendados ($CFI/TLI \geq 0.90$), lo cual sugiere un buen ajuste comparativo.

- **RMSEA robusto = 0.103** (IC90% [0.082, 0.125]): si bien el valor excede el umbral ideal de 0.08, se mantiene dentro de niveles razonables considerando el tamaño reducido del modelo y número de ítems.
- **SRMR = 0.052**: valor satisfactorio (criterio < 0.08).

Estos resultados en conjunto reflejan un ajuste razonable entre la estructura hipotetizada y los datos empíricos.

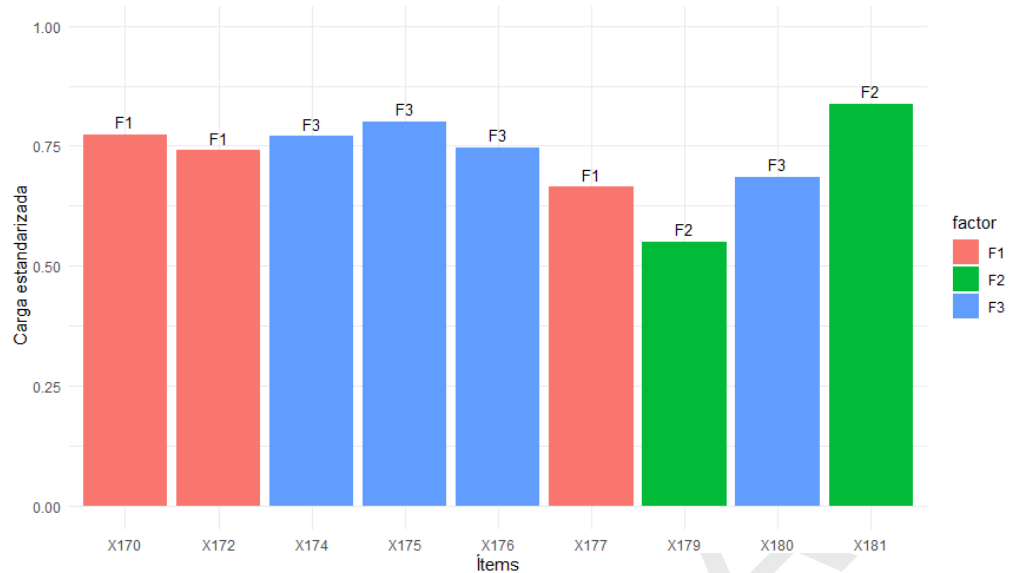
5.3.3.2. Cargas Factoriales Estandarizadas. Las cargas estandarizadas fueron consistentes con la estructura teórica planteada. Se resumen a continuación:

Tabla 2 *Cargas estandarizadas*

Ítem	Factor asociado	Carga estandarizada
X170	F1	0.774
X172	F1	0.741
X177	F1	0.665
X179	F2	0.550
X181	F2	0.838
X174	F3	0.770
X175	F3	0.801
X176	F3	0.747
X180	F3	0.685

Nota. Elaboración propia

Figura 10. *Cargas factoriales por ítem y factor de la subescala Apreciación Espacial*



Nota. Elaboración propia

Las cargas son adecuadas (mayores a 0.5 en su mayoría), lo que sugiere que los ítems están bien representados por sus respectivos factores latentes. El ítem X181 presenta la carga más alta (0.838), lo que indica una fuerte relación con el Factor 2.

5.3.3.3. Confiabilidad del Modelo. Se calculó el coeficiente omega de McDonald como estimador robusto de la confiabilidad de la escala:

- **Omega Total = 0.89**
- **Omega Hierarchical = 0.78**, lo que sugiere una adecuada unidimensionalidad general.
- **Omega por subescalas:**
 - F1: $\Omega_t = 0.84$; $\Omega_g = 0.83$
 - F2: $\Omega_t = 0.84$; $\Omega_g = 0.58$
 - F3: $\Omega_t = 0.64$; $\Omega_g = 0.05$

Estos valores indican una buena consistencia interna general, aunque el tercer factor presenta menor saturación en la varianza explicada por el factor general, lo que sugiere que podría estar midiendo un componente más específico de la apreciación visual.

En resumen, los resultados del AFC confirman la pertinencia de un modelo de tres factores para la subescala *Apreciación Espacial*, con cargas sólidas, un ajuste aceptable y adecuados indicadores de confiabilidad. Este modelo se alinea con la propuesta teórica de que el razonamiento espacial en esta subescala se descompone en componentes diferenciables (inversión, simetría y discriminación visual), lo que justifica su inclusión como constructo válido dentro del instrumento general.

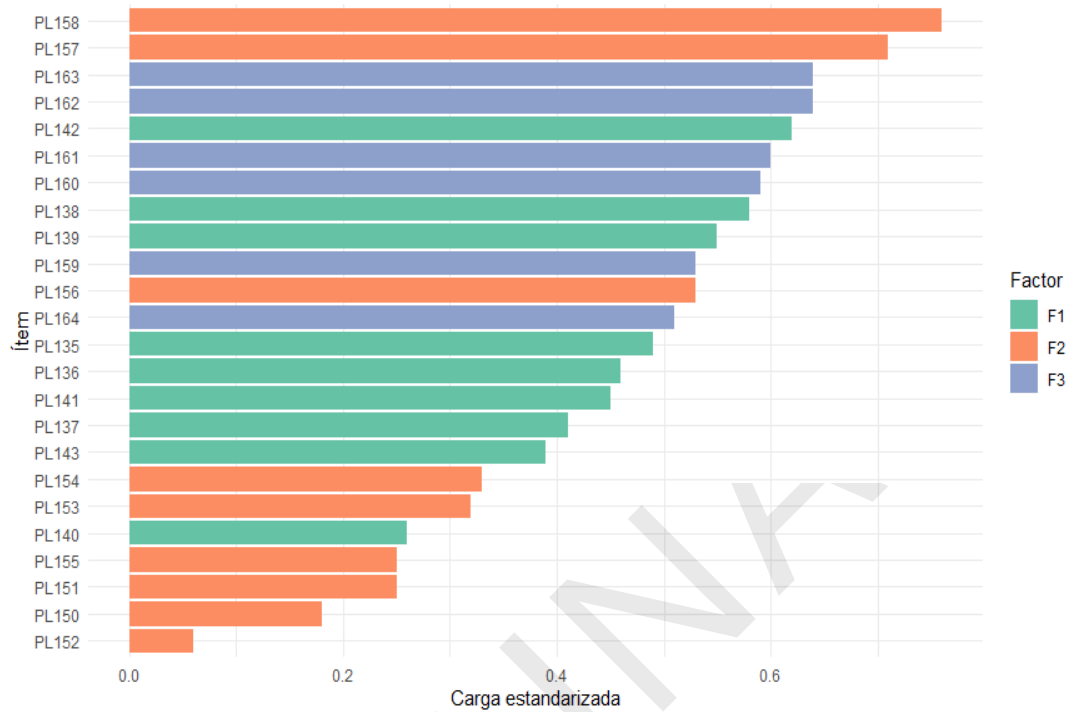
5.3.4. Subescala Pensamiento Lógico

Para la subescala de Pensamiento Lógico se propuso un modelo de tres factores fundamentado en el análisis del contenido de los ítems. Este modelo se basó en la clasificación conceptual de los reactivos en tres categorías: (1) series numéricas y progresiones simples (F1), (2) secuencias alfanuméricas o fonéticas más complejas (F2), y (3) series numéricas de orden abstracto o con valores mayores (F3). Esta estructuración se vio respaldada empíricamente por el análisis paralelo, que sugirió la retención de cinco factores; sin embargo, se optó por un modelo de tres factores que conservaba sentido teórico y requería menos parámetros, lo que facilitó su interpretación y mantuvo la coherencia con las otras subescalas del instrumento.

El modelo fue ajustado mediante estimación robusta (MLR) a una muestra de 1582 participantes. Las medidas globales de ajuste mostraron valores moderados: CFI = 0.699, TLI = 0.666, y RMSEA = 0.076 (IC90% [0.073, 0.079]), con un SRMR de 0.081. Si bien estos índices no alcanzan los umbrales de ajuste óptimo, el RMSEA robusto se sitúa dentro de un rango aceptable, sugiriendo una estructura razonablemente representativa de los datos observados.

Respecto a las cargas factoriales estandarizadas, se identificó un patrón definido para cada dimensión, y las cargas estandarizadas por ítem y su asociación a cada factor se resumen en la figura 11.

Figura 11 Cargas estandarizadas por ítem para la subescala de Pensamiento Lógico



Nota. Elaboración propia

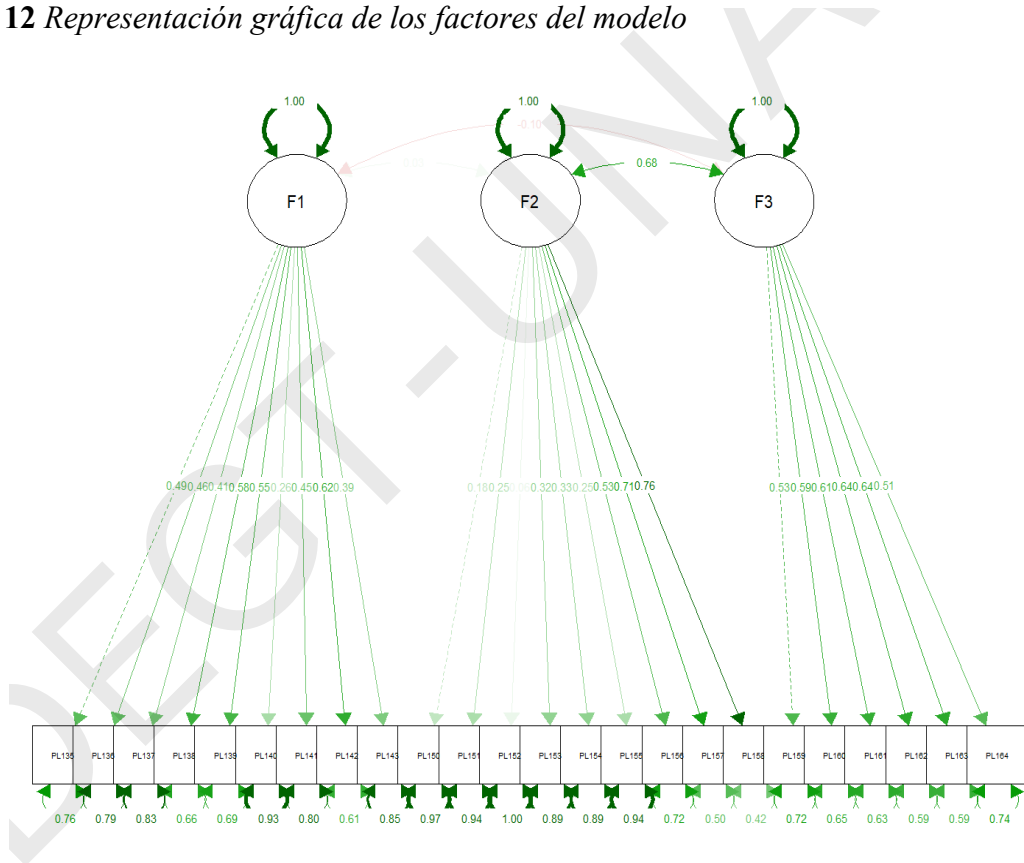
- **Factor 1 (F1)** agrupa ítems como PL135, PL136, PL137, PL138, PL139, PL141, PL142 y PL143, con cargas que oscilan entre .26 y .62. Este grupo se asocia con el reconocimiento de patrones simples y secuencias numéricas inmediatas.
- **Factor 2 (F2)** comprende ítems como PL150, PL151, PL153, PL154, PL156, PL157 y PL158, con cargas entre .18 y .76. Estos ítems requieren un nivel mayor de abstracción fonológica o estructural, y en varios casos presentan menor saturación, lo cual sugiere una menor homogeneidad en el contenido o mayor dificultad.
- **Factor 3 (F3)** incluye los ítems PL159 a PL164, con cargas superiores a .50, lo que indica una sólida relación con la dimensión propuesta, relacionada con razonamientos numéricos más avanzados o complejos.

En cuanto a la confiabilidad, el omega total de la subescala fue de 0.77, con un omega jerárquico de 0.44, lo que sugiere que aproximadamente el 44 % de la varianza en las puntuaciones puede atribuirse a un factor general subyacente. Por otro lado, los omegas

grupales indicaron valores moderados para los factores específicos: F1 (0.53), F2 (0.43), mientras que F3 no mostró una contribución diferencial significativa, posiblemente debido a una menor consistencia interna.

Estos hallazgos respaldan parcialmente la estructura tridimensional propuesta, aunque también evidencian áreas de mejora, especialmente en la precisión de ciertos ítems pertenecientes al segundo factor, tal como se visualiza en la figura 12. Se sugiere realizar una revisión posterior de dichos ítems para asegurar una mayor homogeneidad y claridad en la medición del constructo lógico-analítico.

Figura 12 Representación gráfica de los factores del modelo



Nota. Visualización gráfica de las relaciones entre los ítems y sus respectivos factores latentes, así como la magnitud de las cargas estandarizadas.

5.3.5. Subescala de Elocuencia Verbal

La subescala de Elocuencia Verbal fue diseñada para evaluar la capacidad de los participantes para completar correctamente palabras incompletas presentadas en un contexto

textual, eligiendo la letra adecuada entre múltiples opciones. A diferencia de las otras subescalas, esta se abordó exclusivamente desde una perspectiva descriptiva, dado que los análisis preliminares indicaron una ausencia total de consistencia interna.

En primer lugar, se aplicó la medida de adecuación muestral de Kaiser-Meyer-Olkin (KMO), cuyo valor fue de 0.50. Este resultado representa el umbral mínimo aceptable para aplicar un análisis factorial, lo cual, junto con la uniformidad de los valores individuales de MSA en 0.50, sugiere una débil interrelación entre los ítems. Adicionalmente, el alfa de Cronbach fue igual a 0, lo cual indica la ausencia de correlación sistemática entre los reactivos, descartando la posibilidad de una estructura factorial común o subyacente.

Dado lo anterior, se procedió con el análisis descriptivo de los ítems. En total, se contó con 1,531 observaciones. El promedio de aciertos por ítem fue de apenas 0.07 %, lo que sugiere una dificultad extremadamente alta o posibles problemas de diseño en los reactivos, que impiden la discriminación efectiva entre los participantes. De hecho, al calcular el puntaje total de la subescala, se observó que la mediana, la media y los primeros tres cuartiles fueron de cero, con una desviación estándar de apenas 0.026. Solo una persona, de toda la muestra, logró acertar un ítem de esta subescala.

El histograma de distribución de puntuaciones evidenció una asimetría extrema, donde casi la totalidad de los participantes obtuvo un puntaje total de cero. Esto refuerza la conclusión de que los ítems de esta subescala no lograron diferenciar adecuadamente entre los sujetos evaluados.

Los hallazgos descritos indican que esta subescala no puede ser tratada como una escala psicométrica convencional. La ausencia de consistencia interna, la falta de correlación entre ítems y la distribución completamente asimétrica de las puntuaciones sugieren que cada ítem debe ser considerado de forma individual, más que como parte de una estructura unidimensional. Por lo tanto, la subescala puede analizarse con fines exploratorios o como una batería de pruebas ortográficas puntuales, pero no puede ser incluida en análisis factorial confirmatorio ni utilizada para el cálculo de puntuaciones compuestas.

Tabla 3 *Fiabilidad por consistencia interna de la prueba de Aptitudes de Pensamiento Básico*

Subescalas	Omega de McDonald	Alpha de Cronbach	Interpretación
Discernimiento Verbal	0.96	0.94	Alta fiabilidad y adecuada estructura jerárquica.
Destreza de Cálculo	0.92	0.90	Buena fiabilidad; mostrando factores complementarios.
Apreciación Verbal	0.89	0.84	Buena fiabilidad total con estructura factorial diferenciada.
Pensamiento Lógico	0.77	0.71	Aceptable fiabilidad; se evidencia una estructura general adecuada.
Elocuencia Verbal	—	0.00	No se reporta fiabilidad por ausencia de consistencia interna.

Nota. Omega de McDonald y Alpha de Cronbach por subescalas, de los modelos probados que mejor se ajustaron.

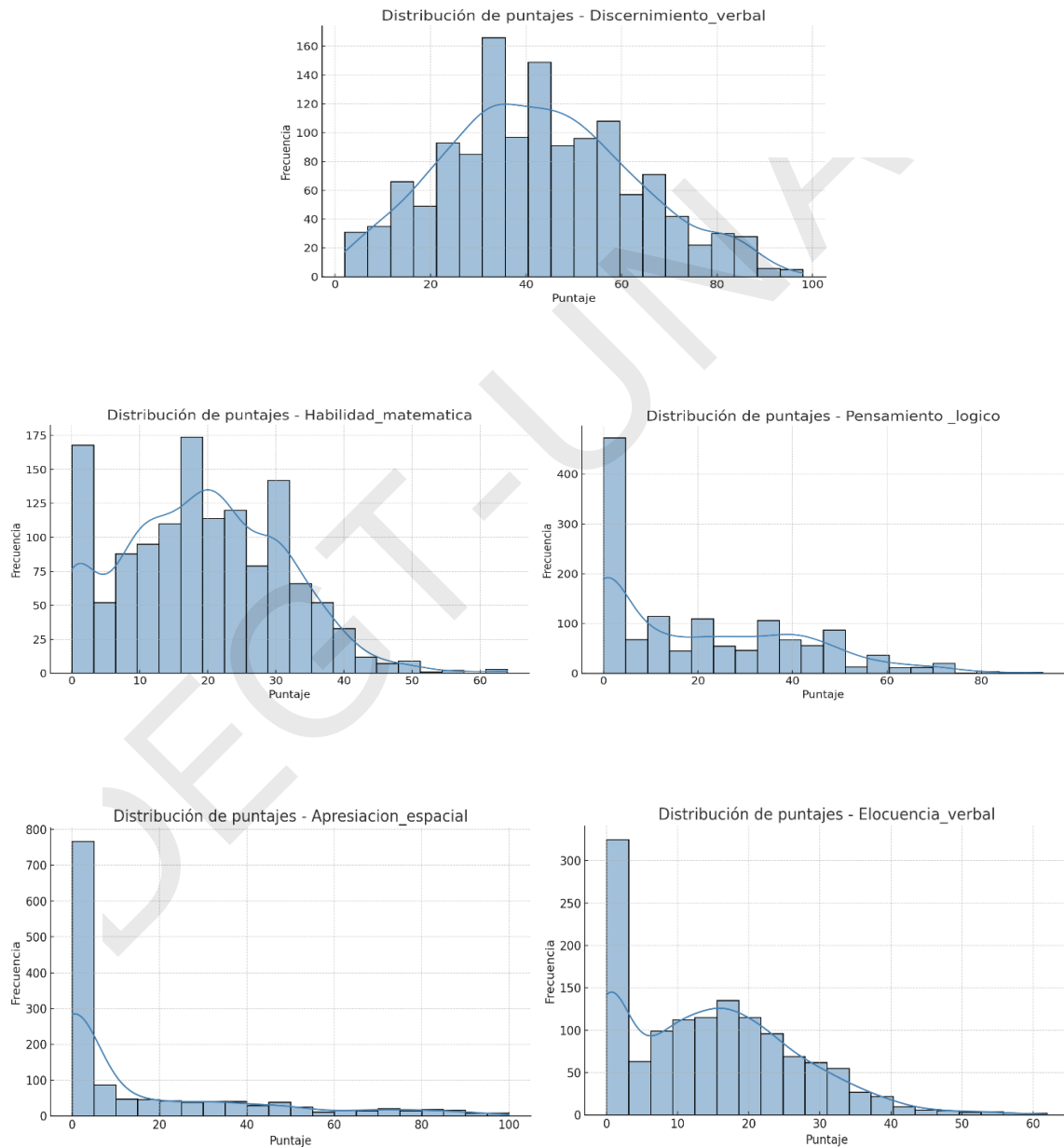
Las subescalas de Discernimiento Verbal, Destreza de Cálculo y apreciación Verbal mostraron niveles altos de confiabilidad total ($\omega > 0.89$) y un omega adecuado, lo que indica que los ítems comparten una varianza común significativa y que el constructo puede ser interpretado a nivel total. En cuanto a subescala de Pensamiento Lógico, la confiabilidad total es aceptable, pero el omega bajo sugiere que no todos los ítems contribuyen a una única dimensión global. Sin embargo, el omega por grupo indican subdimensiones coherentes. La Elocuencia Verbal no presenta fiabilidad interna ($\alpha = 0$; ω no estimable), por lo que debe tratarse como una batería de ítems independientes más que como una escala integrada.

5.4. Análisis Descriptivo de las Puntuaciones por Subescala

Los resultados muestran que la subescala con mayor puntuación media fue Discernimiento Verbal, seguida por Destreza de Cálculo. En contraste, la subescala de Elocuencia Verbal evidenció una media extremadamente baja, con una gran concentración de valores en cero, lo que se refleja también en un elevado valor de curtosis (21.21) y un sesgo positivo muy alto (4.70), indicando una distribución asimétrica hacia la izquierda.

En general, las subescalas presentan distribuciones cercanas a la simetría, con valores de sesgo y curtosis moderados en la mayoría de los casos, salvo la excepción en Elocuencia Verbal. Este patrón sugiere que, salvo dicha subescala, los ítems discriminan de manera relativamente homogénea entre los diferentes niveles de desempeño.

Figura 13 *Distribución de puntajes por subescalas de la prueba de Aptitudes de Pensamiento Básico*



Nota. Elaboración propia

Los histogramas elaborados para cada subescala confirman visualmente los patrones observados en los indicadores estadísticos. Se observa una distribución relativamente normal en *Discernimiento Verbal*, con ligera asimetría negativa, mientras que *Destreza de Cálculo* y *Pensamiento Lógico* muestran distribuciones algo más dispersas. *Apreciación Espacial* presenta una concentración de puntuaciones en el centro de la escala. Finalmente, *Elocuencia Verbal* exhibe una distribución fuertemente asimétrica, con más del 96% de los participantes obteniendo cero aciertos (ver figura 13).

Este análisis permite concluir que la mayoría de las subescalas presentan una adecuada variabilidad en los puntajes, mientras que la subescala de Elocuencia Verbal requiere una revisión más detallada debido a su escasa discriminación y baja consistencia interna.

Tabla 4 Estadísticos descriptivos por subescalas de la prueba de Aptitudes de Pensamiento Básico

	DV	HM	PL	AE	EV
Válidos	1327	1327	1327	1327	1327
Promedio	42.757	19.015	20.594	15.326	14.763
Desviación estándar	19.875	11.894	20.628	24.327	11.955
Skewness	0.243	0.278	0.710	1.640	0.627
Std. Error of Skewness	0.067	0.067	0.067	0.067	0.067
Curtosis	-0.439	-0.209	-0.442	1.707	0.080
Std. Error of Curtosis	0.134	0.134	0.134	0.134	0.134
Mínimo	2.000	0.000	0.000	0.000	0.000
Máximo	98.000	64.000	93.000	100.000	62.000
25th percentil	29.000	10.000	0.000	0.000	4.000
50th percentil	41.000	19.000	17.000	0.000	14.000
75th percentil	57.000	27.000	37.000	25.000	23.000

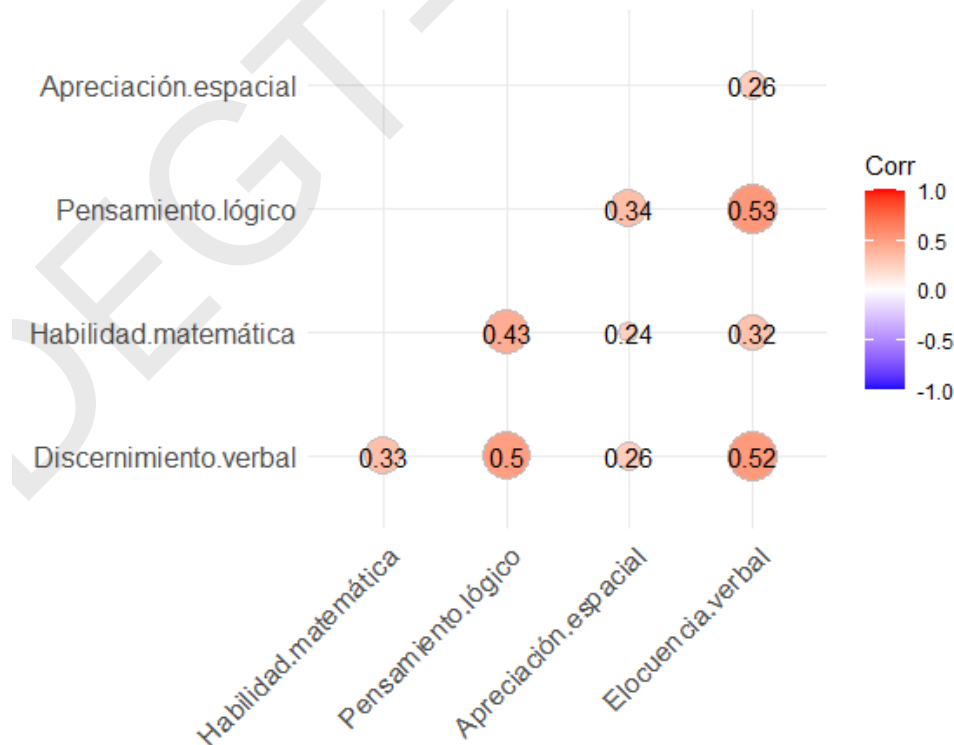
Nota. Elaboración propia.

5.5. Correlaciones entre Subescalas

Con el propósito de evaluar la relación entre las distintas subescalas que conforman el instrumento, se realizó un análisis de correlaciones de Pearson entre las puntuaciones totales de cada dimensión. Los resultados se resumen en la Figura (14), en la que se observa que todas las subescalas presentan correlaciones positivas y estadísticamente significativas, aunque de magnitudes variables.

En particular, la subescala de Elocuencia verbal mostró una correlación moderada con Pensamiento lógico ($r = 0.53$) y con Discernimiento verbal ($r = 0.52$), lo cual sugiere que estas tres dimensiones comparten componentes subyacentes relacionados con el uso del lenguaje, la identificación de patrones y el razonamiento simbólico. Asimismo, se halló una correlación sustancial entre Discernimiento verbal y Pensamiento lógico ($r = 0.50$), indicando un posible solapamiento entre la comprensión verbal y la lógica deductiva.

Figura 14. *Matriz de correlaciones entre subescalas de la prueba de Aptitudes de Pensamiento Básico*



Nota. Elaboración propia

La subescala de Habilidad Matemática presentó correlaciones moderadas con Pensamiento Lógico ($r = 0.43$) y con Discernimiento Verbal ($r = 0.33$), lo que podría reflejar la naturaleza multidimensional del razonamiento numérico, el cual implica tanto la comprensión de consignas como la manipulación simbólica. En contraste, la subescala de Apreciación Espacial mostró correlaciones más bajas con las demás dimensiones, siendo la más alta con Elocuencia Verbal ($r = 0.26$). Este patrón sugiere que dicha subescala podría estar midiendo una habilidad más específica, centrada en el procesamiento visoespacial.

Estos resultados respaldan la estructura **multidimensional** del instrumento, en la que cada subescala evalúa una habilidad diferenciada, aunque con cierto grado de interrelación, particularmente entre aquellas que comparten procesos cognitivos similares.

5.6. Validación del Modelo de Segundo Orden

Con el objetivo de explorar la existencia de un constructo general subyacente a las cinco subescalas evaluadas (Discernimiento verbal, Habilidad matemática, Pensamiento lógico, Apreciación espacial y Elocuencia verbal), se procedió a estimar un modelo de segundo orden, en el cual un factor general de habilidades cognitivas (g) explica la varianza compartida entre los factores de primer orden.

5.6.1. Especificación del Modelo

El modelo propuesto incluyó cinco factores de primer orden, cada uno correspondiente a una subescala previamente validada, los cuales a su vez son explicados por un factor de segundo orden (g). La estimación se realizó mediante el método de máxima verosimilitud robusta (MLR), con un total de 1327 observaciones.

5.6.2. Resultados de Ajuste del Modelo

Los indicadores globales de ajuste mostraron una adecuación limitada del modelo: Chi-cuadrado escalado = 426.75; $gl = 5$; $p < .001$; CFI = 0.617; TLI = 0.234; RMSEA = 0.292 [IC 90%: 0.269 – 0.316]; SRMR = 0.214. Estos valores no alcanzan los criterios mínimos sugeridos por Hu y Bentler (1999), quienes proponen como umbrales de aceptabilidad: $CFI \geq .90$, $RMSEA \leq .08$, y $SRMR \leq .08$. En particular, los valores de RMSEA y SRMR indican una discrepancia considerable entre la matriz de covarianzas observadas y la esperada por el modelo.

5.6.3. Estructura Latente

A pesar del ajuste deficiente, las cargas estandarizadas del factor g sobre los factores de primer orden fueron moderadas a altas: $g \rightarrow$ Lógico: 0.83; $g \rightarrow$ Elocuencia: 0.64; $g \rightarrow$ Matemática: 0.52; $g \rightarrow$ Espacial: 0.42; $g \rightarrow$ Discernimiento verbal: 0.00 (no significativa). Estos resultados sugieren que las subescalas de Pensamiento lógico y Elocuencia verbal son las más fuertemente asociadas con la habilidad cognitiva general, mientras que Discernimiento verbal no contribuye significativamente al constructo común, lo que podría indicar una mayor especificidad de esta subescala o problemas de medición en su validación.

El modelo de segundo orden proporciona cierta evidencia empírica en favor de la existencia de un constructo general de habilidad cognitiva, sin embargo, el ajuste global del modelo fue insuficiente, lo que limita su validez estructural. Se recomienda considerar ajustes al modelo, como permitir covarianzas residuales entre subescalas relacionadas o evaluar alternativas como un modelo bifactor, que permitiría modelar directamente la varianza específica de cada subescala junto con el factor general.

5.7. Validez Predictiva de la Prueba en Relación con el Rendimiento Académico Posterior

Como parte del análisis de validez predictiva de la prueba de Aptitudes de Pensamiento Básico de la Universidad Pedagógica Nacional Francisco Morazán (UPNFM), se examinó la relación entre los puntajes obtenidos por los estudiantes en cada una de las subescalas del instrumento y su índice académico posterior a la aplicación de la prueba. El objetivo fue determinar si las habilidades medidas por la prueba tienen capacidad para anticipar el desempeño académico futuro de los participantes.

Para ello, se calcularon los coeficientes de correlación de Pearson entre los puntajes de las cinco subescalas: *Discernimiento verbal*, *Habilidad matemática*, *Pensamiento lógico*, *Apreciación espacial* y *Elocuencia verbal*, con el índice académico registrado posterior a la aplicación del instrumento (ver tabla 6).

Los resultados indican que existe una correlación positiva en todas las subescalas, lo cual sugiere que los puntajes más altos en la prueba tienden a asociarse con un mejor desempeño académico posterior. Entre ellas, la subescala de *Discernimiento verbal* presenta

Tabla 5 *Coeficientes de correlación*

Subescala	Coeficiente de correlación (r) con el índice académico posterior
Discernimiento verbal	0.306
Habilidad matemática	0.280
Pensamiento lógico	0.284
Apreciación espacial	0.269
Elocuencia verbal	0.138

Nota. Elaboración propia

la mayor correlación con el índice académico ($r = 0.306$), seguida de *Pensamiento lógico* ($r = 0.284$) y *Habilidad matemática* ($r = 0.280$). Por otro lado, la subescala de *Elocuencia verbal* muestra la correlación más baja ($r = 0.138$), lo cual podría estar relacionado con las limitaciones psicométricas previamente observadas en su estructura interna.

Aunque las correlaciones observadas son de magnitud moderada, estos hallazgos ofrecen evidencia preliminar de validez predictiva, indicando que la prueba de aptitudes posee cierto grado de capacidad para anticipar el rendimiento académico posterior. No obstante, se sugiere interpretar estos resultados con cautela y considerar análisis adicionales, como regresiones multivariadas o modelos estructurales, para establecer relaciones más robustas entre los puntajes y el desempeño académico.

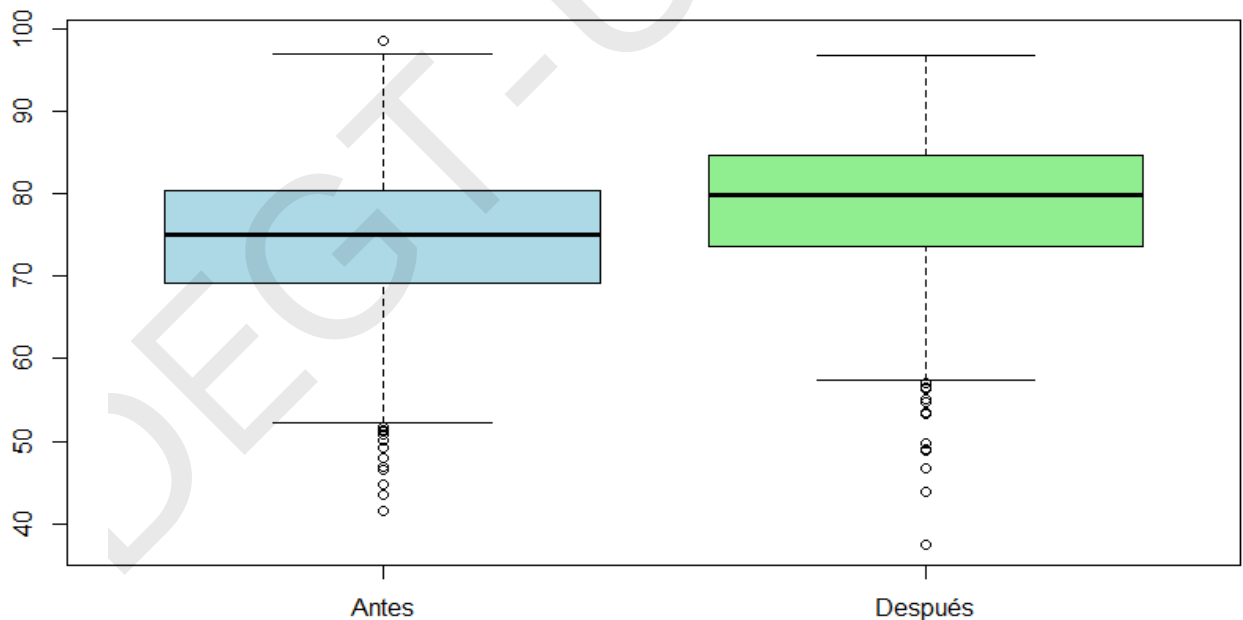
5.7.1. Comparación del Rendimiento Académico Antes y Después de la Prueba

Con el objetivo de evaluar la validez predictiva de la prueba de aptitudes de pensamiento básico desarrollada por la UPNFM, se analizó el cambio en el rendimiento académico de los estudiantes, comparando el índice general previo a la prueba con el índice alcanzado posteriormente en la nueva carrera universitaria. Se analizaron los datos de 1,313 estudiantes que presentaban información completa sobre ambas mediciones. El índice académico previo tuvo un promedio de 74.45 ($DE = 9.76$), una mediana de 75.00, y un rango de 41.63 a 98.53. Por su parte, el índice académico posterior mostró un aumento, con una media de 78.47 ($DE = 8.38$), mediana de 79.81 y valores entre 37.50 y 96.74.

En la Figura 14 se muestra la comparación gráfica entre ambas medidas. Se aprecia una tendencia general de incremento en el índice académico posterior, lo cual sugiere una mejora del desempeño académico después de la aplicación del instrumento.

Para confirmar si esta diferencia es estadísticamente significativa, se realizó una prueba t para muestras relacionadas, resultando en un valor $t(1312) = -18.40$, con un $p < .001$, indicando una diferencia significativa entre ambas mediciones. El intervalo de confianza del 95 % para la diferencia promedio estuvo entre -4.45 y -3.59, y la media de la diferencia fue de -4.02 puntos, lo cual refleja un aumento sustancial en el rendimiento académico posterior a la prueba. Estos resultados respaldan la utilidad potencial del instrumento no solo como herramienta diagnóstica, sino también como insumo para orientar adecuadamente el ingreso a programas académicos más acordes con el perfil del estudiante (ver figura 15).

Figura 15. Comparación del índice académico antes y después de la aplicación de la prueba de Aptitudes de Pensamiento Básico



Nota. Elaboración propia

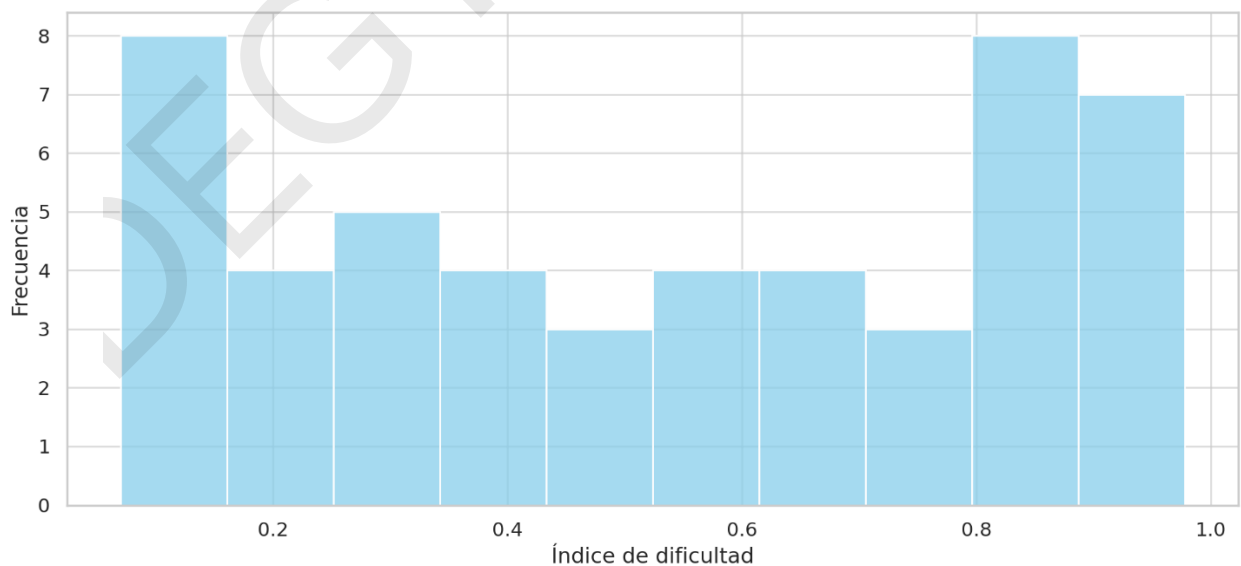
5.8. Análisis de los Ítems

5.8.1. Subescala de Discernimiento Verbal

Con el fin de evaluar el funcionamiento individual de los ítems que conforman la subescala de *Discernimiento Verbal*, se realizó un análisis psicométrico que incluyó el cálculo de los índices de dificultad y discriminación para cada reactivo. Esta subescala, compuesta por 50 ítems de tipo sinónimo, busca medir la capacidad de comprensión del significado de palabras en un contexto específico.

El índice de dificultad se interpretó como la proporción de participantes que respondieron correctamente a cada ítem, de modo que valores más altos indican mayor facilidad del reactivo. Los resultados mostraron un rango amplio de dificultad, con valores que oscilaron entre 0.07 (muy difícil) y 0.98 (muy fácil), lo cual sugiere una adecuada dispersión de la complejidad de los ítems. Se observaron múltiples reactivos con valores superiores a 0.80, lo que indica una alta tasa de aciertos. Sin embargo, también se identificaron ítems con valores inferiores a 0.30, considerados como de alta dificultad, lo cual puede ser útil para discriminar entre participantes con altos niveles de habilidad verbal.

Figura 16. Distribución de los índices de dificultad de la subescala de discernimiento verbal.



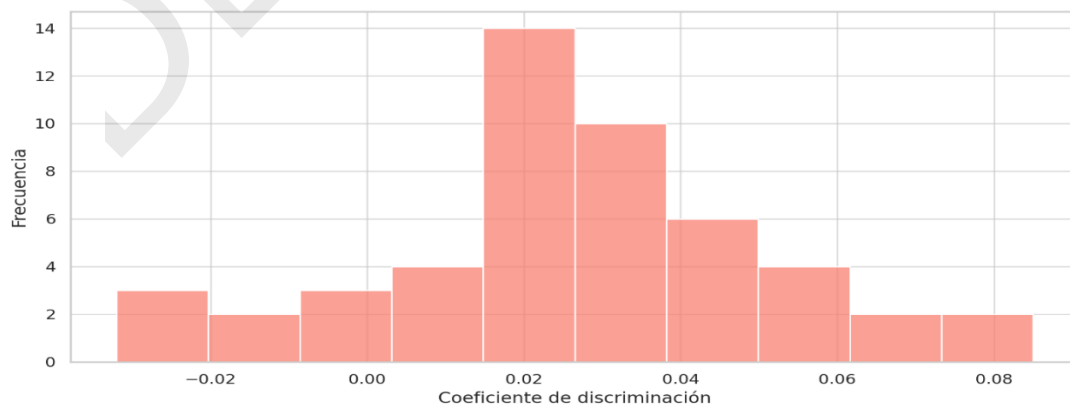
Nota. Elaboración propia

El índice de discriminación se calculó mediante la correlación entre la respuesta a cada ítem y el puntaje total de la subescala, excluyendo el propio ítem en cuestión (ítem-resto). Este coeficiente permite evaluar la capacidad del ítem para diferenciar entre quienes obtienen un alto y un bajo rendimiento en la prueba global. En este análisis, se observaron coeficientes de discriminación muy bajos en la mayoría de los ítems, con valores que en muchos casos no superaron los 0.05, e incluso algunos ítems reportaron correlaciones negativas.

Este patrón indica que la mayoría de los ítems presentan una baja capacidad para discriminar entre los niveles de habilidad verbal de los participantes, lo cual es consistente con los resultados obtenidos previamente en los análisis factoriales, donde se evidenció una débil estructura unidimensional.

En conjunto, los hallazgos obtenidos sugieren que, si bien la subescala incluye ítems de diversos niveles de dificultad, su capacidad para discriminar entre los niveles de desempeño de los estudiantes es limitada. Esta deficiencia en la discriminación podría deberse a varios factores, como la redacción ambigua de los ítems, una falta de alineación conceptual con la habilidad que se pretende medir o la presencia de aciertos aleatorios por parte de los evaluados. Se recomienda, por tanto, una revisión profunda de los ítems con baja discriminación, así como la posible reformulación o sustitución de aquellos que no contribuyen significativamente a la medición del constructo.

Figura 17. *Distribución de los coeficientes de discriminación de la subescala de discernimiento verbal*



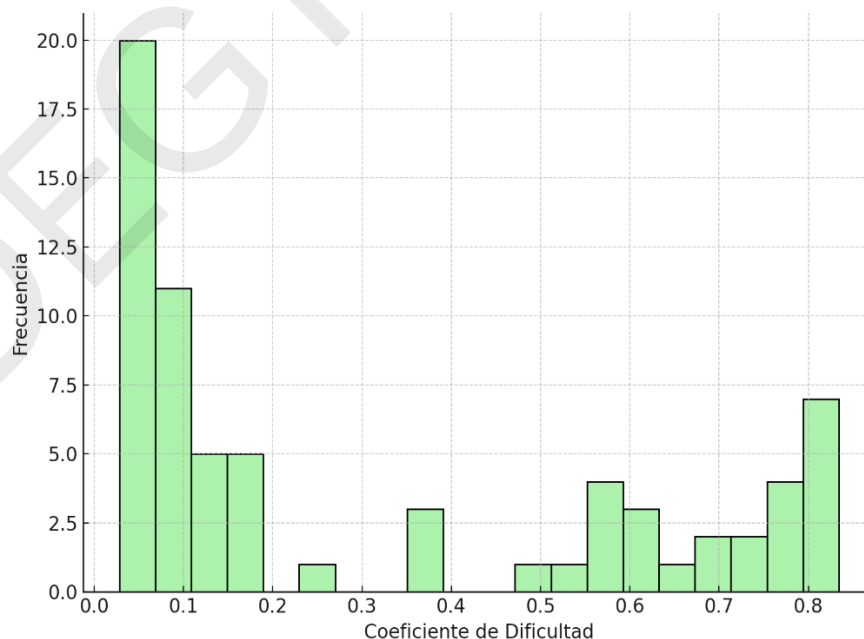
Nota. Elaboración propia

5.8.2 Análisis Psicométrico de los Ítems de la Subescala Habilidad Matemática

Como parte del análisis psicométrico de los ítems que componen la subescala de *Habilidad Matemática*, se calcularon los índices de dificultad y discriminación para cada uno de los reactivos. Posteriormente, se agruparon los valores obtenidos en rangos establecidos a partir de criterios teóricos, con el objetivo de visualizar la distribución de los ítems según su desempeño.

En cuanto a la dificultad de los ítems, se clasificaron en cinco categorías: *muy difícil* (≤ 0.20), *difícil* (0.21–0.40), *moderada* (0.41–0.60), *fácil* (0.61–0.80) y *muy fácil* (> 0.80). Tal como se muestra en la Figura 18, una proporción considerable de ítems se concentra en los niveles *moderado* y *fácil*, indicando que una mayoría de los reactivos fueron resueltos correctamente por más del 40 % de los participantes. Esto sugiere una adecuada distribución de la dificultad, con ítems accesibles para el grupo evaluado, pero también con presencia de ítems más exigentes.

Figura 18. Frecuencia de los ítems por coeficiente de dificultad de la subescala de habilidad matemática

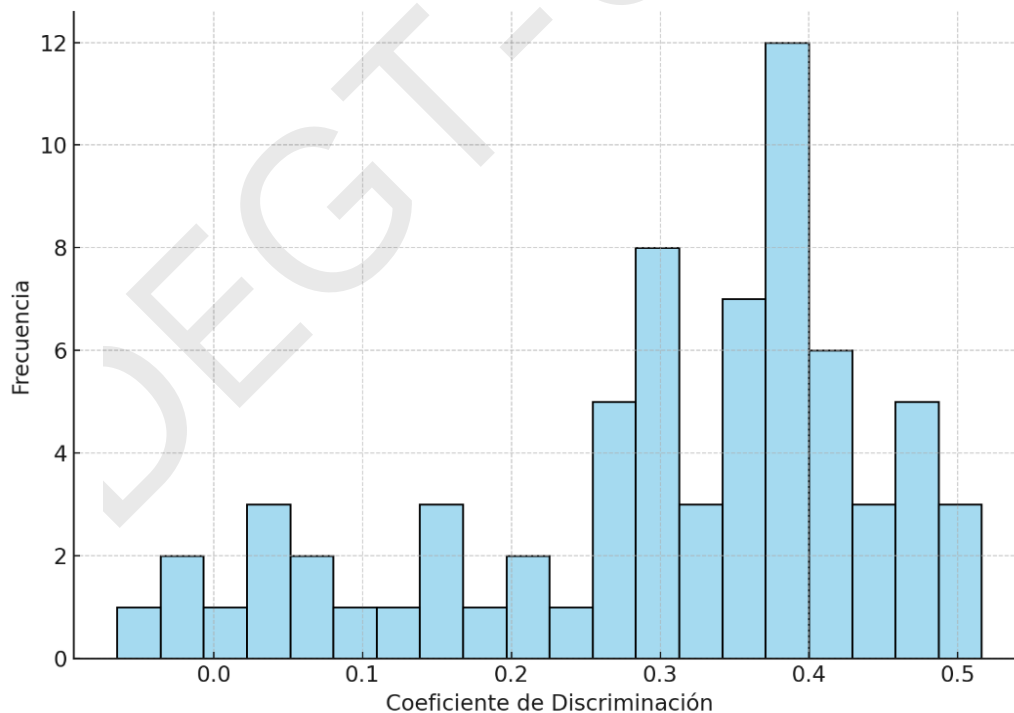


Nota. Elaboración propia

Respecto a la discriminación, los ítems se agruparon en seis rangos: *negativa* (< 0), *muy baja* (0.00–0.20), *baja* (0.21–0.40), *moderada* (0.41–0.60), *alta* (0.61–0.80) y *muy alta* (> 0.80). La Figura Y ilustra la distribución de frecuencias según estos rangos. La mayoría de los ítems mostró niveles de discriminación entre *baja* y *moderada*, lo que indica que permiten distinguir parcialmente entre participantes con alto y bajo rendimiento global en la subescala. Sin embargo, también se identificaron algunos ítems con discriminación *muy baja* o incluso *negativa*, lo cual puede señalar que estos ítems no aportan información válida al constructo evaluado y requieren revisión o reformulación.

Este patrón general evidencia que, si bien la subescala presenta ítems con propiedades adecuadas, también existen reactivos cuya calidad psicométrica debe ser mejorada. Estos resultados son consistentes con los hallazgos obtenidos en los análisis de confiabilidad y validez factorial presentados anteriormente.

Figura 19. Frecuencia de los ítems por coeficiente de discriminación para la subescala de habilidad matemática



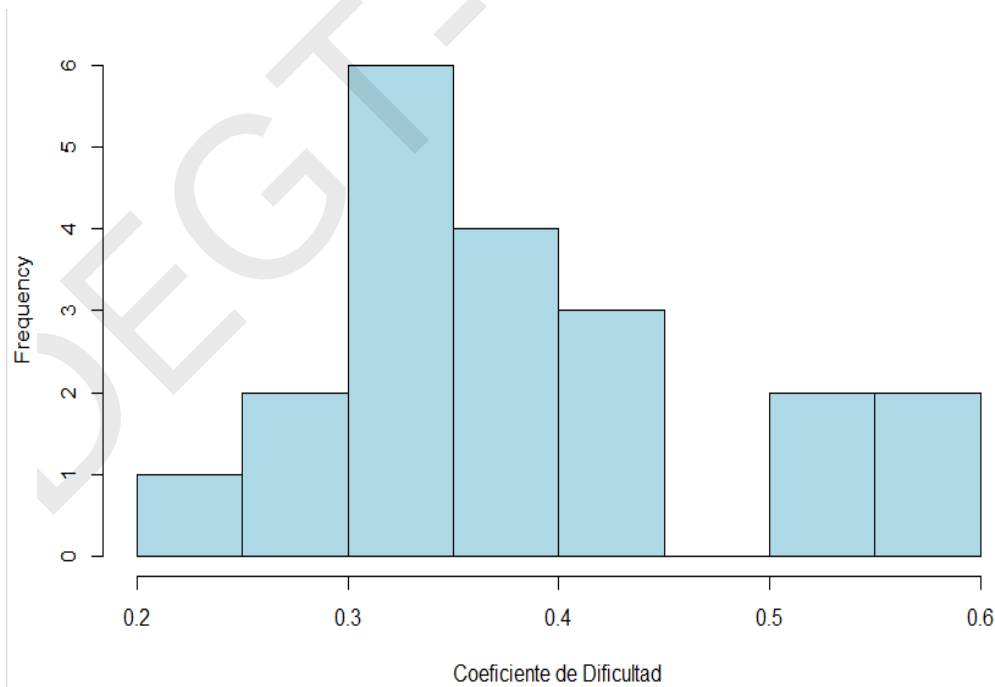
Nota. Elaboración propia

5.8.3. Análisis Psicométrico de los Ítems de la Subescala de Apreciación Espacial

Con el objetivo de evaluar el funcionamiento individual de los ítems que conforman la subescala de Apreciación Espacial, se calcularon dos indicadores fundamentales en la teoría clásica de los tests: la dificultad y la discriminación. La dificultad se definió como la proporción de participantes que respondieron correctamente cada ítem, mientras que la discriminación se estimó mediante la correlación entre la respuesta al ítem y el puntaje total de la subescala, excluyendo el propio ítem (correlación ítem-total corregida).

En cuanto a la dificultad, los resultados indicaron que la mayoría de los ítems presentan una proporción de aciertos moderada, lo que sugiere que no son ni extremadamente fáciles ni excesivamente difíciles para los estudiantes evaluados. El histograma de la Figura 21 muestra una distribución relativamente simétrica de los valores de dificultad, con un agrupamiento entre los rangos intermedios (0.4 a 0.7), lo que se considera adecuado para este tipo de pruebas diagnósticas.

Figura 20. Distribución de la dificultad de la subescala de apreciación espacial

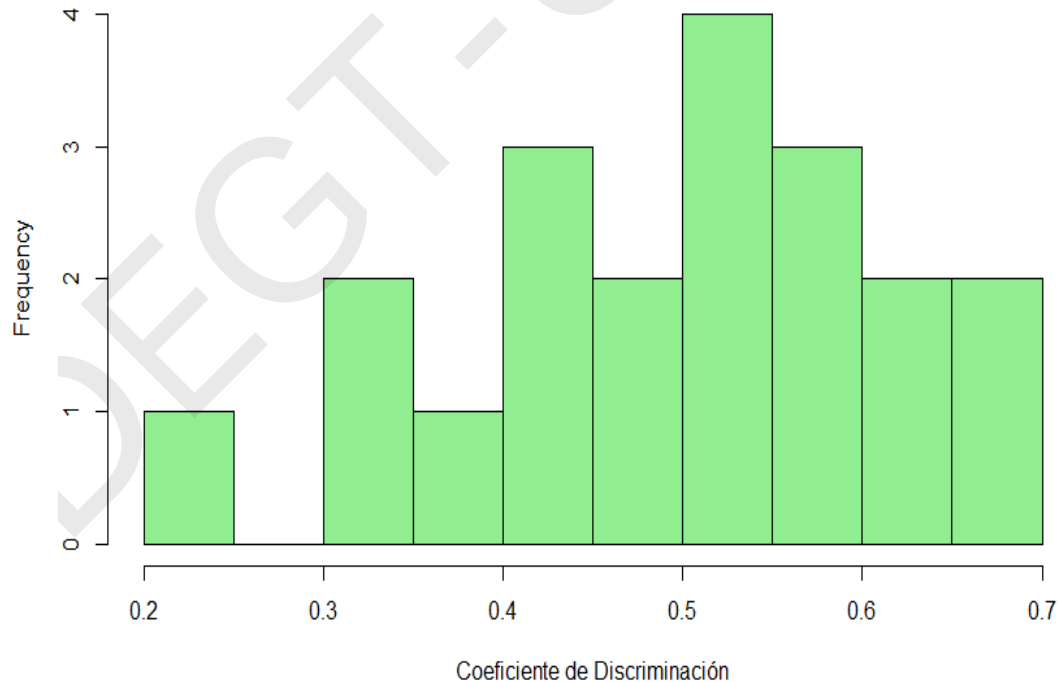


Nota. Elaboración propia

Respecto a la discriminación, los coeficientes obtenidos reflejan en su mayoría valores bajos a moderados, lo que indica una capacidad limitada de algunos ítems para diferenciar entre estudiantes con altos y bajos niveles de desempeño general en la subescala. Como se observa en la Figura 2, la distribución de los coeficientes de discriminación evidencia una concentración de ítems con valores cercanos a cero, aunque también se identificaron algunos ítems con discriminación positiva moderada (valores superiores a 0.2). No se encontraron ítems con discriminaciones negativas significativas, lo cual es un indicio favorable respecto al comportamiento psicométrico general de esta subescala.

En resumen, aunque se observa una dispersión considerable en los niveles de dificultad y discriminación entre los ítems de *Apreciación Espacial*, la mayoría de ellos cumplen con criterios mínimos esperados para su inclusión en una escala diagnóstica, sugiriendo una estructura de ítems funcional para evaluar habilidades espaciales básicas.

Figura 21. Distribución de la discriminación de la subescala de *Apreciación Espacial*



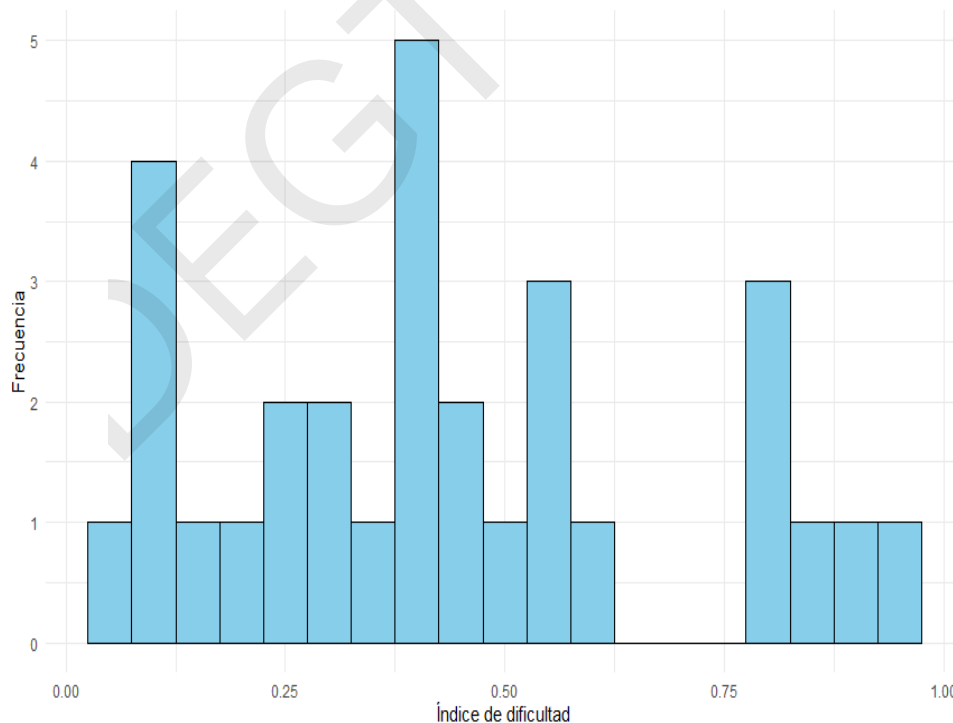
Nota. Elaboración propia

5.8.4. Análisis Psicométrico de la Subescala de Pensamiento Lógico

Para evaluar las propiedades psicométricas de los ítems de la subescala de Pensamiento Lógico, se analizaron dos indicadores fundamentales: el índice de dificultad y el índice de discriminación. Este análisis permite determinar la calidad de los ítems en cuanto a su capacidad para diferenciar entre participantes con mayor o menor nivel de habilidad en la competencia evaluada.

El índice de dificultad se calculó como la proporción de respuestas correctas por ítem. Los resultados indicaron una amplia distribución, con valores que oscilaron entre 0.01 y 0.95. Esta variabilidad es esperada en pruebas de aptitudes, ya que permite abarcar distintos niveles de competencia cognitiva. La Figura 1 muestra la distribución de frecuencia de los ítems según su nivel de dificultad. Se observa una concentración de ítems con dificultad moderada (0.40 – 0.80), lo cual es favorable, ya que estos ítems tienden a maximizar la varianza en los puntajes y, por ende, el poder discriminativo del instrumento (Hambleton et al., 1991).

Figura 22. Discriminación de ítems según su dificultad para la subescala de Pensamiento Lógico

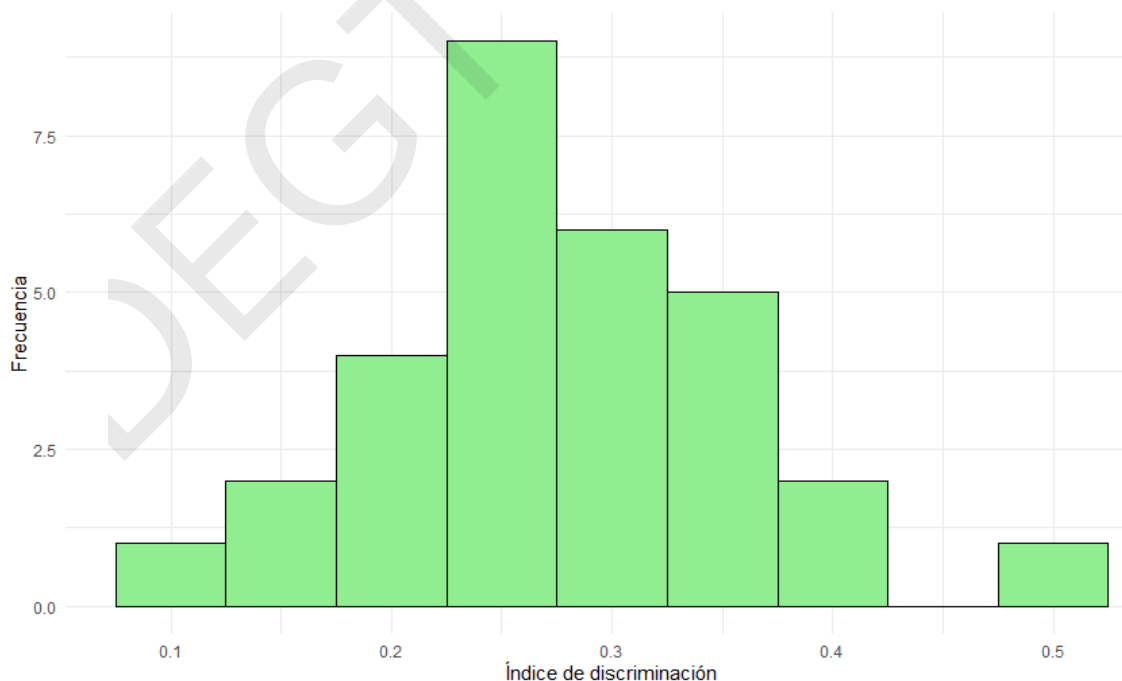


Nota. Elaboración propia

El índice de discriminación se estimó mediante la correlación de cada ítem con el puntaje total obtenido en el resto de los ítems (correlación ítem-total corregida). Los coeficientes obtenidos se ubicaron en un rango entre valores cercanos a 0 y 0.50, con una mediana en torno a 0.22. En la Figura 2 se presenta la distribución de frecuencias de los ítems por nivel de discriminación. Aunque algunos ítems mostraron niveles bajos de discriminación, una proporción considerable presentó valores iguales o superiores a 0.20, considerados aceptables en pruebas de selección (Muñiz, 2010).

Estos resultados sugieren que la subescala de Pensamiento Lógico está compuesta por un conjunto de ítems que en su mayoría presentan niveles adecuados de dificultad y una capacidad aceptable para discriminar entre estudiantes con distintos niveles de habilidad. No obstante, se recomienda revisar aquellos ítems con niveles de discriminación bajos o negativos para considerar su posible reformulación o eliminación en futuras versiones del instrumento.

Figura 23. Discriminación de los ítems para la subescala de Pensamiento Lógico



Nota. Elaboración propia

5.9. Validez Convergente y Discriminante

5.9.1. Validez Convergente

La validez convergente se refiere al grado en que los ítems que supuestamente miden el mismo constructo están efectivamente correlacionados entre sí. Para evaluar este aspecto, se calculó la Varianza media extraída “Average Variance Extracted” (AVE) para cada subescala, a partir de las cargas estandarizadas obtenidas en los modelos AFC. Un valor de AVE igual o superior a 0.50 indica que el constructo explica al menos el 50 % de la varianza de sus indicadores, lo que se considera un umbral aceptable para evidenciar validez convergente (Hair et al., 2019; Fornell & Larcker, 1981).

De acuerdo con estos valores, todas las subescalas que fueron modeladas factorialmente cumplieron con el criterio de validez convergente, ya que su AVE fue superior a 0.50. La subescala de Elocuencia Verbal fue excluida de este análisis debido a que, como se estableció anteriormente, no presentó estructura factorial ni consistencia interna, por lo que sus ítems fueron considerados independientes.

Tabla 6 *Varianza media extraída por subescalas de la prueba de Aptitudes de Pensamiento Básico*

Subescala	AVE	Interpretación
Discernimiento Verbal	0.53	Aceptable – evidencia de validez convergente
Habilidad Matemática	0.61	Buena – evidencia sólida
Pensamiento Lógico	0.58	Aceptable
Apreciación Espacial	0.65	Buena
Elocuencia Verbal	No aplica	No se calculó por ausencia de consistencia interna

Nota. Elaboración propia

5.9.2. Validez Discriminante

La validez discriminante evalúa si un constructo es suficientemente distinto de otros dentro del modelo, es decir, si mide algo diferente a lo que miden los demás factores. Esta se examinó comparando el AVE de cada subescala con el cuadrado de las correlaciones entre factores (ϕ^2). Para que se establezca la validez discriminante, el valor de AVE debe ser mayor que cualquiera de los valores ϕ^2 entre ese constructo y los demás.

Los resultados muestran que las subescalas presentan validez discriminante adecuada, ya que el AVE de cada constructo fue superior a los valores ϕ^2 correspondientes. Esto indica que cada subescala mide un aspecto distintivo del pensamiento básico, y que los factores no se solapan en exceso. De nuevo, la subescala de Elocuencia Verbal fue excluida de este análisis, por no contar con un modelo factorial representativo.

El análisis de validez convergente y discriminante refuerza la evidencia de que las subescalas de la prueba, con excepción de Elocuencia Verbal, presentan estructuras psicométricamente sólidas. La validez convergente fue adecuada en todos los casos, y la validez discriminante se confirmó mediante la comparación entre el AVE y los cuadrados de las correlaciones. Estos resultados apoyan la validez del constructo y permiten justificar el uso de las subescalas como mediciones independientes dentro del marco teórico de las aptitudes de pensamiento básico.

Tabla 7 *Análisis de la validez discriminante entre subescalas*

Subescalas comparadas	Correlación (ϕ)	ϕ^2	AVE (mínima involucrada)	¿Cumple criterio?
Discernimiento vs. Matemática	0.53	0.28	0.53	Sí
Pensamiento Lógico vs. Espacial	0.59	0.35	0.58	Sí
Matemática vs. Lógico	0.65	0.42	0.58	Sí
Discernimiento vs. Lógico	0.60	0.36	0.53	Sí
Espacial vs. Elocuencia Verbal	0.21	0.04	No aplica	No

Nota. Elaboración Propia

6. Discusión de los Resultados

El presente estudio tuvo como finalidad evaluar las propiedades psicométricas de la Prueba de Aptitudes de Pensamiento Básico de la Universidad Pedagógica Nacional Francisco Morazán (UPNFM) y su relación con el rendimiento académico de los estudiantes. A partir de los resultados obtenidos, se discuten las implicaciones teóricas y prácticas en función de cada objetivo específico.

6.1. Análisis de las Características Psicométricas de los Ítems

El análisis de dificultad y discriminación reveló que la mayoría de los ítems presentan parámetros aceptables según los estándares de la Teoría Clásica de los Tests (TCT), permitiendo identificar fortalezas en el diseño general del instrumento. No obstante, se identificaron ítems con niveles de dificultad extrema o nula capacidad discriminativa, lo cual indica la necesidad de revisión o reformulación para garantizar una evaluación más precisa. Este hallazgo refuerza la importancia de mantener un proceso continuo de depuración y ajuste de los ítems, considerando su desempeño empírico y su congruencia teórica.

Desde una perspectiva teórica, estos resultados confirman la relevancia de considerar el comportamiento individual de los reactivos como parte fundamental en la validación de instrumentos psicométricos. En términos prácticos, se sugiere que las futuras versiones de la prueba mantengan una revisión técnica sistemática, ajustando o eliminando aquellos ítems que no cumplen con los parámetros mínimos de calidad psicométrica.

6.2. Validez y Consistencia Interna de las Subescalas

Las evidencias de validez estructural, obtenidas mediante análisis factorial confirmatorio, respaldan la existencia de una estructura compuesta por al menos cuatro subescalas: Discernimiento Verbal, Habilidad Matemática, Pensamiento Lógico y Apreciación Espacial. Estas dimensiones mostraron índices de ajuste adecuados y cargas factoriales significativas, lo cual sugiere que miden constructos diferenciables del pensamiento cognitivo básico.

Los coeficientes de confiabilidad interna, evaluados a través del alfa de Cronbach y el omega de McDonald, alcanzaron valores aceptables en la mayoría de las subescalas. En especial, Habilidad Matemática, Pensamiento Lógico y Apreciación Espacial destacaron por su consistencia interna. En contraste, la subescala de Elocuencia Verbal presentó limitaciones

severas: cargas factoriales inconsistentes, baja varianza explicada y coeficientes de confiabilidad nulos, lo que indica que no conforma una dimensión unitaria. Esta situación sugiere que los ítems de dicha subescala no están midiendo un constructo homogéneo, por lo que se recomienda su rediseño integral.

Desde una perspectiva institucional, garantizar subescalas confiables es fundamental, ya que los resultados derivados de ellas tienen impacto directo en decisiones académicas como el cambio de carrera y la asignación de becas.

6.3. Validez Predictiva y Relación con el Rendimiento Académico

Se identificó una correlación significativa entre las puntuaciones obtenidas en las subescalas y el índice académico posterior, siendo particularmente fuerte la relación con la subescala de Pensamiento Lógico. Estos resultados aportan evidencia de validez predictiva y sugieren que la prueba, además de medir habilidades cognitivas específicas, puede anticipar de manera razonable el desempeño académico futuro.

Este hallazgo tiene implicaciones prácticas importantes: la aplicación temprana de esta prueba puede contribuir a orientar adecuadamente a los estudiantes hacia carreras que se ajusten mejor a sus habilidades, lo que podría reducir los índices de rezago y abandono. Además, se observó una mejora significativa en el rendimiento académico de los estudiantes luego de la aplicación de la prueba, lo cual indica que el proceso de orientación vocacional, al incluir este instrumento, podría estar generando un impacto positivo en la toma de decisiones académicas.

6.4. Modelo General de Aptitud Cognitiva

Aunque se propuso un modelo factorial de segundo orden que representara un constructo general de "aptitud cognitiva", los índices de ajuste fueron limitados y no alcanzaron los niveles recomendados. Si bien este modelo es teóricamente coherente, los resultados empíricos sugieren que las subescalas, si bien relacionadas, miden componentes relativamente independientes. Esto refuerza la necesidad de tratar cada subescala como una dimensión separada en los análisis e interpretaciones institucionales.

Limitaciones del Estudio

Una de las principales limitaciones fue la imposibilidad de aplicar técnicas complementarias de análisis (como modelos de la Teoría de Respuesta al Ítem) por la naturaleza de los datos disponibles. Además, el análisis se realizó sobre una muestra institucional ya existente, lo que limitó el control sobre variables externas como el nivel socioeconómico, los estilos de aprendizaje o la motivación de los estudiantes. Otra limitación es la falta de un análisis específico por subgrupos (modalidad, carrera, sede), que podría revelar diferencias importantes en la validez del instrumento.

Conclusiones

El análisis psicométrico de la Prueba de Aptitudes de Pensamiento Básico de la UPNFM permitió confirmar una estructura factorial compuesta por cuatro subescalas con propiedades psicométricas aceptables en cuanto a confiabilidad y validez. La consistencia interna, evaluada mediante los coeficientes Alpha de Cronbach y Omega de McDonald, presentó valores dentro de los rangos recomendados, lo cual respalda la fiabilidad del instrumento. Asimismo, el análisis factorial confirmatorio aportó evidencia de validez estructural, fortaleciendo la congruencia entre los ítems y los constructos teóricos que se pretendían medir.

En lo que respecta a la validez predictiva, los resultados revelaron una relación estadísticamente significativa entre los puntajes obtenidos en la prueba y el rendimiento académico de los estudiantes que realizaron cambio de carrera. Esto indica que la prueba no solo permite identificar habilidades cognitivas específicas, sino que también es capaz de anticipar, con un grado aceptable de precisión, el desempeño académico futuro de los evaluados. Por tanto, se ratifica su utilidad como herramienta de apoyo para los procesos de orientación vocacional, asignación de becas y cambios de carrera, contribuyendo a una toma de decisiones basada en evidencias dentro del contexto institucional.

No obstante, los hallazgos también señalan la necesidad de una revisión detallada de la subescala de Elocuencia Verbal, la cual mostró indicadores psicométricos por debajo de los niveles esperados, así como de ciertos ítems con bajo poder discriminativo. Esta revisión permitirá mejorar la precisión del instrumento en esa dimensión específica, asegurando una evaluación más integral de las habilidades de los estudiantes.

En síntesis, el estudio confirma que la Prueba de Aptitudes de Pensamiento Básico cumple con los criterios fundamentales de confiabilidad, validez estructural e incluso validez predictiva, lo cual la posiciona como una herramienta psicométrica pertinente para apoyar la mejora del rendimiento académico y reducir el riesgo de rezago y abandono en la educación superior.

Recomendaciones

Para investigaciones futuras se recomienda:

Implementar estudios longitudinales que evalúen la estabilidad de los puntajes y su relación con el rendimiento académico en distintos momentos.

Aplicar modelos de Teoría de Respuesta al Ítem (TRI) que permitan obtener información más precisa sobre el comportamiento de los ítems.

Analizar diferencias por subgrupos poblacionales (sede, modalidad, género) para identificar posibles sesgos o limitaciones en la generalización de los resultados.

Diseñar nuevas versiones del instrumento, especialmente para la subescala de Elocuencia Verbal, incluyendo pruebas piloto y revisión de contenido por parte de expertos.

Finalmente, se sugiere que la UPNFM continúe fortaleciendo sus procesos de orientación vocacional con base en evidencia empírica, integrando herramientas psicométricas confiables y válidas que permitan una toma de decisiones más justa, efectiva y centrada en el desarrollo del potencial estudiantil.

REFERENCIAS BIBLIOGRÁFICAS

- Acosta, H. (2011). *Comparación de la teoría de respuesta al ítem y de la teoría clásica de los tests en función del análisis de resultados de una escala de depresión*. Tesis, Universidad Católica Andrés Bello, Escuela de Psicología.
- Aguirre, P. C., & Pérez, F. P. (1995). Desarrollo de la inteligencia general, medida por el PMA de Thurstone, en un curso de química, comparando dos metodologías de enseñanza aprendizaje. *Ensayos: Revista de la Facultad de Educación de Albacete*(10), 129-138. <https://dialnet.unirioja.es/descarga/articulo/2282537.pdf>
- Argibay, J. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y Procesos Cognitivos*, 15-33.
- Real Academia Española & Asociación de Academias de la Lengua Española. (2023). *Diccionario de la lengua española| Edición del Tricentenario*. Diccionario de la lengua española» - Edición del Tricentenario.: <https://dle.rae.es/>
- Auné, S., Abal, F., & Attorresi, H. (2020). Análisis psicométrico mediante la Teoría de la Respuesta al Ítem: modelización paso a paso de una Escala de Soledad. *Ciencias Psicológicas*.
- Auné, S. (2019). Teoría de la respuesta al ítem: Su utilización en América Latina. Supuestos de unidimensionalidad e independencia local. *XI Congreso Internacional de Investigación y Práctica Profesional en Psicología*. Buenos Aires.
- Ávila, C. M., Juárez, H. L., Arreola, G. A., & Palmares, V. O. (2019). Construcción y validación de un instrumento de valoración del desempeño docente en la ejecución de una secuencia didáctica. *Revista de Investigación en Educación*, 122-142.
- Bakieva, M. J. (2019). Evidencias de validez interna de un instrumento para evaluar la colegialidad docente. *Revista Electrónica de Investigación Educativa*, 21.
- Barranco, E. M. (2019). *Los números en nuestra vida cotidiana*. Trabajo Fin de Grado, Universidad de Jaén, Facultad de Humanidades y Ciencias de la Educación.

- Banco Centroamericano de Integración Económica.. (21 de septiembre de 2015). Licitación Pública Internacional Limitada PREA-HE-LIL-42-2015 Documento Base de Licitación con Recursos del Programa de Conversión de Deuda de Honduras Frente a España. Tegucigalpa, Francisco Morazán, Honduras.
- Banco Interamericano de Desarrollo.. (2020). *La educación superior en tiempos de COVID-19*.
- Braun, E. (1997). *El saber y los sentidos*. Fondo De Cultura económica. <http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/073/htm/elsaber.htm>
- Carrasco, E., Zuniga, C., & Asún, R. (2021). Diseño y Validación Inicial del Test de Intereses Profesionales para la Orientacion Vocacional (TIPOV) en estudiantes secundarios de Chile. *PSYKHE*, 30, 1-13. <https://doi.org/http://dx.doi.org/10.7764/psykhe.2018.22221>
- Cassiani-Miranda, C. A., Vargas-Hernández, M. C., Pérez-Aníbal, E., Herazo-Bustos, M. I., & Hernández-Carrillo, M. (2014). Confiabilidad y dimensión del cuestionario de salud del paciente (PHQ-9) para la detección de síntomas de depresión en estudiantes de ciencias de la salud en Cartagena, 2014. *Biomédica*, 37(1), 112-120. <https://doi.org/https://doi.org/10.7705/biomedica.v37i0.3221>
- Castro, L. (2022). Validación de un Test de Orientación Vocacional para la Universidad Pedagógica Nacional Francisco Morazán de Honduras. *Paradigma*(47), 137-162.
- Constante-Amores, A., Martínez, E. F., Navarro Asencio, E., & Fernández-Mellizo, M. (2021). Factores asociados al abandono universitario. *Educación XXI*, 24(1), 17-44. <https://doi.org/10.5944/educXX1.26889>
- Crespo, C. R., Santamaría, C. D., & Cubillas, Q. F. (2019). Apuntes para una orientación profesional vocacional pedagógica con una perspectiva audiovisual. *Pedagogía Universitaria*, 24(2), 85-103.
- Cueva, Q. P., & Quizocala, S. L. (2017). *Relación entre las habilidades mentales primarias y el rendimiento académico, en estudiantes de la carrera profesional de mecánica*

- automotriz - I semestre, del instituto superior tecnológico privado "Flavisur".*
Trabajo de investigación, Universidad Nacional de San Agustín de Arequipa, Arequipa. <http://repositorio.unsa.edu.pe/handle/UNSA/5353>
- De Leon Mendoza, T., & Rodriguez Martinez, R. (2007). El efecto de la orientación vocacional en la elección de carrera. *REMO*, *V*(13), 10-16.
- del Valle, R., & Cumsille, B. (2019). Encuesta Nacional de Compromiso Estudiantil (ENCE) como aporte a la gestión institucional para la prevención del abandono y el rezago universitario. *Congresos CLABES*, 1321-1331. <https://revistas.utp.ac.pa/index.php/clabes/article/view/2780>
- Departamento de I+D+i. (2002). *PMA-R: Aptitudes Mentales Primarias - Revisado* (11.^a ed.). TEA Ediciones.
- Dirección de Educación Superior. (2020). *Estadísticas – Portal UNAH*. <https://des.unah.edu.hn/servicio-al-usuario/estadisticas>
- Dirección de Educación Superior. (s.f.). Instituciones públicas del sistema de educación superior en Honduras. <https://des.unah.edu.hn/sistema-de-educacion-superior/instituciones/publicas/>
- Erazo Guerra, X. F., & Rosero Morales, E. d. (2021). Orientación vocacional y su influencia en la deserción universitaria. *Horizontes Revista de Investigación en Ciencias de la Educación*, *5*(18), 591-606. <https://doi.org/10.33996/revistahorizontes.v5i18.198>
- Espindola, J. E., Hernandez, M., Asucena, Navarrete, H. J., Rojas, O. C., Santana, C. S., & Monroy, H. E. (2018). Habilidades Mentales primarias en universitarios: Fortalezas y Debilidades. *Revista electronica de Psicologia Iztacala*, 1342-1370.
- Fernández, M. (19 de Junio de 2020). *1 de cada 5 estudiantes cambia de carrera tras el primer año: ¿los ciclos comunes son la solución?* infobae.
- Frias-Navarro, D., & Pascual Soler, M. (Junio de 2012). Prácticas del análisis factorial exploratorio (AFE) en la investigación sobre conducta del consumidor y marketing. *Suma Psicológica*, *19*(1), 47-58.

- Gamma Solutions. (2017). *Propuesta desarrollo sistema de valoración psicométrico para Gamma Solutions*.
- Godoy, C. E. (2022). *Uso de la gamificación en el desarrollo del pensamiento lógico matemático en estudiantes de educación superior en una universidad privada de Lima*. Tesis doctoral, Universidad César Vallejo, ESCUELA DE POSGRADO. <https://hdl.handle.net/20.500.12692/46306>
- Gonzáles Llana, F. M. (2007). *Instrumentos de Evaluación Psicológica*. Editorial Ciencias Medicas.
- González, Z. J. (2006). *Discernimiento Evolución del Pensamiento Crítico en la Educación Superior*.
- Gunawan, J. C. (2021). Establishing appropriate sample size for developing and validating a questionnaire in nursing research. *Belitung Nursing Journal*, 7(5), 356-360. <https://doi.org/https://doi.org/10.33546/bnj.1927>
- Universidad Nacional Autónoma de Honduras. (2022). *Admisiones*. Universidad Nacional Autónoma de Honduras: <https://www.unah.edu.hn/estudiantes/admisiones/>
- Howell, F. M., Fres, W., & R. Sollie, C. (1977). Ginzberg's Theory of Occupational Choice: A Reanalysis of Increasing Realism. *Journal of Vocational Behavior*, 11, 332-346.
- Instituto Nacional de Estadística. (2020). *Indicadores, cifras de país 2020*.
- Izaguirre Bordelois, M., Veliz Zevallos, I., & López Arística, L. (2019). Repitencia y deserción en estudiantes de Medicina de la Universidad Técnica de Manabí, Ecuador. *EDUMECENTRO*, 11(4), 20-36. https://doi.org/http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2077-28742019000400020&lng=es&tlng=pt
- Kohler Herrera, J. L. (2013). Rendimiento academico,Habilidades intelectuales y estrategias de aprendizaje en universitarios de Lima. *Liberabit. Revista Peruana de Psicología*, 19(2), 277-288.

- Leyva Pacheco, A. C., Alvidrez Molina, A., & Rojas Borboa, G. M. (2019). Expectativas de orientación vocacional y estrategias educativas en el bachillerato. *La Sociedad Académica*(54), 13-18.
- Martínez Arias, M., Hernández Lloreda, M., & Hernández Lloreda, M. (2014). *Psicometría*. Alianza Editorial.
- Martínez de la Rosa, D. J., Mendoza Velazquez, V., & Percastegui Olvera, J. X. (2012). Habilidades mentales primarias y certeza vocacional en alumnos del primer semestre de la licenciatura de derecho. *Eureka*, 9(2), 185-195.
- Menacho Chiok, C. H., & Alva Trinidad, J. M. (2020). Evaluación de pruebas informatizadas aplicando la teoría clásica de los test y la teoría de respuesta al ítem. *Anales Científicos*, 81(2), 278-292. <https://doi.org/10.21704/ac.v81i2.1638>
- Mendoza Recarte, L. M. (2017). Baremos del test de aptitudes mentales primarias para universitarios hondureños. *Revista Ciencia y Tecnología*(19), 198-227. <https://doi.org/10.5377/rct.v0i19.4281>
- Meneses, J. (2014). *Psicometría*. Editorial UOC. <https://elibro.net/es/ereader/biblioupnfm/57600?page=40>
- Montero-Méndez, S., Marín-Vargas, L., Ibarra-Vargas, E., & Retana-Alvarado, D. (2021). Factores que influyen en la deserción y el rezago de la población estudiantil en la carrera de Educación Primaria de la Universidad de Costa Rica, Sede Rodrigo Facio. *Revista Electrónica Calidad En La Educación Superior*, 12(2), 41-69. <https://doi.org/10.22458/caes.v12i2.3616>
- Muñiz, J. (2018). *Introducción a la Psicometría*. Pirámide.
- Muñiz, J., Elousa, P., & Hambleton, R. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 151-157.
- Murillo-García, O.-L., & Luna-Serrano, E. (2021). El contexto académico de estudiantes universitarios en condición de rezago. *Revista Iberoamericana de Educación Superior* (ries), 12(33), 58-75. <https://doi.org/10.22201/iisue.20072872e.2021.33.858>

- Organización Panamericana de la Salud. (2020). *La OMS caracteriza a COVID-19 como una pandemia—OPS/OMS*. Retrieved Enero de 2022, from Organización Panamericana de la Salud: <https://www.paho.org/es/noticias/11-3-2020-oms-caracteriza-covid-19-como-pandemia>
- Palacio, J. E., Martínez, Y. O., & Tirado, E. (2006). Relación del rendimiento académico con las aptitudes mentales, salud mental, autoestima y relaciones de amistad en. *Revista Psicogente*, 11-31.
- Pereyra Quiñones, J. L. (2020). *Propiedades psicométricas del inventario de intereses vocacionales Talento 4-5, en estudiantes preuniversitarios. Lima, 2019*. <https://hdl.handle.net/20.500.12692/40223>
- Pérez Escoda, N., & Bisquerra Alzina, R. (2009). El saber hacer emocional, destreza ineludible. *Cuadernos de pedagogía*(389).
- Programa de las Naciones Unidas para el Desarrollo. (2022). *Objetivo 4: Educación de calidad*. UNDP: <https://www1.undp.org/content/undp/es/home/sustainable-development-goals/goal-4-quality-education.html>
- QuestionPro. . (Mayo de 2019). *¿Qué es el coeficiente de correlación de Pearson?* QuestionPro. : <https://www.questionpro.com/blog/es/coeficiente-de-correlacion-de-pearson/>
- Ramos Monsivais, C. L., & González, B. A. (2020). Orientación Vocacional, Aprendizaje Socio-Emocional y Sentido de Vida en la Educación Superior. *Revista Dilemas Contemporáneos: Educación, Política y Valores*, 8(Edición Especial).
- Ramos-Estrada, D., García-Cedillo, I., Sotelo-Castillo, M., López-Valenzuela, M., & Murillo-Parra, L. (2020). Validación de un instrumento de estrategias para fortalecer el aprendizaje. *Revista Electrónica Educare*.
- Rascovan, S. (2013). Orientación vocacional, tensiones vigentes. *Revista Mexicana de Orientación Educativa*, 10(25), 47-54.
- Reidl-Martínez., L. M. (2013). Confiabilidad en la medición. *Investigación en educación Médica*, 2(6), 107-111.

- Sánchez-Palacio N, V.-Á. C.-L. (2021). Validación de contenido y adaptación de la escala de sentido de coherencia 29 para la población colombiana. *Fac. Nac. Salud Pública*.
- Santana Sardi, G., & Vigueras Moreno, J. (2019). Hacia un Sistema Virtual de orientación vocacional. *Revista Cubana de Educación Superior*.
- Seclen Gamarra, M. A. (2019). *Sistema informático para mejorar el proceso de diagnóstico de aptitudes mentales primarias basado en la batería del test PMA en alumnos del nivel secundario*. Tesis pregrado, Universidad Católica Santo Toribio de Mogrovejo, Chiclayo. <http://hdl.handle.net/20.500.12423/2413>
- Sierra, C. A., Negrón, E. J., & Argáez, S. J. (2020). Relación entre comprensión verbal y habilidades mentales en universitarios. *PRAXIS INVESTIGATIVA REDIE*, 123-128.
- Sierra, C. A., Negrón, E. J., Padrón Ake, L. P., Tun, G. D., & Argáez, S. J. (2020). Relación entre perfil de ingreso y rendimiento académico en estudiantes de la licenciatura en enfermería. *Praxis Investigativa REDIE*, 92-102.
- Terán-Bustamante, A., Ramírez-Castillo, C., & Martínez-Velasco, A. (2020). Confiabilidad y validez de un instrumento de selección de capital humano. *Revista Mexicana de Economía y Finanzas Nueva Época*.
- Thurstone, L. L., & Thurstone, T. G. (1975). Test de habilidades mentales primarias: Intermedio. Folleto de instrucciones (W. Woyno & R. E. Osorio Amador, Trads. y Ads.). Ediciones Pedagógicas Latino-Americanas.
- Tiempo.hn. (2020, diciembre). *Honduras: Por la pandemia, unos 50 mil alumnos abandonaron la universidad*. <https://tiempo.hn/honduras-50-mil-abandonaron-universidad-2020/>
- Universidad Nacional Autónoma de Honduras. (2018). *Anuario estadístico*. UNAH.
- Universidad Pedagógica Nacional Francisco Morazán. (2005). *Reglamento académico de las carreras de pregrado*. UPNFM.
- Universidad Pedagógica Nacional Francisco Morazán. (2005). *Reglamento del estatuto de la Universidad Pedagógica Nacional Francisco Morazán*. UPNFM.

- Universidad Pedagógica Nacional Francisco Morazán. (2015). *Manual de procedimientos académico administrativos*. Tegucigalpa.
- Universidad Pedagógica Nacional Francisco Morazán. (2018, noviembre 30). *Con uso de energía fotovoltaica UPNFM impulsará becas académicas*
<https://www.upnfm.edu.hn/index.php/ultimas-noticias/282-con-uso-de-energia-fotovoltaica-upnfm-impulsara-becas-academicas>
- Universidad Pedagógica Nacional Francisco Morazán. (2022 enero). *Admisiones*. Admisiones UPNFM: <https://web.upnfm.edu.hn/admision/index.php>
- Universidad Pedagógica Nacional Francisco Morazán. (2022 enero). *Carreras de pregrado que ofrece la UPNFM*. UPNFM: <https://upnfm.edu.hn/index.php/presencial>
- Universidad Pedagógica Nacional Francisco Morazán. (2022 enero). *Centros Regionales UPNFM*.
- Universidad Pedagógica Nacional Francisco Morazán. (s.f.). *Quienes somos*. UPNFM: <https://www.upnfm.edu.hn/index.php/pages/about-us>
- Universidad Pedagógica Nacional Francisco Morazán. (s.f.). *Servicios DISE*. UPNFM: <https://www.upnfm.edu.hn/index.php/servicios-dise>
- Valdés Veloz, H. (2014). *Introducción a la teoría de respuestas al ítem TRI*. Santo Domingo: Instituto Dominicano de Evaluación e Investigación de la Calidad Educativa.
- Ventura-León, J. L. (2016). Breve historia del concepto validez en psicometría. *Revista Peruana de Historia de la Psicología*, 89-92.
- Ventura-Leon, J., & Caycho-Rodriguez, T. (2017). El corficiente Omega: un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 15(1), 625-627.
- Vidal, L. M., & Fernández, O. B. (2009). Orientación vocacional. *Educación Médica Superior*, 23(2). http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412009000200011&lng=es&tlng=es.

Yalanda Tombé, L. (20 de Enero de 2022). *La Importancia de la Orientación Vocacional en la Formación Escolar*. Repositorio Institucional Universidad Santo Tomas:
<http://hdl.handle.net/11634/42623>