

# Validez y Confiabilidad de los Ítems del Examen Ordinario 2016-I a la Universidad Nacional Pedro Ruíz Gallo.

Eduar Vásquez Sánchez<sup>1</sup>, Gloria María Ortiz Basauri<sup>2</sup>, Robert Edgar Puican Gutiérrez<sup>3</sup>

## Resumen

*El objetivo de la investigación fue determinar la validez y confiabilidad de los ítems de examen ordinario 2016-I de la Universidad Nacional Pedro Ruíz Gallo. Para lograr este propósito se realizó el análisis correspondiente del grado de dificultad y discriminación de la prueba así como la validez y confiabilidad de los 120 ítems, los cuales han sido elaborados tomando en cuenta los tópicos propuestos en el prospecto de admisión. Los ítems tienen un nivel medio de dificultad del 23.47 % haciendo que la prueba se considere como muy difícil; 86 ítems tienen esta característica. Los ingresantes a la universidad obtuvieron un promedio de  $5.75 \pm 2.28$  en la escala vigesimal, estando el 90 % con nota de 0 a 10. El índice promedio de discriminación de la prueba es de  $18.41 \pm 14.17$ , condición que determina la revisión de 31 preguntas y probable eliminación de 56. Considerando la Taxonomía revisada de Bloom, el 55.83 %, 15 %, 23.33 %, 5.83 % de los ítems se ubican en el nivel de memoria, comprensión, aplicación y análisis respectivamente. La confiabilidad de la prueba es de 0.7 valor que indica que el instrumento no es adecuado para la selección de estudiantes a la Universidad.*

**Palabras clave:** Validez, confiabilidad, ítems.

## Abstract

*The objective of the investigation was to determine the validity and reliability of the ordinary exam 2016-I items to the Universidad Nacional Pedro Ruíz Gallo. To achieve this purpose was carried out the corresponding analysis of the validity and reliability of the 120 items, these have been elaborated taking as basic theory the different disciplines of the topics proposed in the prospectus of admission, each one had validity of content. The items have an average level of difficulty of 23.47 % making the test considered very difficult, 86 items have this characteristic. The students entering the university obtained an average of  $5.75 \pm 2.28$  in the vigesimal scale, being 90 % with a score of 0 to 10. The average discrimination rate of the test is of  $18.41 \pm 14.17$ , condition that determines the revision of 31 Questions and probable elimination of 56. Considering Bloom revised Taxonomy, the 55.83 %, 15 %, 23.33 %, 5.83 % of the ítems are at the level of memory, understanding, application and analysis respectively. The alpha coefficient is of 0.7 values which indicate that the instrument is not reliable in the selection of students to the University.*

**Palabras clave:** Validity, reliability, items.

---

1. Profesor Principal D.E. Departamento Académico de Estadística. UNPRG.

2. Profesora Principal D.E. Departamento Académico de Matemáticas. UNPRG.

3. Profesor Principal D.E. Departamento Académico de Ingeniería de Sistemas. UNPRG.

## Introducción

La selección de estudiantes para la universidad no es cuestión de improvisaciones, se hace fundamentalmente desde el punto de vista de sus competencias cognitivas reflejadas en el logro de calificativos que obtengan en los exámenes de admisión, saber cuál es el grado de dificultad de cada ítem y del examen en general, permitirá diseñar las preguntas más adecuadas, toda vez que se trata de selección de estudiantes para la obtención de un grado académico en las diferentes carreras profesionales.

En el trabajo realizado por docentes de la Universidad Mayor de San Marcos titulado “Estudio Psicométrico sobre el examen de admisión 2008-I” concluyen que la mencionada prueba tiene una confiabilidad aceptable, los ítems tienen validez de contenido, los grados de dificultad de los ítems se concentran en los niveles intermedios y difíciles con reducido grupo de preguntas fáciles. (Díaz & Mayaute, 2008)

### ■ Validez de contenido

La validez de contenido se refiere al conocimiento científico de la variable o relación de variables presente en el instrumento, cuyos ítems elaborados para el recojo de datos tomando en cuenta el contenido de la variable que se quiere medir. También se le denomina validez lógica o racional.

Para un examen de admisión a una universidad, se elabora un sistema de preguntas, ítems o reactivos con los que se tendrá conocimiento de la variable o relación de variables y sus dimensiones para luego someterlos a un proceso de validación por tres o cinco expertos quienes certificarán que efectivamente las preguntas seleccionadas y el instrumento en general es adecuado y puede aplicarse en el recojo de datos para la investigación planteada. (Escobar-Pérez & Cuervo-Martínez, 2008)

Para establecer un posible universo de reactivos se requiere tener una adecuada conceptualización y operacionalización del constructo, es decir, el investigador debe especificar previamente las dimensiones a medir y sus indicadores, a partir de los cuales se realizarán los ítems. Los ítems deben capturar las dimensiones que la prueba pretende medir, por ejemplo, en la prueba de procesos lectores (Cuetos, Rodríguez, & Ruano, 2001) este constructo está evaluado en las dimensiones de procesos sintácticos, semánticos y pragmáticos.

### ■ Validez de Constructo

Constructo es un objeto conceptual, es la conceptualización teórica de algo, siendo las proposiciones, los conceptos, los contextos y las teorías sus formas fundamentales. Para que un instrumento tenga validez de constructo es necesaria una conceptualización clara con fundamentación teórica del rasgo característico que se va a medir.

Interpretando a Paul Meehl y Lee Cronbach, se propone tres pasos para evaluar la validez de constructo:

- Determinar la relación teórica entre los conceptos que contiene el instrumento.
- Diseñar métodos y técnicas para medir los constructos propuestos.
- Probar empíricamente las relaciones planteadas en el instrumento.

La validación de un constructo se vincula con la teoría que dio origen su construcción. Saber si un instrumento tiene validez de constructo implica solicitar a expertos que expliciten los conceptos en los que se basan los ítems. (Corral de Franco, 2009)

- **Validez Predictiva o de criterio**

La predicción, normalmente se la asocia con visión o anticipación de futuro. En este sentido, cuando se estudia la validez predictiva de un instrumento lo que interesa es determinar hasta dónde se puede anticipar el desempeño futuro de una persona en una actividad determinada, a partir de su ejecución actual en dicho instrumento; por ejemplo, se podría estudiar, hasta dónde la Prueba de Aptitud del Consejo Nacional de Universidades predice el éxito académico de los estudiantes en los primeros semestres universitarios, o anticipar el desempeño futuro de un vendedor a partir de su ejecución en un test de inteligencia social. En consecuencia, la validez predictiva, también llamada validez de criterio externo o validez empírica, se estudia comparando los puntajes de un instrumento (variable independiente) con una o más variables externas (variables dependientes) denominadas variables criterio. Se asume que tales criterios, indicadores del desempeño futuro, están teóricamente y lógicamente relacionados con el rasgo representado en el instrumento bajo estudio. Esta comparación entre los puntajes de la variable en estudio y los del variable criterio se expresa a través de un coeficiente de correlación, el cual se interpreta como un índice de validez. Entre más alta sea la correlación entre una medida o medidas de aptitud académica y el promedio de notas, tomado como variable criterio, mejor será la validez predictiva de la prueba de aptitud académica. (Ruiz, 1992)

- **Confiabilidad**

La consistencia interna de un cuestionario de preguntas es la característica fundamental de la confiabilidad de un instrumento para tomar datos en la investigación, está dada por el grado de homogeneidad total y covariación que tienen las preguntas entre sí, salvo que se tenga interés por conocer la consistencia interna de dos o más partes del instrumento tales como la primera mitad, los reactivos pares o impares etc., se determinaría por partes el Alpha de Cronbach. Ésta está indicada para instrumentos con alternativas múltiples de respuesta, la variación del coeficiente es entre 0 y 1. Si el valor es mayor que 0.8, el instrumento se considera confiable y adecuado para investigar. El cálculo puede realizarse mediante dos fórmulas, considerando la varianza de los ítems y la matriz de correlación de los ítems. (Frías-Navarro, 2013)

- **Taxonomía de Bloom**

Benjamín Bloom identifica tres dominios de aprendizaje, cognitivo, afectivo y el psicomotor, se considerará solo el dominio cognitivo con sus seis niveles. Memoria: Lo aprendido previamente. Comprende recordar datos específicos, teorías complejas, es traer a la mente la información apropiada. Conocimiento de conceptos básicos, principios, hechos, conocimiento de métodos y procedimientos. Comprensión: Apropiación del significado de elementos o cosas. Pasar material de una forma a otra (palabras a números), explicar, resumir, estimar tendencias predecir consecuencias o efectos. Comprender hechos (realidades) y principios, interpretar material verbal, interpretar cuadros y gráficas, trasladar material verbal a fórmulas matemáticas, justificar métodos y procedimientos.

Aplicación: Capacidad de utilizar el material aprendido a situaciones concretas, nuevas. Incluir la aplicación de reglas, métodos, conceptos, principios, leyes y teorías a situaciones nuevas, a situaciones prácticas, resolver problemas matemáticos, construir cuadros y gráficas, demostrar el uso correcto de un método o procedimiento.

Análisis: Habilidad de separar material en las partes que lo componen, de manera que su estructura organizativa pueda entenderse. Esto incluye la identificación de las partes, el

análisis de la relación entre las partes, y el reconocimiento de los principios de organización implicados. Reconocer suposiciones tácitas, reconocer en el razonamiento errores de Lógica, distinguir entre hechos y deducciones, evaluar la importancia de los hechos, analizar la estructura organizativa de un trabajo (arte, música, escritura).

Síntesis: Habilidad de unir partes diferentes para formar un todo nuevo un conjunto de relaciones abstractas, esquemas. Aprendizaje creativo, nuevos patrones o estructuras. Escribir un ensayo bien organizado, dar un discurso bien estructurado, escribir un cuento corto creativo (o un poema o música), proponer el plan para realizar un experimento, integrar aprendizajes de diferentes áreas en un plan para resolver un problema, formular un nuevo esquema.

Evaluación: Habilidad para juzgar con un propósito determinado. El juicio debe basarse en criterios definidos. Los resultados del aprendizaje en esta área son los más altos de la jerarquía cognitiva involucran juicios de valor reflexivos, basados en criterios claramente definidos. Es este nivel se juzga la consistencia en la Lógica de un material escrito, que tan adecuadamente las conclusiones se soportan con datos, el valor de un trabajo utilizando estándares. (Bloom, 1956)

**Análisis de la dificultad y discriminación de un ítem de examen**

El índice de dificultad permite determinar en qué medida una pregunta de examen es fácil o difícil, es el porcentaje de estudiantes de un grupo determinado que han respondido correctamente a una pregunta de examen, varía de 0 a 100. El cálculo se inicia con:

- a) Ordenar las pruebas corregidas de mayor a menor.
- b) Dividir la totalidad en tres partes.
- c) Se toma la parte de los mejores exámenes (grupo fuerte, 27 %), y la otra parte de los peores exámenes (grupo débil, 27 %) mientras que la parte faltante se elimina.
- d) Se suman las respuestas correctas de ambos grupos (fuerte y débil), luego las multiplicamos por 100 y este resultado se divide entre n, suma total de los dos grupos.

$$ID = \frac{(\text{grupo fuerte} + \text{grupo débil}) * 100}{n}$$

La escala para evaluar los ítems según grado de dificultad está dada por:

Tabla 1: Escala para evaluar los ítems según grado de dificultad

Índice de Dificultad (%)	valoración
Menor o igual a 30	Muy difícil.
Mayor que 30 a 40	Difícil.
Mayor que 40 a 50	Relativamente difícil
Mayor que 50 a 70	Adecuado.
Mayor que 70 a 80	Relativamente fácil
Mayor que 80 a 90	Fácil
Mayor que 90 a 100	Muy fácil

El índice de discriminación de una pregunta permite determinar en qué medida una pregunta es bastante selectiva para distinguir un grupo fuerte (F) de un grupo débil (D) de estudiantes. Para su cálculo se utiliza la siguiente fórmula:

$$IDC = \frac{2*(F-D)*100}{n}$$

Tabla 2: El índice de discriminación se establece según la escala

Discriminación	Acciones por implementar
00 al 15 %	Preguntas para examinar (probablemente eliminar)
15 al 24 %	Preguntas límite para revisar (regular para revisar)
24 al 34 %	Preguntas buenas
34 al 100 %	Preguntas excelentes

Evaluación crítica de una pregunta. La evaluación crítica de una pregunta se hace sobre la base de los índices de dificultad y de discriminación. El índice de dificultad cuanto más elevado es, más fácil es la pregunta, algunos lo llaman índice de facilidad o de éxito, en principio una pregunta que tenga un índice de dificultad comprendido entre 30 y 70 es aceptable, es decir que aumenta la probabilidad de incrementar el índice de discriminación, así su índice promedio será del 50 %. Se ha demostrado que un test con tal índice tiene grandes probabilidades de ser fiable en lo que concierne a su consistencia interna y homogeneidad.

El índice de discriminación cuanto más elevado sea, mejor permite diferenciar un grupo fuerte de un grupo débil, así un test con altos índices de discriminación asegura una clasificación discriminativa de estudiantes según su nivel de actuación. (Ebel, 1965)

## Material Y Métodos

Los documentos guías para llevar a cabo la investigación fueron el prospecto de admisión 2016-I, donde se encuentran los tópicos propuestos sobre las diferentes disciplinas para postular a las carreras universitarias que tiene la Universidad Nacional Pedro Ruíz Gallo; el tema N° 03 del examen ordinario, el mismo que contiene 120 preguntas, 20 de razonamiento verbal, 20 de razonamiento matemático y 80 de conocimientos en las áreas de físico-matemática, humanidades y bio-química respectivamente.

El número de ítems por área académica, asignatura y orden en la prueba se aprecia en la tabla (3). Para su aplicación, las preguntas se aleatorizan, así como sus alternativas de respuesta para configurar cuatro temas, cuyos ítems se han ponderado de acuerdo al grupo de postulación, alcanzando un total de 480 puntos. Los postulantes fueron 1471 y el tiempo de solución de la prueba fue de tres horas.

Se calculó el índice de dificultad, discriminación, así como la confiabilidad de la prueba, también la ubicación de los ítems según a taxonomía de Bloom.

## Resultados y Discusión

Con respecto a la validez de contenido, se constató que las ciento veinte preguntas están dentro de las áreas temáticas que aparecen en el prospecto 2016-I, como se aprecia en la tabla (3). El razonamiento matemático comprende cinco subáreas: aritmética, álgebra, geometría, trigonometría, lógico, se ubicaron siete preguntas de aritmética (21, 22, 24, 25, 26, 27, 34), nueve preguntas de álgebra (29, 30, 32, 33, 35, 36, 37, 38, 40), una de geometría (31), una de estadística (39), dos de razonamiento lógico (23, 28), por lo que infiere que se precisa reformular el contenido de lo que es razonamiento matemático.

Por otro lado, sobre las preguntas de razonamiento verbal casi todo se encuentran en el nivel literal, salvo la 2, 3, 6 y 12 que están en el nivel parafrásico y la 1 que está en el nivel inferencial, por último, no hay ninguna a nivel criterial. (Alfonso & del Pilar Flórez, 2009), así mismo se determinó que las preguntas sobre diptongos, palabras homógrafas, palabras cohipónimas o hiperónimas, hiatos, composición de palabras compuestas, tilde diacrítica, están dirigidas solo a medir ortografía, gramática y no razonamiento verbal.

Tabla 3: Áreas, asignaturas y preguntas que compone la prueba de admisión Del examen ordinario 2016-I de la universidad nacional Pedro Ruíz Gallo.

Área de Examen	Asignatura	$\alpha$ Cronbach	Número Preguntas	Orden Pregunta
Aptitud Académica	Razonamiento Verbal	0.69	20	ene-20
	Razonamiento Matemático	0.77	20	21 - 40
Físico Matemático	Álgebra	0.69	8	41 - 48
	Geometría y Trigonometría	0.43	8	49 - 56
	Física	0.62	8	57 - 64
Humanidades	Lengua y Literatura	0.42	8	65 - 72
	Psicología, Filosofía y Lógica	0.27	8	73 - 80
	Economía y Cívica	0.39	8	81 - 88
	Historia y Geografía	0.41	8	89 - 96
	Anatomía	0.31	8	97 -104
Bioquímica	Biología	0.49	8	105-112
	Química	0.59	8	113-120
Total			120	

Fuente: Prospecto de admisión 2016-I

En la validez de constructo se analiza sintáctica y semánticamente las preguntas, determinándose que las de razonamiento verbal están adecuadamente elaboradas. Para el caso de las de razonamiento matemático se encontró que: En la pregunta 21, los tiempos gramaticales no coinciden (se construye - usaron), en caso de una cuadrícula, las dimensiones son ancho y largo. En la 22, la contextualización fue un gran distractor que pudo obviarse, por cuanto la pregunta para ser resuelta solo se necesitaba de un ordenamiento y una suma. En la 23, la alternativa de respuesta ya está anticipada. En la 24, el enunciado estuvo contextualizado; sin embargo, en los distractores no se tomó en cuenta. En la 25, el nombre propio (Largión Tucumango) fue un distractor. La pregunta 26 estuvo dentro de un contexto, pero fuera de la realidad. Imposible en caída libre, el objeto se estrellaría en 5 segundos, y es la velocidad y no la distancia con que cae

un cuerpo. En la 34, los tiempos no concuerdan (nace - nació). En la 39, la contextualización fue un gran distractor que pudo obviarse, pues para pedir promedio, simplemente se debió presentar la tabla de distribución. Las otras preguntas se consideraron redactadas adecuadamente.

La validez de criterio, según la Taxonomía de Bloom, se puede apreciar en la tabla (4), que el 55.8 % de los ítems se ubican en el primer nivel en donde la operación mental más importante es la memoria, no se cuenta con ítems de síntesis y evaluación, sólo el 5.8 % está en el nivel de análisis.

Tabla 4: Distribución de los ítems del examen de admisión ordinario 2016-I a la Universidad según los niveles de la Taxonomía de Bloom.

Niveles cognitivos	Ítems	Nº	%
Memoria	3,5,7,8,9,10,11,13,14,15,16,17,18,19,20,22,31,32,38,59,60,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,86,87,88,89,90,91,93,94,95,96,97,98,99,101,103,105,106,108,109,110,111,112,113,115,116,117.	67	55.83
Comprensión	2,4,12,21,23,27,29,30,33,34,36,37,39,40,85,92,107,118.	18	15
Aplicación	24,25,26,28,35,41,42,43,44,45,46,47,48,49,52,53,54,55,56,57,8,61,62,63,564,114,119,120.	28	23.33
Análisis	1,6,50,51,100,102,104.	7	5.83
Síntesis		0	0
Evaluación		0	0
<b>Total</b>		<b>120</b>	<b>100</b>

La tabla (5), muestra la distribución de los postulantes según sus puntajes en el examen ordinario 2016-I, como se aprecia el 90 % obtuvieron un puntaje menor o igual a 10 puntos en la escala vigesimal, puntajes que calificarían como deficientes, el puntaje promedio obtenido fue de  $5.75 \pm 2.89$  de cual se deduce que la prueba tuvo un alto nivel de dificultad.

Tabla 5: Distribución de postulantes según sus puntajes en el examen ordinario de admisión 2016-I a la Universidad Nacional Pedro Ruiz Gallo.

Puntaje ponderado	Puntaje vigesimal	Número de postulantes	%	Nivel de conocimientos
00 - 240	00 a 10	1324	90	Deficientes
240 - 312	10 a 13	101	6.87	Regulares
312 - 384	13 a 16	45	3.06	Buenos
384 - 432	16 a 18	1	0.07	Muy buenos
432 - 480	18 a 20	0	0	Excelentes
$\bar{y} \pm S$	$5.75 \pm 2.89$	1471	100	

Fuente: Prueba de examen de admisión 2016-I

En la tabla (6), se presenta la distribución de los ítems, según su índice de dificultad, siendo el 48 % de nivel muy difícil y sólo el 15 % de preguntas son adecuadas. En esta prueba el índice promedio de dificultad es de  $35.75 \% \pm 24.18$ , calculado a partir de la distribución de la tabla (6), catalogándose como prueba difícil.

Tabla 6: Índice de dificultad de los ítems de la prueba de admisión 2016-I

Dificultad %	Preguntas según su dificultad	Ítems	%	Dificultad
00 - 30	3,27,29,30,32,37,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,75,77,87,89,92,95,96,97,99,100,101,102,103,104,105,107,109,110,111,113,114,000,000,000,000	57	48	Muy Difícil
30 - 40	2,20,21,24,26,33,34,35,39,66,67,68,70,72,74,83,98,108,112,117,00	20	17	Difícil
40 - 50	11,19,28,36,38,71,82,84,85,88,91,93.	12	10	Relativamente Difícil
50 - 70	1,7,9,10,13,17,18,22,31,69,76,79,80,81,86,90,94,106	18	15	Adecuado
70 - 80	8,14,16,78.	4	3	Relativamente Fácil
80 - 90	4,5,15,23.	4	3	Fácil
90 - 100	6,12,25,73,118.	5	4	Muy Fácil
Total		120	100	

Fuente: Prueba de admisión 2016-I

La tabla (7), presenta el número de ítems por categorías de discriminación, se aprecia que el 42 % de los ítems son para examinar y probablemente eliminar, el 15 % son para revisar, preguntas buenas y excelentes son solo el 18 y 26 por ciento respectivamente.

Tabla 7: Ítems según Índice de discriminación en el examen ordinario 2016-I

Índice de Discriminación %	Preguntas por categoría del índice de discriminación	Número de Ítems	%	Acciones por implementar
menor de 15	1,2,3,6,10,12,23,25,29,37,42,47,49,50,51,52,55,56,59,60,64,65,67,68,72,73,74,75,77,78,83,84,89,91,92,93,95,96,97,99,100,101,102,103,104,105,111,116,118,119,	50	42	Preguntas para examinar (probablemente eliminar)
15 al 24	4,5,16,36,40,41,45,54,57,58,61,63,70,82,88,109,110,120.	18	15	Preguntas límite para revisar
24 al 34	13,15,19,27,30,33,48,53,62,66,79,87,94,106,107,108,112,113,114,115,117,	21	18	(regular para revisar)
34 al 100	7,8,9,11,14,17,18,20,21,22,24,26,28,31,32,34,35,38,39,43,44,46,69,71,76,80,81,85,86,90,98,	31	26	preguntas buenas
TOTAL		120	100	

Fuente: Prueba de examen de admisión 2016-I

El índice promedio de discriminación determinado a partir de la tabla 5 es de 28.43 % ± 24.16 lo que indica que la prueba tiene un índice de discriminación regular. Para el cálculo de la confiabilidad de la prueba se aplicó la fórmula el alfa de Cronbach que considera la variabilidad de los ítems  $\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum V_i}{V_t} \right)$ . Este cálculo se realizó para cada asignatura, tabla (3).

## Conclusiones

1. Los ítems han sido elaborados tomando como base teórica las diferentes disciplinas de los tópicos propuestos en el prospecto de admisión asegurando de esta manera la validez de contenido.
2. Los estudiantes ingresantes a la universidad obtuvieron un promedio de  $5.75 \pm 2.28$  en la escala vigesimal, estando el 90 % con nota de 0 a 10.
3. Los ítems tienen un nivel medio de dificultad del  $35.75 \% \pm 24.18$  haciendo que la prueba se considere como difícil.
4. El índice promedio de discriminación de la prueba es de  $28.43 \% \pm 14.17$ , condición que determina la revisión de 18 preguntas y probable eliminación de 50 ítems.
5. Considerando la Taxonomía revisada de Bloom, el 55.8 %, 15 %, 23.3 %, 5.8 % de los ítems se ubican en el nivel de conocimiento, comprensión, aplicación y análisis respectivamente.

## Referencias Bibliográficas

- [ 1 ] Alfonso, A. G., & del Pilar Flórez, M. (2009). *Los niveles de comprensión lectora: hacia una enunciación investigativa y reflexiva para mejorar la comprensión lectora en estudiantes universitarios*. Actualidades Pedagógicas(53), 95-107.
- [ 2 ] Bloom, B. S. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive domain*. New York: McKay, 20-24.
- [ 3 ] Corral de Franco, Y. J. (2009). *Validez y confiabilidad de los instrumentos de investigación para la recolección de datos*.
- [ 4 ] Cuetos, F., Rodríguez, B., & Ruano, E. (2001). *PROLECSE*. Procesos Lectores: Madrid: TEA.
- [ 5 ] Díaz, C. R. P., & Mayaute, L. M. E. (2008). *Estudio psicométrico sobre el examen de admisión 2008-I a la Universidad Nacional Mayor de San Marcos*. Revista de investigación en psicología, 11(1), 137-152.
- [ 6 ] Ebel, R. L. (1965). *Measuring educational achievement*: Prentice-hall Englewood Cliffs, NJ.
- [ 7 ] Escobar-Pérez, J., & Cuervo-Martínez, Á. (2008). *Validez de contenido y juicio de expertos: una aproximación a su utilización*. Avances en medición, 6(1), 27-36.
- [ 8 ] Frías-Navarro, D. (2013). *Alfa de Cronbach y consistencia interna de los ítems de un instrumento de medida*. Facultad de Psicología. Universidad de Valencia. Recuperado de <http://www.uv.es/~friasnav/AlfaCronbach.pdf>.
- [ 9 ] Ruiz, C. (1992). *Instrumentos y técnicas de investigación educativa: un enfoque cuantitativo y cualitativo para la recolección y análisis de datos*. Recuperado de: <https://goo.gl/idiZnw>.