

DISEÑO Y CALIBRACIÓN DE UN EXAMEN DE ADMISIÓN PARA NIVEL SUPERIOR

ERNESTO ALONSO CARLOS MARTÍNEZ / RICARDO RUIZ MORENO / CLAUDIA GABRIELA ARREOLA OLIVARRÍA
 Instituto Tecnológico Superior de Cajeme
 Universidad de Sonora

RESUMEN: En este trabajo se realizó un estudio descriptivo con una metodología cuantitativa, a partir del diseño, calibración y aplicación de los reactivos de matemáticas que constituyen parte medular del examen de admisión del

Instituto Tecnológico Superior de Cajeme (ITESCA) para todas las licenciaturas que se ofertan. Asimismo se hizo un análisis de los rasgos o atributos de los alumnos a través de los resultados arrojados por dicho examen.

PALABRAS CLAVE: Aprendizaje, confiabilidad, evaluación y validez.

Introducción

La mayoría de las instituciones de nivel medio y superior del país, tanto de carácter público como privado, han venido aplicando desde tiempo atrás los llamados exámenes de admisión a los alumnos que desean ingresar por vez primera a sus planteles. Estos exámenes han sido utilizados más que nada como filtros para seleccionar a los mejores estudiantes. Con esto las instituciones educativas pretenden asegurarse que las capacidades y competencias con que cuentan sus alumnos de recién ingreso, son apropiadas para la asimilación y generación de conocimientos y habilidades pertinentes al desarrollo humano – científico – tecnológico. [ANUIES], 2007; López, 2008).

Los exámenes de admisión a las universidades no son un fenómeno privativo de México, ya que todas las instituciones de educación superior (IES), incluyendo las de los países desarrollados han establecido procedimientos para seleccionar a los estudiantes con mayores probabilidades de éxito (López, 2008); sin embargo, por obvias razones en países como el nuestro, donde la cobertura es aún tan limitada, una correcta selección de los alumnos es aún más relevante ya que contribuye al ahorro de los recursos existentes. (Carlos, Hernández, Valdés y Castro, 2009) Esta situación ha ido cambiando paulatinamente, sobre todo a partir de la creación en 1994 del Centro Nacional para la

Evaluación de la Educación Superior (CENEVAL) cuyo propósito primordial es elaborar y administrar exámenes para el ingreso a la educación media superior y superior. Algunas IES, como la Universidad Autónoma de Baja California (UABC) y la Universidad Nacional Autónoma de México (UNAM), han desarrollado y validado sus propios exámenes de admisión, adaptándolos a sus necesidades. (Backhoff, Larrazolo y Rosas, 2000). Para el 2002 surge el Instituto Nacional de Evaluación Educativa (INEE) con la finalidad de desarrollar un sistema de pruebas estandarizadas que permitan comparar los resultados de los estudiantes de los diversos niveles del sistema educativo, e impulsar el desarrollo y análisis de pruebas profesionalmente construidas como el Examen de Habilidades y Conocimientos Básicos (EXHCOBA), EXANI I y II, CoSNET, PISA entre otros (Tirado, Backhoff, Larrazolo & Rosas, 1997; Díaz-Barriga, 2006).

Esta evaluación, permite también obtener la información necesaria para mejorar la propuesta formativa que se desarrolla tanto en cada uno de los cursos de plan de estudios, como en la estructura curricular de las licenciaturas. Asimismo coadyuva a que se pueda realizar un seguimiento de los avances que el estudiante va obteniendo desde su ingreso a la institución, hasta la finalización de sus estudios.

Esto lleva a las IES a la necesidad de garantizar a la sociedad por así decirlo, que los exámenes de admisión posean ciertas propiedades o atributos psicométricos que hagan meritoria su utilización en la selección de los mejores estudiantes. Realizar una evaluación correcta del alumno mediante la aplicación de un examen de admisión, no es solo una de tantas actividades académicas de educación superior, sino que se constituye como verdadera exigencia académica a partir de la cual pueden generarse acciones que comprometan y orienten con certeza el esfuerzo institucional hacia la excelencia educativa.

A pesar de los casos anteriores, actualmente todavía existe un gran rezago en nuestro país en materia de evaluación educativa. Esto se muestra por los siguientes hechos (Backhoff, Larrazolo y Rosas, 2000,).

1. La ausencia de organismos nacionales o regionales tanto gubernamentales como civiles que normen sobre los criterios de calidad de las pruebas.
2. La falta de pruebas de calidad de gran escala desarrolladas, estandarizadas y validadas en nuestro país.

3. La escasez de publicaciones de los indicadores psicométricos que poseen el reducido número de pruebas de gran escala desarrolladas en México.

El objetivo de este estudio fue el diseño, calibración y aplicación de un examen de admisión de matemáticas, a partir de lo cual se analizaron las fortalezas y debilidades académicas de los alumnos aspirantes a ingresar al nivel superior.

Específicamente, se hace una acotación al análisis de un instrumento para la evaluación en las ciencias matemáticas, por la importancia que éstas tienen en la detección de alumnos que satisfagan los requerimientos académicos.

Contenido

El instrumento que se utilizó para esta investigación, consta de 33 reactivos de matemáticas, seleccionados conforme a índices aceptables de confiabilidad, dificultad y discriminación. Para ello se contemplaron las siguientes dimensiones: 1.- Trigonometría (ítems: 1,2,3,4,5); 2.- Cálculo integral (ítems 6,7,8, 9); 3.- Geometría analítica (ítems 10,11,12,13); 4.- Cálculo diferencial (ítems: 14,15,16,17); y, 5.- Álgebra (ítems: 18,19,20 y 21).

Para el análisis estadístico y la obtención de tablas se utilizó el paquete computacional SPSS 15.0, para las distintas variables independientes que se analizaron en el estudio, con respecto a una sola variable dependiente.

Las variables independientes analizadas son:

1. Diferencias por género
2. Escolaridad de los padres
3. Estatus familiar
4. Escuela de procedencia
5. Estatus de ciudad de origen
6. Estatus laboral

También se hace un análisis comparativo entre cada una de las dimensiones del examen, para determinar las fortalezas y debilidades de las distintas áreas de la matemática que contempla el instrumento, que son con las que el alumno de bachillerato egresa de manera general. En base a las desviaciones estándar, se establecen cuatro categorías de calificación, para establecer si las calificaciones obtenidas por los alumnos se ajustan a una distribución normal. Las cuatro categorías de calificación que se consideraron son: Deficiente, Insuficiente, Satisfactorio y Sobresaliente. Finalmente se realizó una prueba Post Hoc, para comparar las calificaciones de entre estos cuatro niveles.

El instrumento se aplicó a los alumnos sustentantes de los distintos bachilleratos de la región sur de Sonora, para el ingreso a las distintas carreras de ingeniería del Instituto Tecnológico Superior de Cajeme. Se aplicó al total de alumnos aspirantes, que en este caso fue de 202 alumnos. 74 mujeres y 128 hombres.

Las Características del instrumento son las siguientes:

1. Nivel de medición: Nominal.
2. Reactivos de opción múltiple.
3. Preguntas cerradas.
4. Una sola respuesta.
5. Respuestas precodificadas. a), b), c),d).
6. Distribución proporcional de la respuesta correcta dentro del grupo de alternativas de respuesta.

Los índices de confiabilidad, dificultad y discriminación para los ítems de cada una de las dimensiones arriba mencionadas, resultaron satisfactorios tal como se aprecia en la Tabla 1.

Se obtuvo que en el examen en general, con un valor de $t = 0.666$, no hubo diferencias entre hombres y mujeres en cuanto a la calificación global y al analizar cada una de las dimensiones se tiene lo siguiente: La dimensión de trigonometría, con un valor de $t = 0.456$, en cálculo integral, $t = 0.830$, en geometría analítica $t = 0.840$, en cálculo diferencial $t = 0.815$ y álgebra $t = 0.813$. En todas ellas no existen diferencias significativas

en las calificaciones obtenidas por hombres y mujeres (Ver tabla 2). Con esto podemos decir pues, que el género no es un factor que influya en las calificaciones.

También se planteó como hipótesis de trabajo una posible asociación entre la escolaridad de los padres de los alumnos y las calificaciones que obtuvieron. Para ello se consideró a los padres en tres categorías; los que tienen un nivel básico de estudio, los que tienen un nivel medio, esto es, secundaria y preparatoria y papas de alumnos que tienen un nivel superior de estudios, ya sea profesional o posgrado. Se observó en los resultados que el valor de p asociado a la $ji - cuadrada$, resultó ser igual a 0.113, que es mayor que 0.05, con lo cual podemos afirmar que no hay relación entre la escolaridad de los padres y las calificaciones de los alumnos.

Otra hipótesis de trabajo que se realizó es la idea de ver o comprobar, si la escuela de procedencia de los alumnos está relacionada o no con las calificaciones que estos obtuvieron en el examen. Para esto se consideraron también dos variables, la variable independiente que fue la escuela de procedencia de los alumnos en la cual se consideraron dos niveles que fueron los alumnos de procedencia de escuelas públicas y los alumnos de procedencia de escuelas privadas. La otra variable dependiente, o categoría, nuevamente con dos niveles, que son los alumnos que aprobaron el examen y los alumnos que los reprobaron.

Los resultados arrojaron un valor de p asociado a la $ji - cuadrada$, igual a 0.197, con lo cual se concluye que no existe relación significativa entre estas dos variables. Es decir, que el hecho que los alumnos hayan aprobado o reprobado, no tiene nada que ver con la escuela de procedencia de los alumnos. De esta manera se utilizaron otras variables independientes como son: probar una posible relación que pudiera existir entre el estatus de la ciudad de origen del alumno, es decir, si este es local o foráneo, con el estatus de calificación del examen ya sea que haya salido aprobado o reprobado., la consideración de que si el alumno trabaja o no trabaja tenía que ver con el resultado de sus calificaciones, etc. Resultando que no tienen nada que ver con la calificación del examen de admisión de matemáticas.

Por otra parte, resulta interesante hacer un análisis comparativo entre cada una de las dimensiones que contienen el examen, con el fin de determinar las fortalezas y debilidades en las distintas áreas de la matemática con que el alumno egresa del

bachillerato de manera general. Para esto, se utilizó el modelo estadístico llamado: Anova de medidas repetidas.

Se establece una medida de esfericidad para probar la hipótesis nula de que la matriz de covarianza con error de la matriz hortonormalizada del error, es proporcional a una matriz identidad independiente (Ver tabla 3). El valor de p estimado al estadístico de Mauchly $s W$, con un valor de $p = 0.000$, nos indica que no se cumple con el supuesto de esfericidad, por lo tanto resulta necesario utilizar pruebas de hipótesis robustas que no impliquen el cumplimiento de este supuesto. Así pues, resultó que todos los estadísticos que se utilizan para comprobar posibles diferencias entre los factores del examen, tomaron un valor de $p = 0.000$. Entonces podemos afirmar que si existen diferencias significativas entre las calificaciones que los alumnos obtuvieron en cada una de las dimensiones (Ver tabla 4).

Para determinar qué diferencias existía entre estas dimensiones específicas, se realizó una prueba de comparaciones múltiples de Paiwise, entre estas dimensiones y comparando los valores de p se obtuvo lo siguiente.

De manera general entonces, en base a los resultados podemos decir que los alumnos obtuvieron mejores calificaciones en las áreas de trigonometría y de álgebra. Sin embargo, obtuvieron sus más bajas calificaciones en las áreas de cálculo diferencial y cálculo integral. Esto es, en estas dos áreas es donde los alumnos traen su mayor debilidad y las áreas de trigonometría y álgebra traen su mayor fortaleza.

Con el fin de establecer si las calificaciones obtenidas por los alumnos en el examen se ajustan a una distribución normal se establecieron con base a las desviaciones estándares, 4 categorías de calificación (Ver tabla 5). Con la frecuencia observada realizó un histograma y una prueba de bondad de ajuste de ji-cuadrada, y se encontró a partir de esta información una ji-cuadrada igual a 0.59. Luego se realizó el cálculo de la ji-teórica con un grado de libertad y a un nivel de significancia de t igual a .05 y esta resultó ser un igual a 3.84.

Así pues, en la tabla de bondad de ajuste se puede observar que el valor de la ji-cuadrada de 0.59 es mucho menor al valor de la ji. Teórica de 3.84 con un grado de libertad y un nivel de significancia de 0.05 con lo cual podemos concluir que las

calificaciones que los alumnos obtuvieron se ajustan perfectamente a una distribución normal. Es decir, pueden ser moderados mediante una distribución normal.

Después, se realizó una prueba de análisis de varianza para establecerse si en estos cuatro niveles de calificaciones establecidas, había diferencias significativas o no, y se obtuvo que el valor de p asociado a la F de Fisher, con las diferencias de medias resultó ser igual a 0.000, con lo cual podemos establecer que si existen diferencias significativas entre estos cuatro rangos de calificaciones y donde el gran porcentaje de alumnos obtuvo una calificación satisfactoria.

Conclusiones

Un examen de admisión que no posea validez ni confiabilidad, ocasiona que no se pueda tener un panorama acertado acerca del perfil de ingreso del estudiante. Esto provoca a su vez, que no se tengan proyecciones confiables, con respecto a la eficiencia terminal en un determinado período. Del mismo modo, el no contar con un diagnóstico adecuado, del nivel de conocimientos del alumno de nuevo ingreso, dificultará el diseño y aplicación de estrategias, para disminuir los rezagos académicos de los estudiantes e impulsar sus fortalezas. Por otro lado, una población estudiantil con mejores posibilidades de cursar una carrera universitaria, permitirá que los recursos con que cuenta la institución se utilicen de manera óptima.

Como se pudo observar anteriormente, es importante que se efectúe un análisis continuo de los resultados de la aplicación y de las características del examen de admisión. De este modo, se podrán realizar los ajustes necesarios en cuanto al diseño del mismo y lograr poco a poco, que el instrumento posea la calidad técnica adecuada.

Los resultados obtenidos en el presente trabajo, pueden servir de apoyo para que se efectúen estudios similares, ya sea que estén relacionados con el diseño y la elaboración de instrumentos de evaluación, o que tengan que ver con el análisis de reactivos de un instrumento ya elaborado.

Tablas y figuras

Tabla 1. Índices de dificultad, discriminación y confiabilidad del examen

Número de ítem	Índice de dificultad	Índice de discriminación	Alfa de Cronbach
1	0.57	0.51	.80
2	0.42	0.24	
3	0.60	0.50	
4	0.51	0.50	
5	0.52	0.23	
6	0.72	0.32	.85

Tabla 1. Índices de dificultad, discriminación y confiabilidad del examen (continuación)

Número de ítem	Índice de dificultad	Índice de discriminación	Alfa de Cronbach
7	0.48	0.44	
8	0.41	0.46	
9	0.35	0.23	
10	0.42	0.47	.81
11	0.29	0.36	
12	0.33	0.24	
13	0.38	0.37	
14	0.41	0.30	.84
15	0.45	0.37	
16	.40	.47	
17	.47	.38	
18	.56	.47	.80
19	.39	.47	
20	.60	.50	
21	.42	.48	

Tabla 2. Diferencias por género.

		N	Media	Igualdad	t	P sig.
Cal. global	Femenino	74	47.425	Supuesto de igualdad de varianza	0.433	0.666
	Masculino	128	46.317			
Trigonometría	Femenino	74	62.162	Supuesto de igualdad de varianza	0.747	0.456
	Masculino	128	58.906			
Cal integral	Femenino	74	35.405	Supuesto de igualdad de varianza	0.215	0.830
	Masculino	128	34.531			
Geom. Analítica	Femenino	74	42.342	Supuesto de igualdad de varianza	0.203	0.840
	Masculino	128	41.406			
Cal. diferencial	Femenino	74	32.770	Supuesto de igualdad de varianza	0.234	0.815
	Masculino	128	31.835			
Álgebra	Femenino	74	62.500	Supuesto de igualdad de varianza	-0.237	0.813

Tabla 3. Prueba de Esfericidad

Sujetos en efecto	Mauchlys w	Ji-Cuadrada aproximada	Grados de lib.	P
Conocimientos en matemáticas	0.359	204.280	9	0.000

Tabla 4. Diferencia de medias

Efecto	Error en grados de lib.	P
Pillai's Trace	198.0	0.000
Wilks' Lambda	198.0	0.000
Hotelling's Trace	198.0	0.000
Roys	198.0	0.000

Tabla 5. Prueba de Bondad de ajuste para una distribución normal

Calificación	Rango	Frecuencia observada	Ji calculada	Ji de la tabla
Deficiente	0 – 11.6	2		
Insuficiente	11.6 – 29.2	30		
Satisfactorio	29.2 – 64.2	134		
Sobresaliente	64.2 – 100	36	0.59	3.84

Bibliografía

ANUIES (2007). *Cobertura de Educación Superior en México 1997-1998 a 2006-2007*. México: ANUIES.

Backhoff, E., Larrazolo. N & Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2 (1). Pp. 12-14. Recuperado de <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>

Carlos, E, Hernández, H.; Valdés, A. y Castro A. (2009). Propiedades psicométricas del examen de admisión del COSNET. *Revista de investigación educativa*, 4(8), 23-37.

Díaz-Barriga, A. (2006). Las pruebas masivas, análisis de sus diferencias técnicas. *Revista Mexicana de Investigación Educativa*, 11 (29), 583-615.

López, F. (2008). Tendencias de la educación superior en el mundo y en América Latina y el Caribe. *Avaliacao*, 13 (2), 267-291..

Tirado, F.; Backhoff, E.; Larrazolo. N. & Rosas, M. (1997). Validez predictiva del Examen de Habilidades y Conocimientos Básicos (EXHCOBA), *Revista Mexicana de Investigación Educativa*, 2 (3), 67-84.