

PROPIEDADES PSICOMÉTRICAS DEL EXAMEN DE ADMISIÓN DEL CoSNET¹

Ernesto Alonso Carlos Martínez
Humberto Hernández Vaca Gastelum
Angel Alberto Valdés Cuervo
Antelmo Castro López

Resumen: En este trabajo se realizó un estudio descriptivo con una metodología cuantitativa para determinar la validez predictiva, la confiabilidad y el índice de dificultad y poder de discriminación de los reactivos del examen de admisión del Consejo del Sistema Nacional de Educación Tecnológica (CoSNET) en el Instituto Tecnológico Superior de Cajeme en el Sur de Sonora (ITESCA). Se tomó una muestra no probabilística de 291 estudiantes que presentaron exámenes de admisión en el período 2005-2006 y que habían cursado hasta el 5to. semestre de la carrera. Aunque se encontró que el instrumento en apariencia no tiene validez predictiva, esto debe tomarse con cierta reserva, dado que al considerarse sólo las calificaciones del quinto semestre, es posible que otras variables no consideradas en el estudio (familiar, escolar e individual, entre otras) hayan influido en los resultados. Se evidenció un alto grado de dificultad de los ítems y un débil poder discriminatorio de los ítems. Estos resultados sugieren que el examen CoSNET no responde satisfactoriamente a las necesidades de la institución en estudio.

Palabras clave: validez, confiabilidad, índice de dificultad, examen de admisión.

Abstract. In this work, a descriptive study was made through a quantitative methodology to determine the predictive validity, the reliability, the index of difficulty and the discrimination power of the items from the entrance examination of the Council of the National System of Technological Education (CoSNET) from the Higher Technological Institute from Cajeme at Southern Sonora (ITESCA). A nonprobability sampling was used composed of 291 students who had taken the entrance examination during 2005-2006 and had coursed until the 5th semester of their degree. It was found that apparently the examination instrument didn't have predictive validity, however, this situation must be taken carefully as when considering only the 5th semester grades, it is possible that other variables that weren't considered in this study such as family, schooling, etc., have influenced on the results. It was shown a high index of difficulty but a low discrimination power. These results suggest that the CoSNET's entrance examination does not respond satisfactorily to the needs of the studied institution.

Keywords: validity, reliability, index of difficulty, entrance examination.

¹ Consejo del Sistema Nacional de Educación Tecnológica

Introducción

Desde hace tiempo existen en México presiones por parte de la sociedad civil e instituciones académicas sobre el Gobierno Federal y los Gobiernos Estatales para que generen acciones destinadas a incrementar la oferta educativa y, especialmente, la matrícula en las instituciones públicas de nivel superior.

La situación antes descrita dista mucho de ser solucionada, lo cual se evidencia en el hecho de que la tasa de cobertura de la Educación Superior en México hasta 2007 fue de 24.1%. Este porcentaje se encuentra lejos de los que presentan los países desarrollados donde los índices de cobertura en educación superior son superiores a un 50%, e incluso del de países con niveles de desarrollo parecidos como Argentina (54%) y Chile (48%) (Asociación Nacional de Universidades e Instituciones de Educación Superior [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; sin embargo, por obvias razones, en países como el nuestro, donde la cobertura es aún tan limitada, una correcta selección de los estudiantes resulta aún más relevante ya que contribuye al ahorro de los recursos existentes.

El examen de admisión forma parte de estos procedimientos de selección, cuyo propósito es evaluar, en los estudiantes, las habilidades y conocimientos adquiridos en la educación media superior que se consideran relevantes para cursar los estudios de educación superior; se parte de la idea de que estos conocimientos y habilidades, permiten a los candidatos aprovechar mejor las oportunidades que se les brindan y con ello aumentar las posibilidades de éxito escolar (Chain, Cruz, Martínez & Jácome, 2003; Tirado, Backhoff, Larrazolo & Rosas, 1997; Padilla, 2007).

Hasta la década de los 90, la mayor parte de los exámenes de admisión que se utilizaban en las instituciones de educación superior del país (IES), se desarrollaban, por lo general, por las mismas instituciones con poco rigor metodológico e incluso no se realizaban estudios serios para determinar las propiedades psicométricas de los mismos.

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.

Para 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).

En el contexto anterior, se espera que los exámenes de admisión no sólo contribuyan a predecir el comportamiento académico de los alumnos durante su trayectoria escolar, sino también a detectar en los estudiantes las áreas curriculares de mayor deficiencia que requieren ser subsanadas mediante cursos remediales al momento de ingresar a la universidad. La eficacia y el impacto de los exámenes de admisión en los contenidos de los aprendizajes adquiridos por los estudiantes en estos cursos de nivelación académica, permiten seleccionar a los estudiantes que por sus conocimientos y habilidades previas se encuentran en mejor situación para sobresalir en los estudios universitarios.

Esta evaluación permite también obtener la información necesaria para mejorar la propuesta formativa que se desarrolla en cada uno de los cursos de plan de estudios y 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.

De lo anterior se desprende que realizar una evaluación correcta del alumno mediante la aplicación de un examen de admisión válido y confiable, no es sólo una de tantas actividades académicas de la educación superior, sino que se constituye una condición esencial para que se puedan generar acciones que comprometan y orienten con certeza el esfuerzo institucional hacia la excelencia educativa.

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 hacen meritoria su utilización en la selección de los mejores estudiantes. Desde esta perspectiva, algunas de las propiedades más importantes de estos instrumentos son las siguientes:

1. El grado de dificultad de los ítems no debe ser tan elevado que casi todos los estudiantes no puedan responderlo correctamente, o tan fácil que la mayoría acierte en su respuesta, es decir, deben tener un grado de dificultad mediano.
2. Los ítems deben tener la capacidad de discriminar a los estudiantes de mayor puntaje obtenido en el examen versus los de menor puntaje.
3. Los distractores empleados en las opciones de respuesta debe ser adecuados a sus finalidades.
4. El instrumento debe predecir el comportamiento académico de los estudiantes durante sus estudios universitarios.

Indicadores empíricos empleados en el análisis del instrumento

Una vez que se ha aplicado un examen y se conocen sus resultados, resulta útil y necesario obtener información acerca de la eficiencia que cada reactivo mostró para satisfacer los propósitos para los que fue diseñado. Así pues, desde la perspectiva de la normatividad de la evaluación, una orientación primordial de los reactivos es su capacidad para diferenciar el desempeño de los estudiantes en el examen.

Para ello, es preciso estimar el grado de dificultad y discriminación de los reactivos. La dificultad de un reactivo se relaciona con la proporción de alumnos que respondieron de manera correcta al mismo; mientras que la discriminación se refiere a la capacidad de un ítem para diferenciar al grupo de alumnos que obtuvieron una puntuación alta en el examen, versus el grupo de los que obtuvieron una puntuación global baja.

Índice de dificultad. Representa la cantidad de alumnos examinados que respondieron adecuadamente al ítem. Se simboliza con la letra p y puede expresarse en proporciones o porcentajes, por lo que sus valores pueden fluctuar entre .00 y 1.00, o entre 0% y 100%, respectivamente.

De acuerdo con la psicometría, lo ideal es que cada reactivo posea un mediano índice de dificultad igual a .5 ó 50%; sin embargo, para considerarse como aceptable, éste puede oscilar entre .35 y .65 (Santibáñez, 2001). Este índice se calcula de la siguiente manera:

$$p_i = \frac{A_i}{N_i}$$

donde:

p_i = Índice de dificultad del reactivo i

A_i = Número de aciertos en el reactivo i

N_i = Número de aciertos más número de errores en el reactivo i

Índice de discriminación. Se define como el grado de diferenciación que un reactivo es capaz de captar mediante la forma correcta o incorrecta en que los alumnos responden; de tal manera que sea respondido mayoritariamente correcto por quienes obtuvieron un puntaje global alto en el examen; y viceversa, mayoritariamente incorrecto por los alumnos con puntuaciones globales bajas en el examen. En síntesis, se espera que los reactivos sean capaces de distinguir, con la máxima nitidez posible, entre los alumnos de alto y bajo rendimientos académicos. (Backhoff, Larrazolo & Rosas, 2000).

En este trabajo utilizaremos la siguiente fórmula:

$$D_i = \frac{GA_{aciertos} - GB_{aciertos}}{N_{grupomayor}}$$

donde:

D_i = Índice de discriminación del reactivo i

$GA_{aciertos}$ = Número de aciertos en el reactivo i del 27% de personas con las puntuaciones más altas en el examen.

$GB_{aciertos}$ = Número de aciertos en el reactivo i del 27% de personas con las puntuaciones más bajas en el examen.

$N_{grupomayor}$ = Número de personas en el grupo más numeroso (GA o GB).

Correlación del punto biserial (r_{pbis}). Se utiliza para saber si las personas “adecuadas” son las que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo, y cómo puede contribuir a las predicciones. Henrysson (1971) sugiere que el r_{pbis} es una medida que combina la relación entre el criterio del reactivo y el nivel de dificultad.

Ebel y Frisbie (1986) afirman que el r_{pbis} describe la relación entre las respuestas a un reactivo (0 o 1) y las calificaciones en el examen de todas las personas. La ecuación para obtener este indicador, de acuerdo a Glass y Stanley (1986), es la siguiente:

$$r_{pbis} = \frac{\bar{X}_1 - \bar{X}_0}{S_x} * \sqrt{\frac{n_1 n_0}{n(n-1)}}$$

donde:

$\bar{X}_1$ = Media de las puntuaciones totales de aquéllos que respondieron correctamente el ítem.

$\bar{X}_0$ = Media de las puntuaciones totales de aquéllos que respondieron incorrectamente el ítem.

S_x = Desviación estándar de las puntuaciones totales.

n_1 = Número de casos que respondieron correctamente el ítem.

n_0 = Número de casos que respondieron incorrectamente el ítem.

$n = n_1 + n_0$

Índice de confiabilidad. Los errores en la medición pueden provenir de dos fuentes: la variabilidad en la ejecución de reactivo a reactivo dentro de un examen; y cuando se emplean versiones diferentes de un examen ya sea en el

mismo momento o en momentos diferentes. La primera fuente de error puede ser controlada satisfactoriamente por el coeficiente Alpha de Cronbach (Nunnally & Bernstein, 1999), empero, no resulta eficiente para la segunda fuente de error. El examen del COSNET, motivo del presente estudio, por presentarse en una misma versión y en un mismo momento, se adecua a la primera fuente de error, y por ende, el Alpha de Cronbach resulta un estimador adecuado de su nivel de confiabilidad.

La adivinación de las respuestas por parte de los alumnos, yerros en la revisión manual del examen, redacción confusa de los ítems, cansancio mental de los alumnos, equivocaciones al marcarse una respuesta diferente a la que se deseaba señalar, son algunos de los diferentes errores que pueden presentarse dentro de una prueba. Todos estos factores situacionales influyen en menor o mayor medida en la manera en que un alumno responde cualquier reactivo del examen.

En los errores anteriores —aunque tienden a disminuir la correlación promedio entre los reactivos dentro del examen—, la correlación promedio proporcionada por el Alpha de Cronbach sigue siendo un excelente indicador de la confiabilidad del examen; esto debido a la sensibilidad que muestra a la variabilidad de la medición atribuida a factores situacionales, así como de contenido de los reactivos (Nunnally & Bernstein, 1999). El cálculo de este coeficiente se realizó con la siguiente ecuación:

$$r_{\alpha} = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

donde:

K = Número total de ítems

S_i^2 = Varianza de las puntuaciones del reactivo i

S_t^2 = Varianza de las puntuaciones totales obtenidas por cada alumno

Coeficiente de correlación de Pearson. Permite analizar la relación entre dos variables medidas en un nivel por intervalos o razón. Se simboliza por la letra “ r ”. Es importante hacer notar, que la prueba no considera a una variable independiente y a la otra dependiente, ya que no se está evaluando causalidad. Es decir, la noción de causa-efecto se puede establecer teóricamente, pero la prueba no considera dicha causalidad. Este coeficiente se calcula a partir de las puntuaciones obtenidas en dos variables observadas en los mismos sujetos. Las

puntuaciones obtenidas de las dos variables se relacionan entre sí. El coeficiente puede variar desde -1.00 hasta +1.00.

Este coeficiente es uno de los mejores y más utilizados para analizar el grado de relación lineal existente entre dos variables cuantitativas; un coeficiente alto de correlación no implica necesariamente causalidad, ya que dos variables pueden estar muy relacionadas entre sí, sin que una sea causa de la otra (Pardo & Ruiz, 2000). Su cálculo se realiza con la siguiente ecuación:

$$r_x = \frac{\sum x_i y_i}{n S_x S_y}$$

donde:

r_x = coeficiente de correlación de Pearson

x_i e y_i = puntuaciones diferenciales de cada par

n = número de casos

S_x y S_y = desviaciones típicas de cada variable

Establecimiento del problema

El Instituto Tecnológico Superior de Cajeme (ITESCA), al igual que la mayoría de los tecnológicos de México, ha utilizado el examen desarrollado por el Consejo del Sistema Nacional de Educación Tecnológica (CoSNET), para apoyar su proceso de selección de aspirantes de nuevo ingreso. El resultado obtenido de este examen es uno de los criterios importantes que se utiliza para decidir quiénes serán aceptados en los programas de licenciatura y quiénes no. Ahora bien, ¿qué tan válidos y confiables son los resultados obtenidos por el CoSNET?

Desafortunadamente, no se han puesto, hasta el presente, a disposición del público reportes de investigación que demuestren qué tan válidos y confiables son los resultados que se obtienen de la aplicación de dicho examen. Dado que el examen mismo se utiliza para tomar decisiones acerca de si un estudiante ingresa o no a alguna de las instituciones de educación superior del país, surge entonces la necesidad de realizar un estudio empírico para evaluar la validez y la confiabilidad de este instrumento.

Objetivo General

Analizar las propiedades psicométricas del examen elaborado por el CoSNET, que se aplica durante el proceso de selección de aspirantes de nuevo ingreso del ITESCA.

Objetivos Específicos

1. Evaluar la validez predictiva del examen.
2. Analizar la confiabilidad del examen utilizando el coeficiente de Kuder-Richardson.
3. Analizar los indicadores que describen las características psicométricas de los reactivos del examen de CoSNET (índice de dificultad y poder de discriminación).

Método

El presente trabajo se realizó con un enfoque cuantitativo de tipo no experimental y de corte transversal.

Población

Se consideró como población a todos los estudiantes que presentaron exámenes de admisión en el período 2005-2006 para ingresar a las diversas carreras de ingeniería que ofrece el ITESCA. En total presentaron exámenes 1200 estudiantes.

Muestra

Se tomó una muestra no probabilística de 291 estudiantes que presentaron exámenes de admisión en el período 2005-2006 y que habían finalizado el 5to semestre de la carrera.

Instrumento

El examen de admisión del ITESCA es de carácter multiescalar objetivo, fue elaborado por el Consejo del Sistema Nacional de Educación Tecnológica (CoSNET), órgano descentralizado de la Secretaría de Educación Pública (SEP). El examen se divide básicamente en dos secciones: Prueba de habilidades matemáticas y verbales, y Examen de conocimientos. La primera sección se subdivide en dos grandes áreas temáticas que son las habilidades matemáticas y verbales; mientras que la segunda sección evalúa los conocimientos en las áreas de matemáticas, física y química.

El examen consta de un total de 194 ítems, de los cuales 96 corresponden a la primera sección y 98 a la segunda sección. Todos los ítems tienen cinco opciones de respuesta y sólo una de ellas es correcta.

La habilidad matemática contiene 42 ítems y a la habilidad verbal pertenecen 54 ítems. El tiempo de respuesta para cada una de estas habilidades es de una

hora. La sección de conocimientos consiste en tres áreas: matemáticas, física y química. Se otorga al alumno un tiempo máximo de dos horas y media para responder a esta sección. El instrumento es administrado por un conjunto de docentes previamente capacitados.

Resultados

En este apartado se presentan los resultados correspondientes a la validez predictiva y a la confiabilidad del examen de CoSNET. Del mismo modo, se muestra la información relativa al nivel de dificultad y al poder de discriminación de los reactivos de dicho instrumento.

Validez predictiva

Para establecer la validez predictiva del examen de admisión, se utilizó el coeficiente de correlación de Pearson “r”. Se correlacionó la puntuación obtenida por cada alumno en el examen, respecto al promedio de calificación que dicho alumno tenía en el sexto semestre de su carrera. Aunque lo deseable para el análisis hubiera sido considerar las calificaciones promedio de los primeros semestres, no se tuvo acceso a las mismas.

Sin embargo, cabe señalar que en estudios similares, en los que se tomaron en cuenta los promedios de calificaciones y los resultados del examen de admisión, se encontró que las correlaciones fueron mayores en el segundo semestre (.45) con relación a los del primero (.37) (Tirado *et. al.*, 1997); por otra parte, Chain *et. al.* (2003) encontró, en un estudio realizado en una universidad pública mexicana, que las calificaciones bajas en las áreas de razonamiento verbal y español del examen de admisión, predicen trayectorias escolares pobres de los estudiantes hasta el tercer semestre. Esto puede ser indicador de que la validez predictiva de los puntajes de los exámenes de admisión se extiende más allá de los primeros semestres.

Los valores obtenidos del coeficiente de correlación se presentan en la tabla 1.

Tabla 1
Correlación entre los resultados del CoSNET y los promedios de los estudiantes

Áreas	r	p*
Prueba de habilidades matemática y verbal	.102	.210
Examen de conocimientos	.192	.110
Resultados globales	.168	.165

*significativa a un nivel $\alpha = .05$

Se pudo observar que el grado de correlación (r) entre los resultados obtenidos en el examen de admisión por los aspirantes y sus promedios de calificaciones es

muy bajo, además de que éste no resultó ser estadísticamente significativo para ninguna de las dos áreas, ya que, en ambos casos (.21 y .11, respectivamente), los valores “p” resultaron mayores a .05.

Los resultados anteriores sugirieron, desde la perspectiva del modelo estadístico empleado para el análisis, que el examen de admisión no presenta validez predictiva. Sin embargo, esto debe ser tomado con cierta reserva ya que el resultado pudo verse afectado también por otros factores no considerados en el presente estudio, tales como el dinamismo de los contextos social-escolar-familiar y personal, entre otros.

Confiabilidad

Al efectuarse el cálculo de confiabilidad por medio del coeficiente de Kuder-Richardson, para cada una de las secciones del examen de admisión, se obtuvieron los siguientes coeficientes contenidos en la tabla 2.

Tabla 2
Coefficientes de correlación por área del examen del CoSNET

Dimensión	Coefficiente de confiabilidad
Prueba de habilidades matemática y verbal	0.86
Examen de conocimientos	0.81

Al resultar ambas dimensiones con coeficientes superiores a .8, se evidencia que la confiabilidad del examen es buena.

Análisis de reactivos

La tabla 3 presenta los promedios de las medias y desviaciones estándar de los valores del índice de dificultad (p), índice de discriminación (D) y correlación del punto biserial (r_{pbis}), para cada uno de los reactivos de las cinco áreas temáticas del examen CoSNET. Estos valores fueron obtenidos mediante la elaboración de un programa *ad hoc* en una hoja de cálculo Excel.

Para calcularse el r_{pbis} , se dicotomizaron las respuestas del examen; para ello, aunque cada ítem constaba de cinco opciones de respuesta, se asignó un valor 0 a las respuestas incorrectas y, otro valor 1, a las respuestas correctas. De esta manera, este índice se utilizó para expresar la relación existente entre una variable continua, en este caso la puntuación total del examen; y otra variable dicotómica artificial consistente en las respuestas correctas o incorrectas de los alumnos. Se trató con este indicador de obtener la correlación entre el desempeño de los estudiantes en cada ítem y el desempeño en el examen completo (Santibáñez, 2001).

Tabla 3
*Medias (μ) y desviaciones estándar (σ) del índice de facilidad (p),
 índice de discriminación (D) y correlación del punto biserial (r_{pbis})
 del examen del CoSNET ($n=291$)*

Secciones	Área temática	Índice de facilidad (p)		Índice de discriminación (D)		Correlación del punto biserial (r_{pbis})	
		μ	σ	μ	σ	μ	σ
Habilidades	Matemáticas	.19	.13	.19	.19	.19	.16
	Verbales	.33	.14	.35	.24	.30	.17
	Promedio	.26	.13	.27	.21	.25	.17
Conocimientos	Matemática	.21	.12	.13	.16	.14	.12
	Física	.32	.15	.29	.16	.26	.13
	Química	.30	.16	.25	.14	.24	.10
	Promedio	.27	.14	.23	.15	.21	.12
Promedio Global		.27	.14	.24	.18	.23	.14

Como puede observarse en la Tabla 3, el promedio general de los índices de dificultad de los reactivos correspondientes a la primera sección (habilidades matemática y verbal), fue de 0.26, con una desviación estándar de 0.13. Si se considera que cuanto más se aproxime el índice al valor 0, mayor es el grado de dificultad que presenta al alumno para que éste pueda responderlo de manera correcta; lo anterior significa que el 74% de los alumnos respondieron de forma incorrecta la mayoría de los ítems pertenecientes a esta sección; desde una perspectiva psicométrica puede afirmarse que el examen presentó un índice de dificultad alto, ya que teóricamente la media del índice de dificultad debe estar alrededor del .5 (Santibáñez, 2001).

Asimismo, el índice promedio general, correspondiente al nivel de dificultad de la segunda sección (Examen de conocimientos), es de .27, por lo que también cabe señalar que, debido a que el 73% de los examinados respondió erróneamente estos ítems, el índice de dificultad resultó alto y se encuentra al igual que en el caso anterior, más arriba de lo recomendable de la media teórica de .5 (Santibáñez, 2001).

Los índices promedios de discriminación y de correlación del punto biserial, correspondientes a la primera sección, resultaron .27 y .25 (Ver Tabla 3), respectivamente; con respecto a la segunda sección, los índices fueron .23 y .21. Los bajos índices de discriminación indican que en el examen no se pueden distinguir claramente a los estudiantes con altos y bajos puntajes.

Asimismo en correspondencia, los bajos índices promedios de correlación del punto biserial con signo positivo sugieren que los ítems no discriminan bien, ya que el nivel relativo de éxito alcanzado en los ítems bien respondidos es igualmente alto, tanto para el grupo de estudiantes con puntajes altos en el examen completo, como para el grupo con puntajes bajos en el examen.

Por otra parte, se apreció en la tabla 3, que tanto las habilidades matemáticas correspondientes a la primera sección, como los conocimientos matemáticos evaluados en la segunda sección, presentaron índices promedios de dificultad superiores a las demás áreas; contrario a lo anterior, sus índices promedios de discriminación son inferiores a las otras áreas.

La tabla 4 contiene los porcentajes de ítems que por área temática fueron rechazados ya que presentaron índices de dificultad menores a .35 (Santibáñez, 2001):

Tabla 4
Reactivos aceptados y rechazados por área temática

Secciones	Área temática	Total de reactivos por sección	Reactivos Aceptados		Reactivos Rechazados	
			Número Total	% del Total	Número Total	% del Total
Habilidades	Matemáticas	42	7	16.7%	35	83.3%
	Verbales	54	28	51.9%	26	48.1%
	Total	96	35	36.5%	61	63.5%
Conocimientos	Matemática	34	6	17.6%	28	82.4%
	Física	34	18	52.9%	16	47.1%
	Química	68	12	40%	18	60%
	Total	98	36	36.7%	62	63.3%
Total General		194	71	36.6%	123	63.4%

Como se pudo observar, los 123 reactivos que no cumplieron con las mínimas normas de calidad, se distribuyen en todas las áreas temáticas. Sin embargo, fueron más frecuentes en el área de habilidad matemática (83.3%) y en la de conocimientos matemáticos (82.4%).

El desglose para mostrar los motivos que originaron que se rechazara el reactivo se presenta en la tabla 5.

Tabla 5
Motivos del rechazo de los reactivos

Motivos de rechazo								
Secciones	Área temática	Reactivos por sección	Nivel de dificultad		Poder de discriminación		Nivel de dificultad y poder de discriminación	
			No.	% Total	No.	% del Total	No.	% del Total
Habilidades	Matemáticas	42	8	19 %	1	2.4%	26	61.9%
	Verbales	54	11	20.4%	4	7.4%	11	20.4%
	Total	96	19	19.8%	5	5.2%	37	38.5%
Conocimientos	Matemática	34	2	5.9%	2	5.9%	24	70.6%
	Física	34	6	17.6%	1	2.9%	9	26.5%
	Química	30	5	16.7%	1	3.3%	12	40.0%
	Total	98	13	13.3%	4	4.1%	45	45.9%
	Total General	194	32	16.5%	9	4.6%	82	42.3%

De la información anterior, se observa que hay un mayor porcentaje de reactivos rechazados por causa de ambos indicadores (42.3%).

Discusión y conclusiones

Una forma fundamental de evaluar la calidad de un examen de admisión multiescalar objetivo, es a través del análisis de sus poderes predictivos y discriminatorios, así como del grado de dificultad de sus reactivos. Dos indicadores apropiados para realizar este análisis son los índices de dificultad (p) y discriminación (D y r_{pbis}).

En cuanto al índice de dificultad, los reactivos del examen de CoSNET, aplicado en el ITESCA en julio de 2005, presentaron un promedio de 0.73. Este valor está muy alejado del índice esperado para un examen normativo con cinco opciones de respuesta, el cual es de 0.5 (Santibáñez, 2001). De acuerdo a lo anterior, los reactivos del examen de CoSNET tienen un grado de dificultad muy superior a lo deseable.

Es importante señalar que se encontraron 114 reactivos con niveles inapropiados de dificultad (demasiado fáciles o demasiado difíciles), que representan el 58.8% del total de 194 reactivos de que consta el examen. El porcentaje anterior, se encuentra distante del 10% esperado teóricamente (Santibáñez, 2001).

El alto grado de dificultad de los ítems aunado a su débil poder discriminatorio, sugieren la necesidad de rehacer, o al menos modificar, la mayoría de

los ítems que componen el examen de admisión del COSNET para poder ser utilizado con efectividad en el ITESCA. Esto no significa necesariamente que el examen del COSNET no posea de forma intrínseca estas propiedades; se trata más bien de que quizá no responda a las necesidades del contexto.

Empero, la validez predictiva del examen no fue verificada de manera definitiva en el presente estudio, los bajos índices de correlación que resultaron, al menos parecen consistentes con los valores insatisfactorios de los demás indicadores obtenidos.

La importancia de verificar que el examen del COSNET posea una validez predictiva aceptable, radica sobre todo en la posibilidad de que el ITESCA pueda tener un panorama acertado acerca del perfil de ingreso del estudiante. Esto coadyuva a realizar proyecciones confiables respecto a la eficiencia terminal en un determinado período, con la consiguiente optimización de los recursos humanos y materiales.

Contar con un instrumento evaluativo que posea las propiedades psicométricas requeridas para no sólo tener una adecuada selección de los mejores estudiantes, sino también para tener un diagnóstico certero de las fortalezas y debilidades académicas de los estudiantes de nuevo ingreso, constituye una invaluable aportación para que la institución aplique estrategias, para abatir de manera significativa, el rezago escolar y, en última instancia, los índices de deserción universitaria. 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.

Esto es sumamente importante ya que dada la situación actual de existencia de cupos limitados dentro de las IES es necesario que:

- a. se pueda garantizar que realmente ingresen a las instituciones educativas los estudiantes que mejor puedan aprovechar las experiencias que se brindan en las mismas.
- b. que la información obtenida de los exámenes de admisión brinde una información fidedigna acerca de las fortalezas y debilidades de los estudiantes de nuevo ingreso, de manera tal que:
- c. se puedan generar efectivos programas de apoyo y tutoría con vista a mejorar su desempeño académico y con ello:
- d. disminuir los índices de reprobación y deserción dentro de las IES.

Sin duda alguna, los resultados obtenidos en el presente trabajo resaltan la relevancia y necesidad por parte de las instituciones educativas de todos los niveles, en especial del superior, de impulsar en forma continua, el desarrollo de estudios similares al presente. Las escuelas son instancias cuyo compromiso social primordial es la certificación profesional de los conocimientos y habilidades de quienes a través de su paso en un sistema escolar escalonado,

son declarados como aptos para incorporarse con éxito a la sociedad; en este contexto, la aplicación de instrumentos de evaluación válidos y confiables es herramienta imprescindible para alcanzar tan altos fines.

Referencias

- 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)” [En línea]. *Revista Electrónica de Investigación Educativa*, 2 (1). Disponible en <http://redie.uabc.mx/vol2no1/contenido-backhoff.html> [consulta el 9 de marzo de 2009].
- Chain, R., Cruz, N., Martínez, M. & Jácome, N. (2003). “Examen de selección y probabilidad de éxito escolar en estudios superiores. Estudio en una universidad pública estatal mexicana” [En línea]. *Revista Electrónica de Investigación Educativa*, 5 (1) Disponible en <http://redie.uabc.mx/vol5no1/contenido-chain.html> [Consulta 5 de febrero del 2009].
- Díaz-Barriga, A. (2006). “Las pruebas masivas, análisis de sus diferencias técnicas.” *Revista Mexicana de Investigación Educativa*, 11 (29), 583-615.
- Ebel, R. & Frisbie, D. (1986). *Essentials of Education Measurement*. Englewood Cliffs, NJ: Prentice Hall.
- Glass, G. & Stanley, J. (1986). *Métodos estadísticos aplicados a las ciencias sociales*. México: Prentice Hall Hispanoamericana.
- Henrysson, S. (1971). “Gathering, Analyzing and Using Data on Test Items”, en Thorndike, R. (Comp.). *Educational Measurement*. Washington, DC: American Council on Education, 120-135.
- López, F. (2008). “Tendencias de la educación superior en el mundo y en América Latina y el Caribe”. *Avaliacao*, Campinas, Sofocaba, 13 (2), 267-291.
- Nunnally, J. & Bernstein, I. (1999). *Teoría Psicométrica*. México: Mc Graw Hill.
- Padilla, R. (2007). “El sentido del examen en la educación superior ¿Reproducción o no demostración de lo aprendido?” *Reencuentro*, 48, 27-33.
- Pardo, M. & Ruíz, D. (2002). *SPSS 11. Guía para el análisis de datos*. España: McGraw Hill.
- Santibáñez, J. (2001). *Manual para la evaluación del aprendizaje infantil: Conceptos, análisis e interpretación para el proceso de evaluación*. México: Trillas.
- 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.