

Aplicación de características psicométricas para evaluar la calidad de los parciales de la asignatura Redes de Área Extendida

Carlos Albaca Paraván¹, Romina Paola Nahas¹

¹Facultad de Ciencias Exactas y Tecnología – Universidad Nacional de Tucumán, S. M. de Tucumán, Argentina

{calbaca, mahas}@herrera.unt.edu.ar

Resumen

La evaluación de la adquisición de conocimientos es el mecanismo regulador del proceso enseñanza-aprendizaje, el cual, nos posibilita su retroalimentación, dirección y perfeccionamiento. Por ello, un punto importante en este proceso, es la determinación de la calidad de los instrumentos evaluativos. En este contexto, la cátedra de Redes de Área Extendida (FACET-UNT), en el año 2015, inició la medición cuantitativa de la calidad de sus exámenes parciales por medio de la aplicación de los índices de dificultad y discriminación, relación discriminativa y norma discriminativa a las preguntas que componen los mismos. Este trabajo propone mostrar y analizar los resultados obtenidos de dicha evaluación, y sugerir modificaciones que permitan mejorar la calidad de los instrumentos de evaluación de la cátedra.

Palabras clave: Calidad, Instrumentos de Evaluación, Características Psicométricas.

Abstract

Evaluating knowledge acquisition serves as a regulatory mechanism of the teaching-learning process, allowing its feedback, direction and improvement. Therefore, an important step in this process is the determination of the quality of the grading instruments. In this context, the professors of the course Wide Area Networks (FACET-UNT), in 2015, have started to evaluate the quality of its grading instruments by applying difficulty and discrimination indexes, discriminative ratio and norm, to the questions of its mid-term and second exams. This paper tries to show and analyze the results obtained and suggests modifications to improve the quality of the grading instruments used during the mentioned course.

Keywords: Quality, Evaluation Instruments, Psychometric Characteristics.

1. Introducción

En el transcurso de los últimos años, los sistemas de evaluación en los niveles de grado y posgrado han alcanzado un gran protagonismo, llegando a convertirse en uno de los aspectos centrales de discusiones, reflexiones y debates pedagógicos.

¿El motivo?... Halcones menciona [1] que pocas tareas provocan tantas dudas y contradicciones en los docentes, como las relacionadas con las evaluaciones y las actuaciones o decisiones asociadas a ellas. Por otra parte, todos los profesores están de acuerdo que la evaluación es un proceso necesario y es una actividad docente ineludible.

AFACIMERA [2] reconoce la importancia de una buena evaluación, ya que la misma tiene implicancias para el alumno debido a que a través de la aprobación de los exámenes, puede ir acercándose a su meta profesional. También tiene implicancias para los docentes, ya que a través de sus resultados se puede inferir y analizar la calidad de la enseñanza. Además, la evaluación tiene una importancia social, puesto que con ella, los docentes certifican responsablemente la competencia de los egresados (profesionales que se insertarán en la comunidad para ayudar a resolver algunos de sus incontables problemas).

Por lo tanto, se puede decir que el principal rol de la evaluación consiste en proporcionar evidencias apropiadas que ayuden a estudiantes y docentes a ver en qué medida:

- Se han logrado los objetivos de aprendizaje.
- Resultaron efectivas las estrategias puestas en juego para el logro de los objetivos.

Sacristán [3] propone también considerar a la evaluación como una comprobación de la validez de las estrategias didácticas, y Palmer [4] complementa esta idea declarando que la misma, sirve para buscar datos que nos ayuden a decidir si la estrategia metodológica desarrollada fue adecuada o no, y en qué medida lo fue, de manera de guiar un proceso de enseñanza que desemboque en la obtención de los resultados de aprendizaje previamente propuestos. Por lo que, desde el punto de vista de este enfoque, uno de

los grandes valores de la evaluación es el de ser un instrumento de investigación didáctica [5].

La cátedra de Redes de Área Extendida de la Facultad de Ciencias Exactas y Tecnología de la Universidad Nacional de Tucumán, en el marco de procesos de mejora continua de la calidad de la enseñanza, empezó a desarrollar desde 2015 (tal como lo hizo la cátedra de Protocolos de Comunicación TCP/IP en 2014 [6]) mecanismos de evaluación de calidad de los exámenes parciales para poder abordar estos procesos de mejora.

Al igual que otras evaluaciones de calidad llevadas a cabo por ejemplo por la Universidad Autónoma de Baja California en México en su Examen de Habilidades y Conocimientos Básicos [7], la carrera de especialidad de Medicina General Integral de la Universidad de Ciencias Médicas de Santiago en Cuba [8] y la Universidad Autónoma Metropolitana-Azcapotzalco en México [9], esta evaluación se realiza teniendo en cuenta las características psicométricas: Índices de Dificultad p [10], Índice de Discriminación [11] y Relación Discriminativa [8] de los ítems que componen los tres exámenes parciales de la asignatura, ya que estos son los indicadores básicos que se utilizan para el análisis en un examen objetivo [12] [13] [14].

El propósito de este trabajo es analizar los resultados de los indicadores psicométricos de los parciales de la cátedra como una herramienta más para llegar a mejorar los instrumentos evaluativos que se utilizan en ella.

2. Materiales y métodos

Se realizó un estudio cuantitativo, retrospectivo de tipo observacional. Se empleó como insumo, las respuestas de los tres parciales de los 32 alumnos que cursaron la asignatura durante el período lectivo 2015.

La asignatura pertenece al módulo IX (quinto año) de la Carrera de Ingeniería en Computación y tiene como objetivo capacitar al alumno en conceptos relacionados a las redes de área extendida, tales como protocolos de enrutamiento, mecanismos de control de congestión, modos de transferencia de datos, conceptos de VoIP y aspectos de seguridad.

El método de corrección de los exámenes se realizó de manera manual por los docentes de la cátedra según un patrón de respuestas realizado con anterioridad.

Cada pregunta del examen contaba con un puntaje máximo preestablecido, y a cada respuesta se le asignó un puntaje de acuerdo al grado en que la respuesta se acercaba a la respuesta correcta y completa. Para dar por correcta una respuesta se exigió al menos un puntaje igual o mayor al 60% del puntaje máximo asignado, y para dar por aprobado un examen una nota mínima de 40/100.

Los datos recabados de los exámenes se transcribieron a una planilla de cálculos y se calcularon p , ID, ND y RD.

El Índice de Dificultad (p) expresa la proporción de alumnos que respondieron correctamente un ítem de un examen. Se puede establecer el grado de dificultad de una pregunta en función de cómo la respondieron los alumnos utilizando el criterio de interpretación indicado en la Tabla 1 [2], sugiriéndose elaborar evaluaciones adoptando una

distribución de los ítems con valores de p según se indica en la Tabla 2 [7].

Índice de Dificultad (p)	% que responden correctamente
Muy Fácil (MF)	$85 \leq p \leq 100$
Relativamente Fácil (RF)	$69 \leq p \leq 84$
Dificultad Media (DM)	$32 \leq p \leq 68$
Relativamente Difícil (RD)	$16 \leq p \leq 31$
Muy Difícil (MD)	$0 \leq p \leq 15$

Tabla 1. Criterio de Interpretación de p .

Índice de Dificultad (p)	% de preguntas en el examen
MF	5
RF	20
DM	50
RD	20
MD	5

Tabla 2. Valores Sugeridos de Distribución de p .

El Índice de Discriminación (ID) se utiliza para diferenciar a los alumnos que tienen un mayor dominio de los temas evaluados, de los que no lo tienen. Para ello, se usa el criterio mostrado en la Tabla 3 [2].

Ítem	ID %	Conducta a seguir
MB	$ID \geq 40$	Conservar
B	$30 \geq ID \geq 39$	A mejorar
R	$20 \geq ID \geq 29$	Debe mejorarse
D	$ID \leq 19$	Debe descartarse o revisarse

Tabla 3. Criterio de Interpretación de ID.

Se sugiere lograr evaluaciones en donde se obtenga una distribución de los ítems con valores de ID como se indica en la Tabla 4.

Ítem	% de preguntas en el examen
MB	Mayor o igual a 25
B	Entre 16 y 24
R	Menor a 15
D	Menor a 5

Tabla 4. Valores Sugeridos de Distribución de ID.

Debe tomarse en consideración que el máximo valor del Índice de Discriminación de un ítem tiene estrecha relación con su Índice de Dificultad, ya que un ítem con valor muy alto o muy bajo de p en realidad no discrimina, pues es resuelto por la mayoría de los estudiantes o por la minoría de ellos respectivamente.

A partir de este análisis, Díaz Rojas [8] entre otros, propone utilizar la Norma Discriminativa (ND). Este índice tiene como objetivo determinar el valor óptimo que debe presentar un ítem en su ID de acuerdo al p del mismo, y con ella elaborar otro indicador que muestre si un ítem es aceptable o no según su valor de Índice de Discriminación. Este índice se conoce como Relación Discriminativa (RD) y se lo puede usar según el criterio de la Tabla 5.

Relación Discriminativa (RD)	Conducta a seguir
$RD \geq 1$	Ítem aceptable
$0,6 \leq RD < 1$	Analizar el ítem
$RD \leq 0,6$	Descartar el ítem

Tabla 5. Criterio de Interpretación de RD.

3. Resultados

En el Gráfico 1 se puede apreciar la comparación entre las distribuciones reales de p en cada uno de los parciales con respecto a la óptima ideal sugerida. En el parcial 1 se puede notar una leve tendencia a preguntas entre muy fáciles y de dificultad media, mientras que en el segundo parcial se puede observar una gran concentración de preguntas de dificultad media, y en el tercer parcial se puede advertir una pequeña desviación hacia preguntas muy difíciles.

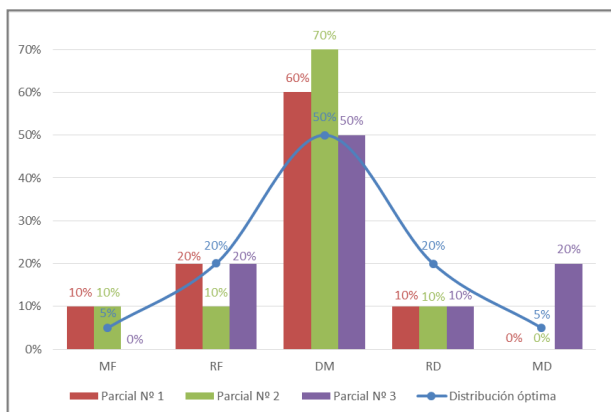
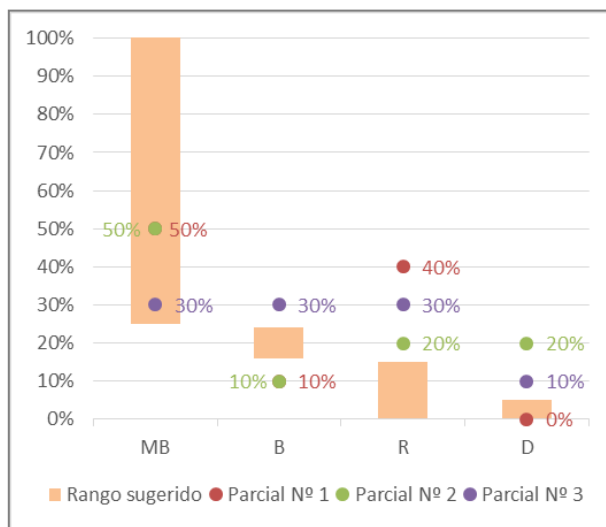


Gráfico 1. Comparación de la distribución de p en los tres exámenes parciales respecto de la distribución ideal.

En el Gráfico 2 se puede observar que el porcentaje de preguntas con Índice de Discriminación igual a MB en

todos los parciales, junto con el porcentaje de preguntas con ID igual a D en el Parcial Nº 1 están acorde a lo sugerido, mientras que para el resto de los casos, esos



valores están alejados de los rangos óptimos.

Gráfico 2. Comparación de la distribución de ID en los tres exámenes parciales respecto a los rangos sugeridos.

Por último, hay que recordar que el ID tiene estrecha relación con p, por lo que en el Gráfico 3 se puede denotar que solo el 10% de las preguntas del Parcial Nº 2 deben ser revisadas, mientras que para el resto, la RD indica que el ID es acorde respecto de p, sin que haya pregunta alguna que deba ser descartada.

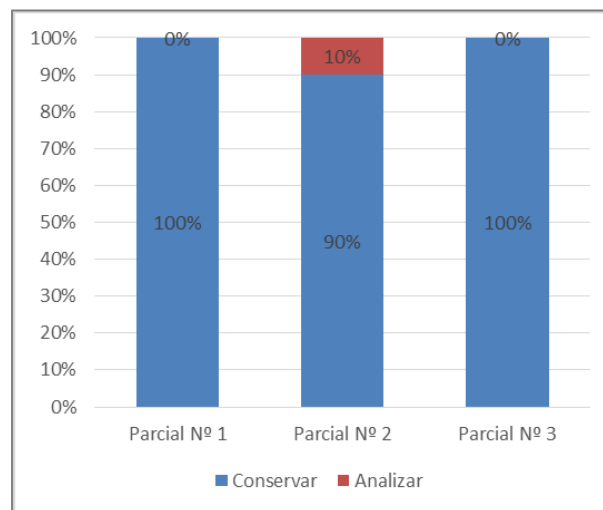


Gráfico 3. Comparación de la distribución de RD en los tres exámenes parciales.

Conclusiones

Este estudio funciona como inicio de un trabajo sistemático en la evaluación de la calidad de los instrumentos con los cuales la cátedra mide el nivel de conocimientos de los estudiantes, y así lograr una

verificación más objetiva del cumplimiento del proceso enseñanza-aprendizaje.

En primera medida, la cátedra tiene que realizar un esfuerzo en mejorar la distribución de p en los exámenes parciales, tratando de acercarse aún más a la distribución óptima.

En segundo lugar, a pesar de que la Relación Discriminativa de la mayoría de las preguntas es adecuada, es conveniente revisar y reformular las preguntas que poseen Índice de Discriminación Regular y Deficiente como forma de mejorar la calidad de los instrumentos evaluativos, intentando siempre llegar a la excelencia.

En tercer lugar, cabe destacar que los resultados de realizar la tarea de análisis continuo de los instrumentos de evaluación serán motivo de futuras revisiones en las que se espera modificar de manera positiva las características psicométricas mencionadas ut supra, sin olvidar que, estos índices se hayan influenciados por el grupo de alumnos que rinde cada examen.

Además, con los resultados de la aplicación de estos índices a los exámenes parciales, la cátedra tiene como objetivo forjar un banco de preguntas validado, el cual sirva de base para generar mejores instrumentos evaluativos. Sobre este punto, cabe destacar que a la fecha, ya se encuentra en proceso de desarrollo un proyecto final de graduación por parte de un alumno de la carrera bajo la tutoría de uno de los autores, el cual tiene como finalidad la gestión de un banco de preguntas validado por medio de la aplicación de las características psicométricas descritas en el presente trabajo.

Por último, este trabajo realizado en la cátedra de Redes de Área Extendida es la continuación de una experiencia llevada a cabo por la asignatura Protocolos de Comunicación TPC/IP [6], siendo el propósito a mediano plazo, la incorporación de este proceso de evaluación de calidad a todos los exámenes de las asignaturas del ciclo superior de la carrera de Ingeniería en Computación.

Referencias

- [1] Halcones, M. Á. G. & González, N. P. (2004). La evaluación del proceso de enseñanza-aprendizaje. Fundamentos básicos. Docencia e Investigación, 4.
- [2] AFACIMERA (2002). Evaluación Educacional. Módulo 1 y 2. ISBN: 987-98237-6-1, Buenos Aires, Argentina.
- [3] Sacristán, J. G. (1992). Teoría de la enseñanza y desarrollo del currículo. 4^o edición; Rei, Buenos Aires, Argentina.
- [4] Palmer, E. J. & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. BMC Medical Education, 7(1), 49-55.
- [5] Córdoba Islas, A. M. (2010). Evaluación Educativa. Congreso Iberoamericano de Educación, Buenos Aires, Argentina.
- [6] Albaca Paraván, C.; Saade, S. D. & Lutz, F. H. (2015). Evaluación de la calidad de los parciales de la asignatura Protocolos de Comunicación TCP/IP mediante la aplicación de diferentes índices. Investigaciones en Facultades de Ingeniería del NOA, 458-461.
- [7] Escudero, E. B.; Reyna, N. L. & Morales, M. R. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básico (EXHCOBA). Revista Electrónica de Investigación Educativa, 2(1).
- [8] Díaz Rojas, P. A. & Leyva Sánchez, E. (2013). Metodología para determinar la calidad de los instrumentos de evaluación. Educación Médica Superior, 27(2), 269-286.
- [9] González Cuevas, Oscar M.; (2003). Evaluación de opción múltiple v.s. evaluación tradicional. Un estudio de caso en ingeniería. *Ingeniería*, mayo-agosto, 17-37.
- [10] Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart and Winston, Orlando.
- [11] Ebel, R. L. (1986). Essentials of educational measurement, Englewood Cliffs, NJ: Prentice Hall.
- [12] Pérez Tapia, J. H.; Acuña Aguilar, N. & Arratía Cuela, E. R. (2008). Nivel de dificultad y poder de discriminación del tercer y quinto examen parcial de la cátedra de cito-histología 2007 de la carrera de medicina de la UMSA. Cuadernos Hospital de Clínicas, 53, 16.
- [13] De los Santos Lázaro, G. (2010). Desarrollo, operación y evaluación de un módulo para capacitar a docentes en servicio para que mejoren sus exámenes de opción múltiple mediante el análisis gráfico de ítems. Tesis de Maestría en Ciencias Educativas. Universidad Autónoma de Baja California, México.
- [14] Mercau, G.; Coccioli, M. V.; D'Urso, M.; Siciliani, M.; del Castillo, M. y Valverde, M. (2013). Impacto de la Calificación Regular en la Validación de Instrumentos de Evaluación Tradicionales.

Dirección de Contacto del Autor/es:

Carlos Albaca Paraván
 Facultad de Ciencias Exactas y Tecnología (UNT)
 Av. Independencia N° 1800
 San Miguel de Tucumán
 Argentina
 e-mail: calbaca@herrera.unt.edu.ar

Romina Paola Nahas
Facultad de Ciencias Exactas y Tecnología (UNT)
Av. Independencia N° 1800
San Miguel de Tucumán
Argentina
e-mail: rnahas@herrera.unt.edu.ar

Carlos Albaca Paraván

Ing. en Computación (UNT), tesista de Maestría en Ingeniería de Software (UNLP) y Maestría en Administración (UNT). Jefe de Trabajos Prácticos de Protocolos de Comunicación TCP/IP y Redes de Área Extendida (FACET – UNT).

Romina Paola Nahas

Ing. en Computación (UNT), Prof. Adjunta de Informática y Diseño Gráfico Informatizado (EUCVyTv – UNT), Jefe de Trabajos Prácticos de Transmisión de Datos y Redes de Área Extendida (FACET – UNT) y consultora de Sistemas en el Ministerio de Educación de Tucumán.