

## ANÁLISIS SOBRE LA VALIDEZ DE RESULTADOS DE UN CUESTIONARIO DE OPINIÓN SOBRE EL DESEMPEÑO DOCENTE

P. A. Medina Mora Escalante<sup>1</sup>  
A. P. Cabrera Meza<sup>2</sup>

### RESUMEN

En este trabajo se presenta, el análisis de validez de criterio de los resultados del *Cuestionario de opinión del desempeño académico del profesor y del alumno* (CODAPA), de la Facultad de Ingeniería de la Universidad Nacional Autónoma de México (UNAM). Este análisis forma parte de la revisión del CODAPA, dirigida a renovar su diseño, aplicación y usos. Para este análisis de validez de criterio se establecieron tres grupos de comparación, de acuerdo con la calificación obtenida por los profesores en el CODAPA, en el semestre 2018-1, y se contrastaron sus resultados en tres medidas del aprendizaje, una basada en la percepción del alumno, otra basada en la percepción del profesor, y la tercera basada en las calificaciones finales asentadas en actas. Los resultados obtenidos mediante estos análisis son positivos, lo que además de la validez, refuerza la utilidad de los resultados del CODAPA, en la medida en que se confirma que sí están asociados al aprendizaje de los alumnos, objetivo principal del desempeño docente.

### ANTECEDENTES

El Cuestionario de Opinión Estudiantil o SET por sus siglas en inglés (*Student Evaluation of Teaching*), es una de las herramientas más utilizadas como instrumento de medida para evaluar el desempeño docente en las instituciones de educación superior (IES). En su mayoría son instrumentos de lápiz y papel, aunque en años recientes se ha migrado a cuestionarios en línea, en los cuales los estudiantes responden con base en una escala numérica a reactivos que reflejan características de la enseñanza y del desempeño de los docentes que, de acuerdo con expertos, pueden ser apreciados o juzgados por ellos (d'Apollina y Abrami, 1997).

Spooren, Brockx y Mortelmans (2013) mencionan que los SET tienen tres propósitos (1) mejorar la calidad de la enseñanza a través de la retroalimentación que se brinda a los profesores, producto de los resultados de la aplicación de los SET; (2) aportar información para la toma de decisiones y (3) proveer evidencia de la responsabilidad de la institución, en lo que se refiere a contar con procesos adecuados para asegurar la calidad de la enseñanza (Kember, Leung y Kwan, 2002, citados en Spooren *et al.*, 2013).

Es un hecho que muchos docentes cuestionan la validez y la confiabilidad de los SET (Ory, 2001 citado en Spooren *et al.*, 2013) por diversas razones, algunas debido a las diferencias de percepción de la eficacia de la enseñanza entre estudiantes y docentes, pero también por el mal diseño y uso de los cuestionarios. Por estas razones, es fundamental realizar la validación de los SET y precisamente, este es el propósito del presente trabajo: presentar un análisis de la validez de los resultados del *Cuestionario de opinión del desempeño académico del profesor y del alumno* (CODAPA) de la Facultad de Ingeniería.

Onwuegbuzie, Daniel y Collins (2009) citados en Spooren *et al.* (2013), desarrollaron un modelo que comprende tres tipos de validación: validez de constructo, de contenido y de

<sup>1</sup> Técnico Académico Titular de la Coordinación de Evaluación Educativa de Facultad de Ingeniería de la Universidad Nacional Autónoma de México. pabme@unam.mx

<sup>2</sup> Psicóloga adscrita al Proyecto PAPIME 102517 de la Facultad de Ingeniería de la Universidad Nacional Autónoma de México. ana.cabrera.meza@gmail.com

criterio. En torno al CODAPA se han realizado estudios de validez de contenido a través de jueces y de validez de constructo, mediante los análisis factorial exploratorio y confirmatorio de sus resultados (FI-UNAM, 2018); se presentará ahora el estudio de la *validez de criterio*, lo que se hará comparando los resultados del CODAPA con otras medidas objetivas.

## METODOLOGÍA

Para proceder a este estudio se identificaron tres grupos de profesores de acuerdo con sus calificaciones en la sección “sobre el profesor” del CODAPA, los cuales se denominaron como grupo bajo (B), grupo medio (M) y grupo alto (A); esto con el fin de comparar sus resultados en tres medidas del aprendizaje de sus alumnos (criterio), una producida por el propio CODAPA, de las preguntas “sobre el alumno”, otra obtenida del *Cuestionario para profesores*, que se aplica simultáneamente al de los alumnos, y la tercera emanada de las calificaciones de los alumnos asentadas en actas.

### Grupos

El estudio se realizó con los resultados del CODAPA obtenidos en el período escolar 2018-1, que se aplicó a 2,741 grupos-asignatura del rango de 6 a 65 alumnos. De estos grupos, para el presente estudio, se seleccionaron 1,917, que son todos aquellos para los que, además de sus resultados en el CODAPA se contó con la información de sus calificaciones asentadas en actas.

De estos 1,917 grupos-asignaturas, se identificaron tres subconjuntos de acuerdo con sus calificaciones obtenidas en el CODAPA. La definición de estos tres grupos se realizó dividiendo el total en tres tercios (o terciles) de similar tamaño, con los necesarios desempates. En la Tabla 1 se presentan las características de estos tres grupos de comparación. El Grupo B (puntuaciones más bajas) corresponde a los grupos en los que el profesor obtuvo un rango de calificación de 52 a 83, el Grupo M (puntuaciones medias) donde obtuvo calificaciones entre 84 y 90 y el Grupo A (puntuaciones más altas) donde obtuvo entre 91 y 100.

**Tabla 1.** Características de los grupos de comparación conformados a partir de las calificaciones promedio “sobre el profesor” del CODAPA: tamaño (N), rango de calificación (mínimo a máximo), media y desviación estándar

| Tercios        | N   | Mínimo | Máximo | Media | Desviación Estándar |
|----------------|-----|--------|--------|-------|---------------------|
| <b>Grupo B</b> | 607 | 52     | 83     | 75.76 | 6.686               |
| <b>Grupo M</b> | 648 | 84     | 90     | 87.20 | 1.978               |
| <b>Grupo A</b> | 662 | 91     | 1000   | 93.76 | 2.295               |

### Criterios

Definidos estos tres grupos, la pregunta consecuente es, ¿estos tres grupos difieren en cuanto a sus resultados de aprendizaje?, ya que como señalan “es necesario medir el éxito del proceso educativo a partir de los resultados de aprendizaje” (Vásquez, Martínez y Benítez, 2015). Como indicios de resultados de aprendizaje, en este estudio se acudió a las siguientes tres fuentes:

- 1) Dentro del CODAPA, en la sección de preguntas sobre el alumno, la número 26 inquiriere: *“Mi aprendizaje sobre los temas de la asignatura ha sido”*, ofreciéndose las siguientes opciones de respuesta: “excelente”, “bueno”, “regular” y “deficiente”
- 2) En el *Cuestionario para profesores* que el profesor responde frente al grupo, al momento en que se les aplica el CODAPA, la pregunta 8 formula, *“En la mayoría de los estudiantes del grupo, el aprendizaje de los temas de la asignatura ha sido”*.
- 3) Las calificaciones finales de cada grupo-asignatura asentadas en las actas por los profesores, de donde se conoce la frecuencia y frecuencia relativa de las calificaciones 10, 9, 8, 7, 6, 5 y NP, por cada grupo.

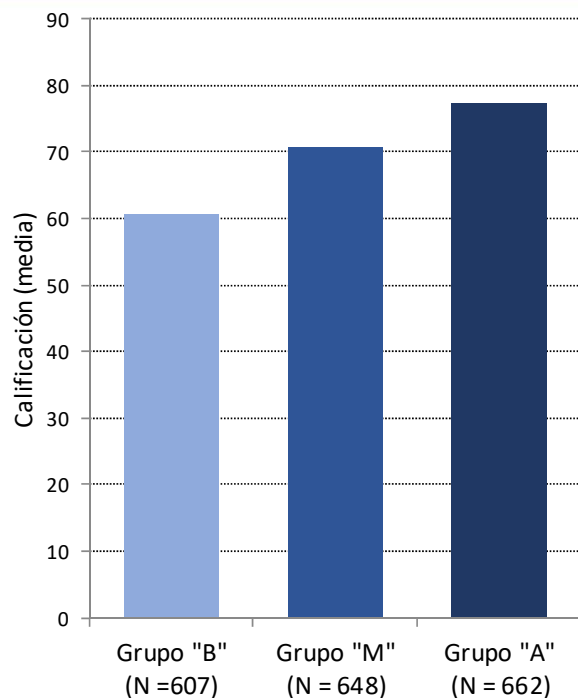
### Procedimiento

Acto seguido, se procedió a organizar los datos, efectuar la comparación de las medias y estimar la significatividad de las diferencias, para probar las siguientes hipótesis:

- Existe una relación directamente proporcional entre los resultados “sobre el profesor” del CODAPA y los resultados de la pregunta, *“Mi aprendizaje de los temas de la asignatura ha sido”* del propio CODAPA
- Existe una relación directamente proporcional entre los resultados “sobre el profesor” del CODAPA y los resultados de la pregunta, *“En la mayoría de los estudiantes el aprendizaje de los temas de la asignatura ha sido”*: del *Cuestionario para profesores*.
- Existe una relación directamente proporcional entre los resultados “sobre el profesor” del CODAPA y los resultados de aprendizaje reflejados en las calificaciones finales de las asignaturas asentadas en actas.

### RESULTADOS

Los resultados en el reactivo, *“mi aprendizaje de los temas de la asignatura ha sido”* del mismo CODAPA varían de manera directamente proporcional con respecto a las características de los grupos de comparación (bajo, medio y alto), como se observa en la Figura 1.



**Figura 1.** Calificación media en el reactivo “Mi aprendizaje sobre los temas de la asignatura ha sido” del CODAPA de los grupos de comparación.

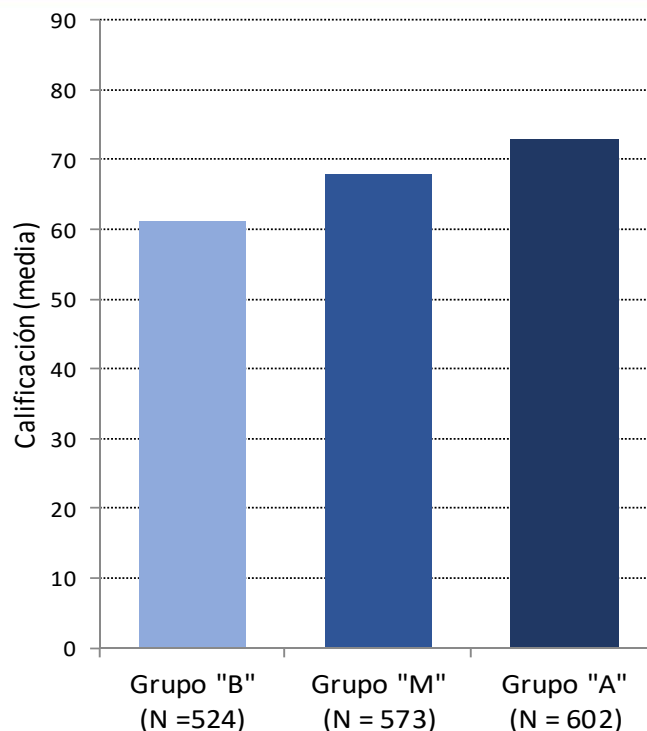
En este caso, las diferencias entre los grupos son significativas, como se comprueba en los resultados de ANOVA, calculados mediante el SPSS y presentados en la Tabla 2, con una F de 656.4 y una  $p < 0.0001$

**Tabla 2.** Resultados del análisis de varianza de las calificaciones de los grupos de comparación en el reactivo “Mi aprendizaje sobre los temas de la asignatura ha sido” del CODAPA

ANOVA Table

|                                                             |                |            | Sum of Squares | df   | Mean Square | F       | Sig. |
|-------------------------------------------------------------|----------------|------------|----------------|------|-------------|---------|------|
| Aprendizaje (p26) *<br>Tercios por<br>calificación profesor | Between Groups | (Combined) | 87362.332      | 2    | 43681.166   | 656.042 | .000 |
|                                                             | Within Groups  |            | 127439.6       | 1914 | 66.583      |         |      |
|                                                             | Total          |            | 214801.9       | 1916 |             |         |      |

De la misma manera, los resultados en el reactivo, «En la mayoría de los estudiantes del grupo, el aprendizaje de los temas de la asignatura ha sido» de la Encuesta de profesores varían de manera directamente proporcional con respecto a los grupos de comparación (alto, medio y bajo), como se observa en la Figura 2. Cabe observar que, en este caso el número de grupos-asignatura (la N) de cada grupo de comparación descende, esto debido a que en algunos grupos los profesores no respondieron la encuesta respectiva.



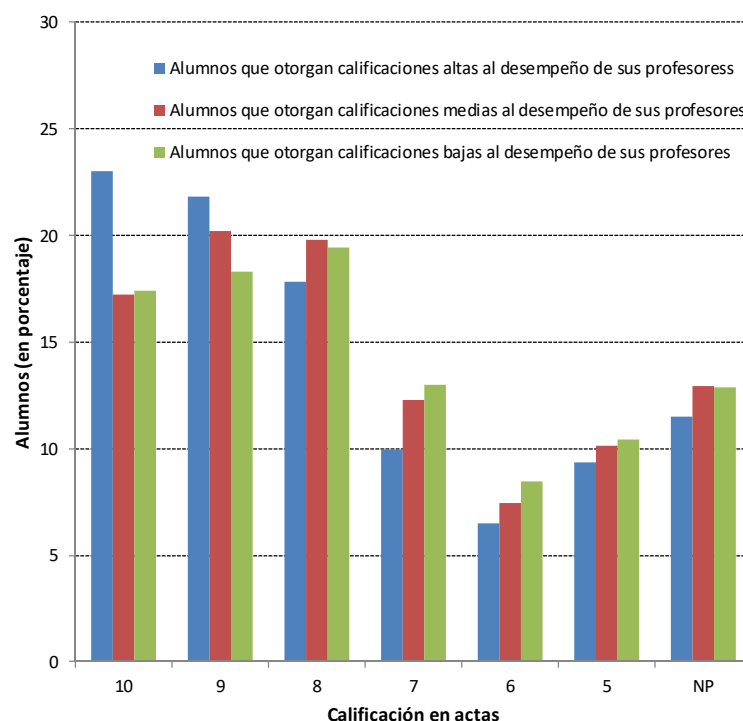
**Figura 2.** Calificación media en el reactivo “En la mayoría de los estudiantes del grupo, el aprendizaje de los temas de la asignatura ha sido” de la Encuesta de profesores

Aquí también las diferencias entre los grupos son significativas, como se comprueba en los resultados de ANOVA, calculados mediante el SPSS y presentados en la Tabla 3, con una F de 7.029 y una  $p = 0.001$

**Tabla 3.** Resultados del análisis de varianza de las calificaciones de los grupos de comparación en el reactivo «En la mayoría de los estudiantes del grupo, el aprendizaje de los temas de la asignatura ha sido» de la Encuesta de profesores.

|                                                              |                |            | Sum of Squares | df   | Mean Square | F     | Sig. |
|--------------------------------------------------------------|----------------|------------|----------------|------|-------------|-------|------|
| Aprendizaje (pp 8) *<br>Tercios por<br>calificación profesor | Between Groups | (Combined) | 4.714          | 2    | 2.357       | 7.029 | .001 |
|                                                              | Within Groups  |            | 568.756        | 1696 | .335        |       |      |
|                                                              | Total          |            | 573.470        | 1698 |             |       |      |

Por último, la evidencia también es afirmativa respecto a la relación entre los tres grupos de profesores definidos a partir de sus calificaciones en el CODAPA y las calificaciones de los alumnos asentadas en las actas: en la Figura 3 puede observarse que hay un mayor número relativo de “dieces” y “nueves” en el grupo *Altos*, un mayor número de “ochos” en el grupo *Medios* y un mayor número de “sietes” y “seises” en el grupo *Bajos*.



*Figura 3.* Distribución de la frecuencia relativa de las calificaciones asentadas en actas en cada uno de los grupos de comparación

En este caso, no todas las diferencias son significativas, como se puede observar en la Tabla 4, que son los resultados de ANOVA calculados mediante el SPSS. Las diferencias son significativas en cuanto a “dieces” y “nueves” y en cuanto a “sietes” y “seises”, no así en cuanto a “ochos” ni en cuanto a “cincos” y “enepés”

**Tabla 4.** Resultados del análisis de varianza de las calificaciones asentadas en actas de los grupos de comparación.

|                                                       |               |            | Sum of Squares | df   | Mean Square | F      | Sig. |
|-------------------------------------------------------|---------------|------------|----------------|------|-------------|--------|------|
| C10p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 6896.598       | 2    | 3448.299    | 6.072  | .002 |
|                                                       | WithinGroups  |            | 1086891.492    | 1914 | 567.864     |        |      |
|                                                       | Total         |            | 1093788.090    | 1916 |             |        |      |
| C09p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 4254.429       | 2    | 2127.215    | 7.786  | .000 |
|                                                       | WithinGroups  |            | 522898.617     | 1914 | 273.197     |        |      |
|                                                       | Total         |            | 527153.046     | 1916 |             |        |      |
| C08p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 810.540        | 2    | 405.270     | 1.972  | .139 |
|                                                       | WithinGroups  |            | 393255.780     | 1914 | 205.463     |        |      |
|                                                       | Total         |            | 394066.320     | 1916 |             |        |      |
| C07p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 3495.548       | 2    | 1747.774    | 13.666 | .000 |
|                                                       | WithinGroups  |            | 244781.524     | 1914 | 127.890     |        |      |
|                                                       | Total         |            | 248277.072     | 1916 |             |        |      |
| C06p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 1366.119       | 2    | 683.060     | 6.965  | .001 |
|                                                       | WithinGroups  |            | 187717.880     | 1914 | 98.076      |        |      |
|                                                       | Total         |            | 189083.999     | 1916 |             |        |      |
| C05p *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 181.040        | 2    | 90.520      | .436   | .647 |
|                                                       | WithinGroups  |            | 397809.679     | 1914 | 207.842     |        |      |
|                                                       | Total         |            | 397990.719     | 1916 |             |        |      |
| CNPp *<br>Tercios<br>por<br>calificació<br>n profesor | BetweenGroups | (Combined) | 209.381        | 2    | 104.690     | .465   | .628 |
|                                                       | WithinGroups  |            | 431292.212     | 1914 | 225.336     |        |      |
|                                                       | Total         |            | 431501.593     | 1916 |             |        |      |

## CONCLUSIONES

Los resultados obtenidos brindan argumentos en favor de la validez de criterio del CODAPA, lo que complementa los fundamentos de validez de contenido y empírica previamente realizados, además de ser consonante con los estudios realizados en otros contextos (por ejemplo, Marsh, 1984; Howard, 1985; y Cashin, 1995, citados en Greenwald, 1997), que concluyen en la existencia de la correlación entre los puntajes de los profesores y las calificaciones. Así mismo, se comprueba que los SET son instrumentos estadísticamente confiables y validos (Greenwald, 1997).

Por otro lado, al ser este un estudio de campo, en donde no se manipulan las variables, no es posible formular causalidades, tan solo se conoce que existen correlaciones o asociaciones



entre las variables. Estudios similares como los de Cohen (1987) citado en Marsh y Roche (1997), demuestran que los SET reflejan el aprendizaje de los estudiantes en componentes específicos del cuestionario, y de igual forma que pueden estar involucrados otros criterios como el historial del estudiante, su motivación, estrategias de estudio, etc.

Respecto a la relación entre la percepción de los alumnos acerca de su profesor y la percepción de los alumnos respecto a su propio aprendizaje, podemos decir que el supuesto propuesto por d'Apollina y Abrami (1997) que en aquellos docentes que son calificados como más eficientes en su labor, los alumnos, de igual manera, autocalifican su desempeño e involucramiento de forma favorable, se confirma.

En cuanto a la relación entre la percepción de los alumnos acerca de su profesor y la percepción del profesor acerca del aprendizaje de sus alumnos, se puede concluir que aquellos profesores que son evaluados de forma favorable también tienen una percepción favorable respecto a su grupo, no solo en cuanto a aprendizaje sino también en cuanto a actitudes y motivación, lo que se refleja en una correlación positiva entre todas estas variables.

Y sobre la relación entre la percepción de los alumnos acerca de su profesor y sus calificaciones finales en la asignatura, puede afirmarse que efectivamente de acuerdo con d'Apollina y Abrami (1997) quienes argumentan que las calificaciones altas implican un mayor aprendizaje, como que aquellos profesores quienes obtienen una calificación alta en su desempeño, son quienes generan mayor aprendizaje en los alumnos, se puede comprobar al observar que los alumnos quienes obtuvieron una calificación favorable califican de la misma forma su aprendizaje.

Como ya se mencionó en un comienzo, los resultados del cuestionario se utilizan para efectos de retroalimentación, y en la actualidad, también para la promoción y permanencia de los profesores, uso muy controversial que nuestros resultados respaldan, considerando que, como Acevedo Álvarez y Olivares Miranda (2010) y Spooren *et al.* (2013) señalan que, el aprendizaje de los estudiantes es la razón de ser de la docencia. Pero siempre los resultados de los SET deben ser una de múltiples fuentes para la toma de decisiones.

## BIBLIOGRAFÍA

- Acevedo, R. & Olivares, M. (2010). Fiabilidad y validez en la evaluación docente universitaria. *Revista Electrónica Actualidades Investigativas en Educación*, volumen (10) pp. 1-30. Recuperado de: <http://www.redalyc.org/articulo.oa?id=44713068009>
- Barrera, F., González, L., Medina, P., Núñez, G., Reyes, I. & Santos, J. (2018). *Actualización del Cuestionario de opinión del desempeño académico del profesor y del alumno (CODAPA)*. Ponencia presentada en el Coloquio Académico por la Equidad de Género en la FI-2018.
- D'Apollina, S. & Abrami, P. (1997). Navigating student ratings of instruction. *American Psychologist*, volumen (52) pp. 1198-1208.



- Facultad de Ingeniería- UNAM (2018). *Proyecto de renovación del cuestionario de opinión sobre el desempeño académico del profesor y del alumno (CODAPA)- Línea de acción 3 del programa 1.3 del Plan de Desarrollo de la Secretaría de Apoyo a la Docencia y Proyecto PAPIME 102517, abril de 2018.* México: UNAM
- Greenwald, A. (1997) Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, Volumen (52) pp. 1182-1186.
- Marsh, H. & Roche, L. (1997) Making student's evaluations of teaching effectiveness effective. *American Psychologist*, Volumen (52) pp. 1187-1197.
- Spooren, P., Brockx, B. & Mortelmans, D. (2013) On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*. volumen (83) 4 pp. 598-642
- Vásquez, J., Martínez, R. y Benítez, L. (2015). La evaluación de profesores de ingenierías desde la percepción de los alumnos. *Revista ANFEI Digital*, Año 1, No. 2, Ene-jun. 2015. Recuperado de <http://anfei.org.mx/revista/index.php/revista/article/view/121/469>