

A PROPÓSITO DE LA CALIDAD DE LA EVALUACIÓN

LAS CALIFICACIONES Y LAS COMPARACIONES



M.C. LUZ MARÍA NIETO CARAVEO

Profesora Investigadora de la UASLP

Lmnieto@uaslp.mx

*Publicado en Pulso, Diario de San Luis
Sección Ideas, Pág. 4a del jueves 25 de julio de 2002
San Luis Potosí, México.*

URL: <http://ambiental.uaslp.mx/docs/LMNC-AP020725.pdf>

UNA ANÉCDOTA

Hace muchos años, en una de las escuelas donde mis hijos estudiaron, el director hizo una reunión a la que convocó a todos los padres de familia del primer año de secundaria. Cual sería mi sorpresa al darme cuenta de que se trataba de regañarnos por los pésimos resultados obtenidos por la mayoría de nuestros hijos en los exámenes que recién se habían aplicado, en forma estandarizada, a todos los grupos. Nos explicó que, como una de las acciones de la escuela rumbo a la "excelencia", se estaban aplicando este tipo de exámenes que permitían comparar el desempeño de todos. Para lograrlo, nos aclaró, la evaluación final se basaría 100% en la calificación de tales exámenes, sin tomar en cuenta participaciones, tareas, ni otros criterios "subjetivos". Todo ello con el propósito de que nuestros hijos aprendieran a ser "competitivos". Luego procedió a hacer un recuento de los "errores" que nosotros, los papás, estábamos cometiendo y que permitían pronosticar un camino seguro hacia el fracaso y la mediocridad de nuestros hijos, claro, si seguíamos así.

En parte para ahorrarme el sermón y en parte por experiencia y "ojo clínico", me puse a leer los exámenes que mi hija había respondido y que yo tenía en mis manos. Francamente, muy pocas veces los revisaba con cuidado y jamás hacía comentarios en las reuniones mensuales del grupo, al que asistía como madre de familia. Pero esa tarde me puse a revisar. Tenía el examen de matemáticas, de español y de otros temas que ya no recuerdo bien. A la mitad de la primera hoja ya me estaba preguntando ¿Y dónde quedó la calidad de la evaluación? Porque el

director estaba hablando de evaluar la calidad del desempeño de nuestros hijos, pero nunca se planteó el problema de la calidad de las evaluaciones realizadas, es decir, la calidad intrínseca del instrumento y los criterios subyacentes. Varias preguntas del examen estaban pésimamente redactadas, y algunas eran imposibles de responder por la forma como estaban planteadas. El texto estaba mal impreso en algunas secciones, de manera que se tenía que adivinar su contenido. Había también faltas de ortografía y errores de sintaxis. Y por supuesto había preguntas que no parecían tener que ver con el tema bajo estudio, sino con el desciframiento de mensajes crípticos y lógicas booleanas.

Mientras los papás comenzaban a hacer preguntas y comentarios, me vi nuevamente en la disyuntiva de hablar o no hablar. Podía resultar contraproducente. Vi que la mayoría de las intervenciones mostraban padres sintiéndose culpables por no poder atender a sus hijos todo el tiempo que requerían las tareas y los trabajos encargados en la escuela. Finalmente levanté la mano y llegó mi turno. Dije lo que tenía que decir y expliqué, punto por punto, algunas cuestiones técnicas que, según yo recordaba, tienen que resolverse para hacer exámenes estandarizados¹. Mostré algunos ejemplos de los errores que ya comenté. Y finalmente pregunté ¿Qué técnicas de validación de reactivos se habían usado? ¿Cuál era el índice de validez predictiva² del "fracaso-éxito" de la vida futura de mi hija? ¿Y la validez de contenido? ¿Cuál era el coeficiente de confiabilidad y el margen de error de las evaluaciones? ¿Cuál era la distribución de frecuencias de las calificaciones dentro de grupos y entre grupos? ¿Cuáles eran las principales fuentes de variación? ¿Qué capacitación habían recibido los profesores que diseñaron los exámenes? y así... Obviamente las consecuencias fueron funestas y ese fue el último año que mis hijos estuvieron en esa escuela.

LAS COMPARACIONES

La anécdota da para muchas reflexiones, pero viene a cuento porque el martes pasado por la mañana, mientras leía algunos párrafos de la columna "de Todos" en Pulso, la recordé inmediatamente. Los párrafos a que me refiero, comienzan con una afirmación con la que estoy en completo acuerdo: "Es evidente que el sistema educativo necesita una revisión a fondo".

¹ Ver por ejemplo: http://www.chasque.net/gamolnar/evaluacion_educativa/evaluacion.03.html

² Sobre el tema de la validez predictiva, ver por ejemplo:
<http://redie.ens.uabc.mx/vol3no1/contenido-tirado.pdf>;
<http://www.ceneval.edu.mx/3nivel/3.2otroserv/g-sivp/sivp1.htm>

Mis desacuerdos vienen después, cuando los argumentos subsiguientes comparan los resultados de los exámenes de admisión de la UASLP con los de "... *otros planteles de educación superior...*". También aluden a una relación casi mecánica entre dicha calificación, el prestigio de los egresados y sus posibilidades de empleo, lo que parece apoyarse más en ciertas ideas que circulan por ahí que en datos o investigaciones concretas³.

Mi propósito aquí no es discutir dicho prestigio, ni cuestionar las calificaciones que obtienen los estudiantes de ingreso a las instituciones mencionadas. Mi propósito con la anécdota que acabo de mencionar es explicar que esos datos sobre los exámenes y sobre las percepciones de los egresados no son comparables, mientras no se usen los instrumentos adecuados y se hagan las investigaciones apropiadas. Es decir, quizás sea cierto lo que se dice, pero quizás no. ¿Cómo saberlo, si es que en realidad se trata de un tema relevante?

EL TEMA DE LOS EXÁMENES

La UASLP desde hace muchos años ha tratado de resolver el problema de los diversos criterios que se pueden utilizar para asignar calificaciones en los exámenes de admisión, aplicando exámenes estandarizados. Como ya se sabe⁴ (2), el examen de admisión de la UASLP está formado por tres componentes: A) Un examen de conocimientos, específico de cada escuela, facultad y/o carrera; B) Un examen psicométrico, diseñado dentro de la UASLP y que se aplica a todos los aspirantes; y C) El EXANI II, es decir, el Examen Nacional de Ingreso a la Educación Superior⁵, diseñado y aplicado por el Centro Nacional de Evaluación (CENEVAL) en muchas instituciones a lo largo y a lo ancho de nuestro país. Al generar una calificación final con estos tres componentes ponderados, la UASLP busca compensar o balancear los sesgos inherentes a cada uno de esos niveles de evaluación, y por lo tanto, equilibrar los criterios para ingreso de aspirantes a cada carrera. Esta es una de las mejoras que la UASLP ha introducido en el procedimiento de ingreso, ganando mucho en transparencia.

Pues bien, los resultados de los exámenes de conocimiento (tipo A) sólo son comparables entre aspirantes que van a la misma escuela y/o carrera.

³ Como podría ser un seguimiento sistemático de egresados o las investigaciones realizadas por el Dr. Ángel Díaz Barriga, del Dr. Carlos Muñoz Izquierdo, entre otras. Ver por ejemplo la investigación de Ángeles Valle sobre contratación de profesionistas en: <http://www.cesu.unam.mx/iresie/revistas/perfiles/perfiles-ant/64-07.htm>. Independientemente de los comentarios específicos que yo tendría a este artículo, me parece que necesitamos más estudios similares.

⁴ Ver procedimiento de admisión en: <http://www.uaslp.mx/admisiones/index.html>

⁵ Ver sitio con más información: <http://www.ceneval.edu.mx/4nivel/2exani2/exani2-3.htm>

Estrictamente hablando, no pueden ser usados para comparar aspirantes de diferentes escuelas. No sirve de nada que yo diga que un estudiante sacó 60 puntos en el examen de conocimientos de la Fac. de Medicina y que otro sacó 90 en el examen correspondiente de la Facultad de Ingeniería. Son instrumentos diferentes, aunque la escala sea la misma. Por su parte, el examen psicométrico (tipo B) de la UASLP, sólo puede ser usado para hacer comparaciones dentro de la UASLP, y no sirve para compararse con otros exámenes psicométricos diseñados con otros fines en otras instituciones.

Así, en el contexto actual, las únicas calificaciones de ingreso que son comparables entre diferentes instituciones, son las del CENEVAL (Tipo C). Y hasta donde yo sé, los *otros planteles* de educación superior del Estado a que se refiere el texto comentado, no lo aplican (aunque podrían tener sus propios instrumentos estandarizados). Dicho de otra manera, si un aspirante presentó el EXANI II en la Universidad Autónoma de Yucatán, puede comparársele con un aspirante a la UASLP que presentó el mismo examen; pero las calificaciones de los exámenes de ingreso de los "otros planteles" locales no pueden compararse con las de la UASLP, mientras los instrumentos no sean compatibles entre sí.

En síntesis: A mí nunca me han convencido los exámenes estandarizados como único mecanismo de evaluación del aprendizaje; pero siguen siendo necesarios cuando se trata de comparar, evaluar y tomar decisiones. 50 puntos en la calificación total de ingreso a la UASLP no son comparables con 70 de *otros planteles*.

¿Y dónde quedó la relación de todo esto con el problema del egreso y del prestigio de los egresados? Eso da para más artículos, así que lo veremos otro día.



Visita la página de la
Agenda Ambiental
de la Universidad Autónoma de San Luis Potosí

<http://ambiental.uaslp.mx/>

La información y opiniones contenidas en los artículos y demás publicaciones disponibles en las páginas de la Agenda Ambiental de la UASLP, son responsabilidad exclusiva de los autores, y se publican con base en el principio universitario del libre examen y discusión de las ideas.