

HACIA UNA NUEVA ERA EN LA EVALUACIÓN. EL DISEÑO DE REACTIVOS DIGITALES EN EL EXHCOBA

FELIPE TIRADO

Universidad Nacional Autónoma de México

EDUARDO BACKHOFF / NORMA LARRAZOLO

Universidad Autónoma de Baja California

RESUMEN: En este trabajo se hace el planteamiento de una nueva estrategia de evaluación basada en las ventajas que ofrecen los recursos digitales mediados por cómputo. Se presenta un planteamiento conceptual en torno al conocimiento considerando una visión taxonómica y una estrategia de respuesta construida, las cuales sirven como sustento para el diseño de los *patro-*

nes generadores de los nuevos *reactivos digitales* que se emplean en el *Examen de Habilidades y Conocimientos Básicos (EXHCOBA)*. Para ilustrar el planteamiento se ofrecen distintos ejemplos en diferentes áreas del conocimiento.

PALABRAS CLAVE: evaluación, reactivos digitales, campo cognoscitivo.

Introducción

La evaluación es un componente sustancial de todo proceso educativo, porque todo proyecto educativo tiene propósitos, por lo tanto es necesario poder valorar si se logran o en qué medida se alcanzan esos propósitos. La evaluación es un requerimiento fundamental para promover la calidad de la educación, considerando como criterios de calidad educativa cuando hay: "... una buena selección de objetivos educativos (*relevancia*) y los objetivos educativos son alcanzados (*efectividad*). También debe hacer énfasis sobre la justa distribución de los recursos educativos (*equidad*) y la preocupación específica del uso económico de los recursos (*eficiencia*)" (Jaap, Cees y Sall, 2003, p.5). Asimismo se debe cumplir con ciertos criterios de calidad en la evaluación para que sus conclusiones sean válidas. Jaap, Cees y Sall señalan que se requiere tener validez de contenido al evaluar lo que se pretende evaluar; validez de constructo teórico al estar apoyado en un cuerpo lógico integral y congruente de razonamientos; validez de contenidos por ser relevantes y representativos de la estructura del dominio que se evalúa; validez de criterio predictivo o concurrente al corresponder de manera coherente los indicadores que presentan otros

referentes o pruebas que son análogas; confiable en tanto es recurrente y consistente en sus aplicaciones y resultados; apropiado y viable en términos de los tiempos y costos requeridos; transparente en cuanto es pública la cobertura de los contenidos y de los procedimientos utilizados, tanto para evaluar como para calificar, de manera tal que se pueda orientar la toma de decisiones para definir políticas educativas bien sustentadas (pp. 100-102).

Los requerimientos y procedimientos para hacer evaluación cambian en función de los propósitos y las condiciones específicas con que se cuenta, por ello se requiere de manera permanente adecuar e innovar procedimientos que permitan aprovechar las múltiples ventajas que ofrecen las nuevas tecnologías, para encontrar mejores maneras que permitan evaluar los logros educativos de forma más rápida, eficaz, eficiente, objetiva, fiable y válida.

Las evaluaciones de gran escala, en las que se necesita evaluar a decenas o centenas de miles de estudiantes en tiempos cortos, precisan de procedimientos estandarizados y automatizados que sean objetivos, válidos y confiables. De aquí que tradicionalmente los procedimientos utilizados en evaluaciones de gran escala se han basado en el modelo de preguntas de opción múltiple que ofrecen las ventajas enunciadas, no obstante este modelo ha sido objeto de diversos señalamientos críticos, tales como; limitaciones para valorar ciertos niveles de complejidad de la taxonomía del conocimiento; la distancia que guardan con el uso auténtico del conocimiento en condiciones donde se aplica; las desviaciones que se introducen con las respuestas correctas dadas por azar; la facilidad de copiar respuestas; la caducidad de las preguntas en tanto éstas empiezan a ser conocidas y divulgadas a los nuevos sustentantes antes del examen. De aquí el que haya el interés por poder encontrar nuevos procedimientos que permitan solventar, o si no al menos atenuar, algunos de los problemas referidos.

La revolución tecnológica que más ha impactado la primera década del milenio ha sido sin duda el desarrollo de los medios electrónicos, sus posibilidades son enormes. La tecnología digital ofrece condiciones que abren oportunidades que apenas hace una década eran impensables, las cuales están dando lugar a una nueva era en la evaluación educativa. Entre las múltiples ventajas a las que se puede recurrir con las tecnologías de la información y la comunicación (*TICs*) están los enlaces en línea; los recursos multimedia que integran texto, imágenes, sonido y video en formatos de alta resolución; el hipertexto,

las ligas web, la interacción dinámica, la interconexión, la operación a distancia y, las base de datos relacionadas, entre otras ventajas.

Valiéndose de las ventajas de las nuevas tecnologías y con una nueva visión conceptual más constructivista, se desarrolló un nuevo paradigma para generar *reactivos digitales* en el *Examen de Habilidades y Conocimientos Básicos (EXHCOBA)*, en este trabajo se presenta esta nueva concepción y se proporcionan algunos ejemplos para ilustrar su metodología.

El *EXHCOBA* es el primer examen de admisión computarizado del país, se usa desde 1992, se aplica en 18 instituciones de educación media-superior y superior, en el año 2010 más de 125 mil aspirantes fueron evaluados, el examen ha sido estudiado sistemáticamente dando lugar a diversas publicaciones y tesis de doctorado (Backhoff y Tirado, 1992; Tirado y Backhoff, 1993; Backhoff y Tirado, 1994; Backhoff, Ibarra y Rosas, 1995; Backhoff, Tirado, Larrazolo y Antillón, 1996; Backhoff, Tirado, Larrazolo, 2001; Backhoff, Larrazolo y Rosas, 2002; Backhoff, 2003; González, M., 2004)

Planteamiento

La naturaleza del conocimiento que poseen las personas constituye un fenómeno complejo, hay un continuo de complejidad creciente que va de no saber nada, pasando por diferentes niveles, hasta llegar al dominio del experto. La diferenciación de los distintos niveles de dominio del conocimiento llevaron a plantear taxonomías, como la de Bloom (1956), en la cual se reconoce niveles de dominio de complejidad creciente: *conocimiento* (recordar información), *comprensión* (entendimiento del significado), *aplicación* (resolución de problemas), *análisis* (desglosar los contenidos), *síntesis* (reorganización de conocimientos) y *evaluación* (juzgar). Esta taxonomía fue actualizada por Anderson (2001) bajo una serie de adecuaciones que ponen énfasis en los procesos educativos en lugar de los productos, ahora las categorías son referidas a acciones del estudiante y procesos metacognitivos (comprender la regulación del propio pensamiento), la categoría de *conocimiento* cambia por *recuerdo*, *comprensión* por *entender*, y *síntesis* por *creatividad*, la cual se toma como el nivel más complejo.

En la metodología de evaluación utilizada por *PISA 2009* (OCDE, 2009; pp. 35-46) también se usa una diferenciación de niveles de dominio del conocimiento empleando criterios de proceso definidos como: *acceder* y *recuperar*, *integrar* e *interpretar*, *reflexionar* y

evaluar, al igual que en las otras taxonomías la categoría de *evaluar* es referida en los niveles más complejos, se argumenta que en tanto cada juicio de evaluación requiere consultar la propia experiencia, comparar y explicar bajo hipótesis que implican conocimientos específicos para construir una respuesta propia, donde lo relevante es la calidad del proceso de pensamiento más que la conclusión en sí referida en la respuesta (pp. 45-46).

Los modelos llamados de *evaluación auténtica* (Ahumada, 2005) de igual manera ponen énfasis en la apreciación del dominio teniendo en cuenta el proceso más que el logro del resultado, atiende a las actividades que se realizan dentro de un contexto auténtico para integrar una respuesta, lo significativo es la manera en que se construye la respuesta, en lugar de centrarse en el carácter reproductivo se pone el acento en el carácter productivo de la construcción de la respuesta. El modelo de “aprender haciendo” (“*learning by doing*”) se transfiere a “evaluar haciendo”.

La nueva generación de reactivos del *EXHCOBA* está orientada a evaluar las competencias cognoscitivas más complejas reflejadas en la elaboración de juicios, se trata de poder valorar la comprensión de las ideas, plantear dilemas que evoquen procesos de pensamiento en los que se requiera entender, juzgar y crear respuestas por parte de los sustentantes. Los sustentantes son requeridos a construir respuestas en formato abierto o reconstruir respuestas al tener que colocar una serie de elementos dentro de un campo en el cual se debe estructurar de una manera determinada los componentes del reactivo.

El *EXHCOBA* explora conocimientos básicos estructurales, básicos en tanto son las nociones fundamentales que sirven para la comprensión de diferentes campos de conocimientos, y estructurales porque forman las relaciones semánticas que permiten la organización del conocimiento. En términos de Piaget (1983), se trata de explorar las estructuras de asimilación del conocimiento.

La estrategia basada en recursos digitales se sustenta en aprovechar las ventajas de los recursos computarizados para evaluar respuestas construidas, en su posibilidad de presentar imágenes dinámicas interactivas, donde los sustentantes deben construir respuestas, ya sea por texto escrito o al acomodar por medio de arrastres (drag) los componentes que se presentan en las posiciones correspondientes en que deben ser colocados.

En reactivos de esta naturaleza se tiene un contexto de campo en el que se construyen las respuestas. Los conocimientos se integran en campos semánticos estructurales, como

lo plantea Luria (1981), se constituyen redes de significado donde unos conceptos son más cercanos a otros, por ejemplo el concepto *vaca* tiene consonancia con conceptos estrechamente relacionados como pueden ser: *leche*, *cuernos*, *establo* o *mugidos*. Por relación con la *leche* en un segundo nivel de asociación se tiene consonancia con *queso*, *crema*, *mantequilla* o *yogurt*. Y por relación con el *yogurt* en un tercer nivel se tiene asociaciones con *fermentación*, *bacteria*, *lactosa* o *ácido láctico*, pudiendo así seguir a otros niveles la expansión del campo semántico estructural.

Cuando se presentan reactivos dentro de un campo en el que se emiten diversas respuestas, las relaciones estructurales que se reconstruyen entre los elementos son parte misma de la respuesta. Por ejemplo, en un reactivo sobre historia universal donde se estructura un campo definido por una línea del tiempo y áreas geográficas, la ubicación de cada personaje o evento histórico guarda relaciones y por lo tanto requiere de congruencia con los demás, como el regreso de la capital a Tebas por Tutankamon y la conquista de Persia por Alejandro Magno.

Otro ejemplo, en un mapa los sustentantes deben colocar los componentes de arrastre que se les presentan, considérese acontecimientos históricos como la toma de la Bastilla o el asesinato de John Kenedy, los cuales deben ser arrastrados y colocados en sitios como París y Dallas respectivamente. Para esto, se delimita un área como respuesta correcta con cierto diámetro alrededor del lugar determinado de la respuesta correcta, de forma tal que si se coloca la etiqueta del acontecimiento relativamente cercana al círculo correspondiente ésta se da por buena, en tanto el sustentante hace explícito que tiene una noción. De esta manera la precisión para definir el área de la respuesta correcta queda en función del tamaño del diámetro que se desee establecer. Cuando se es posible ubicar y ordenar eventos históricos en el espacio y el tiempo es porque se tienen necesariamente otros conocimientos sobre el acontecimiento, de manera que se valora un campo de conocimientos de forma indirecta.

En otro ejemplo, pero ahora con reactivos de química, se muestra la estructura atómica de diferentes elementos y el sustentante debe arrastrar cada uno de ellos para posicionarlos en su lugar correspondiente dentro de la tabla periódica, o bien al presentar la estructura atómica de moléculas de diferentes compuestos éstas deben ser colocadas (arrastrarlas) en las categorías y orden de los diferentes compuestos (orgánicos/inorgánicos) que se presentan. En el área de lenguaje se presentan fragmentos de un texto que requiere ser complementado colocando (arrastrando) un conjunto de categorías gramaticas-

les (sustantivos, verbos, adverbios, adjetivos, conjunciones) a los espacios correspondientes, de forma tal que tenga congruencia y cobre sentido el escrito.

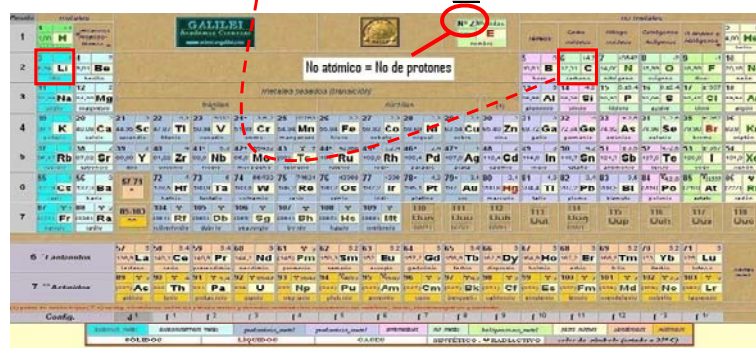
Ejemplo de reactivo de Química:

Coloca (arrastra) la representación de los elementos (núcleos atómicos), en la casilla que les corresponde en la *Tabla Periódica*, y escribe su nombre:

Representación del núcleo atómico con sólo protones

Contiene 6 protones Contiene 8 protones Contiene 3 protones Contiene 30 protones

Nombre: Carbono _____ Litio _____



La nueva generación del EXHCOBA está sustentada en bases de datos relacionadas que permiten tener *patrones generadores* de reactivos a partir de múltiples alternativas para complejizar un campo de conocimientos con diferentes respuestas que son equivalentes. Las alternativas son combinadas con otros grupos de respuestas que guardan congruencia, de manera que se pueden construir combinaciones de alternativas seleccionadas aleatoriamente por la computadora para cada caso, lo que permite tener cientos de reactivos diferentes, con lo cual se resuelve el problema de tener que estar haciendo nuevas versiones de reactivos. Asimismo los reactivos son aleatoriamente combinados con otros, de manera que se construyen un número cuasi infinito de posibles versiones, dando lugar a tener prácticamente cuestionarios únicos para cada aspirante. De este modo se resuelven dos importantes problemas de los exámenes objetivos de aplicación a gran escala, que son el desgaste que se tiene cuando se presenta repetidamente una pregunta, porque da lugar a que éstas sean conocidas por los nuevos sustentantes antes de la aplicación, o bien, a que siendo iguales las preguntas en formato de opción múltiple

se facilite el copiado de las respuestas. Ambas situaciones son problemas muy relevantes porque invalidan los resultados de la evaluación.

Al crearse grupos de respuestas integradas por múltiples componentes se pueden neutralizar ciertos sesgos culturales que ocurren en los reactivos convencionales. Por ejemplo, en un mapa de la república mexicana se solicita colocar (arrastrar) el nombre de 6 estados y 6 capitales en el lugar correspondiente. Dos estados y dos capitales son seleccionados aleatoriamente de un grupo de entidades del norte del país, de igual modo otros dos estados y dos capitales son seleccionados de entidades del centro y otros dos del sur, de manera tal que si el sustentante es del norte del país probablemente le será fácil reconocer los estados y capitales por ser de la región, más no los del centro o sur, de igual forma si el sustentante es del sur tendrá facilidad en identificar los de esta región más no los del centro o norte, neutralizándose así algunos posibles efectos de sesgo.

Gracias al formato electrónico en que operan los reactivos éstos se pueden formular en una gran diversidad de arreglos con imágenes digitales interactivas, en las que al momento en que el sustentante va resolviendo el examen se cambian colores o niveles de transparencia para advertir las respuestas que se han dado, las que quedan pendientes, los niveles de avance que se llevan, las respuestas en las que quedan dudas, así como el tiempo transcurrido y restante.

Conclusión

Las ventajas que ofrecen los recursos digitales en la nueva generación del *EXHCOBA* son muchas, en tanto se pueden evaluar habilidades cognoscitivas del pensamiento más complejas a partir de respuestas construidas, tales como resolución de problemas matemáticos, de escritura, lectura y pensamiento crítico. En el nuevo constructo del *EXHCOBA* se constituyen *patrones generadores* de reactivos en los que se requiere construir un complejo de respuestas, de manera que se transforman muchos aspectos tradicionales de las tecnologías de la evaluación educativa de gran escala, en tanto mejora procedimientos y superar problemas que se presentan en los métodos basados en lápiz y papel.

En el nuevo formato el sustentante puede cambiar su respuesta fácilmente sin tener que borrar, es más fáciles de operar, los reactivos son dinámicos, más atractivos (multimedia), seguros, se califica al instante y ofrece resultados inmediatamente (transparencia), para las

nuevas generaciones resultan ser más apropiados en tanto están familiarizadas con las nuevas tecnologías digitales, las cuales tienden a ser dominantes.

Referencias

- Anderson, L. W. and Krathwohl, D. R. (Editors); Airasian, P. W.; Cruikshank, K. A.; Mayer, R. E.; Pintrich, P. R.; Raths, J. and Wittrock, M. C. (Contributors). (2001) *A Taxonomy for Learning, Teaching, and Assessing A Revision of Bloom's: Taxonomy of Educational Objectives*. New York: Addison Wesley Longman.
- Ahumada, P. (2005). *Hacia una evaluación auténtica del aprendizaje*. México: Paidós.
- Backhoff E. y Tirado F., (1992). "Desarrollo del Examen de Habilidades y Conocimientos Básicos", en: *Revista de la Educación Superior*, México: ANUIES, Vol. XXI, No 3 (83), pp. 95-118
- Backhoff, E. y Tirado, F. (1994). Estructura y lógica del Examen de Habilidades y Conocimientos Básicos. *Revista Sonorense de Psicología*, 8(1), 21-33.
- Backhoff, E., Ibarra, M. y Rosas, M. (1995). Sistema Computarizado de Exámenes (SICODEX). *Revista Mexicana de Psicología*, 10(1), 55-62.
- Backhoff, E., Tirado, F., Larrazolo N. y Antillón L.E. (1996). Desigualdad de la educación básica en México. *Revista Latinoamericana de Estudios Educativos*, XXVI(3), 73-88.
- Backhoff, E., Larrazolo, N. y Rosas, M. (2000). Nivel de dificultad y discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1), 1-16.
- Backhoff, E. (2003). *Desarrollo y validación de un modelo para la selección de estudiantes de la Universidad Autónoma de Baja California*. Tesis doctoral. Universidad Aguascalientes, México.
- Bloom, B. S. (ed.), M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. (1956) *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*. New York: David McKay.
- González, M. (2004). *Defining and measuring academic standards for higher education: a formative study at the University of Sonora*. Tesis doctoral. The University of Arizona. USA.
- Jaap S., Cees G. and Sally M.T. (2003). *Educational evaluation, assessment, and monitoring: a systemic approach*. The Netherlands, Lisse, Swets & Zeitlinger Publishers (SZP)
- Luria, A. R. (1981); *Language and Cognition*, New York: Wiley & Sons.
- OECD (2009). *PISA 2009 Assessment Framework - Key Competencies in Reading, Mathematics and Science*. Paris: OECD.
- Piaget, J. (1983). "Piaget's Theory", en: Mussen, P. H. (editor), *Cari Michael's Manual of Child Psychology*, 4TH edition, Vol. I, pp. 103-128. New York: John Wiley & Sons.

- Tirado S. F. y Backhoff E. E.; (1993). "El Valor Psicopedagógico de los Errores, Un Recurso Desaprovechado en la Evaluación", en: *Revista Mexicana de Psicología*, Vol. 10, No. 2, pp, 175-182.
- Tirado S. F. y Backhoff E. E.; (1995). "El uso de un mismo examen para evaluar distintos niveles educativos: El caso del EXHCOBA". Memorias del III COMIE.
- Tirado, F., Backhoff, E., Larrazolo, N. y Rosas, M. (1997). Validez predictiva del Examen de Habilidades y Conocimientos Básicos. *Revista Mexicana de Investigación Educativa*, Vol. II No 3.
- Tirado, F. y Backhoff, E.; (1999). "La Compleja Elaboración de Exámenes, 16 Razones para Utilizar la Opción 'No sé' ", en: *Revista Mexicana de Investigación Educativa*, México: pp. 13-26, COMIE, número 7.