## "Implementación de un Modelo Psicométrico para el Nivel Medio Superior".

Luis Alberto Álvarez Aldaco

Asesor y Tutor Instituto Investigación y Desarrollo Educativo Universidad Autónoma de Baja California Iuis.alvarez.aldaco@gmail.com



### Resumen

El presente trabajo culmina con el proceso de desarrollo y validación de un modelo metodológico para elaborar exámenes de gran escala de referencia criterial, alineados con el curriculum en el ámbito de las competencias. Los antecedentes del trabajo son un conjunto de procedimientos desarrollados previamente para diseñar exámenes, elaborar especificaciones de reactivos v construir los ítems correspondientes. Así, la mejora del modelo consistió en desarrollar, operar y evaluar dos conjuntos de procedimientos complementarios: uno para validar mediante juicios los ítems y versiones de un test de logro educativo, otro para analizar los datos derivados de su aplicación y para elaborar reportes de resultados a los destinatarios de la información. El modelo se representó a través de una tabla de cuádruple entrada, en la que interactúan cuatro dimensiones o planos: temporal, metódico, de producto y calidad técnica.

Palabras clave: Psicometría, Evaluación del aprendizaje, Diseño de pruebas.

### **Abstract**

The present work culminates with the process of development and validation of a methodological model to elaborate large-scale of criterion-referenced tests, aligned with the curriculum in the field of competences. The antecedents of the work are a set of previously developed procedures to design exams, develop item specifications and construct the corresponding items. Thus, the improvement of the model consisted of developing, operating and evaluating two sets of complementary procedures: one to validate the items and versions of an educational achievement test through judgments, another to analyze the data derived from its application and to prepare results reports to the recipients of the information. The model was represented through a quadruple entry table, in which four dimensions or categories interact: temporal, methodical, product and technical

**Keywords:** Psychometrics, Assessment, Test development.

### ¿Demanda o necesidad de un modelo psicométrico?.

La siguiente propuesta está inmersa en la línea de investigación referida a la gestión y planificación, en particular en el ámbito de la evaluación educativa. Es un trabajo de intervención de carácter prescriptivo, el cual adapta un modelo que permita desarrollar exámenes criteriales a gran escala dentro del ámbito de las competencias, mismos que fue implementado en el Nivel Medio Superior (NMS) del Estado de Baja California, a partir del 2003. Tiene un carácter innovador el cual permite avanzar en el ámbito de la evaluación educativa a través de un diseño experimental.

En los últimos 20 años la historia de la educación en México ha estado caracterizada por una expansión de la matrícula en todos los niveles, en una atmósfera financiera reducida v casi solo pública. La reducción del gasto público en educación conlleva la necesidad de hacer más eficiente el Sistema Educativo por la vía de mejorar el desempeño académico y administrativo de los planteles, lo cual requiere de criterios, procedimientos de evaluación más elaborados y una mayor capacidad de supervisión; es decir, mediante un monitoreo permanente de la calidad del servicio que ofrecen.

En este contexto, cabe señalar que, después de un largo periodo de desinterés generalizado por los asuntos de corte evaluativo en el ámbito educativo, durante los últimos años se ha observado un creciente interés por la evaluación educativa y en particular por la evaluación de la calidad de la educación (Martínez, 1995). Lo anterior, es resultado de un cambio del énfasis ocurrido en las décadas de los 60's y 80's relativo a la expansión en la matrícula, que privilegió los aspectos cuantitativos de la educación, a uno que destaca la importancia de políticas explícitamente concebidas para dar prioridad a los objetivos de calidad y equidad Organización para la Cooperación y Desarrollo Económico (OCDE, 1997, 2007).

En diversos documentos que forman parte de la literatura reciente relativa a la investigación educativa nacional, se comenta la falta de una cultura de la evaluación generalizada y una escasa investigación sobre la evaluación del aprendizaje (Martínez, 1993, Jackson, 1993). Por lo cual resulta extraño si se considera la importancia que representa el enorme subsistema del NMS como insumo para el resto del Sistema Educativo Nacional y en general, para el desarrollo del país.

Así, el Colegio de Bachilleres del Estado de Baja California (CBBC) es una Institución Educativa con más tradición en el Estado, se tiene un egreso de 1984 al 2020 aproximado de 200 mil alumnos preparados para seguir con sus estudios superiores, así, el ingreso a la Universidad Autónoma de Baja California (UABC), en los últimos años, se ha mantenido en promedio del 60 por ciento de ingreso, con un aproximado de 20 mil alumnos admitidos. Sin embargo, es necesario incrementar estos porcentajes, lo cual garantizaría en los egresados un mejor nivel de competitividad (Álvarez y Miramontes, 2009).

En la actualidad, el 100 por ciento de los programas de estudios se encuentran inmersos en competencias. Esto significa que cada asignatura contenga la asignación didáctica, la cual se integra por programación semestral, secuencias didácticas, plan de evaluación. acreditación e instrumentos de evaluación. Por otro lado, el total de la población estudiantil deberá contar con el servicio de orientación educativa v tutorías (RIEMS, 2008). Aunado a lo anterior. la necesidad de evaluar el aprendizaje en la educación media superior (EMS), se pone de manifiesto ante la perspectiva del reciente proceso de reforma curricular. Tal política educativa requiere urgentemente de procedimientos e instrumentos, válidos y confiables, que apoyen un principio de diversificación educativa el cual debe garantizar la unidad del subsistema.

En México, la EMS se ha caracterizado por tener una gran diversidad de currículum, sin embargo, con la puesta en marcha de la Reforma Integral para la Educación Media Superior (RIEMS) a partir del 2009 al 2011 en el ámbito nacional, la Secretaria de Educación Pública (SEP) realizó un profundo cambio curricular en esta propuesta educativa. cuyos propósitos principales fueron el fortalecer este nivel educativo con un Marco Curricular Común (MCC), incrementar la cobertura nacional. mejorar los indicadores educativas, movilidad estudiantil y la profesionalización del docente del 100 por ciento en el ámbito de las competencias para el 2012. No obstante, poco se sabe acerca de la operación de tal currículum. De hecho, las modificaciones curriculares efectuadas forman parte de un contexto más amplio. El movimiento mundial de "educación para la vida" y el cambio de enfoque en cuanto a la enseñanza en competencias. Dicho entusiasmo ha traído apareiado un interés generalizado por la evaluación de dichas competencias (conceptos, habilidades, actitudes y valores).



# ¿Qué hace diferente a México de otros países en cuanto a contenidos curriculares o en materia de medición del aprendizaje?.

En el primer caso, parece que el país ha tenido la ventaja de contar siempre con una política curricular coherente, por lo menos en lo relativo a la planeación de los contenidos de la educación; tal es la situación actual que se orienta al desarrollo de competencias básicas y competencias disciplinares para el aprendizaje de los contenidos.

En el segundo caso, México ha estado prácticamente ausente de la historia de la psicometría moderna, misma que se dio durante casi todo el Siglo XX. Tal ausencia ha limitado la capacidad de la SEP para planear congruentemente el desarrollo de este ciclo educativo y de la sociedad en general en cuanto al conocimiento de lo que sucede en las escuelas. En este sentido, el presente trabajo constituye un esfuerzo para reducir en alguna medida dicha brecha.

En cuanto a las implicaciones que tiene, para la evaluación del aprendizaje a gran escala, el hecho de tener en el país un currículum nacional, resulta claro que también debe ser la base racional sobre la que se sustente la evaluación educativa (Nitko, 1994, 2004). Así, los instrumentos diseñados serán concebidos de bajo impacto. Ello significa que la aplicación no compromete las expectativas de los estudiantes en cuanto a la certificación del aprendizaje o al ingreso a nivel educativo posterior (LIX Legislatura, 2006).

**Tabla A.** Consideraciones para desarrollar el módulo de validación de ítems y módulo de análisis de datos y reportes de salida.

Módulo	Propósitos	Aspectos a considerar	Operatividad
Validación de	-Validar cada etapa	-Posterior a cada etapa	Panel de expertos
cada una de las	del modelo	-Previo a la prueba	multidisciplinarios que
etapas y	-Garantizar justicia	empírica de ítems.	emiten juicios en cada
validación de	a los examinados.	-Alineamiento del	etapa del modelo y
Îtems por juicio de	-Contar con	currículum con los ítems.	para los ítems y
expertos	evidencias basadas	-Ausencia de errores.	modelos de examen.
(docentes en	en el juicio de	-Detección de sesgo.	
servicio).	expertos.	-Legibilidad.	
		-Validez cultural.	
Análisis de datos	-Hacer significativa	-Análisis de ítems a	Expertos en análisis
y elaboración de	la información	través de datos	psicométrico.
informes para los	obtenida.	empíricos.	Técnicos en
usuarios del	-Informar de	-Elaboración de informes	informática.
instrumento.	manera apropiada	para: autoridades, jefes	
	al usuario.	de área, directivos,	
		docentes, padres de	
		familia y estudiantes.	

Diseñada por Álvarez (2010).

No obstante que esta metodología fue útil para desarrollar pruebas con una mínima calidad técnica en los niveles de educación básica, media superior y superior en Baja California, a la luz de los estándares psicométricos nacionales e internacionales para este tipo de instrumentos (AERA, APA y NCME, 1985,1993, 1999, 1994; Backhoff, Contreras y Larrazolo, 1999, INEE, 2005 y PREAL, 2010), se vislumbra la necesidad, de generar al menos dos módulos más: uno que permita evidenciar la validación de los reactivos y modelos de examen mediante juicios de expertos, previo a la prueba empírica; y otro modulo para el análisis de los datos obtenidos tras la aplicación a gran escala del instrumento y la elaboración de los correspondientes informes de resultados para los diversos usuarios de la prueba (Tabla A.).

### Impacto del Modelo Psicométrico.

El alcance de la investigación es de corte innovador ya que el modelo psicométrico se complementa con los dos módulos propuestos, y no se ha desarrollado un modelo similar para el nivel medio superior en el estado y en el país. Su implementación permite al CBBC, estar a la vanguardia en el ámbito evaluativo y la utilidad permitirá a la institución integrarse al bachillerato nacional, ya que uno de los requerimientos es que contenga como sistema de evaluación instrumentos estandarizados con indicadores nacionales e internacionales

Así, a través del modelo, es posible certificar el logro de los contenidos curriculares durante el proceso de enseñanza-aprendizaje, con el objetivo de conocer aquellas lagunas que son recurrentes en los diversos niveles educativos y por ende, la puesta en marcha de estrategias enfocadas a solventar las deficiencias identificadas.

La presente investigación se desarrolla en el CBBC. Para lo cual, se utilizó el modelo psicométrico propuesto por Nitko (1994, 2004) y complementado con los dos módulos diseñados por Álvarez (2010). Es una investigación mixta, se fundamenta en un esquema deductivo y lógico; formula supuestos para solucionar problemas, se caracteriza por ser de tipo experimental, mide ciertos rasgos y los analiza a través de técnicas estadísticas, por otro lado, utiliza instrumentos con preguntas cerradas para obtener evidencias durante el desarrollo del modelo.

Ambos paradigmas emplean procesos cuidadosos, sistemáticos y empíricos para generar conocimiento; sin embargo, el proceso cualitativo es secuencial y probatorio, cada etapa precede a la siguiente con orden riguroso y pretende intencionalmente acotar la información. El método utilizado para la selección de la muestra de estudio fue de tipo no probabilística, ya que representa el universo de la población que integra la comunidad estudiantil del CBBC. La muestra fue intencionada, posterior a la selección, se efectúa la aplicación al total de la población y posteriormente, se procede al análisis estadístico de los datos generados con los cuales se elaborar los diversos informes para los usuarios del modelo.

En las diversas etapas de la investigación se trabajó con 500 profesores de los diferentes planteles del Colegio, a quienes se les convocó a través del departamento de evaluación del aprendizaje para su participación en las actividades que implican el desarrollo de los tres módulos del modelo psicométrico. Cabe mencionar que para validar cada una de las etapas, se utilizaron evaluadores expertos, quienes

dieron cuenta de la efectividad del proceso a través de un cuestionario ex profeso.

A partir del 2009 se complementó el modelo con dos módulos más, hasta la actualidad se han integrado 90 especialistas, siendo profesores en servicio, mismos que han participado en los diversos comités como evaluadores distribuidos de tal manera que se integraron de uno o dos especialistas por asignatura.

Los criterios que se utilizaron para la selección de los profesores fueron: ser maestro asignado a las materias en proceso, estar en activo; tener amplio conocimiento del currículum, contar con un alto grado de compromiso institucional y responsabilidad para el trabajo académico. Además de estos requisitos, se consideraron otras características que les permitieron participar como diseñadores o especialistas de las pruebas criteriales. A pesar de que los profesores que participan en el proceso no son especialistas en el diseño de exámenes criteriales, si cuentan con gran experiencia en el ámbito educativo, con una reconocida trayectoria laboral y docente por parte de la institución. La mayoría han participado en la coordinación de las académicas del área de la especialidad, por otro lado, casi todos los docentes han seguido formándose dentro y fuera de la institución.

La población de estudio pertenece al CBBC, incluye Zona Costa (municipios de Tijuana, Tecate, Rosarito y Ensenada) y Zona Valle (municipio de Mexicali), con el total de los planteles oficiales. El tamaño de la muestra de estudio se encuentra conformada por 500 especialistas (diseñadores de la prueba); 120 mil alumnos (análisis cuantitativo) y 90 jueces (evaluadores en cada una de las etapas del proceso) que participaron en el análisis cualitativo del estudio realizado.

## Conclusiones y recomendaciones derivadas de la aplicación del modelo psicométrico.

Los dos nuevos módulos fueron desarrollados y operados, se obtuvieron evidencias sobre la funcionalidad; posteriormente, el contenido se desagregó con detalle, junto los tres módulos iníciales Nitko (1994, 2004); finalmente, el contenido de todos ellos se integró para formar el modelo completo para desarrollar exámenes de logro. Una fortaleza del modelo, es haber desarrollado, documentado con detalle una metodología integral para elaborar y validar exámenes criteriales de gran escala alineados con el currículum en competencias para el nivel medio superior. Hasta donde se sabe, ello es inédito en el contexto educativo (Martínez, 1993, 2005).

Otra fortaleza que se percibe, es la incorporación e integración de diferentes especialistas (docentes en servicio) en los diversos comités para la elaboración del examen, así cómo, la correspondiente capacitación para realizar las tareas solicitadas Ravela (2008); Solano y cols. (2006). Ello ha implicado sensibilizar, capacitar en aspectos básicos del análisis curricular, la psicometría y el desarrollo de pruebas, a un gran número de docentes, personal técnico, sobre todo al grupo de elementos que conforman el departamento de evaluación del aprendizaje del CBBC, con lo cual ha sido posible desarrollar los exámenes con cierta calidad técnica. Así, al reintegrarse los docentes participantes a sus labores académicos tienen el potencial de extender a los compañeros, tanto los materiales que les fueron entregados para apoyar la capacitación, como los conocimientos, habilidades adquiridas; coadyuvando así a consolidar una cultura evaluativa en los planteles e institución.

De manera relacionada, una consecuencia importante tras la aplicación del modelo, ha sido el creciente interés por la evaluación de autoridades educativas, directores de plantel y cuerpos colegiados, lo cual ha cubierto las necesidades evaluativas mediante la aplicación de los procedimientos, módulos, materiales que ahora constituyen el modelo, así como el acercamiento a docentes y alumnos.

También se considera un logro el esfuerzo que se hizo por relacionar la teoría, la investigación cognitiva, la psicometría al desarrollo de las pruebas; es decir, el haber operacionalizado el modelo con las principales conceptualizaciones, procedimientos y criterios técnicos propuestos por los autores (Nitko, 2004); Verheslt, 2004 y Popham 2000), en cuanto a los aspectos psicométricos de la teoría clásica de los test involucrados en el modelo, así como Bruner (1988) y Miller (1990), en cuanto a los aspectos cognitivos implicados. Así, fue posible elaborar múltiples documentos, formatos, ejemplos a partir de un marco conceptual fundamental, para apoyar la planeación de la prueba, el desarrollo de las especificaciones de ítems, la construcción de la tarea evaluativa, entre otras actividades necesarias para el desarrollo del examen.

Se propone que el modelo psicométrico desarrollado se pueda instrumentar a todos los subsistemas del nivel medio superior, lo cual fortalecería sus esquemas de evaluación y certificarían de manera científica el logro de los aprendizajes de los estudiantes.

Así como, crear estrategias diversas para la formación de recursos humanos en temas psicométricos, psicopedagógicos y evaluativos, que impulsen el desarrollo de exámenes en pequeña y gran escala. De hecho, es una necesidad que se hace sentir cada vez más en el país, y que ha sido una solicitud reiterada de las autoridades educativas CBBC.

Se considera necesario realizar un estudio orientado a explorar sistemáticamente el valor de los análisis gráficos de ítems, versiones de las pruebas, que se han desarrollado a través del tiempo en el CBBC, de tal manera que sea posible llevar a cabo interpretaciones menos intuitivas, pero si, con mayor precisión de las formas de representación gráfica que resulten asequibles, significativas para profesores y otros tipo de personal interesado.

Se sugiere que a partir de las técnicas estadísticas multinivel y los modelos causales, se lleven a cabo investigaciones sobre los posibles factores que inciden en los resultados de los exámenes, de modo que la metodología que aquí se propuso pueda contar con un modelo de variables contextuales que permitan configura un marco explicativo, con ello sea posible contar con políticas, programas y acciones efectivas para promover la mejora de los aprendizajes.

#### Referencias.

Alvarez, L., y Miramontes, A. (2009). Comportamiento de la Deserción y Reprobación en el Colegio de Bachilleres del Estado De Baja California: caso Plantel Ensenada. En: Memoria del X Congreso Nacional de Investigación Educativa. México: COMIE.

American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME). (1985, 1993, 1999). Standard for educational and psychological testing. Washington, D. C.: American Educational Research Association.

American Psychological Association (APA). (2010). APA Publication Manual. Sixth Edition. Publication Manual of the American Psychological Association, Sixth Edition American Psychological Association

Backhoff, E., Contreras, Ly Larrazolo, N. (1999). Estándares de Calidad de los Exámenes de Gran Escala: el Caso del EXHCOBA. Simposio: Estándares de Calidad de Gran Escala. Memorias del V Congreso Nacional de Investigación Educativa. México: COMIE.

Bruner, J. (1988). Desarrollo Cognitivo y Educación. Madrid: Morata.

Instituto Nacional para la Evaluación de la Educación. (INEE) (2005). Estándares nacionales ciclo 2003-3004, base de datos logro y contexto, alumno, docentes y directores México. INEE.

Jackson, M. (1993). Desarrollo del Español como Primera Lengua. En II Congreso Nacional de Investigación Educativa. La Investigación Educativa en los Ochenta, Perspectiva para los Noventa. Estados del Conocimiento.

LIX Legislatura/Gaceta Parlamentaria. (2006). Ley del sistema nacional para la evaluación de la educación y del sistema educativo. Revista Dossier Educativo. No. 56, pp. 2

Martínez, F. (1993). Evaluación del Aprendizaje. Presentación del Estado del Conocimiento. En II Congreso Nacional de Investigación Educativa. Síntesis del Congreso Nacional Temático. D.F.

Martínez, F. (1995). Diseño e Implementación de un Sistema Integral de Monitoreo de la Calidad de la Educación Básica. Universidad Autónoma de Aquascalientes. Documento.

Miller, G. (1990). Nouns in WordNet: a lexical inheritance system. In: international Journal of Lexicography 3 (4), pp. 245-264. Disponible en: ftp://ftp.cogsci.princeton.edu/pub/wordnet/5papers.

Nitko, A. (1994). A Model for Developing Curriculum-Driven Criterion-Referenced and Norm-Referenced National Examinations for Certification and Selection of Students. Ponencia presentada en la Conferencia Internacional sobre Evaluación y Medici6n Educativas, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).

Nitko, A. (2004). Educational Assessment of Students. Ohio: Prentice-Hall.

OCDE (1997), OECD Report on Regulatory Reform, Vol. II: Thematic Studies, OCDE, París.

OCDE (2002). Definition and Selection of Competences (DeSeCo): Theoretical and conceptual foundations. Strategy Pape, http://www.portal-stat.admin.

OCDE (2007), Economic Policy Reforms: Going for Growth 2007, OCDE, París.

Popham, W. (2000). Modern educational Measurement: practical guidelines for Educational. Por una educación de buena calidad para todos. Un enfoque educativo para el siglo. Press, pp. 8-28.

Programa de Promoción de la Reforma Educativa en America Latina y el Caribe. (2010). Seis lecciones internacionales sobre estándares nacionales. Siposis edcative. Selecciones que iluminan el cambio educativo. Partnership for Educational Revitalization in the Americans Santa Magdalena. No. 27

Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez, F., Aylwin, M., y Wolff, L. (2008). Las evaluaciones Educativas que América Latina Necesita. Documento de trabajo No. 40. Santiago de Chile: PREAL.

Reforma Integral de la Educación Media Superior. (RIEMS) (2008). Secretaría de Educación Pública, Subsecretaría de Educación Media Superior y ANUIES. Encuentro Nacional. "Avances y desafíos de la Reforma Integral de la Educación Media Superior". Consultado el 22 de Agosto del 2009 en: http://www.reforma-iems.sems.gob.mx/.

Solano, F., Contreras, N., y Backhoff, E. (2006). Traducción y adaptación de pruebas: Lecciones aprendidas y recomendaciones para países participantes en timss, pisa y otras comparaciones internacionales. Revista Electrónica de Investigación Educativa, 8 (2). Consultado el día de mes de año en: http://redie.uabc.mx/vol8no2/contenido-solano2.html.

Verhelst, N. (2004). Suplemento de referencia a la versión piloto preliminar del Manual para relacionar los exámenes de lenguaje al marco de referencia europeo común para lenguajes: aprendizaje, enseñanza, evaluación.