

---

# PROCEDIMIENTO ANALÍTICO-INFORMÁTICO PARA EL PROCESAMIENTO DE DATOS DERIVADOS DE PRUEBAS DE REFERENCIA CRITERIAL ALINEADAS CON EL CURRÍCULUM

---

CITLALLI SÁNCHEZ ÁLVAREZ / LUIS ÁNGEL CONTRERAS NIÑO

## RESUMEN:

En el presente trabajo se describe un componente metodológico para el análisis de resultados derivados de la evaluación criterial a gran escala y la correspondiente elaboración de informes de resultados a los destinatarios de la información, mismo que complementa un modelo psicométrico para elaborar pruebas desarrollado en la Universidad Autónoma de Baja California y que han utilizado varias instituciones educativas del estado para desarrollar sus exámenes. Así, se describen e ilustran los principales procedimientos, materiales e instrumentos desarrollados para llevar a cabo el procesamiento analítico – informático de los resultados de los exámenes. En particular, el manual técnico necesario para la operación, los formatos y listas de cotejo para el análisis de resultados y la elaboración de informes, y ejemplos que ilustran la aplicación de tales elementos; Además, se describe el esquema conceptual que guió las actividades de diseño de los procedimientos y materiales antes mencionados.

**PALABRAS CLAVE:** exámenes criterios, análisis de resultados, elaboración de informes

## INTRODUCCIÓN

En 2001 se publicó el Programa Nacional de Educación (PNE) 2001-2006, el cual proponía entre sus objetivos llevar a cabo un fortalecimiento del sistema educativo nacional mediante la consolidación del sistema de evaluación. Para ello propuso crear el Instituto Nacional para la Evaluación de la Educación

---

(INEE), instituir el organismo de evaluación para la educación media superior, y establecer criterios relativos al uso y la difusión de las evaluaciones (SEP, 2001).

El PNE destacaba la preocupación reciente que ha surgido en torno al perfeccionamiento de la medición de los indicadores de eficiencia y, especialmente, la medición del aprovechamiento escolar; y que aún prevalece:

[...] el uso insuficiente de los instrumentos y sistemas de evaluación, y la inexistencia de mecanismos claros y sistemáticos de rendición de cuentas – relativos al sistema y su organización, al aprendizaje de los alumnos, al impacto de los programas y el desempeño del personal, entre otros, los cuales contribuyen a desalentar el funcionamiento eficiente del sistema (SEP, 2001:121).

En respuesta a lo anterior, diversas instituciones educativas de nivel medio superior y superior desarrollaron estrategias y acciones para atender la problemática planteada por la SEP. En el estado de Baja California, las principales acciones institucionales emprendidas para mejorar la evaluación del aprendizaje a gran escala fueron realizadas principalmente por el Colegio de Bachilleres de Baja California (COBACH), el Colegio de Estudios Científicos y Tecnológicos de Baja California (CECyTE) y la Universidad Autónoma de Baja California (UABC); instituciones que para elaborar sus exámenes utilizaron un modelo psicométrico propuesto originalmente por Nitko (1994) para desarrollar pruebas criterioles de gran escala alineadas con el currículum, y que desde 1997 fue adaptado y ha sido desarrollado por el Instituto de Investigación y Desarrollo Educativo (IIIDE) de la UABC. (Contreras, 2000; Contreras y Backhoff, 2004)

## **MODELOS PARA EL DESARROLLO DE EXÁMENES CRITERIALES**

Durante los últimos 25 años se han desarrollado diversos modelos psicométricos para construir y validar instrumentos destinados a evaluar el logro educativo, los cuales pueden ser clasificados en dos grandes categorías:

---

los que se encuentran orientados a elaborar pruebas referidas a una norma, y los que buscan diseñar pruebas referidas a un criterio.

Con las pruebas normativas se pretende interpretar la ejecución del examinando en relación con la de otros que también respondieron a la prueba. Así, se ubica su ejecución por arriba, abajo o dentro de lo normal o promedio, a partir de la medida que se toma de una muestra representativa o de toda la población de estudiantes. En consecuencia, este tipo de pruebas proporciona una estimación de la habilidad relativa del estudiante. Por su parte, las pruebas criterioles interpretan la ejecución del examinando en relación a su estatus respecto a un criterio o dominio evaluativo bien definido y comunican lo que puede y no puede hacer el examinando en un campo de conocimiento, en relación a la medida que se toma de todo el dominio de contenido o de una muestra representativa de tareas obtenidas de ese dominio; usualmente para la evaluación de programas educativos, certificar competencias académicas, etc. (Berk, 1984; Nitko, 1994; Popham, 2000).

Los principales modelos para desarrollar pruebas criterioles fueron desarrollados a principios de los años ochenta del siglo pasado y quedaron plasmados en una guía general para la construcción de pruebas referidas a un criterio, cuyo editor (Berk, 1984) compiló a partir de los trabajos de investigadores como James Popham, Anthony Nitko, Ronald Hambleton y otros, quienes posteriormente continuaron desarrollando sus propios modelos psicométricos para apoyar la evaluación del aprendizaje.

Entre estos modelos destaca el de Nitko (1994), quien propuso construir exámenes nacionales referidos a un criterio orientado por el currículum, y en el cual se establece que el desarrollo de una prueba criterial debe considerar como punto de partida al currículum, y que las decisiones respecto a qué evaluar y cómo hacerlo están determinadas por los resultados de aprendizaje que establece dicho currículum. El modelo plantea nueve etapas para construir el

---

instrumento: 1) definir el dominio de resultados que pretende el currículum; 2) analizar el currículum; 3) desarrollar un plan de evaluación; 4) desarrollar la especificación de ítems; 5) producir y validar ítems; 6) ensamblar los exámenes; 7) establecer estándares de ejecución; 8) realizar un análisis primario de los ítems elaborados; y 9) realizar un análisis secundario de los resultados para elaborar reportes o informes dirigidos a distintos tipos de usuarios.

En el caso de México, a fines del siglo pasado se adaptó el modelo de Nitko para desarrollar una prueba de español para la educación primaria, reduciendo el número de etapas del modelo original, de nueve a seis para hacerlo viable y funcional en nuestro contexto. Dichas etapas incluyeron: 1) La definición del dominio de resultados que pretende el currículum. En esta primera fase se selecciona y capacita al grupo de especialistas que conformarán al comité diseñador del examen, quienes centran su análisis sobre los resultados que se pretenden obtener a través del currículum, de tal forma que el examen pueda quedar alineado con éste. Para ello deben revisarse y sintetizarse una variedad de fuentes de información que den luz acerca de aspectos teóricos y metodológicos subyacentes al programa educativo y currículum. 2) El análisis curricular para detectar el contenido importante a evaluar. Aquí se hace un esquema o “mapeo” del currículum de tal forma que queden estructurados sus contenidos, a fin de detectar aquellos que son particularmente relevantes. 3) Desarrollar un plan de evaluación. El plan debe incluir la descripción de lo que se espera evaluar, una selección o muestreo del contenido a evaluar, la elaboración de las especificaciones de ítems mediante las cuales se elaborarán las tareas evaluativas, y la selección y capacitación del comité elaborador de ítems. 4) Producir y validar ítems. Es aquí cuando se construyen los ítems en estricto apego a las especificaciones elaboradas. Concluida la prueba, debe efectuarse un piloteo y un análisis empírico de los puntajes obtenidos por los estudiantes, así como una revisión de los ítems y de la prueba, para entonces

---

proceder a su aplicación a gran escala. 5) Hacer un análisis primario para revisar la congruencia con los estándares de calidad del contenido y calidad técnica de los ítems y versiones de la prueba, así como diseñar y aplicar los procedimientos para calificar el examen y analizar los resultados. 6) Calificar la prueba y analizar sus resultados, con la finalidad de generar reportes de resultados apropiados a las necesidades informativas de diferentes destinatarios (Contreras, 2000).

### **ANÁLISIS DE DATOS Y ELABORACIÓN DE INFORMES DE RESULTADOS**

Tras varios años de desarrollo, dicho modelo psicométrico ha producido una documentación considerable, se ha implementado con finalidades diversas, ha operado en diferentes condiciones, con distintas poblaciones, y los procedimientos de capacitación y asesoría de los distintos comités de expertos han sido adaptados para satisfacer un amplio rango de necesidades institucionales. No obstante, el modelo aún no está consolidado. Hay varias áreas de trabajo que faltan por desarrollarse y existen fases y procedimientos que contempla el modelo original que todavía no se trabajan. Por ejemplo, se cuenta con pocos datos empíricos derivados de la aplicación a gran escala de los instrumentos elaborados hasta el momento, que permitan realimentar la metodología mediante la que fueron desarrollados los ítems y modelos; tampoco están documentadas las experiencias educativas tras la aplicación del modelo, ni se ha efectuado un proceso de evaluación formal de su operación, a pesar de que ha sufrido numerosos cambios al adaptarse a diversas condiciones.

Una notable limitación del modelo radica en que si bien considera la selección, capacitación, asesoría y realimentación de los trabajos de los comités de examen que coordinan la construcción de las pruebas, de quienes las diseñan, de los que elaboran las especificaciones de ítems y de quienes finalmente elaboran los

---

reactivos y los calibran, una vez que las versiones de la prueba son estructuradas y se aplican a gran escala, no se han desarrollado procedimientos o materiales específicos para proporcionar un entrenamiento, asesoría y realimentación a los responsables institucionales de los exámenes, de modo que puedan analizar los datos derivados de su aplicación o elaborar informes especializados destinados a los distintos usuarios de la información obtenida.

Así, cada institución ha definido de manera más o menos intuitiva qué datos le interesa procesar y analizar y el tipo de informes que elaborará para documentar los resultados de la aplicación de los exámenes. Sin embargo, por comentarios aportados por las instituciones con las que se ha colaborado, se sabe que los análisis de datos así realizados tienden a ser los más simples y directos; usualmente solo aquéllos que permiten calificar los exámenes y asignar las notas correspondientes a los estudiantes y elaborar un informe de resultados que contiene una estadística muy general sobre porcentajes de acreditación y reprobación por plantel y materia.

En consecuencia, el análisis de la información lograda mediante la aplicación de los instrumentos y la subsiguiente elaboración de los informes de resultados, no han resultado todo lo técnicos y significativos que podrían y debieran ser, sobre todo si se considera la enorme inversión en recursos humanos, financieros y tiempo hecha por las instituciones que han elaborado las mencionadas pruebas de gran escala.

Lo anterior planteó la necesidad de desarrollar una metodología explícita que incluya procedimientos y materiales específicos para proporcionar capacitación, asesoría y realimentación a los diferentes usuarios del modelo a fin de que puedan analizar, interpretar y generar informes significativos a partir de los resultados obtenidos tras la aplicación a gran escala de los instrumentos de evaluación diseñados. Así, el presente trabajo aporta una metodología para

---

atender las necesidades en dichas áreas a fin de continuar el desarrollo del mencionado modelo psicométrico.

## **METODOLOGÍA PARA ANALIZAR DATOS Y ELABORAR INFORMES DE RESULTADOS**

El proceso metódico que se propone en el presente trabajo se elaboró a partir de un esquema analítico-informático bastante detallado, mismo que se presenta de manera sintética en la figura 1. Como se observa, el esquema incluye un conjunto de elementos interrelacionados que están constituidos por procedimientos y materiales específicos para apoyar tanto el análisis e interpretación de resultados derivados de la evaluación criterial a gran escala, como la elaboración de los correspondientes informes de resultados.

La lógica general de las relaciones entre los elementos que presenta la figura 1 (ubicada al final del documento, al igual que todas las tablas y figuras de este trabajo) considera que cualquier análisis (esquema analítico en la figura) de resultados derivados del examen (base de datos) debe considerar primero las evidencias recabadas sobre la calidad técnica del instrumento (análisis psicométrico), antes de proceder a las operaciones necesarias para relacionar la ejecución de los examinados con su dominio del currículum (estructura curricular), desagregado por área o línea de formación evaluada en la prueba. Todo ello a la luz de las necesidades informativas de los usuarios (destinatarios de informes), a fin de que los correspondientes informes de resultados (esquema informático) les resulten útiles y significativos.

La población destinataria de la metodología son los responsables institucionales de los exámenes desarrollados. Se considera que la mayor parte de ellos fueron miembros del comité que coordinó la elaboración de exámenes en su institución, y quienes acompañaron y coordinaron las actividades de los demás comités de examen durante el desarrollo de cada instrumento. Por ello se

---

considera que poseen el contexto más apropiado para llevar a cabo las actividades de análisis y reporte de resultados. Para apoyar sus trabajos se elaboraron los siguientes materiales e instrumentos:

- Manual técnico para el análisis de datos derivados de la aplicación a gran escala de los exámenes criterios elaborados, así como para la elaboración de los correspondientes informes de resultados del examen.
- Formato general para elaborar el informe de resultados del examen.
- Ejemplos para ilustrar la aplicación de los procedimientos y criterios tanto para el análisis de datos, como para la elaboración de los informes de resultados.
- Cuestionario para evaluar la operación institucional de los materiales producidos, de los procedimientos operados y de la asesoría y realimentación proporcionadas.
- Documento de justificación del cuestionario que indica los factores e indicadores evaluativos que se juzgó importante incluir en el instrumento y la manera en que se operacionalizaron en las preguntas del instrumento.
- Listas de cotejo para realimentar los análisis e informe elaborados por los participantes.

Enseguida se describen brevemente los principales productos mencionados:

- 1) El resultado más importante del diseño de la metodología fue el manual técnico, mismo que incluye los aspectos conceptuales y metodológicos básicos necesarios para llevar a cabo el análisis de datos y la elaboración de informes especializados, según las necesidades de los destinatarios de la información. Los elementos incluidos en el manual están alineados con

---

aquellos que se especifican, para el análisis de resultados y la elaboración de informes, en los principales estándares internacionales que norman el desarrollo de instrumentos psicológicos y educativos (AERA-APA-NCME, 1999); En particular, el manual incluye diversos indicadores que deben conocer los destinatarios, estrategias analíticas para la búsqueda de información específica y relevante dependiendo del uso que se dará a los resultados, métodos de análisis de información e interpretación de la misma, y referencias técnicas a diversas fuentes respecto a la elaboración de informes. La figura 2 muestra la portada y el índice del manual.

- 2) Formatos para registrar los resultados de la aplicación a gran escala del examen, a fin de efectuar el análisis psicométrico de reactivos y el examen ya con suficientes datos empíricos, así como para consignar los resultados de los análisis desagregados de la ejecución de los examinandos en las áreas curriculares que explora el examen, y aquellos necesarios para registrar la información requerida para la elaboración de informes, según las necesidades de los destinatarios de la misma. La tabla 1 presenta una sección de uno de los formatos para apoyar la elaboración del reporte de resultados.
- 3) Ejemplos de aplicación de las técnicas de análisis de datos y de elaboración de informes de resultados, mediante los cuales se ilustran diferentes procesos analíticos de casos reales en los que se aplicaron ciertos aspectos de la metodología, y a partir de los cuales se crearon informes de resultados. Incluyen también informes dirigidos a distintos destinatarios, a partir de los mismos resultados. La figura 3 ilustra un análisis desagregado por línea de formación y nivel de dominio de uno de los exámenes elaborados.
- 4) Durante el desarrollo del proceso analítico-informático resulta necesario un acompañamiento, asesoría y realimentación de las actividades y

---

productos elaborados. Asimismo, es necesario evaluar la operación de los procedimientos utilizados y los materiales elaborados. Para ello se elaboraron un cuestionario y listas de cotejo para los análisis de datos efectuados y los informes de resultados elaborados por los responsables del examen. (la tabla 2 ilustra una sección de una de las listas de cotejo elaboradas)

## **CONCLUSIONES**

Hasta el momento, la transferencia de la tecnología contenida en el modelo para elaborar pruebas criterioles de gran escala, se ha llevado a cabo mediante tres módulos cuyos focos han sido, respectivamente, el diseño de las pruebas, la elaboración de especificaciones de reactivos y la elaboración de ítems, así como su posterior estructuración en modelos de examen y calibración psicométrica. Sin embargo, la transferencia de esta tecnología resulta incompleta y objetable si se deja a los responsables institucionales de los exámenes sin apoyos apropiados para que puedan analizar los resultados de sus exámenes y elaborar los correspondientes informes para los destinatarios institucionales de la información, de manera que resulten útiles y significativos.

Se considera que los procedimientos y materiales presentados en este trabajo son una contribución a llenar el vacío mencionado del modelo, y que eventualmente podrán formalizarse en un módulo especializado para llevar a cabo el entrenamiento, asesoría y realimentación necesarias para que los responsables institucionales de los exámenes elaborados estén en condiciones de analizar los resultados que se obtienen después de aplicarlos a la población objeto, así como para que puedan construir los informes de resultados especializados que den cuenta, de manera diferencial, de la ejecución que tuvieron los examinandos en las pruebas. Lo anterior permitirá contar no solo con un complemento deseable y necesario del modelo para desarrollar pruebas

---

criteriales de gran escala, sino que de hecho permitirá que los responsables del desarrollo del modelo puedan asumir plenamente el compromiso de transferir a las instituciones que la utilizan, la metodología para elaborar esta clase de instrumentos y sacar el debido provecho tras su aplicación.

## REFERENCIAS

- American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Berk, R. A. (ed.). (1984). *A guide to criterion-referenced test construction*. Baltimore: The Johns Hopkins University.
- Contreras, L. (2000). *Desarrollo y pilotaje de un examen de español para la educación primaria en Baja California, tesis para optar por el grado de maestro en Ciencias Educativas*. Ensenada, B.C.: IIDE. UABC. Disponible en: <http://eduweb.ens.uabc.mx/egresados/Tesis/indicetesis.htm>
- Contreras, L. y Backhoff, E. (2004). "Metodología para elaborar exámenes criteriales alineados al currículo", en Castañeda, S. (ed.), *Educación aprendizaje y cognición, teoría en la práctica*, México: Manual Moderno.
- Nitko, A. J. (1984). "Defining 'Criterion-referenced Test'". En Berk, R. A. (ed.). *A guide to criterion-referenced test construction*. Baltimore: The Johns Hopkins University Press, pp. 8-28.
- Nitko, A.J. (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 Medición Educativas, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).
- Popham, W. J. (2000). *Modern educational Measurement: practical guidelines for educational leaders* (3rd ed.). Needham Heights, MA: Allyn y Bacon, pp. 64-116.
- Secretaría de Educación Pública (SEP) (2001). *Programa Nacional de Educación 2001-2006. Por una educación de buena calidad para todos. Un enfoque educativo para el siglo XXI*. México: Secretaría de Educación Pública. pp. 100-184.

## CUADROS Y ESQUEMAS

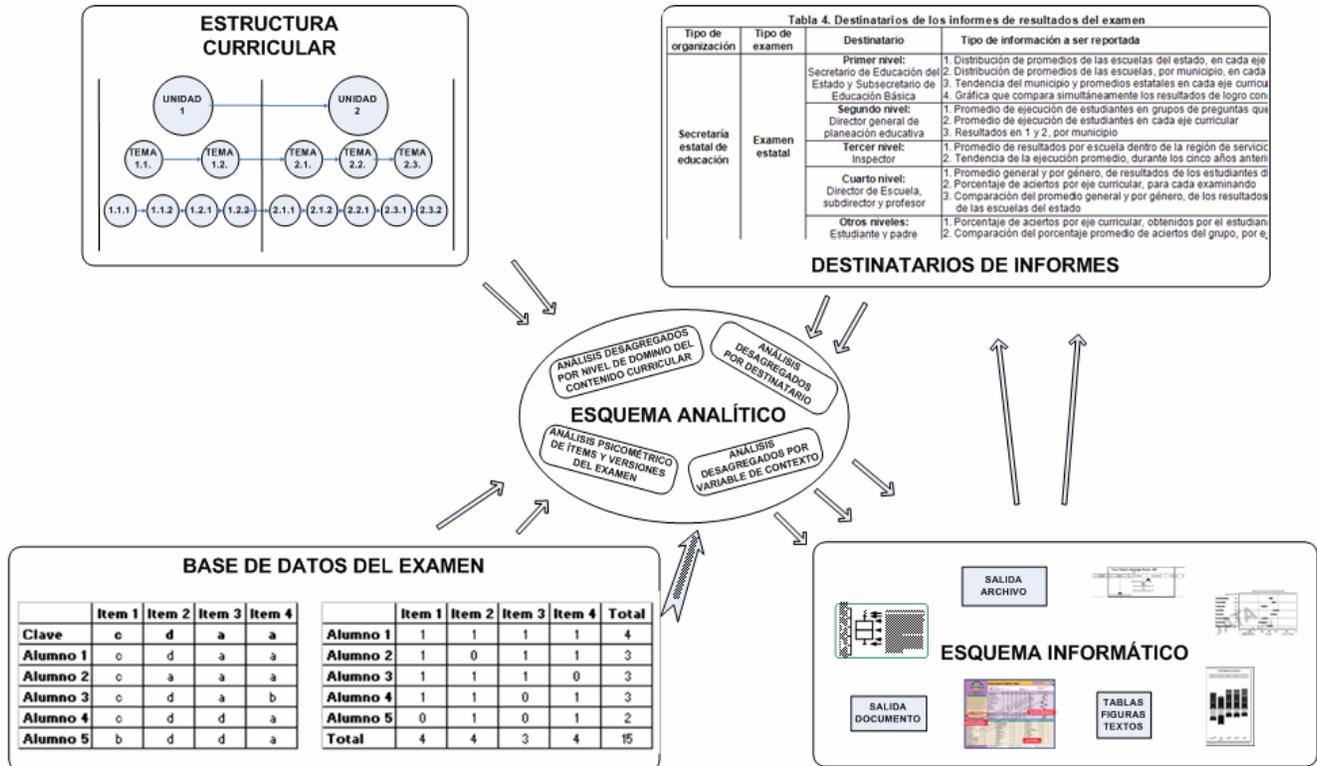


Figura 1 Esquema analítico - informático

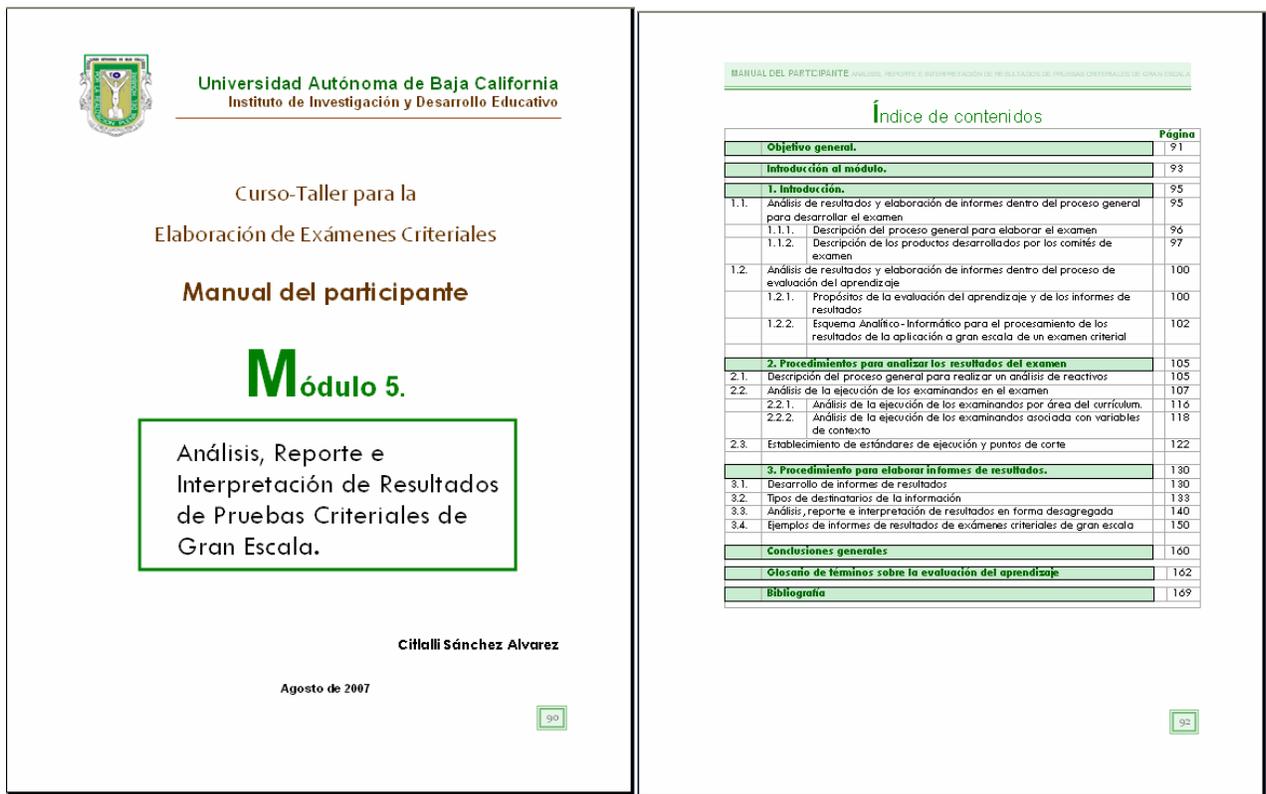


Figura 2 Portada e índice del manual para el análisis y reporte de resultados del examen

Figura 5. Resumen de los resultados de la evaluación por municipio y en el estado, de la línea de comprensión y creación de géneros populares.

La figura muestra el porcentaje de examinados cuyo dominio general de la comprensión y creación de géneros populares resultó alto, intermedio o bajo. El análisis porcentual muestra que el municipio con mayor porcentaje de estudiantes que tuvieron un dominio alto de esta línea de formación del currículum de español fue Ensenada, seguido por Rosarito, Mexicali, Tijuana y Tecate.

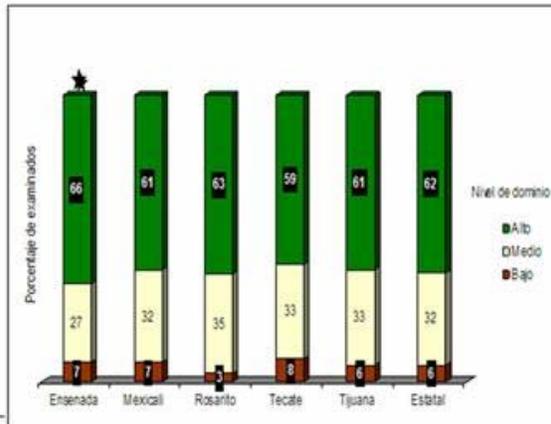


Figura 3 Ejemplo de análisis y reporte de resultados por línea de formación y nivel de dominio

Tabla 1 Sección del formato para elaborar informes de resultados del examen

ANÁLISIS DE LA EJECUCIÓN DE LOS EXAMINADOS
Resultados del ítem análisis:
Distribución de porcentajes de aciertos por área curricular:
Tendencias por área curricular respecto al año anterior:
Análisis de la relación entre la ejecución de los examinados y variables de contexto:
INFORMACIÓN GRÁFICA
Gráficas, figuras y tablas:
CONCLUSIONES
Limitaciones encontradas:
Recomendaciones a partir de los resultados (Usos de la información):

**Tabla 2 Sección de la lista de cotejo de componentes del análisis de la ejecución del examinando**

COMPONENTE	SI	NO
Describe las metas específicas que persigue el informe de resultados?		
Comunica los resultados de la evaluación en forma clara, precisa y significativa a los destinatarios?		
Explica la forma en que los resultados individuales y grupales ayudarán a la escuela en el mejoramiento del proceso enseñanza – aprendizaje?		
Explica y describe los propósitos que orientaron la desagregación de la información?		
Explica y describe los propósitos y técnicas utilizadas al reportar la información?		
Explica y describe el contexto dentro del cual se llevó a cabo la evaluación?		
Incluye información de:		
• Nombre de la escuela		
• Características del centro educativo		
• Descripción socioeconómica de la ciudad en la que se ubica,		
• Cantidad de alumnos que se atienden,		
• Características socioeconómicas del alumnado		
• Características culturales del alumnado		
• Grados académicos que se imparten		
• Materias o áreas que se evaluaron		
• Otros		
Explica y describe la naturaleza de la evaluación (prueba)?		
Contextualiza al usuario respecto al tipo de examen que se aplicó?		
Contextualiza al usuario respecto a la forma en que fue elaborado el examen (planeación, construcción, aplicación, y análisis de resultados)?		
Explica y describe quiénes son los destinatarios de la información?		
• Estudiantes		
• Profesores		
• Autoridades educativas		
• Padres de familia		
• Público en general		
• Otro		
Explica y describe qué información se les proporcionará a los destinatarios?		
Explica y describe qué usos podrán darle los destinatarios a la información?		
Explica y describe un análisis fundado, respecto a las razones por las cuales se deben reportar determinados resultados de la evaluación a estos destinatarios en particular?		
Si es destinatario de primer nivel, incluye:		
• Resultados obtenidos respecto a la ejecución a nivel estatal y municipal		