



CENTRO NACIONAL
DE EVALUACIÓN PARA
LA EDUCACIÓN SUPERIOR, A.C.

CENEVAL®

ANÁLISIS MULTINIVEL DE PREDICTIVIDAD DEL EXANI-I

Resultados de los estudios de validez
concurrente y predictiva de pruebas objetivas

Agustín Tristán López

1

ESTADÍSTICAS



CENTRO NACIONAL
DE EVALUACIÓN PARA
LA EDUCACIÓN SUPERIOR, A.C.

CENEVAL®

ANÁLISIS MULTINIVEL DE PREDICTIVIDAD DEL EXANI-I

Resultados de los estudios de validez concurrente
y predictiva de pruebas objetivas

INVESTIGACIONES

Agustín Tristán López

1

- *Análisis multinivel de predictividad del EXANI-I*
Resultados de los estudios de validez concurrente y predictiva de pruebas objetivas
Investigaciones 1

Agustín Tristán López

Revisión técnica:

Miguel Herrera Ortiz

Análisis multinivel de predictividad del EXANI-I
Resultados de los estudios de validez concurrente y predictiva de pruebas objetivas
Investigaciones 1

D.R. © 2009, Centro Nacional de Evaluación
para la Educación Superior, A.C. (Ceneval)
Av. Camino al Desierto de los Leones 19,
Col. San Ángel, Deleg. Álvaro Obregón,
C.P. 01000, México, D.F.
www.ceneval.edu.mx

Diseño y formación: Mónica Cortés Genis

Diciembre de 2009

Impreso en México • Printed in México

CONTENIDO

- Resumen 5
- Introducción. 7
- Metodología 17
- Análisis. 25
- Conclusiones 49
- Referencias 53
- Anexos. 57

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

RESUMEN • EL “Proyecto 12x2” es un ambicioso trabajo hecho por el

Ceneval a lo largo de varios años para dar seguimiento a los resultados de las pruebas de ingreso a la educación media superior y al nivel licenciatura (EXANI-I y II), con grupos controlados en escuelas públicas y privadas de 12 estados de la República Mexicana. El seguimiento incluye las calificaciones de los estudiantes en las asignaturas cursadas y aplicaciones repetidas de las pruebas del Ceneval. De este proyecto se tomaron los resultados de la primera generación (2005-2007) con la intención de estudiar la validez de criterio del EXANI-I (predictiva y retrospectiva) y obtener indicadores de la calidad de la prueba para los propósitos de selección de las instituciones usuarias.

La investigación se hizo con modelos multinivel, una metodología poco usual para los estudios de predictividad pero con la que es posible hacer análisis en muestras que no son rigurosamente aleatorias, como es el caso de la muestra del proyecto, en la cual los anidamientos por institución afectan la interpretación de las correlaciones simples, empleadas tradicionalmente en este tipo de análisis.

El objetivo primordial del trabajo consiste en la verificación de las cualidades predictivas del EXANI-I en comparación con las calificaciones del ciclo previo; también son analizadas las relaciones del éxito escolar con otras variables socioeconómicas recogidas en los cuestionarios contextuales del proyecto.

En la investigación se hicieron algunas constataciones que deberán explorarse en estudios posteriores, especialmente en relación con las correlaciones intraclase que se tienen en las instituciones y en el cambio de escala del Índice Ceneval, en comparación con la escala de calificaciones que otorgan las instituciones (generalmente de 6 a 10).

Con objeto de aclarar el funcionamiento de los modelos se comparan los resultados obtenidos por modelos multinivel y por correlaciones simples, haciendo evidentes las diferencias de aproximación entre ambos métodos. En las conclusiones se proponen nuevas líneas de trabajo para estudios de predictividad que podrán realizarse en el futuro con este tipo de herramientas.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

INTRODUCCIÓN • EL Ceneval emprendió un estudio longitudinal con estudiantes de nivel medio superior en varios estados de la República Mexicana, utilizando las dos pruebas desarrolladas hasta la fecha: una para la selección al ingreso medio superior (EXANI-I) y la otra para la selección al ingreso universitario (EXANI-II). Son varios los propósitos del “Proyecto 12x2”, pero en este estudio solamente se trabaja con uno: realizar análisis de validez de criterio (predictiva y retrospectiva o concurrente) del EXANI-I, contra otras formas de evaluar que practican las instituciones educativas (en particular de la escuela de procedencia y del primer año de la institución donde estudian los sustentantes).

El “Proyecto 12x2” se hizo en convenio con 24 instituciones educativas distribuidas en todo el territorio nacional, con el compromiso de reportar las calificaciones de los estudiantes desde su ingreso hasta la terminación de los estudios en la institución. Se trata de un proyecto ambicioso, con un enfoque de largo plazo (tres años por lo menos) y una estrecha coordinación entre las partes para disponer de los resultados de las asignaturas, en cada semestre, además de los resultados de aplicaciones de los EXANI a lo largo del tiempo.

La misma complejidad logística y de coordinación explica la falta de algunos datos en ciertas instituciones, así como también el hecho de tener valores parciales que indican omisiones, afectando a los modelos de validez predictiva objeto de este estudio. Independientemente de estos problemas el proyecto es muy atractivo y, si se procede con precaución, puede proporcionar líneas de trabajo y conclusiones interesantes para otras instituciones usuarias de las pruebas del Ceneval.

Se pretende realizar un estudio de validez de criterio del EXANI-I, enfocado en la predictividad de resultados en el primer semestre o el primer año del bachillerato y en la validez concurrente o retrospectiva de los resultados de secundaria.

Quedan como línea de continuidad de este trabajo los análisis de valor agregado o de ganancia académica, la identificación de aspectos relacionados con la equidad y la eficacia institucionales, o el diagnóstico de competencias desarrolladas o movilizadas en las pruebas y en las asignaturas.

Modelo general de la investigación

Para verificar la capacidad predictiva del EXANI-I con los resultados en el nivel medio superior de los estudiantes, se utiliza la base de datos disponible del “Proyecto 12x2”. A diferencia de los estudios tradicionales que emplean correlaciones simples, múltiples o parciales, en este proyecto se usan los modelos jerárquicos lineales o modelos multinivel, y se trata de dimensionar la predictividad respecto del promedio del primer semestre de educación media superior (PSEM1), del promedio de secundaria (PROMSEC) y del puntaje en la prueba objetiva del Ceneval (ICNE).

Se formulan entonces varias preguntas de investigación que hay que abordar con la metodología de modelos multinivel:

1. ¿Son igualmente predictivos PROMSEC e ICNE en el desempeño escolar de los primeros semestres de la educación media superior?¹
2. ¿Hay algunas variables socioeconómicas recabadas en el cuestionario de contexto que influyen en la predictividad respecto del PSEM1?
3. ¿Existe una relación concurrente o retrospectiva entre el promedio de secundaria y el puntaje Ceneval y otras variables socioeconómicas, que explique su intervención como variables predictivas del PSEM1?
4. ¿Qué tan predictivas son las secciones del EXANI-I (habilidades y conocimientos)?
5. ¿Existen posibles relaciones funcionales entre las variables socioeconómicas procedentes del cuestionario de contexto?

Descripción y antecedentes del proyecto

El EXANI-I es una prueba de selección para los estudiantes que van a ingresar al nivel medio superior, aunque las instituciones lo utilizan también como prueba diagnóstica de conocimientos y habilidades de los estudiantes provenientes de la educación secundaria. La prueba está construida con 128 reactivos objetivos de opción múltiple (cinco opciones),² con los cuales se produce un resultado global, dos valores de habilidades intelectuales (razonamiento matemático y verbal) y resultados de ocho áreas de conocimiento disciplinario (Español, Historia, Geografía, Formación cívica y ética, Matemáticas, Física, Química, y Biología). La calificación se hace en forma directa como proporción de aciertos y posteriormente se obtiene un puntaje –denominado Índice Ceneval– en forma normativa, asignando 1,000 puntos a la media y escalando a 100 la desviación estándar.

Para el “Proyecto 12x2”, la prueba producida por el Ceneval se aplicó en forma repetida a los estudiantes de las instituciones participantes. El propósito era hacer una investigación longitudinal desde el ingreso hasta el egreso de la educación media superior, con lo que se formó una base de datos que incluye los resultados globales y parciales de los EXANI-I y II, junto con datos académicos del estudiante y algunos datos socioeconómicos. Un antecedente de este tipo de estudios es el análisis longitudinal de validez predictiva (Lamp y Krohn, 2001), que podría servir de apoyo para construir otros modelos con los datos del “Proyecto 12x2”.

El “Proyecto 12x2” –tal como fue planteado originalmente por el Ceneval– tuvo un criterio de inclusión por cohortes anuales de 24 instituciones (de 12 estados de la República, una escue-

¹ Obsérvese que se pretende comparar las predictividades de la prueba objetiva de selección que produce el Ceneval y de las calificaciones de secundaria, pero no se intenta demostrar que la prueba es predictiva en algunos aspectos particulares del desempeño de los estudiantes a lo largo del bachillerato, lo cual requeriría, además de las calificaciones escolares, de otro tipo de insumos como los modelos educativos de cada institución, datos sobre la forma de integrar las calificaciones de cada asignatura, entre otros elementos.

² A partir de 2009, los reactivos del EXANI-I incluyen cuatro opciones de respuesta.

la privada y una pública, de ahí el nombre del proyecto), debiendo someterse a cuatro aplicaciones de los EXANI-I y II (una aplicación anual de cada prueba). Las cohortes de interés para este estudio conforman una muestra de 1,373 personas, provenientes de la generación 2005-2007.

Junto con las calificaciones de los sujetos, se cuenta con datos procedentes del cuestionario de contexto, en el que se incluyeron aspectos socioeconómicos y descriptivos de la persona, su familia y su entorno.

La tabla 1 presenta un resumen de las 20 instituciones conservadas para este estudio, en particular el número de casos y las medias y desviaciones estándar del Índice Ceneval y del promedio de secundaria. Obsérvese que para cada entidad federativa se pretendió contar con dos instituciones, pero esto fue imposible en cuatro estados: Campeche, Chihuahua, Querétaro y Tamaulipas. Esta limitante no perjudica la calidad de los resultados que se obtienen con el análisis multinivel (que se presentará posteriormente).

Tabla 1. Resultados generales de las instituciones de la muestra

Estado	Código de la institución	Índice Ceneval		Promedio de secundaria		Número de casos
		Media	Desviación estándar	Media	Desviación estándar	
AGUASCALIENTES 1	01DBP0001U	1020.7	99.3	8.79	0.77	99
AGUASCALIENTES 2	01PBH0036F	997.1	82.7	8.48	0.52	25
BAJA CALIFORNIA 1	02DBH0001K	979.1	87.9	8.7	0.75	99
BAJA CALIFORNIA 2	02PBH3096N	996.7	81.6	9.33	0.63	44
CAMPECHE	04ECB0001M	1004.0	90.1	8.36	0.71	70
CHIHUAHUA	08SBC2036O	1034.7	83.4	8.50	0.81	86
COLIMA 1	06DCT0019Z	1029.5	81.9	8.13	0.66	50
COLIMA 2	06PBH0007F	1001.3	97.6	7.96	0.81	20
DISTRITO FEDERAL 1	09DCB0021W	1032.5	88.2	8.67	0.80	90
DISTRITO FEDERAL 2	09PBH0124S	1046.9	91.4	8.23	0.70	89
ESTADO DE MÉXICO 1	15PBH6072B	1038.6	70.5	8.46	0.68	39
ESTADO DE MÉXICO 2	15UBH0094K	1068.2	93.4	8.96	0.74	91
MORELOS 1	17ECB0001Q	1057.4	86.2	9.07	0.66	123
MORELOS 2	17PBH0027Z	1084.6	78.4	8.49	0.63	66
QUERÉTARO	22PCT0007R	1094.4	75.3	7.98	0.83	78
QUINTANA ROO 1	23ECB0002Z	1096.0	77.0	8.87	0.77	99
QUINTANA ROO 2	23PBH3135L	1110.9	84.5	8.64	0.70	60
TABASCO 1	27ECB0031R	1129.8	75.7	8.58	0.71	94
TABASCO 2	27PCB0023O	1157.7	74.3	8.57	.81	46
TAMAULIPAS	28PBH0177L	1104.0	26.4	8.28	0.90	5
Total						1,373

Una revisión general de las medias del Índice Ceneval y del promedio de secundaria no revela una tendencia específica, por lo que son necesarios análisis más finos, hechos con los datos individuales y no con las medias, que permitan identificar las posibles relaciones o asociaciones entre las variables de desempeño de los estudiantes.

Validez de criterio

Una de las preguntas de investigación se refiere a la predictividad que tiene el EXANI-I respecto del éxito de los estudiantes en el nuevo ciclo escolar (primer semestre o primer año). Interés motivado por las reflexiones y los resultados que provienen de numerosos trabajos efectuados en el ámbito de la literatura sobre el tema, especialmente con pruebas de los Estados Unidos (SAT, ACT, GMAT, entre otras).

Este concepto de predictividad definido por la validez de criterio es una de las formas de recabar evidencias sobre la validez de una prueba. Para la validez de criterio se espera que la variable objetivo o focal del análisis correlacione con los resultados de otra prueba que representa una variable criterio. Se supone que la prueba “criterio” mide el propósito previsto y mientras más alta sea la correlación de la prueba focal, se dice que hay mayor predictividad con la prueba objetivo. La validez de criterio se puede manifestar en diversas modalidades de interpretación, dependiendo del tipo de prueba “criterio”, de modo que se habla de:

- a) validez predictiva, cuando la prueba focal correlaciona con los resultados que obtendrían en un futuro los sustentantes (Prueba de admisión a la licenciatura *vs.* Calificaciones del primer año de estudios profesionales)
- b) validez concurrente, cuando la prueba focal correlaciona con otra prueba que se utiliza para dictaminar algún desempeño o competencia de los sustentantes, pudiendo ocurrir en cualquier momento en el tiempo y bajo diversas condiciones (Prueba de competencias laborales de obreros *vs.* Calificaciones o puntajes asignados por su supervisor en función del desempeño en el trabajo)
- c) validez retrospectiva, cuando la prueba focal correlaciona con resultados de una prueba administrada en el pasado (Prueba de admisión a una licenciatura *vs.* Calificaciones de egreso del bachillerato)

Estas tres formas se complementan con otros análisis, a saber: la validez incremental (que identifica la intervención de cada variable en la predictividad), la validez de constructo (enfocada a identificar los aspectos medidos por el instrumento) y la validez discriminante (que se interesa por los aspectos no medidos por la misma prueba).

El análisis clásico de validez de criterio (en cualquiera de sus modalidades) se efectúa por medio de correlaciones simples (una variable objetivo *vs.* una variable criterio); en algunos casos se han hecho análisis de correlación múltiple o de correlación parcial (varias variables objetivo

vs. una variable criterio). En cualquier caso, el interés ha sido reportar la correlación como coeficiente de predictividad, dejando de lado la expresión de la regresión propiamente dicha.

El juicio sobre el valor de correlación que demuestra validez de criterio sigue reglas muy diversas, incluyendo algunas sugeridas en forma genérica, como las de Cohen (1988) donde se dice que $r=0.5$ indica una buena predictividad y $r=0.2$ es aceptable; en contraposición con los modelos de significancia estadística (que generalmente dependen del número de sujetos, a menos que se usen modelos en los que no se tenga influencia del efecto de tamaño), o esquemas comparativos contra trabajos equiparables de validez de criterio (Tristán y Vidal, 2006).

Validez predictiva de las pruebas de selección

Talento-Miller y Rudner (2008) recuerdan en forma coloquial la definición de validez predictiva como la relación entre un predictor o la combinación de predictores (que pueden ser los puntajes en una prueba o las calificaciones escolares) y un resultado (que pueden ser las calificaciones en un programa de cursos).

Ha sido sumamente interesante poder determinar si las pruebas de selección son predictivas, para que sean útiles en la admisión de los estudiantes cuya probabilidad de éxito en la institución es mayor.

Una de las preocupaciones relativas a la predictividad se refiere a la calidad de las pruebas objetivo y criterio. Weber revisa los trabajos sobre validez predictiva y compara las pruebas de selección contra el GPA (Grade Point Average o promedio de calificaciones en los primeros semestres), y llega a las mismas conclusiones que otros autores respecto de la mayor predictividad del GPA que el puntaje de las pruebas SAT y ACT para la selección. En un meta-análisis, Kuncel *et al.* (2005) sostienen que el GPA es un predictor robusto de los resultados escolares en el nivel siguiente.

En el informe del “Proyecto 12x2” (Arce, R.D. *et al.*, 2008) se citan algunas investigaciones de validez predictiva realizadas en Latinoamérica y en el Ceneval, siguiendo los modelos de estudios hechos por otras agencias con pruebas internacionales. De los trabajos citados son rescatables los de Moreira (2006), Musayón (2001), Serrano *et al.* (2002) y Tirado *et al.* (1997), entre cuyas conclusiones es destacable aquella según la cual las calificaciones del nivel previo (secundaria) son mejores predictores de las calificaciones del primer semestre y hasta el primer año el siguiente nivel (nivel medio superior), que el resultado de una prueba objetiva de selección.

La conclusión a la que llegan los autores citados sobre la predictividad resulta en buena medida contradictoria. Numerosas voces ponen en tela de juicio la validez de los resultados escolares –que son discutibles y dudosos por la falta de objetividad– pero, al mismo tiempo, se reportan altas correlaciones entre las calificaciones que otorgan los docentes en los diferentes niveles escolares. Por otro lado, los diseñadores de las pruebas objetivas estandarizadas esperan

altas correlaciones positivas entre dichas pruebas y los resultados escolares. De los dos insumos descritos (resultados escolares previos y pruebas objetivas estandarizadas) resulta claro que deberían tenerse correlaciones más altas entre las calificaciones de un ciclo escolar previo y el siguiente, por tratarse de insumos similares; en tanto que hay mayor heterogeneidad y, por lo tanto, menor correlación esperada entre calificaciones escolares y pruebas objetivas. De ahí a justificar la eliminación o una mínima utilización de las pruebas objetivas, hay un gran trecho.

Para atender la inquietud que produce la calidad de las pruebas, Downey (1979) revisa formas de ponderar *ítems* y opciones para obtener valores de validez predictiva, concurrente y retrospectiva entre SAT verbal y matemático y las calificaciones de los dos primeros semestres, en tanto que Sadler y Tai (2007) consideran la predictividad una vez ponderado el valor GPA. En este estudio no se hace una ponderación a priori de los puntajes de la prueba objetiva, pero el modelo que se construye permite obtener los coeficientes de la combinación lineal más apropiada para conseguir la mayor predictividad posible entre variables. Este enfoque objetivo parece ser el más apropiado para un estudio como el que se reporta en este trabajo.

Puede notarse que la mayoría (si no es que todos) los estudios de validez predictiva se limitan a reportar las correlaciones entre las variables, y dejan de lado las ecuaciones de regresión (Nunnally y Bernstein, 1995). Kobrin *et al.* (2008) y Talento-Miller y Rudner (2008) reportan valores brutos de la correlación y valores ajustados o corregidos por la restricción de rango, con lo cual se obtienen valores de correlación más altos.

Anastasi (1985) comenta el problema de sesgo presente en los estudios de validez predictiva porque las correlaciones sobredimensionan el valor de predictividad de la prueba objetivo, especialmente en los grupos más desfavorecidos socialmente. Este problema de sesgo ha inducido a varios autores a discutir la utilidad de las pruebas de selección (por ejemplo Murray, 2007), aunque no ofrecen soluciones técnicas de aplicación inmediata. Esta veta puede ser de mucho interés para el Ceneval si la explora de manera similar a como lo ha hecho el College Board (Mattern y Col, 2008).

Bell y Lumsden (1980), tras revisar la predictividad de las pruebas una vez eliminados *ítems* deficientes, encuentran que no se tiene una degradación en la correlación entre pruebas, inclusive con altos porcentajes de *ítems* eliminados. Esta constatación comprueba lo que se ha postulado en numerosas ocasiones en relación con la eliminación de reactivos deficientes: se mejora la calidad de la prueba, se refleja en una mejor discriminación de los sujetos, se produce un incremento en la desviación estándar de las puntuaciones y un menor error de medida; todas son ventajas nada despreciables comparativamente respecto del error de origen que se tiene con una prueba donde hay *ítems* deficientes.

Con respecto al uso de las correlaciones como modelo de validez de criterio, el Ceneval hizo análisis de predictividad de los EXANI-I y II desde 1996 y hasta 2002, por medio del Servicio de Información y Validez Predictiva del Ceneval. Se firmaba un convenio con las instituciones interesadas y bajo un principio de muestreo de los estudiantes que ingresaron por medio de la

prueba de selección diseñada por el Ceneval se obtenían correlaciones de 0.2 a 0.5, dependiendo del año y de la institución. Para la investigación objeto de este estudio se hizo el análisis de regresión con tres variables: (a) el promedio de secundaria, (b) el promedio del primer semestre de nivel medio superior y (c) el Índice Ceneval, obteniéndose los resultados del coeficiente de predictividad que se presentan a continuación.

La figura 1 proporciona la nube de dispersión de datos del promedio de secundaria vs. el resultado global del primer semestre de educación media superior, en tanto que la figura 2 contrasta el promedio de la secundaria y el Índice Ceneval obtenido en el EXANI-I. Los resultados se concentran en la tabla 2.

Los dos valores de correlación (0.44 y 0.51) son altos y difieren de cero en forma estadísticamente significativa. Debe notarse que las nubes de puntos y sus correlaciones no hacen una distinción entre cohortes, sino que se trata de resultados de todo el conjunto de datos. Evidentemente, se pueden tener instituciones con una mayor predictividad que otras, en términos de correlaciones, lo cual no es discernible con los datos tal y como se presentan en este momento.

El análisis de regresión lineal por cada cohorte (formada por los estudiantes de una institución) muestra la relación funcional entre el promedio de secundaria y el Índice Ceneval (ambos estandarizados con media 0 y desviación estándar 1). La nube de dispersión de puntos contiene las 1,373 parejas de valores estandarizados; la recta de trazo más grueso establece la regresión entre ambas variables en todo el conjunto, en tanto que las otras rectas se obtuvieron para cada institución.

Las rectas de la figura 3 tienen diferente pendiente e intercepto, por tratarse de regresiones específicas para la cohorte de cada institución. Este tipo de regresión se evitó en esta investigación al utilizar modelos multinivel, que toman en cuenta los anidamientos al momento de hacer el cálculo de los coeficientes de las variables independientes.

Para terminar esta sección, se comentarán dos puntos adicionales: las prácticas fraudulentas que se desarrollan con las pruebas de selección y los efectos de validez incremental que se tienen con las variables socioeconómicas y académicas.

Uno de los problemas de las pruebas estandarizadas de selección es la gama de prácticas fraudulentas, cuyo propósito es lograr puntuaciones más altas para ser admitido en una institución que tiene mucha demanda y poca oferta para nuevos estudiantes. Powers (1985) hace observaciones muy pertinentes sobre los problemas que produce la preparación de los estudiantes para una prueba de selección. En este estudio no se tomó en cuenta esta posible variable, por tratarse de la primera aplicación del EXANI-I como prueba de selección; pero debe observarse que el “Proyecto 12x2” cuenta con los resultados de las aplicaciones en años consecutivos sobre la misma población, y con eso se podría hacer un análisis de valor agregado y un comparativo que pudiera encontrar alguna pista en relación con la preparación de los estudiantes para la prueba. Este tipo de análisis queda fuera del propósito de este estudio.

Figura 1. Nube de dispersión y regresión PromSec vs. Sem1

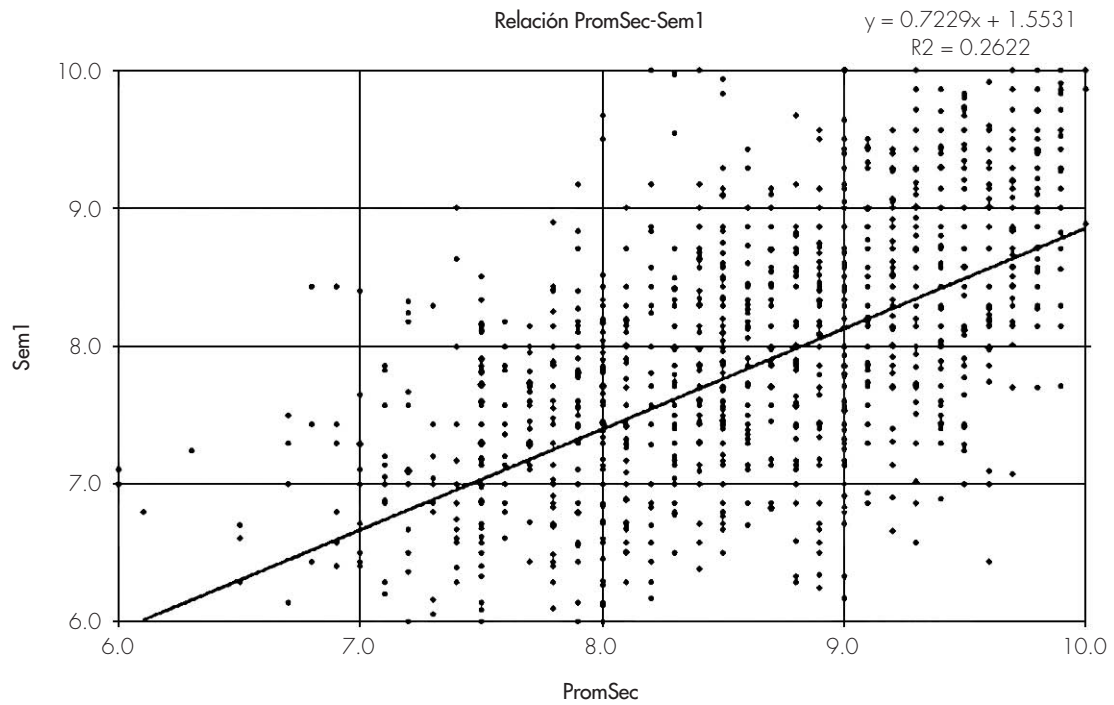


Figura 2. Nube de dispersión y regresión PromSec vs. Índice Ceneval

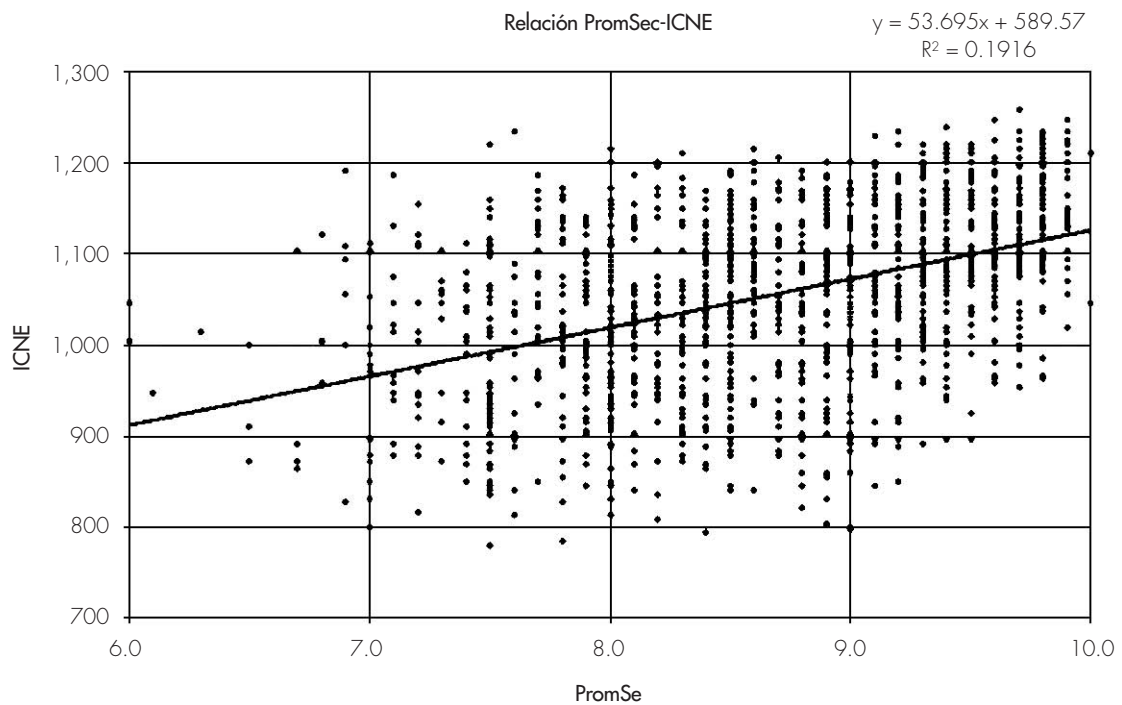
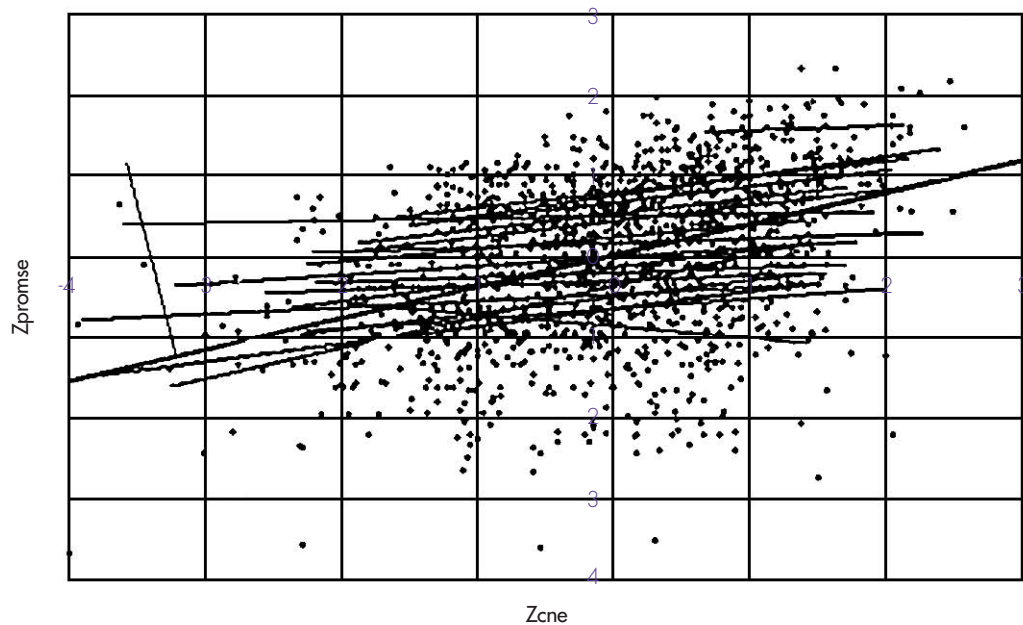


Tabla 2. Ecuaciones de regresión

Variable X	Variable Y	Ecuación de regresión	Coefficiente de predictividad (Correlación de Pearson)
PROM_SEC	SEM1	$Y = 0.7299X + 1.5531$	0.51
PROM_SEC	ICNE	$Y = 53.695X + 589.57$	0.44

Figura 3. Nube de dispersión y regresiones por cohorte



El análisis de validez predictiva puede incluir diversas variables, que conforme se combinan producen lo que se denomina validez incremental (por ejemplo, Hale *et al.*, 1992; Hararin *et al.*, 2007). Prácticamente todos los resultados de los estudios de validez predictiva se refieren a cálculos de correlaciones, pero nada impide realizar modelos multinivel, como los empleados en el presente estudio, con la ventaja de que se puede reportar el efecto de anidamiento por muestreo.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

METODOLOGÍA •

Modelos jerárquicos lineales para el análisis de predictividad

La correlación de las variables tiene sentido cuando el interés se centra en la asociatividad entre variables, pero el valor r de Pearson no permite identificar qué tanto predice una variable (por ejemplo promedio de secundaria) contra otra (por ejemplo resultado en una prueba objetiva); más bien se asocia con la “fuerza” de dicha asociación dentro de un concepto de tipo “gradiente predictivo” (Willms, 2006). Podría tenerse el interés de producir una ecuación de regresión combinando el promedio de secundaria y el Índice Ceneval, pero hay la objeción de que no se garantiza que los datos sean realmente aleatorios de una población muestra, ya que se trata de estudiantes de ciertas instituciones cuya correlación intra-institucional es más alta, debido a los anidamientos en los datos. Por ello, en lugar de trabajar con correlaciones múltiples se deberían utilizar modelos multivariados (Morrison, 1967) y sobre todo los modelos multinivel que permiten tomar en cuenta las correlaciones intra-grupales e inter-grupales, encaminándose a modelos más realistas, que -además de brindar la expresión que relaciona a las variables- permiten reportar indicadores de significancia estadística. La forma de combinar las dos variables en un modelo de regresión múltiple o por medio de un modelo de máxima correlación se explica en Tristán y Vidal (2006).

Pocos estudios como los de Schmidt (2000) o Holden *et al.* (2006) utilizan modelos multinivel en análisis de predictividad; este es el enfoque que se sigue en este trabajo, en virtud de las características de la muestra y porque proporciona principalmente la relación funcional entre variables, junto con los componentes de varianza del modelo. Se trata de un modelo más complejo que el de la correlación simple, pero, al mismo tiempo, es un modelo más ágil para estudiar la intervención de las variables, sus anidamientos, e identificar la significancia en la predictividad.

Cuando los datos procedentes de exploraciones en que las unidades de estudio no son independientes, sino que están anidadas o agrupadas –implicando una correlación entre los elementos de cada grupo–, se vuelve necesario referirse a los modelos jerárquicos lineales o modelos multinivel (Bryk y Raudenbush, 1992; Hox, 1995; Kreft y DeLeeuw, 1998; Goldstein, 1999). En particular se justifica el uso de los modelos multinivel porque los estudiantes que intervienen en el “Proyecto 12x2” no constituyen una muestra estrictamente aleatoria: puede considerarse que los estudiantes de una escuela correlacionan entre sí porque comparten factores grupales, de aula y de la escuela; del mismo modo las escuelas de un estado correlacionan mucho más entre sí que con las escuelas de otro estado. Estas unidades de alta correlación intragrupal se denominan anidamientos. Los cálculos indispensables para determinar un modelo jerárquico lineal o modelo multinivel requieren de software especializado, y se eligió el software HLM6.0 (Bryk, S., Raudenbush, A. y Congdon, R., 2005).

La presencia de la correlación intra-grupal hace que las técnicas de regresión clásicas sean inapropiadas, por lo que se introducen los modelos multinivel que facilitan la separación de las

relaciones de dependencia (o funcionales) en expresiones similares a regresiones lineales, incluyendo partes fijas (que se traducen en interceptos y coeficientes de regresión) y aleatorias (que reportan los errores o diferencias entre el dato observado y el resultado teórico que se obtiene con los valores fijos) (Gaviria, Martínez y Castro, 2004); el modelo da cuenta de las participaciones de cada variable junto con la varianza intra-grupal y la inter-grupal, simultáneamente.

Dentro de las ventajas de este tipo de modelos es posible obtener las relaciones entre las variables de los diferentes niveles, corregir el efecto de auto-correlación que se tiene por el anidamiento de los datos y favorecer la identificación del impacto de cada nivel en el modelo de predicción entre las variables explicativas y la variable dependiente. El “Proyecto 12x2” se presta claramente para emplear esta técnica, atendiendo a que los sujetos de cada cohorte (unidades del primer nivel) tienen características propias de la institución (unidades del segundo nivel) en la cual están siendo anidados. Inclusive se puede asumir que el grupo tiene características propias de los criterios de selección empleados por la institución.

Un modelo multinivel para una variable dependiente se expresa como una función lineal de variables independientes del primer nivel (también denominadas variables explicativas). Los coeficientes de la función se pueden expresar a su vez como modelos lineales de variables explicativas del segundo nivel. Los coeficientes aditivos se denominan interceptos y los multiplicativos se corresponden con las pendientes de la función. Cada dato al ser sustituido en el modelo multinivel produce un valor teórico, cuya diferencia contra el valor observado es aleatoria, pero generalmente se acepta que tiene una distribución normal de media nula y con una varianza que el programa reporta en cada nivel. Los criterios de convergencia están definidos dentro del programa en términos de la varianza del modelo, máxima verosimilitud y otras opciones configurables por el usuario.

Un modelo multinivel que relaciona al promedio de la secundaria y el puntaje Ceneval se expresa de esta manera:

$$P_SEM1 = B0 + B1*(PROM_SEC) + B2*(ICNE_1) + R \quad [1]$$

Donde las variables independientes o explicativas de primer nivel (estudiante) son PROM_SEC (promedio de la secundaria), ICNE_1 (Índice Ceneval), en tanto que la variable dependiente o criterio es P_SEM1 (Promedio del primer semestre). Todas estas variables provienen de la base de datos disponible del proyecto.

Los coeficientes de esta ecuación de primer nivel son las incógnitas que hay que resolver por medio del modelo multinivel. El valor B0 se conoce como intercepto, B1 y B2 son las pendientes relativas a las variables explicativas. Finalmente, R es el residuo o error aleatorio de ese primer nivel de anidamiento.

Los coeficientes de primer nivel son relaciones funcionales de los datos del segundo nivel (institución):

$$B0 = G00 + U0 \quad [2]$$

$$B1 = G10$$

$$B2 = G20$$

Los valores G00, G10 y G20 son los interceptos del modelo en el segundo nivel de anidamiento y U0 es el residuo o error aleatorio que proporciona las diferencias en el segundo nivel. El modelo de segundo nivel puede ser más complejo si se incluyen variables explicativas que corresponden con anidamientos de los datos dentro de las instituciones. Si se sustituyen los valores del nivel 2 en el nivel 1, se llega a la expresión explícita:

$$P_SEM1 = G00 + G10*(PROM_SEC) + G20*(ICNE_1) + R + U0 \quad [3]$$

En esta expresión, el intercepto G00 representa un valor base para el cálculo del promedio del primer semestre; los coeficientes de las variables proporcionan el impacto que un cambio unitario en la variable explicativa produce en la variable dependiente. Estos coeficientes pueden interpretarse como “ponderadores” o “pesos” con los cuales puede intervenir la variable independiente para predecir el resultado observado en la variable dependiente.

Finalmente, esta ecuación puede expresarse también de esta manera (las letras griegas son usuales en esta metodología):

$$P_SEM1_{ij} = \gamma_{00} + \gamma_{10}*(PROM_SEC_{ij}) + \gamma_{20}*(ICNE_1_{ij}) + r_{ij} + u_{0j} \quad [4]$$

El subíndice i indica las unidades del primer nivel (estudiantes), el subíndice j se corresponde con las unidades del segundo nivel (instituciones), el coeficiente aleatorio r_{ij} representa los errores para el primer nivel y el coeficiente aleatorio u_{0j} reporta los errores del segundo nivel.

Al hacer el cálculo, el programa HLM reporta una tabla con los resultados de los coeficientes, que son las incógnitas del modelo y un conjunto de parámetros que permiten dictaminar la calidad del modelo.

Tabla 3. Ejemplo de resultados de modelo multinivel

Efecto fijo	Coficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	7.980627	0.142004	56.200	19	0.000
PROM_SEC pend., B1					
INTRCPT2, G10	0.657064	0.040289	16.309	1370	0.000
ICNE_1 pend., B2					
INTRCPT2, G20	0.003706	0.000316	11.733	1370	0.000

La primera columna contiene el identificador o etiqueta del coeficiente que se calculó para los datos disponibles. La segunda y la tercera columnas dan el valor de los coeficientes y su error estándar. A partir de esta información se puede comparar el coeficiente contra el valor 0, para ello se emplea el valor *t* de Student, para un cierto número de grados de libertad. La última columna muestra la probabilidad (*p*-valor) de que el coeficiente calculado sea nulo y resulta que los coeficientes obtenidos ($G_{10}=0.657$, $G_{20}=0.0037$) difieren significativamente de cero por tenerse $p < 0.05$.

Si se sustituyen los coeficientes o términos fijos en la ecuación [4] del modelo multinivel, se llega a la expresión:

$$P_SEM_{1ij} = 7.98 + 0.657*(PROM_SEC_{ij}) + 0.0037*(ICNE_1_{ij}) + r_{ij} + u_{0j} \quad [5]$$

El modelo es útil para determinar la influencia de cada variable en función de los coeficientes. Cuando las variables son dicotómicas (como GÉNERO, TIPO DE ESCUELA), el cambio en la variable se calcula multiplicando el coeficiente por el paso de una categoría a otra; cuando las variables son continuas (como PROMEDIO DE SECUNDARIA) el cambio debe verse como relacionado con un incremento de una unidad en la variable (como cuando se pasa de $PROMSEC=6.0$ a $PROMSEC=7.0$).

El signo de cada coeficiente indica la “dirección” en la cual interviene el cambio. En el modelo mostrado, todos los coeficientes son positivos, de modo que un cambio de una unidad en el Índice Ceneval se traduce en un cambio positivo de 0.0037 puntos en la calificación del estudiante en el primer semestre. Este valor puede verse muy pequeño, pero hay que tomar en cuenta que un punto Ceneval es una unidad muy pequeña. Normalmente se habla de cambios de 100 puntos Ceneval, con lo que se tendría un impacto de 0.37 unidades en la calificación del primer semestre del estudiante. Del mismo modo, un cambio de una unidad en el promedio de secundaria se traduce en 0.65 unidades en la calificación del primer semestre del estudiante en bachillerato.

Para cada coeficiente se reporta su error estándar, que permite definir su intervalo de confianza y, con ello, su posible aceptación como valor significativamente diferente de cero. Generalmente se define el intervalo de confianza al 95%, que es aproximadamente igual a dos errores estándar alrededor del valor del coeficiente. Por ejemplo, para el valor base (intercepto) de 7.98 se tiene un intervalo de 7.69 a 8.26 ($7.98 - 2 \times 0.142$, $7.98 + 2 \times 0.142$); el intervalo está suficientemente alejado de cero, confirmando que el intercepto difiere del valor nulo.

La significancia del coeficiente se corrobora con el valor *t* ($t = \text{Coeficiente} / \text{Error estándar}$), que es de 56.2, con una probabilidad de que el coeficiente sea nulo (*p*-valor) de 0.00, para 19 grados de libertad (el número de grados de libertad, en este caso, es equivalente a 20 unidades de segundo nivel menos un parámetro calculado), que es altamente significativo ($p \leq 0.05$); es decir, se confirma que el intercepto difiere significativamente de cero.

En la segunda parte de la tabla de resultados se tienen los componentes de varianza.

Tabla 4. Componentes de varianza y ajuste del modelo

Estimación final de componentes de varianza:

Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.64220	0.41242	19	1206.83047	0.000
Residuo Nivel 1 R	0.70569	0.49800			

Deviance = 3036.138182

Los errores aleatorios tienen una media de 0 y una varianza reportada en la tabla, la cual también se dictamina respecto del valor nulo. La varianza total es la suma de las varianzas de R y U, en los dos niveles calculados: 0.412 es la varianza del nivel 2 y 0.498 la del nivel 1, lo cual se interpreta en el sentido de que 45% de la varianza es ocupada por el residuo U0 del nivel 2 y 55% por el residuo R de nivel 1. Estos valores de varianza son “grandes”, atendiendo a que una asociación perfecta sería explicada íntegramente por el modelo y las varianzas serían nulas. No es posible obtener dicha asociación perfecta a partir de datos reales, pero si se incluyen más variables “explicativas” se puede reducir la varianza y mejorar el modelo.

Finalmente, esta tabla proporciona el coeficiente denominado “Deviance” que es una medida del desajuste del modelo a los datos. Este valor no tiene interpretación por sí solo, sino que sirve de comparación entre modelos, de tal modo que la presencia de otras variables explicativas puede reducir el desajuste. En esta investigación no se pretende comparar entre los modelos, sino verificar la validez de criterio entre las variables explicativas, por lo que sólo se hará referencia a este coeficiente cuando sea pertinente.

En los estudios multinivel es costumbre presentar al inicio un modelo donde no hay variables explicativas en ninguno de los dos niveles (denominado modelo “nulo” o “vacío”). Esta práctica tiene el propósito de “explorar” la participación de las variables, pero no es utilizada aquí, ya que es preferible usar el modelo nulo sólo con el propósito de establecer un valor de base o de inicio para juzgar la bondad de los modelos donde se incluyen otras variables explicativas (Tristán *et al.*, 2008).

Criterio de inclusión de registros e imputación de datos faltantes

La base original consta de 2,202 registros asociados con las 24 instituciones que forman la muestra del proyecto. Para el estudio de predictividad del EXANI-I resultan indispensables tres datos “críticos” para el estudio: la calificación del EXANI-I (en Índice Ceneval o porcentaje de aciertos), los resultados del primer semestre o del primer año del estudiante (disponibles en escala de 5 a 10) y el dato de referencia que es el promedio de secundaria (en escala de 5 a 10).

Los registros en los cuales no se contaba con alguno de los datos citados no cumplían el criterio de inclusión y por lo tanto se eliminaron de la base.

Para el análisis de predictividad de esta investigación se conservaron los registros de estudiantes que contaron con promedio de secundaria, calificación en el EXANI-I y calificación en el primer semestre o en el primer año. La muestra de personas es tomada para esta investigación tal y como se obtuvo de las instituciones participantes, por lo que no se planteó la necesidad de hacer un rediseño usando, por ejemplo, los métodos Jackknife y Bootstrap (Fan y Wang, 1995; Willms y Smith, 2005; OECD, 2003-2005), ya que un rediseño muestral de este tipo sería motivo de otro estudio.

A su vez, la eliminación de registros puede hacerse por varios procedimientos y en distintos momentos del análisis: en primera instancia se pueden eliminar desde el manejador de la base de datos estadísticos (SPSS en este caso), siguiendo un modelo de eliminación general o particular (el método *listwise* hace la eliminación íntegramente en el archivo, en tanto que el método *pairwise* va eliminando registros *ad hoc*, dependiendo del análisis por realizar); una segunda opción se tiene al momento de hacer el archivo de estadísticas básicas con el programa de análisis multinivel (HLM), siguiendo el modelo *listwise*; en tercer lugar se puede hacer la eliminación al momento del análisis, partiendo de una base con datos faltantes, en este caso el programa de análisis multinivel (HLM) elimina conforme se eligen los campos en un modelo dado, siguiendo el modelo *pairwise*. Por las características de la muestra, se consideró poco útil hacer la depuración con el modelo *pairwise* y aceptar el criterio de inclusión para los datos críticos del análisis, eliminando los registros desde el manejador de la base de datos, quedando un total de 1,373 registros, representativos de 20 instituciones.

El impacto de registros eliminados al seguirse el criterio de inclusión puede ser una debilidad de la base de datos, pero debe observarse la importancia de eliminar en lugar de imputar las variables “críticas”, lo cual pudo haberse hecho por regresión u otro proceso similar, para tratar de evitar el posible sesgo que se produce con el modelo de eliminación, como advierten la APA (1994-2001) y Thompson (1996). La eliminación es una práctica recomendada y usual para HLM y para ecuaciones estructurales, especialmente en casos como éste, donde se cuenta con grupos de instituciones que definen anidamientos de los sujetos. Es importante hacer notar que la imputación carecería de sentido al asignar a todos los estudiantes de la institución el mismo valor para el Índice Ceneval o para el promedio de la secundaria.

En la nueva base depurada quedaron datos faltantes en alguno de los campos potencialmente útiles para el estudio, como es el caso del servicio de Internet, la escolaridad del padre o de la madre, entre otros. Estos campos son importantes para el estudio, pero no son “críticos” dentro del criterio de inclusión, por lo que en este caso se siguió el criterio de imputar los datos faltantes. De la gama de métodos de imputación disponibles en la literatura (véase por ejemplo la taxonomía propuesta por Little y Rubin, 2002), se eligió para este proyecto el modelo de donantes de un solo valor o “Hot Deck”, que consiste en ordenar los datos por institución, elegir sub-

grupos de donantes similares (que corresponden con los registros semejantes al que tiene el dato faltante) y asignar el valor más parecido entre dichos registros. El registro faltante se toma, por lo tanto, de los donantes más parecidos, por lo que no se trata de una media grupal o global, sino de la asignación más “razonable” o semejante al tipo de población que se tiene en cada institución. Otras formas de imputación y la manera de justificarlas se tiene en Tristán *et al.* (2008).

Los modelos de regresión en general y los modelos multinivel en particular se ven afectados por la escala de las variables, su rango, su origen y su densidad; los modelos pueden carecer de sentido si la escala de alguna variable es ordinal y, peor aun, si es categórica. El carácter ordinal de una variable no garantiza que las unidades sean equidistantes, razón por la cual los resultados del análisis deben revisarse con cuidado.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

ANÁLISIS • AL

abordar las preguntas de investigación planteadas para este trabajo se construyeron modelos multinivel donde se hacen intervenir las variables explicativas contenidas en la base de datos. Cada análisis se presenta con una organización que permite hacer comparaciones entre ellos, con estos elementos:

- a) *Caso*. Presenta el problema por resolver con relación a la pregunta de investigación.
- b) *Modelo*. Contiene las ecuaciones del modelo multinivel propuesto con las variables explicativas del caso.
- c) *Resultados*. Incluye la tabla de resultados procedente del análisis numérico hecho con el programa HLM.
- d) *Interpretación*. Presenta comentarios sobre los resultados obtenidos y algunas consideraciones que favorecen su interpretación.

1. Primer modelo predictivo sobre los resultados escolares

1.a) Caso

La primera pregunta de interés –como problema más inmediato de origen– se refiere a la predictividad que se puede tener sobre los resultados escolares del primer semestre P_SEM1 o del primer año y dos variables explicativas: el promedio de secundaria PROM_SEC y el resultado de la prueba de selección ICNE_1.

La pregunta de investigación para el primer modelo se expresa de esta forma: ¿Son igualmente predictivos PROMSEC e ICNE en el desempeño escolar de los primeros semestres de la educación media superior?

1.b) Modelo

Se consideran solamente las dos variables explicativas en el primer nivel (estudiante). Los coeficientes en el segundo nivel no tienen asociada ninguna variable. El modelo reportará los errores aleatorios en los niveles 1 y 2.

Modelo del nivel 1:

$$P_SEM1 = B0 + B1*(PROM_SEC) + B2*(ICNE_1) + R \quad [6]$$

Modelo del nivel 2:

$$B0 = G00 + U0 \quad [6bis]$$

$$B1 = G10$$

$$B2 = G20$$

1.c) Resultados

Tabla 5. Coeficientes, componentes de varianza y ajuste del primer modelo

Confiabilidad del INTRCPT1, B0 0.971					
Efecto fijo	Coefficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	7.980627	0.142004	56.200	19	0.000
PROM_SEC pend., B1					
INTRCPT2, G10	0.657064	0.040289	16.309	1370	0.000
ICNE_1 pend., B2					
INTRCPT2, G20	0.003706	0.000316	11.733	1370	0.000
Estimación final de componentes de varianza:					
Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.64220	0.41242	19	1206.83047	0.000
Residuo Nivel 1 R	0.70569	0.49800			
Deviance = 3036.138182					

1.d) Interpretación

El modelo brinda un ajuste bastante bueno, de acuerdo con la confiabilidad obtenida (0.971), y puede verse que las dos variables explicativas son significativamente distintas de cero (p-valor = 0.000). La varianza del residuo aleatorio U0 se corresponde con el 45% de la varianza total. El valor de desajuste 3036.1 se emplea como referencia para los siguientes modelos.

Con el modelo realizado se puede establecer que las dos variables explicativas son estadísticamente predictivas ($t_{\text{PROM_SEC}}=16.3$, $t_{\text{ICNE}_1}=11.7$) y, fuera de que los coeficientes son muy diferentes, no hay elementos para afirmar que una variable sea notablemente mejor que la otra.

Aquí se puede observar claramente que los valores de los coeficientes están asociados con la escala de las calificaciones, por lo que dan la apariencia de intervenir de manera “desproporcionada”. Se tiene, por lo tanto, que el promedio de secundaria interviene con un coeficiente de 0.657 y el Índice Ceneval de 0.0037. Estos valores hacen que el promedio de secundaria parezca notablemente más importante para predecir el resultado del primer semestre de la educación media superior, y de manera simplificada alguien podría interpretar el resultado en el sentido de que el Índice Ceneval participa con un coeficiente casi nulo (menos de cuatro milésimas), en tanto que el promedio de la secundaria interviene 170 veces más que lo que hace el Índice Ceneval.

Una interpretación así es errónea pues no debe olvidarse que el Índice Ceneval es un valor que corre de 700 a 1,300 puntos, en tanto que el promedio de secundaria es un valor que fluc-

túa de cinco a 10 puntos (y más probablemente de seis a 10 puntos, o inclusive en rangos menores en algunos sistemas escolares donde sólo se admite a estudiantes con promedios superiores a 7). Debe además tomarse en cuenta que no solamente es el rango de valores, sino también su densidad: en teoría, la escala Ceneval cuenta con 600 puntos (sólo enteros) en el rango, mientras que la escala de promedios de secundaria cuenta con 40 o 50 puntos (asumiendo que se tiene un entero y un decimal), lo cual produce esa intervención notablemente desproporcionada entre las dos variables. Este punto asociado con la escala y su densidad será revisado más adelante en esta investigación.

2. Segundo modelo predictivo en los resultados escolares

2.a) Caso

La primera pregunta de investigación se puede plantear una vez que se hace una transformación de las variables. En este caso se trata de resolver el problema de la escala por medio de una normalización o estandarización de las variables. Las variables estandarizadas tienen la ventaja de referirse a una nueva distribución que tiene media 0 y desviación estándar 1.

La pregunta de investigación es la misma del modelo previo: ¿Son igualmente predictivos PROMSEC e ICNE en el desempeño escolar de los primeros semestres de la educación media superior?

2.b) Modelo

Se consideran solamente las dos variables explicativas en el primer nivel (estudiante). Los coeficientes en el segundo nivel no tienen asociada ninguna variable. El modelo reportará los errores aleatorios en los niveles 1 y 2.

El normalizado se hace transformando los valores del promedio de secundaria (ZPROMSEC) y el Índice Ceneval (ZCNE) para cada estudiante dentro de la institución a la que pertenece, con la expresión clásica:

$$Z = (X-M)/S \quad [7]$$

donde X es el resultado de la persona, M es la media y S la desviación estándar de las puntuaciones de la población en la institución de interés. La media se ubica en $Z=0$ y la desviación estándar es 1. Los valores extremos se ubican aproximadamente a 3 desviaciones estándar por arriba y por debajo de la media, haciendo un rango aproximado de 6 desviaciones estándar entre casos extremos.

La transformación se realizó directamente en la base de datos, definiendo dos nuevas variables: ZCNE para el Índice Ceneval y ZPROMSEC para el promedio de secundaria estandarizado.

Con el normalizado, las escalas de las dos variables explicativas son similares en origen y en rango, aunque no en densidad, como se discutirá más adelante.

Modelo del nivel 1:

$$P_SEM1 = B0 + B1*(ZCNE) + B2*(ZPROMSEC) + R \quad [8]$$

Modelo del nivel 2:

$$B0 = G00 + U0 \quad [8bis]$$

$$B1 = G10$$

$$B2 = G20$$

2.c) Resultados

La tabla es muy similar a la anterior, pero hay diferencias en los coeficientes que están asociados con la escala de cada variable.

Tabla 6. Coeficientes, componentes de varianza y ajuste del segundo modelo

Confiabilidad del INTRCPT1, B0 0.960					
Efecto fijo	Coefficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	7.884527	0.121605	64.837	19	0.000
ZCNE pend., B1					
INTRCPT2, G10	0.294180	0.026637	11.044	1370	0.000
ZPROMSEC pend., B2					
INTRCPT2, G20	0.482411	0.033101	14.574	1370	0.000
Estimación final de componentes de varianza:					
Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.54551	0.29758	19	1017.35065	0.000
Residuo Nivel 1 R	0.70786	0.50107			

Deviance = 3030.336393

2.d) Interpretación

Puede constatarse que la calidad de la aproximación dada por el modelo es muy similar al del anterior, porque el coeficiente de desajuste (Deviance) es 3030.3 y en el anterior era 3036.1. El cambio notable se tiene en los coeficientes del modelo, ya que una vez normalizadas las variables, el promedio de secundaria participa con un coeficiente 0.48 y el Índice Ceneval con 0.29; la proporción no es tan notable como antes (sólo 1.65 en lugar de 177 veces que se tenía antes).

Nótese que los coeficientes siguen siendo significativos (esto no cambia con la normalización) y que los componentes de varianza sí se modifican; ahora la varianza del componente aleatorio de segundo nivel U_0 representa 37% de la varianza total.

Con el segundo modelo se corrobora que las dos variables explicativas son estadísticamente predictivas ($t_{\text{PROM_SEC}}=14.5$, $t_{\text{ICNE_1}}=11.0$); puede volverse a afirmar que ambas variables son igualmente predictivas sin una diferencia notable entre ellas.

Con los valores estandarizados se pueden hacer interpretaciones muy sencillas, por ejemplo un estudiante que se ubique en la media de calificaciones en secundaria y en el EXANI-I (cuyos valores estandarizados son: $Z_{\text{PROMSEC}}=0$, $Z_{\text{CNE}}=0$) tendría una calificación esperada de 7.88 en el primer semestre de educación media superior.

Todos los cambios deben hacerse referidos a fracciones de desviación estándar. Un estudiante que esté por arriba de la media en una desviación estándar del Índice Ceneval tendría un incremento esperado de 0.29 puntos en la calificación de primer semestre, en tanto que un cambio de una desviación estándar en el promedio de secundaria influye teóricamente en 0.48 puntos de mejora en el promedio del primer semestre.

Ya se indicó previamente que los casos extremos tienen una diferencia aproximada de 6 desviaciones estándar; en este caso, los sujetos con mayor diferencia de puntaje en el Índice Ceneval tienen una diferencia teórica en la calificación del primer semestre de 1.8 puntos ($6 \times 0.2941=1.765$), en tanto que los estudiantes con mayor diferencia de calificación promedio de secundaria tienen una diferencia esperada de casi 3 puntos ($6 \times 0.4824 = 2.8944$).

Si se combinan los casos extremos, se llega a una diferencia total teórica de más de 4.6 puntos ($1.765+2.894=4.659$), lo que marca la distancia que hay entre un estudiante que tiene el mayor promedio de secundaria y el más alto puntaje de Índice Ceneval, respecto del estudiante con el peor promedio de secundaria y el más bajo Índice Ceneval.

3. Tercer modelo predictivo en los resultados escolares

3.a) Caso

El modelo anterior atacó el problema de la escala y en este tercer modelo se va a atender además el problema de la densidad. Ya se comentó que el promedio de secundaria debe tener aproximadamente de 40 a 50 puntos cuando más en el rango de calificaciones, en tanto que el Índice Ceneval cuenta con 600 puntos. Se tiene la necesidad de cambiar la escala de la variable normalizada del Índice Ceneval, de manera tal que se pueda reducir su densidad a un valor similar al del promedio de secundaria.

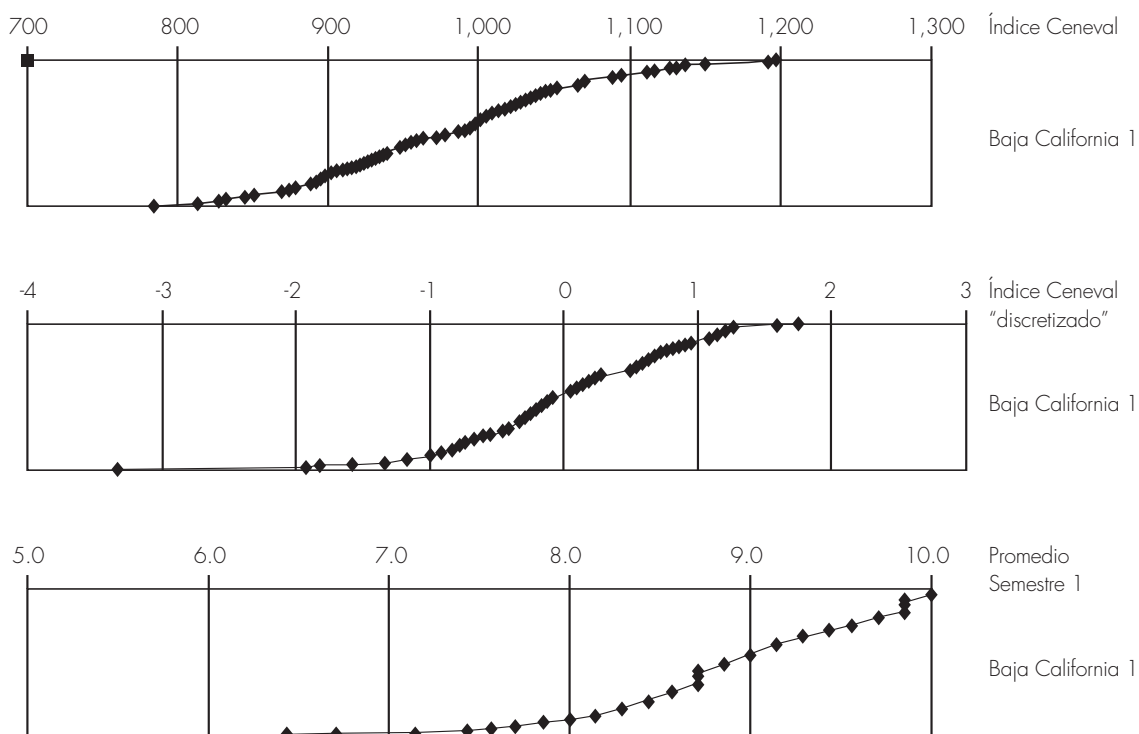
La pregunta de investigación es la misma de los modelos anteriores: ¿Son igualmente predictivos PROMSEC e ICNE en el desempeño escolar de los primeros semestres de la educación media superior?

3.b) Modelo

Se consideran solamente las dos variables explicativas en el primer nivel (estudiante). Al igual que en los modelos anteriores, los coeficientes en el segundo nivel no tienen asociada ninguna variable y se reportan los errores aleatorios en los niveles 1 y 2. La ecuación del modelo emplea la variable de baja densidad Z_2 para el Índice Ceneval.

Para hacer el cambio a la nueva variable se procede cortando los valores decimales y agrupándolos para formar aproximadamente 50 intervalos en lugar de 600, con lo cual se tiene una nueva variable Z_2 que presenta el Índice Ceneval en lo que pueden denominarse “intervalos discretos” (o Índice Ceneval “discretizado”), o variable de baja densidad. El cambio de variable a una variable de baja densidad produce un cambio notable en la distribución de los valores, a diferencia de lo que se tenía originalmente con el uso de la estandarización. Se hizo el procesamiento de los datos y se presenta aquí un ejemplo tomado de los datos de Baja California 1, que se representan en forma de diagrama de frecuencias acumuladas (puntaje en el eje X y sujetos en el eje Y) en la figura 4. Se aprecia claramente el cambio debido a la nueva densidad (60 puntos discretos), la cual ahora tiene una distribución muy similar a la que tienen las calificaciones escolares en la escala de 5 a 10 (teóricamente 40 a 50 puntos discretos, pero se aprecia que sólo se alcanzan a emplear menos de 25 en el promedio de semestre 1).

Figura 4. Ejemplo comparativo de las densidades de las distribuciones de los datos



Modelo del nivel 1:

$$P_SEM1 = B0 + B1*(ZPROMSEC) + B2*(Z_2) + R \quad [9]$$

Modelo del nivel 2:

$$B0 = G00 + U0 \quad [9bis]$$

$$B1 = G10$$

$$B2 = G20$$

3.c) Resultados

La diferencia de resultados entre las tablas 6 y 7 es producto del coeficiente de la variable de baja densidad Z_2 en lugar de ZCNE.

Tabla 7. Coeficientes, componentes de varianza y ajuste del tercer modelo

Confiabilidad del INTRCPT1, B0 0.960					
Efecto fijo	Coeficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	7.884844	0.121647	64.817	19	0.000
ZPROMSEC pend., B1					
INTRCPT2, G10	0.482496	0.033173	14.545	1370	0.000
Z_2 pend., B2					
INTRCPT2, G20	0.350878	0.031779	11.041	1370	0.000
Estimación final de componentes de varianza:					
Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.54579	0.29789	19	1018.16314	0.000
Residuo Nivel 1 R	0.70788	0.50110			
Deviance = 3030.088734					

3.d) Interpretación

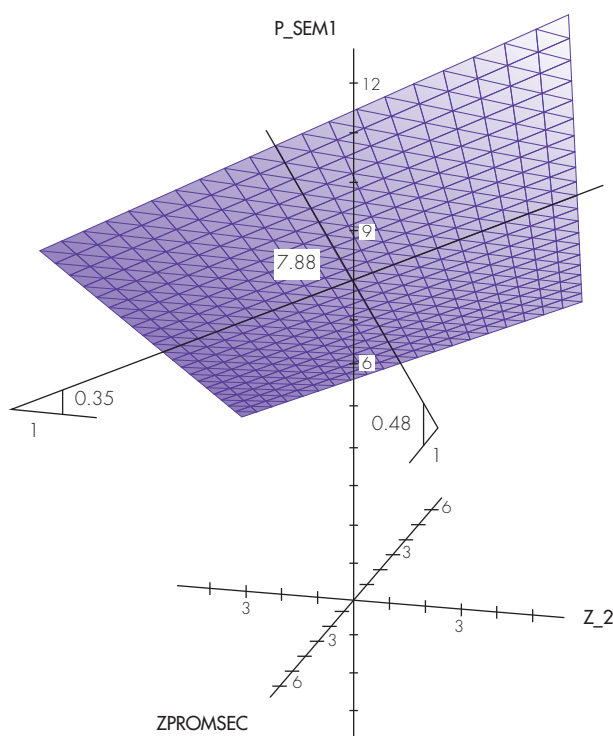
De nuevo se aprecia que la calidad del modelo no se ve alterada con los cambios de variables, porque se obtiene un coeficiente de desajuste (Deviance) muy similar a los anteriores, de 3030.0. Asimismo se observa que (como era de esperarse) al no haber un cambio en los datos del promedio se secundaria, su coeficiente es prácticamente el mismo del caso anterior (0.482). En cambio, el coeficiente referido al Índice Ceneval con “intervalos discretos” o de baja densidad se incrementa a 0.35, dando una proporción respecto de la calificación de secundaria mucho más comparable (1.37 veces). La significancia de los coeficientes y la de los componentes de varianza ya no se modifican respecto del modelo anterior ($t_{PROM_SEC}=14.5$, $t_{ICNE_1}=11.0$, $chi2U0=1018.1$).

La respuesta a la pregunta de investigación es que ambas variables explicativas son igualmente predictivas. Además se cuenta con la recomendación importante para futuros estudios de validez predictiva que consiste en estandarizar las calificaciones y los puntajes Ceneval, haciendo las transformaciones necesarias para tener una densidad similar entre las variables explicativas.

La interpretación de los valores estandarizados es sencilla al hacerse comparaciones en función de la media en 0 y la desviación estándar de 1. Por ejemplo, un estudiante que se ubica en la media de calificaciones en secundaria y en el EXANI-I (cuyos valores estandarizados son: $Z_{PROMSEC}=0$, $Z_2=0$) tiene una calificación esperada de 7.88 en el primer semestre. A su vez, un cambio de una desviación estándar en el valor normalizado del Índice Ceneval (o de la calificación de secundaria) produce un cambio de 0.35 unidades (o de 0.48, respectivamente) en la calificación del primer semestre de un estudiante. Fácilmente se puede obtener la diferencia entre sujetos extremos de la distribución, ya que estarían separados aproximadamente por 6 desviaciones estándar; es decir, la diferencia entre sujetos extremos de acuerdo con el EXANI-I sería de 2.1, en tanto que para el promedio de secundaria sería de 2.88 puntos, dando una diferencia acumulada de casi cinco puntos para los sujetos extremos.

El modelo creado hasta ahora es un plano en el espacio tridimensional y puede representarse cómodamente sin necesidad de recurrir a las habituales rectas de regresión que sólo se ubican en el plano cartesiano de dos dimensiones.

Figura 5. Plano del tercer modelo multinivel



La figura 5 muestra el plano del modelo, donde la variable dependiente se ubica en el eje vertical y las dos variables independientes corresponden con los ejes cartesianos en el plano horizontal. El intercepto es la cota donde se cruzan el eje vertical y el plano; las dos rectas de intersección del modelo con los planos cartesianos verticales del espacio modelado permiten ilustrar los coeficientes de cada variable independiente.

Los coeficientes de las variables en el modelo son las pendientes de las rectas de intersección con los planos verticales; dichas pendientes se indican sobre la misma figura, que se interpretan como el cambio en la variable dependiente debido a un incremento de una unidad en la variable independiente.³

4. Predictividad de otras variables

4.a) Caso

Una vez que se sabe que los resultados del EXANI-I son predictivos y comparables con los resultados de la secundaria, puede ser útil identificar otras variables que pudieran producir una validez incremental en el resultado del estudiante en el primer semestre de educación media superior. En este caso se va a aprovechar la base de datos que contiene las respuestas del cuestionario de contexto que explora aspectos socioeconómicos de los estudiantes.

La pregunta de investigación para este caso es: ¿Hay algunas variables socioeconómicas recabadas en el cuestionario de contexto que influyen en la predictividad respecto del PSEM1?

4.b) Modelo

El análisis de sensibilidad de las variables es de tipo exploratorio, por tratarse de determinar la validez incremental de cada una de ellas sobre la variable dependiente. Este estudio se vuelve complejo porque las combinaciones que pueden tenerse de variables contextuales son muy variadas. Por ello se optó por hacer el análisis siguiendo un procedimiento con dos etapas:

- 4.b.1) La primera etapa consiste en incluir explícitamente una de las variables, eligiéndose la variable SEXO, que es dicotómica, asumiendo que hay una asociación con el resultado que obtienen en las asignaturas del primer semestre.
- 4.b.2) La segunda etapa consiste en desarrollar un modelo exploratorio para variables de primer nivel, enriqueciendo el modelo con más variables independientes explicativas.
- 4.b.3) La tercera etapa se realiza por medio de un modelo exploratorio para variables de segundo nivel con ayuda de herramientas del software HLM.

³ Estas representaciones tridimensionales pueden hacerse en todos los modelos multinivel, pero cuando se tienen más de dos variables es necesario hacer “cortes” del hiper-plano en el espacio de n dimensiones, lo cual puede resultar confuso para el lector, por lo que no se acostumbra hacer este tipo de gráficas.

Primera etapa. Incluyendo variable sexo

Modelo del nivel 1:

$$P_SEM1 = B0 + B1*(SEXO) + B2*(ZPROMSEC) + B3*(Z_2) + R \quad [10]$$

Modelo del nivel 2:

$$B0 = G00 + U0 \quad [10bis]$$

$$B1 = G10$$

$$B2 = G20$$

$$B3 = G30$$

La variable SEXO es discreta, con los valores: 1= hombre y 2= mujer. Las variables ZPROMSEC y Z_2 fueron explicadas en los análisis previos. Este nuevo modelo debe mejorar la predictividad si el sexo funciona como variable explicativa.

4.c) Resultados

La tabla proporciona los resultados que obtiene el programa para el modelo propuesto.

Tabla 8. Coeficientes, componentes de varianza y ajuste del cuarto modelo

Confiabilidad del INTRCPT1, B0 0.961

Efecto fijo	Coefficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	7.671236	0.130125	58.953	19	0.000
SEXO pend., B1					
INTRCPT2, G10	0.140120	0.044832	3.125	1369	0.002
ZPROMSEC pend., B2					
INTRCPT2, G20	0.463381	0.034418	13.463	1369	0.000
Z_2 pend., B3					
INTRCPT2, G30	0.359043	0.031116	11.539	1369	0.000

Estimación final de componentes de varianza:

Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.54519	0.29723	19	1024.69022	0.000
Residuo Nivel 1 R	0.70503	0.49707			

Deviance = 3020.865686

4.d) Interpretación

Ahora se tiene un modelo mejor que los anteriores, con un desajuste (Deviance) de 3020.8. Varias constataciones se desprenden de estos resultados.

En primer lugar, el coeficiente de la variable SEXO es significativamente diferente de cero, pero que puede interpretarse como una diferencia pequeña entre hombres y mujeres, al traducirse en 0.14 unidades en la calificación media para el primer semestre. El valor es positivo, lo cual indica que las mujeres (SEXO=2) tienden a salir ligeramente mejor que los hombres (SEXO=1) en la calificación del primer semestre de la educación media superior. Si se sustituyen los datos de la variable dicotómica, la calificación promedio base para los hombres es 7.8 ($7.67 + 1 \times 0.14012$), en tanto que para las mujeres es 7.95 ($7.67 + 2 \times 0.14012$).

Una vez que se incluye SEXO como variable explicativa se observa que los coeficientes para las variables ZPROMSEC y Z_2 se modifican respecto del modelo anterior que no incluía SEXO. Si se comparan los nuevos valores con los anteriores, se tiene una diferencia relativa de casi 4% menos en ZPROMSEC al incluir la variable SEXO y un incremento un poco mayor a 2% en Z_2.

Tabla 9. Diferencia de coeficientes al intervenir la variable SEXO

Variable	Coeficiente del modelo (Error estándar)		Diferencia relativa (%)	t
	Sin incluir SEXO	Incluyendo SEXO		
ZPROMSEC	0.482496 (0.033173)	0.463381 (0.034418)	-3.96%	-0.40
Z_2	0.350978 (0.031779)	0.359043 (0.031116)	+2.29%	+0.18

Aunque los cambios son pequeños (3.96 y 2.29%) y no significativos estadísticamente (valores de t menores a 2), resulta interesante observar que la variable SEXO interviene en dirección diferente en ambas variables (una aumenta y otra disminuye en su coeficiente).

Segunda etapa exploratoria con variables de primer nivel

Modelo

Se tomó el modelo que incluye SEXO y se agregan otras variables del cuestionario de contexto, además de SEXO. Se incluyeron: Régimen de la escuela de procedencia (REG_PROC=Pública o privada), el tipo de secundaria (TIPO_SEC=Técnica, General, etc.), Ingreso familiar mensual (INGR_FAM=Menos de 1000, menos de 2000, etc.), Dominio del idioma inglés (DOM_ING=Nulo, básico, etc.).

Resultados

Se presentan los resultados con los coeficientes de las variables elegidas.

Tabla 10. Coeficientes del análisis exploratorio con variables del primer nivel

Confiabilidad del INTRCPT1, B0 0.962					
Efecto fijo	Coefficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	1.833200	0.404881	4.528	19	0.000
SEXO pend., B1					
INTRCPT2, G10	0.138164	0.039395	3.507	1365	0.001
REG_PROC pend., B2					
INTRCPT2, G20	0.170483	0.064478	2.644	1365	0.009
TIPO_SEC pend., B3					
INTRCPT2, G30	-0.054518	0.092508	-0.589	1365	0.555
PROM_SEC pend., B4					
INTRCPT2, G40	0.646345	0.041811	15.459	1365	0.000
INGR_FAM pend., B5					
INTRCPT2, G50	0.006603	0.004350	1.518	1365	0.129
DOM_ING pend., B6					
INTRCPT2, G60	0.034938	0.016500	2.117	1365	0.034
Z_2 pend., B7					
INTRCPT2, G70	0.342500	0.030526	11.220	1365	0.000

Interpretación

Se puede notar que algunas de las variables de primer nivel enriquecen el modelo predictivo. En el modelo propuesto son significativas: SEXO, REG_PROC, PROM_SEC, DOM_ING y Z_2; en cambio no son de interés significativo TIPO_SEC ni INGR_FAM ($p=0.555$ y $p=0.129$, respectivamente).

Hay una posible veta de exploración con las variables de primer nivel, lo cual se podrá realizar posteriormente para identificar todo el conjunto de elementos que puedan influir significativamente en los resultados del primer semestre en el nivel medio superior.

Tercera etapa exploratoria con variables de segundo nivel

Modelo

Hasta este momento no se han incluido variables de segundo nivel y los modelos sólo proporcionan la varianza en cada nivel, lo cual es una información interesante para juzgar la importancia de los anidamientos en la predictividad que se tiene sobre las calificaciones del primer

semestre, pero no proporciona elementos de referencia sobre otras variables explicativas que se asocien con el desempeño escolar. Se consideraron como variables independientes: SEXO, REG_PROC, TIPO_SEC, INGR_FAM, que ya se describieron antes, y también Expectativa de ingreso a institución en educación superior (INST_ING=Normal, técnica, UNAM, IPN, etc.), ocupación de padre y madre (OCU_PADR y OCU_MADR= Ninguna, hogar, campo, etc.) y si cuenta con teléfono (SER_TELE=1 si tiene el servicio).

Por lo anterior y con el objeto de identificar el impacto que tienen las variables explicativas de segundo nivel, se diseñó un modelo exploratorio con ayuda del software HLM. En un modelo exploratorio no se tiene que plantear un modelo específico, sino que se utiliza una herramienta del programa que incluye una a una las variables deseadas y calcula el efecto que produce en la variable dependiente. Debe tomarse en cuenta que al no contar con un modelo explícito, el valor obtenido con esta exploración puede diferir del valor real que se tendría al diseñar un modelo explícito para la variable. A pesar de esta limitación, el esquema exploratorio conduce de manera rápida y eficiente a buenas aproximaciones de los coeficientes.

Resultados

Para cada variable de segundo nivel se presenta el coeficiente que potencialmente se tendría en un modelo que la incluya explícitamente, así como el error estándar y el valor de t.

Tabla 11. Resultados del análisis exploratorio con variables del segundo nivel

Coeficientes de nivel 1	Predictores potenciales en el nivel 2					
	SEXO	REG_PROC	TIPO_SEC	INST_ING	INGR_FAM	OCU_PADR
Coeficiente	0.377	0.093	-0.012	0.067	-0.011	0.011
Error estándar	0.229	0.250	0.175	0.042	0.025	0.035
t valor	1.646	0.374	-0.069	1.581	-0.439	0.300
	OCU_MADR	SER_TELE				
Coeficiente	-0.019	-0.099				
Error estándar	0.029	0.408				
t valor	-0.659	-0.243				

Interpretación

La interpretación de los valores del modelo exploratorio es similar a la anteriormente presentada, es decir, para que una variable sea predictora del desempeño escolar en el primer semestre se necesita que su coeficiente sea significativamente diferente de cero, lo cual se puede juzgar si t es, en general y para fines prácticos, superior a 2.

Resulta interesante observar que ninguna de las variables socioeconómicas en el segundo nivel tiene influencia significativa en el resultado del estudiante. Queda claro, por las proporcio-

nes de los valores de t , que la variable SEXO (1.646) y la variable INSTITUCIÓN DE INGRESO (1.581) son las que tienen una mayor influencia, pero no proporcionan valores estadísticamente significativos.

En conclusión, para responder a la pregunta de investigación, se puede afirmar que sí se encontraron algunas variables de primer nivel (individual) que presentan una influencia significativa en los resultados del primer semestre, en particular SEXO, DOM_ING y REG_PROC. En cambio, no se encontraron variables de segundo nivel (institución) con intervención significativa en los resultados del primer semestre.

Dado que se trata de modelos exploratorios, sería necesaria una búsqueda exhaustiva de variables predictoras y sus combinaciones –lo que rebasa el propósito de este estudio–, pero puede ser planeado como un trabajo específico que justifique al cuestionario de contexto aplicado y que además conduzca a posibles modificaciones para incluir solamente las variables que sean predictoras en forma significativa.

5. Modelo exploratorio de validez retrospectiva del promedio de secundaria

5.a) Caso

Así como interesa la predictividad de la prueba objetiva que produce el Ceneval, en algunos casos se ha visto que las instituciones utilizan la prueba para corroborar el nivel de desempeño que tienen los alumnos al egreso de la secundaria. Esto obliga a pensar en un estudio de validez del EXANI-I con un resultado obtenido previamente, lo cual se interpreta como estudio de validez retrospectiva o de validez concurrente, teniendo como variable criterio el promedio de secundaria.

Junto con el resultado del EXANI-I se tiene el interés de identificar variables socioeconómicas que pueden intervenir en el promedio de secundaria.

La pregunta de investigación es: ¿Existe una relación concurrente o retrospectiva entre el promedio de secundaria y el puntaje Ceneval, y otras variables socioeconómicas, que explique su intervención como variables predictivas del PSEM1?

5.b) Modelo

Para este análisis se procedió con dos etapas:

- 5.b.1) En primer lugar se hizo un análisis exploratorio con el promedio de secundaria como variable dependiente y varias variables socioeconómicas como explicativas.
- 5.b.2) El modelo formalmente se construyó con las variables exploradas y el puntaje Ceneval Z_2 .

Primera etapa exploratoria con variables de primer nivel

Resultados

Interesa identificar variables socioeconómicas explicativas para el promedio de la secundaria (PROM_SEC), por lo que se incluyeron SEXO, REG_PROC, TIPO_SEC, INGR_FAM, OCU_PADR y OCU_MADR y DOM_ING, ya descritas antes.

Tabla 12. Resultados de primer análisis exploratorio de variables de primer nivel

Efecto fijo	Coficiente	Error estándar	Valor t	p-valor
INTRCPT1, B0 INTRCPT2, G00	7.955678	0.205602	38.694	0.000
SEXO pend., B1 INTRCPT2, G10	0.353618	0.050209	7.043	0.000
REG_PROC pend., B2 INTRCPT2, G20	-0.271847	0.073106	-3.719	0.000
TIPO_SEC pend., B3 INTRCPT2, G30	0.118437	0.050086	2.365	0.018
INGR_FAM pend., B4 INTRCPT2, G40	0.007106	0.007399	0.960	0.337
OCU_PADR pend., B5 INTRCPT2, G50	-0.005463	0.008436	-0.648	0.517
OCU_MADR pend., B6 INTRCPT2, G60	-0.005068	0.004424	-1.146	0.253
DOM_ING pend., B7 INTRCPT2, G70	0.084089	0.021363	3.936	0.000

Interpretación

En este caso el ingreso familiar y las ocupaciones de los padres parecen no tener una participación en el resultado que obtienen los estudiantes en el ciclo previo ($t=0.337$, 0.517 y 0.253), pero en cambio el sexo, el régimen de la escuela y el dominio del inglés sí se asocian de manera significativa.

En particular, el régimen de la escuela de procedencia (REG_PROC) es una variable dicotoma (1 = Pública, 2 = Privada) que está interviniendo con signo negativo, esto indica que las escuelas privadas tienen una media inferior en 0.27 unidades (en la escala 0-10) en el promedio de secundaria. El resultado es inesperado por otros análisis que se han realizado en México en los que las escuelas privadas obtienen valores más altos que los de las escuelas públicas. Esto es motivo de investigación ulterior para corroborar o refutar el hallazgo.

Este modelo exploratorio se trató de enriquecer al incluir el ingreso familiar como explicativo del promedio de secundaria, pero resultó ser una variable no significativa ($p=0.569$) y hace que el tipo de secundaria se ubique en una condición límite ($p=0.05$):

Tabla 13. Resultados de segundo análisis exploratorio de variables de primer nivel

Efecto fijo	Coficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	8.137729	0.194410	41.859	19	0.000
SEXO pend., B1					
INTRCPT2, G10	0.348633	0.050420	6.915	1366	0.000
REG_PROC pend., B2					
INTRCPT2, G20	-0.331233	0.064164	-5.162	1366	0.000
TIPO_SEC pend., B3					
INTRCPT2, G30	0.108381	0.055391	1.957	1366	0.050
INGR_FAM pend., B4					
INTRCPT2, G40	0.003636	0.006389	0.569	1366	0.569
DOM_ING pend., B5					
INTRCPT2, G50	0.042204	0.016295	2.590	1366	0.010
Z_2 pend., B6					
INTRCPT2, G60	0.328263	0.030454	10.779	1366	0.000

Segunda etapa, modelo con variables de primer nivel

Modelo

Aprovechando la identificación de variables explicativas procedentes del análisis exploratorio, se construyó un modelo explicativo para el promedio de secundaria, incluyendo las variables identificadas como significativas y se agregó la puntuación “discretizada” del Índice Ceneval (Z_2).

Modelo del nivel 1:

$$\text{PROM_SEC} = B0 + B1*(\text{SEXO}) + B2*(\text{REG_PROC}) + B3*(\text{TIPO_SEC}) + B4*(\text{DOM_ING}) + B5*(Z_2) + R \quad [11]$$

Modelo del nivel 2:

$$B0 = G00 + U0 \quad [11\text{bis}]$$

$$B1 = G10$$

$$B2 = G20$$

$$B3 = G30$$

$$B4 = G40$$

$$B5 = G50$$

Resultados

En la tabla de resultados del modelo se incluyen los coeficientes y los componentes de varianza.

Tabla 14. Coeficientes, componentes de varianza y ajuste del quinto modelo

Confiabilidad del INTRCPT1, B0 0.917					
Efecto fijo	Coeficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	8.158741	0.193338	42.199	19	0.000
SEXO pend., B1					
INTRCPT2, G10	0.346595	0.050521	6.860	1367	0.000
REG_PROC pend., B2					
INTRCPT2, G20	-0.326449	0.064320	-5.075	1367	0.000
TIPO_SEC pend., B3					
INTRCPT2, G30	0.108061	0.055556	1.945	1367	0.052
DOM_ING pend., B4					
INTRCPT2, G40	0.044958	0.017081	2.632	1367	0.009
Z_2 pend., B5					
INTRCPT2, G50	0.328691	0.030327	10.838	1367	0.000
Estimación final de componentes de varianza:					
Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	0.31882	0.10164	19	364.40442	0.000
Residuo Nivel 1 R	0.63688	0.40562			

Deviance = 2733.547459

Interpretación

Es notable la participación significativa de todas las variables y, junto con ello, que el Índice Ceneval esté asociado significativamente con el promedio de secundaria, en este caso se toma como variable explicativa pero de tipo “retrospectiva”, al intervenir en la relación de una variable ocurrida previamente en el tiempo. Con esto se responde a la pregunta de investigación en el sentido de que sí puede considerarse como relacionadas ambas calificaciones.

También se confirma la participación negativa del régimen de la escuela de procedencia, que redimensiona el intercepto o valor esperado de PROMSEC, con un coeficiente muy comparable al de SEXO y a Z_2, que es el valor del Índice Ceneval “discretizado”.

6. Modelo predictivo de las secciones del EXANI-I

6.a) Caso

En los modelos abordados se analizó la predictividad del EXANI-I para compararla contra el promedio de secundaria. Debe anotarse que ni el puntaje Ceneval de toda la prueba ni el promedio de secundaria son variables unidimensionales, porque engloban varias áreas del conocimiento, habilidades y ciertas cualidades académicas y actitudinales ante los diversos estímulos o situaciones en las cuales se evalúa al estudiante. De hecho, el promedio de secundaria involucra muchas más dimensiones que la prueba objetiva, ya que es el producto de las acciones académicas que realiza el estudiante a lo largo del tiempo, interactuando con los docentes y con otros compañeros, a lo cual se agregan componentes de tipo subjetivo muy difíciles de controlar o de comparar entre instituciones e, inclusive, entre docentes de una misma institución.

Para revisar el efecto de las dimensiones del EXANI-I en el promedio del primer semestre, se plantea la pregunta de investigación: ¿Qué tan predictivas son las secciones del EXANI-I (habilidades y conocimientos)?

6.b) Modelo

Se diseñó un modelo exploratorio con los resultados de las áreas o disciplinas reportadas con el EXANI, en lugar de trabajar con el Índice Ceneval que es un valor global.

Como ya se probó previamente, las variables pueden reportar influencia significativa en la ecuación multinivel, independientemente de que estén normalizadas o no, por ello se decidió hacer el análisis exploratorio con los datos de origen. El análisis se hizo con el promedio del primer semestre y se extendió hasta los del primer año.

6.c) Resultados

Los resultados obtenidos se presentan en las siguientes tablas (los casos significativos están sombreados). Los resultados del primer modelo se realizaron con el promedio de calificaciones del primer semestre, en tanto que para el segundo modelo se hicieron los mismos cálculos pero tomando el promedio de calificaciones del primer año (P_ANO1).

Interpretación

Además del sexo del estudiante (que interviene más favorablemente en las mujeres que en los hombres), sólo las dos habilidades (“Verbal” y “Matemática”), junto con “Formación cívica y ética” son predictivas hasta el primer año; además, el resultado de “Matemáticas” es predictivo hasta el primer semestre, aunque está muy cerca del intervalo de aceptación en el primer año ($p=0.057$). Igualmente, con un criterio un poco más relajado se podría pensar que la variable “Español” proporciona un resultado predictivo en el primer semestre.

Tabla 15. Resultados de primer modelo exploratorio de las secciones del EXANI-I

Efecto fijo	Coficiente	Errorestándar	Valor t	p-valor
INTRCPT1, B0				
INTRCPT2, G00	3.733334	0.404019	9.240	0.000
SEXO pend., B1				
INTRCPT2, G10	0.132684	0.046494	2.854	0.005
IHV_1 pend., B2				
INTRCPT2, G20	0.000485	0.000198	2.444	0.015
IESP_1 pend., B3				
INTRCPT2, G30	0.000488	0.000283	1.725	0.084
IHIS_1 pend., B4				
INTRCPT2, G40	0.000119	0.000257	0.463	0.643
IGEO_1 pend., B5				
INTRCPT2, G50	0.000243	0.000225	1.080	0.281
IFCE_1 pend., B6				
INTRCPT2, G60	0.000572	0.000251	2.277	0.023
IHM_1 pend., B7				
INTRCPT2, G70	0.000584	0.000153	3.820	0.000
IMAT_1 pend., B8				
INTRCPT2, G80	0.000471	0.000213	2.205	0.028
IFIS_1 pend., B9				
INTRCPT2, G90	0.000239	0.000224	1.065	0.288
IQUI_1 pend., B10				
INTRCPT2, G100	0.000450	0.000211	2.130	0.033
IBIO_1 pend., B11				
INTRCPT2, G110	0.000114	0.000151	0.756	0.450
ZPROMSEC pend., B12				
INTRCPT2, G120	0.455047	0.033279	13.674	0.000

Tabla 16. Resultados de segundo modelo exploratorio de las secciones del EXANI-I

Efecto fijo	Coficiente	Errorestándar	Valor t	p-valor
INTRCPT1, B0				
INTRCPT2, G00	3.430962	0.406073	8.449	0.000
SEXO pend., B1				
INTRCPT2, G10	0.161021	0.047577	3.384	0.001
IHV_1 pend., B2				
INTRCPT2, G20	0.000476	0.000217	2.198	0.028
IESP_1 pend., B3				
INTRCPT2, G30	0.000383	0.000252	1.517	0.129
IHIS_1 pend., B4				
INTRCPT2, G40	0.000156	0.000252	0.620	0.535
IGEO_1 pend., B5				
INTRCPT2, G50	0.000277	0.000194	1.423	0.155
IFCE_1 pend., B6				
INTRCPT2, G60	0.000628	0.000262	2.398	0.017
IHM_1 pend., B7				
INTRCPT2, G70	0.000807	0.000199	4.052	0.000
IMAT_1 pend., B8				
INTRCPT2, G80	0.000430	0.000226	1.904	0.057
IFIS_1 pend., B9				
INTRCPT2, G90	0.000241	0.000216	1.116	0.265
IQUI_1 pend., B10				
INTRCPT2, G100	0.000465	0.000215	2.162	0.031
IBIO_1 pend., B11				
INTRCPT2, G110	0.000126	0.000135	0.934	0.351
ZPROMSEC pend., B12				
INTRCPT2, G120	0.464889	0.034415	13.508	0.000

Análisis complementario

Para decidir qué tan asociado está el resultado general del EXANI con las variables medidas, se construyó un modelo exploratorio complementario para el Índice Ceneval como variable dependiente y las diferentes secciones del EXANI-I como variables explicativas independientes.

Con este modelo se obtienen estos resultados que (como era de esperarse) indican participaciones estadísticamente significativas:

Tabla 17. Coeficientes, componentes de varianza y ajuste del sexto modelo multinivel

Efecto fijo	Coefficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	0.145178	0.087525	1.659	19	0.113
IHV_1 pend., B1					
INTRCPT2, G10	0.125061	0.000078	1602.452	1362	0.000
IESP_1 pend., B2					
INTRCPT2, G20	0.093560	0.000092	1011.969	1362	0.000
IHIS_1 pend., B3					
INTRCPT2, G30	0.093853	0.000090	1041.477	1362	0.000
IGEO_1 pend., B4					
INTRCPT2, G40	0.093621	0.000102	919.157	1362	0.000
IFCE_1 pend., B5					
INTRCPT2, G50	0.093695	0.000101	923.726	1362	0.000
IHM_1 pend., B6					
INTRCPT2, G60	0.125285	0.000060	2081.686	1362	0.000
IMAT_1 pend., B7					
INTRCPT2, G70	0.093616	0.000096	974.217	1362	0.000
IFIS_1 pend., B8					
INTRCPT2, G80	0.093751	0.000122	765.861	1362	0.000
IQUI_1 pend., B9					
INTRCPT2, G90	0.093622	0.000068	1370.866	1362	0.000
IBIO_1 pend., B10					
INTRCPT2, G100	0.093791	0.000069	1363.747	1362	0.000

Puede verse que el único valor no significativo es el intercepto B0, con un valor $p=0.113$, lo cual corrobora que la media de puntajes Ceneval estandarizados se ubica en el valor central 0.

En cambio, algo interesante es que se pueden distinguir dos grupos de variables por el peso que tienen los coeficientes en el modelo. Por un lado se tienen dos variables que intervienen con coeficientes del orden de 0.125 y que son las dos “habilidades” ($IHV=0.1250$, $IHM=0.1252$, que aparecen sombreadas en la tabla). Por otro lado se tiene el grupo de variables que intervienen con coeficientes del orden de 0.09, todas corresponden con “variables disciplinares”. La presencia de estos dos grupos resulta interesante como una veta de análisis de la dimensionalidad de la prueba.

En conclusión se puede afirmar que, como era esperado por el diseño, las secciones del EXANI-I son predictivas del resultado global de la prueba.

De cualquier modo, el hecho de que todas las variables participen de manera significativa en el Índice Ceneval tiene implicaciones interesantes para estudiar la dimensionalidad del EXANI-I. Además, debe anotarse que la presencia combinada de variables explicativas altamente correlacionadas es inconveniente en los modelos multinivel, pudiendo indicar variables de un mismo componente factorial⁴ (véase Tristán *et al.*, 2008, p. 56 y ss.). Es motivo de interés realizar posteriormente los análisis factoriales de los EXANI, no solamente para confirmar la dimensionalidad de las pruebas, sino para verificar la calidad de las agrupaciones.

7. Relación funcional entre variables socioeconómicas

7.a) Caso

Un análisis de interés para este estudio -y para otros que pudieran plantearse en el futuro con los datos de la base del “Proyecto 12x2”- se enfoca en la asociación que tienen los datos socioeconómicos entre sí. Estas variables están referidas a los datos de contexto y resulta atractivo compararlas contra los indicadores socioeconómicos desarrollados en otros proyectos, como los de la OCDE para el proyecto PISA (véase OCDE, 2003-2005) o los diseñados por el INEE (véase Robles y Martínez, 2006); pero esto queda fuera del propósito del presente estudio y puede dar lugar a un desarrollo posterior.

La pregunta de investigación es: ¿Existen posibles relaciones funcionales entre las variables socioeconómicas procedentes del cuestionario de contexto?

7.b) Modelo

En este caso se desarrolló solamente un modelo para reportar el INGRESO FAMILIAR como variable dependiente en función de estas variables:

ESCO_PAD= escolaridad del padre	SER_TELE= servicio de teléfono en casa
ESCO_MAD= escolaridad de la madre	SER_PC= dispone de computadora personal en casa
OCU_PADR= ocupación del padre	SER_INTE= servicio de internet en casa
OCU_MADR= ocupación de la madre	DOM_ING= dominio del inglés

⁴ Vale la pena comentar que en la base de datos del “Proyecto 12x2” se incluyen variables que engloban las habilidades y los conocimientos (IHIB, ICL, etc.). Con esta base se diseñó un modelo con la siguiente expresión:

$$Y = B_0 + B_1*(IHV_1) + B_2*(IESP_1) + B_3*(IHIS_1) + B_4*(IGEO_1) + B_5*(IFCE_1) + B_6*(IHM_1) + B_7*(IMAT_1) + B_8*(IFIS_1) + B_9*(IQUI_1) + B_{10}*(IBIO_1) + B_{11}*(IHIB_1) + B_{12}*(ICL_1) + B_{13}*(ILM_1) + R$$

Este modelo pierde la convergencia, debido a la presencia de variables colineales (por ejemplo habilidad verbal y matemática y la variable global IHIB)

Modelo del nivel 1:

$$\begin{aligned} \text{INGR_FAM} = & B0 + B1*(\text{ESCO_PAD}) + B2*(\text{ESCO_MAD}) \\ & + B3*(\text{OCU_PADR}) + B4*(\text{OCU_MADR}) + B5*(\text{SER_TELE}) \\ & + B6*(\text{SER_PC}) + B7*(\text{SER_INTE}) + B8*(\text{DOM_ING}) + R \end{aligned} \quad [12]$$

Modelo del nivel 2:

$$\begin{aligned} B0 = & G00 + U0 && [12\text{bis}] \\ B1 = & G10 \\ B2 = & G20 \\ B3 = & G30 \\ B4 = & G40 \\ B5 = & G50 \\ B6 = & G60 \\ B7 = & G70 \\ B8 = & G80 \end{aligned}$$

7.c) Resultados

Tabla 18. Coeficientes, componentes de varianza y ajuste del séptimo modelo multinivel

Confiabilidad del INTRCPT1, B0 0.961					
Efecto fijo	Coeficiente	Error estándar	Valor t	Aprox. gdl	p-valor
INTRCPT1, B0					
INTRCPT2, G00	2.800558	0.802247	3.491	19	0.003
ESCO_PAD pend., B1					
INTRCPT2, G10	0.249474	0.042953	5.808	1364	0.000
ESCO_MAD pend., B2					
INTRCPT2, G20	0.126146	0.047306	2.667	1364	0.008
OCU_PADR pend., B3					
INTRCPT2, G30	0.015685	0.028655	0.547	1364	0.584
OCU_MADR pend., B4					
INTRCPT2, G40	-0.006019	0.028841	-0.209	1364	0.835
SER_TELE pend., B5					
INTRCPT2, G50	0.446028	0.419054	1.064	1364	0.288
SER_PC pend., B6					
INTRCPT2, G60	0.992028	0.218756	4.535	1364	0.000
SER_INTE pend., B7					
INTRCPT2, G70	1.152593	0.250852	4.595	1364	0.000
DOM_ING pend., B8					
INTRCPT2, G80	0.421473	0.088200	4.779	1364	0.000

Estimación final de componentes de varianza:					
Efecto aleatorio	Desviación estándar	Componente de varianza	gdl	chi2	p-valor
INTRCPT1, U0	2.33947	5.47313	19	757.38619	0.000
Residuo Nivel 1 R	3.01405	9.08451			

Deviance = 7014.214655

7.d) Interpretación

El modelo planteado permite afirmar que el INGRESO FAMILIAR está asociado con las variables explicativas propuestas, excepto las OCUPACIONES DE LOS PADRES ($p=0.584$ y $p=0.835$) y el SERVICIO DE TELÉFONO EN CASA ($p=0.288$); estas variables pueden eliminarse del modelo y deberán trabajarse independientemente para otros fines dentro del estudio.

¿Qué utilidad puede tener esta relación? Una de sus aplicaciones es que puede conducir a modelos más parsimoniosos, en lugar de emplear la escolaridad del padre o de la madre como explicativa del resultado del estudiante, el modelo indica que es preferible emplear el INGRESO FAMILIAR, que proporciona un valor que engloba a las otras variables de tipo socioeconómico.

La relación propuesta por este modelo es muy útil porque permite afirmar que basta con incluir el ingreso familiar en un modelo para reportar la intervención del conjunto de variables explicativas que fueron significativas en este modelo. Adicionalmente, el modelo indica que sería erróneo incluir el ingreso familiar y la escolaridad del padre como variables explicativas en forma simultánea, ya que una “engloba” a la otra, lo cual influiría en asociaciones ficticias, eventualmente sobredimensionadas y pueden hacer que el modelo no sea convergente o lo haga de manera deficiente.

En conclusión, puede responderse a la pregunta de investigación afirmando que se tienen relaciones funcionales entre las variables y esto da lugar a una posible línea de trabajo para mejorar y depurar el cuestionario de contexto.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES



CONCLUSIONES •

1. Conclusiones sobre los datos utilizados

Con la ayuda de los modelos multinivel desarrollados para este estudio, a partir de los datos disponibles del “Proyecto 12x2”, se obtiene un resultado de gran importancia en relación con la predictividad, ya que intervienen de manera muy similar en el promedio del primer semestre de educación media superior (PSEM1), tanto el promedio de secundaria (PROMSEC) como el puntaje en la prueba objetiva del Ceneval (ICNE).

No se trata de desmentir que las calificaciones del nivel de secundaria son más predictivas que los puntajes de la prueba objetiva, sino de dimensionar el concepto de ser “más predictivas”. Ya se mostró que –de acuerdo con los modelos multinivel– tanto PROMSEC como ICNE son significativamente “no nulas” y, en consecuencia, son igualmente explicativas desde el punto de vista estadístico.

Los coeficientes de ambas variables, para los datos disponibles, difieren tan sólo en 37% ($1.37=0.4824/0.3508$), lo cual implica que la calificación de una prueba objetiva no representa valores despreciables para predecir el resultado en el primer semestre. La nube de dispersión se puede comparar claramente entre los valores observados, pues son muy similares al ojo del observador y comparables de manera objetiva con el coeficiente de determinación ($R^2=0.26$ y 0.19 , respectivamente).

La correlación debe ser mejor entre calificaciones de procesos comparables, es decir: deben tenerse mayores correlaciones entre las calificaciones de secundaria y las calificaciones de bachillerato que entre los puntajes de una prueba objetiva y calificaciones de bachillerato. Esta diferencia de correlaciones no es una razón para desechar las pruebas objetivas; por lo contrario, se deben reconocer las múltiples ventajas de tener una correlación comparable a partir de un solo evento de evaluación, cuya duración y condiciones de aplicación son muy restringidas por cuestiones de logística, precio, facilidades de administración, tiempo de calificación, entre otros múltiples aspectos.

Una variable de interés en la predictividad es el sexo del estudiante, por lo menos con un coeficiente de 0.14 puntos a favor de las mujeres.

También se obtuvo un resultado importante en relación con la validez retrospectiva del puntaje obtenido con el EXANI-I y el promedio de secundaria. El modelo incluye, además del sexo, el régimen de la escuela de procedencia y el dominio del idioma inglés. Al hacer este análisis los valores p de los coeficientes de las variables “ocupación de los padres” e “ingreso familiar” indican que no influyen en un cambio significativo en las calificaciones del estudiante.

Otra agrupación que mostró una relación muy interesante es el Índice Ceneval con las variables que se miden en la prueba (habilidades y disciplinas o asignaturas), las cuales pueden ser consideradas como sub-escalas (Darakjian *et al.*, 1984). Resulta muy atractivo observar que los coeficientes de las habilidades son similares entre sí (0.125) y, por otro lado, las variables asocia-

das con contenidos obtienen coeficientes similares entre sí (del orden de 0.09). Esto indica una posible línea de investigación en el futuro, verificando la dimensionalidad de la prueba por métodos factoriales exploratorios o por lo menos confirmatorios. La significancia de las variables es un indicio, así sea en primera instancia, para considerar que se trata de variables no correlacionadas entre sí, pero que pueden definir dimensiones para integrar, en su conjunto, la variable global denominada Índice Ceneval.

Una relación no menos interesante es la de las variables socioeconómicas de la base de datos. Sin ser definitivo, porque solamente se dispone de los datos de la base del “Proyecto 12x2”, se pudo construir un modelo de ingreso familiar, que depende de algunas variables explicativas (escolaridad de los padres, el servicio de Internet y, notablemente, el conocimiento del idioma inglés). Este modelo debe ser corroborado en el futuro al disponerse de otros datos.

2. Conclusiones generales sobre el uso de modelos de predictividad

Generalmente se reportan los análisis de predictividad con base en correlaciones; en particular se reportan los valores de r de Pearson, que permiten comparar lo que Willms (2006) denomina “fuerza” de la asociatividad entre variables.

Una vez que se tienen las correlaciones, se acostumbra hacer comparaciones entre los valores de r de diferentes variables, de tal modo que –para los resultados obtenidos en este estudio– normalmente se concluiría que la correlación entre PROMSEC y PSEM1 es mejor que la correlación entre ICNE y PSEM1 (0.51 y 0.44, respectivamente). Esta conclusión sería muy limitada, ya que si “ambas son estadísticamente significativas”, carece de sentido interpretar a una como “más significativa” que la otra. Resulta más apropiado y útil establecer la relación de asociatividad a partir de los coeficientes de la variable en el modelo multinivel, una vez garantizada su significancia. Por consecuencia, es preferible juzgar la predictividad de las variables por la combinación de los coeficientes en el modelo multinivel y la significancia estadística proporcionada por el mismo modelo.

La conveniencia de utilizar modelos multinivel en estudios de predictividad está seguramente en desarrollo, porque no han sido explotados más que en muy pocos estudios. Para conjuntos de datos anidados, como ocurre en el presente trabajo, el uso de modelos multinivel es el más apropiado.

Los modelos multinivel proporcionan los coeficientes de las variables explicativas, así como su significancia en relación con la variable dependiente. Se trata de un cálculo que toma en cuenta los anidamientos y reporta las varianzas intra-grupos e inter-grupos, favoreciendo la interpretación de diversos comportamientos de las variables latentes.

El uso de los modelos multinivel, además de sus ventajas inherentes, favoreció la identificación de algunos detalles finos, que hay que considerar en estudios ulteriores. Se trata de aspec-

tos implícitos en la metodología, similares a otros que se encuentran en los análisis de regresión, pero que no se toman en cuenta en muchas ocasiones. En particular, debe atenderse la necesidad de realizar cambios de variable para hacer comparables los coeficientes del modelo multinivel. Se mostró que las transformaciones de variable no perjudican la calidad del modelo (no se alteran los coeficientes de desajuste o “Deviance”) ni la significancia estadística del coeficiente, pero sí su valor absoluto por estar ubicada en la escala de la propia variable.

Se pudo observar que hay un mejor comportamiento en la convergencia y ajuste de los modelos multinivel cuando todas las variables explicativas tienen escalas similares en tres cualidades: origen, rango y densidad. En particular pudo apreciarse que la similitud de densidad de las variables mejora la predictividad que se reporta con el coeficiente del modelo multinivel.

En especial, el uso de variables normalizadas o estandarizadas (media 0 y desviación estándar 1), facilita mucho la comparación entre variables.

3. Extensiones del estudio

Sobre la base de las conclusiones ya presentadas, se pueden sugerir algunas extensiones del presente trabajo.

- a) Revisar la dimensionalidad de las variables contenidas en el EXANI-I, por modelos factoriales. Este tema, ya estudiado por diversos investigadores, es un imperativo de la validez inherente de la prueba. La determinación factorial o su revisión por modelos confirmatorios permitirá que los usuarios definan las mejores combinaciones de variables para sus propósitos de admisión.
- b) Completar las muestras y así corroborar los aspectos descubiertos hasta este momento. Se sabe que la recopilación de las muestras es un trabajo muy complejo y largo, pero conviene al Ceneval entablar este tipo de proyectos para garantizar a los usuarios la calidad de sus pruebas.
- c) Continuar el análisis longitudinal de predictividad del EXANI-I conforme avanzan los estudiantes por la educación media superior. Este estudio se debe complementar con la revisión de las posibles prácticas de aprendizaje para la prueba y el entrenamiento que los estudiantes van recibiendo con las aplicaciones sistemáticas año con año. En la bibliografía se encuentran referencias de estudios en este sentido, porque es un motivo de preocupación que la validez no se vea afectada por un sobre-entrenamiento hacia la prueba.
- d) Realizar los análisis de ganancia o de valor agregado, que es un concepto diferente al del entrenamiento para la prueba. Es importante identificar que la escuela aporta una preparación en términos de competencia, para lo cual se deben desarrollar modelos que distingan este tipo de rasgo. Este aspecto es un tópico de mucho interés para un estudio con los datos de la base, pero también representa un área de exploración o un tema de tesis futura.
- e) Comparar los resultados de EXANI-I y EXANI-II para establecer un análisis diferencial entre ambas pruebas, recopilando los modelos que se puedan desarrollar en los puntos anteriores.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

REFERENCIAS •

- Allison P. (2002). *Missing Data*. SAGE University paper, núm. 136. Series: quantitative applications in the social sciences. California: Sage University. pp. 9-10.
- American Psychological Association (1994). *Publication manual of the American Psychological Association*, 4a ed. Washington D.C.: APA. p. 11.
- American Psychological Association (2001). *Publication manual of the American Psychological Association*, 5a ed. Washington D.C.: APA.
- Anastasi, A. (1985). "Some emerging trends in psychological measurement: a fifty-year perspective". *Applied Psychological Measurement*; 9(2)121-138.
- Arce, D. et al. (2008). *Estudio de validez predictiva "12 x 2"*. Área de estudios y análisis estadísticos, DGA-EXANI. Documento interno Ceneval, abril 2008; 46 p.
- Bell, R. y Lumsden, J. (1980). "Test length and validity". *Applied Psychological Measurement* 4(2)165-170.
- Bryk, S.A. y Raudenbush, S.W. (1992). *Hierarchical linear models. Applications and data analysis methods*. London: Sage Publications. p. 265.
- Bryk, S., Raudenbush, A. y Congdon, R. (2005). *HLM Versión 6.03*. User's Guide to the Software. London: Sage Publications.
- Cohen, J. y Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2a. ed.) Hillsdale, NJ: Erlbaum.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2a. ed.) Hillsdale, NJ: Lawrence Erlbaum.
- Darakjian, G.P. et al. (1984). "The predictive validity of subscales of fan academia self-concept measure administered nine semestres prior to acquisition of criterion data reflecting school achievement". *Educational and Psychological Measurement*. 44(3)175.
- Díaz G.M.A. et al. (2007) *PISA 2006 en México*. México: INEE. p. 343.
- Downey, R.G. (1979). *Item-Option Weighting of Achievement Tests: Comparative Study of Methods*. *Applied Psychological Measurement*; 3(4)453-461.
- Fan, X. y Wang, L. (1995). *How comparable are the jackknife and the bootstrap results: an investigation for a case of canonical correlation analysis*. Annual Meeting of the American Educational Research Association. 31p. ERIC DOCUMENT: ED387509.
- Gaviria, J. L. et al. (2004). *Un estudio multinivel sobre los factores de eficacia escolar en países en desarrollo. El caso de los recursos en Brasil*. Education Policy Analysis Archives, 12 (20) Recuperado el 10/18/2006 de <http://epaa.asu.edu/epaa/v12n20/>
- Goldstein, H. (1999). *Multilevel statistical models*. London: Edward Arnold. 163 p.
- Hale, G.A. et al. (1992). *A comparison of the predictive validity of the current SAT and an experimental prototype*. Research report. Educational Testing Service, Princeton, N.J. 61 p.

- Hararin, E.C. *et al.* (2007). *Predictive and Incremental Validity of Global and Domain-Based Adolescent Life Satisfaction Reports*. *Journal of Psychoeducational Assessment* 25(2)127-138.
- Holden, G. *et al.* (2006). An Assessment of the Predictive Validity of Impact Factor Scores: Implications for Academic Employment Decisions in Social Work. *Research on Social Work Practice* 16(6) 613-624.
- Hox, J.J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties., p. 118.
- Kobrin, J.L. *et al.* (2008). *Validity of the SAT for predicting First-Year College Grade Point Average*. Research report No.2008-5. The College Board. New York. 16 p.
- Kreft, I. y De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage Publications.
- Kuncel, N.R. *et al.* (2005). "The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature". *Review of Educational Research*. 75(1)63-82.
- Lamp, R.E. y Krohn, E.J. (2001). *A Longitudinal Predictive Validity Investigation of the Sb:Fe and K-Abc With At-Risk Children*. *Journal of Psychoeducational Assessment*. 19(4)334-349.
- Little, R.J.A. y Rubin, D.B. (2002). *Statistical analysis with missing data*. Hoboken, New Jersey, Wiley.
- Martínez, R.F. *et al.* (2007). *Propuesta y experiencias para desarrollar un Sistema Nacional de Indicadores Educativos*. Instituto Nacional para la Evaluación de la Educación. México: INEE. p. 280.
- Mattern, K.D. *et al.* (2008). *Differential validity and prediction of the SAT*. Research report No.2008-4. The College Board. New York. 18 p.
- McKnight, P.E. *et al.* (2007). *Missing data. A gentle introduction*, pp. 169-170.
- Moreira M., T. (2006). *La validez predictiva de las pruebas de bachillerato en educación media. Actualidades en Psicología*, 20, 121-140 Costa Rica: Departamento de Supervisión Nacional. (<http://www.dcc.mep.go.cr/validezpredictiva.html>) Consultado el 3 de marzo de 2008.
- Morrison, D.F. (1967). *Multivariate statistical methods*. New York: McGraw-Hill Book Co. p. 338.
- Murray, C. (2007). "Abolish the SAT". *The American, A Magazine of Ideas*. July/August 2007 Issue. Disponible en Internet:<http://american.com/archive/2007/july-august-magazine-contents/abolish-the-sat>
- Musayón O., F. (Julio-Diciembre, 2001). "Relación entre el puntaje de Ingreso y el Rendimiento Académico en el Segundo año de las alumnas de Enfermería ingresantes entre los años 1994-1997 en una Universidad Peruana". *Revista Universidades*, núm. 22. Perú: Centro de Información y Documentación Universitaria (CIDU) de la Unión de Universidades de América Latina y El Caribe (UDUAL) (<http://www.udual.org/CIDU/Revista/22/RelaEnfermeria.htm>) Consultado el 20 de febrero de 2008.
- Nunnally, J.C. y Bernstein, I.J. (1995). *Teoría psicométrica*. 3a ed. México: McGraw Hill. p. 843.
- OECD-PISA (2003). *PISA 2003 Technical Report*. Organization for Economic Cooperation and Development, p. 114.
- OECD-PISA (2005). *PISA 2003 Data analysis Manual: SPSS users*. Cap. 3 "Replicate weights". Organization for Economic Cooperation and Development. p. 32-52.

- Powers, D.E. (1985). "Effects of Test Preparation on the Validity of a of a Graduate Admissions Test". *Applied Psychological Measurement* 9(2) 179-190.
- Raudenbusch, S. *et al.* (2005) *HLM for windows*, V.6.0.3. User manual help.
- Robles, V.H.V. y Martínez, R.F. (2006). *Panorama educativo de México 2006. Indicadores del Sistema Educativo Nacional*. México: Instituto Nacional para la Evaluación de la Educación, p. 605.
- Rubin, D.B. (1976). "Inference and missing data". *Biometrika* (63), pp. 581-592.
- Sadler, P. y Tai, R.H. (2007). "Weighting for Recognition: Accounting for Advanced Placement and Honors Courses When Calculating High School Grade Point Average". *NASSP Bulletin* 91(1)5-32
- Schmidt A.E. (2000). "An approximation of a hierarchical logistic regresión model used to establish the predictive validity of scores on a nursing licensure exam". *Educational and Psychological Measurement* 60(3)463.
- Serrano García, J.M. *et al.* (2002). "Validez del Examen Nacional de Ingreso a la Educación Superior (EXANI-II) en la Facultad de Ciencias de la Conducta de la UAEM". Mesa de trabajo 1: Evaluación de la Educación Superior. En: *Memorias del Tercer Congreso Nacional y Segundo Congreso Internacional Retos y Expectativas de la Universidad*. México: UAEM. (<http://www.congresoretosyexpectativas.udg.mx/FrameSetmemorias.htm>) Revisado el 10 de enero de 2008.
- Talento-Miller, E. y Rudner, L.M. (2008). "The Validity of Graduate Management Admission Test Scores: A Summary of Studies Conducted From 1997 to 2004". *Educational and Psychological Measurement* 68(1)129-138.
- Thompson, B. (1996). AERA "Editorial policies regarding statistical significance testing: three suggested reforms". *Educational Researcher*, vol. 25, núm. 2, pp. 26-30.
- Tirado, F. *et al.* (1997). "Validez predictiva del Examen de Habilidades y Conocimientos Básicos (EXHCOBA)". *Revista Mexicana de Investigación Educativa*, enero-junio 1997, 2, 3, pp. 67-84.
- Tristán, L.A. y Vidal, U.R. (2006). *Manual de fórmulas de correlación*. Trafford. Victoria, Canadá.
- Tristán, L.A. y Vidal, U.R. (2007). *Linear model to assess the scale's validity of a test*. AERA Meeting 2007, New Developments in Measurement Thinking", SIG-Rasch Measurement) Disponible en ERIC: DOCUMENT: ED501232
- Tristán, L.A., Flores-Cervantes, I., Díaz, G.M.A., Flores V.G., Solis, G.R., Canales, S.D., Morelos, M.P. De la Cruz H.Y. (2008) *Análisis multinivel de la calidad educativa en México ante los datos de PISA 2006*. INEE, 188 p.
- Weber, J. (1985). *Assessment and Placement: A Review of the Research*. Community College Review 13(3)21-32.
- Willms J.D. y Smith T. (2005). *A Manual for Conducting Analyses with Data from TIMSS and PISA*. pp. 9-10.
- Willms, D. (2006). *Las brechas de aprendizaje: diez preguntas de la política educativa a seguir en relación con el desempeño y la equidad en las escuelas y los sistemas educativos*. Instituto de Estadística de la UNESCO. Montreal, 95 p.

ANÁLISIS MULTINIVEL
DE PREDICTIVIDAD DEL EXAMEN
INVESTIGACIONES 1

ANEXOS •

A1. Resumen estadístico de las variables del “Proyecto 12x2”

a) Estadísticas descriptivas del nivel 1

Nombre de variable	Descripción	N	Media	Desviación estándar	Mínimo	Máximo
SEXO	Sexo	1373	1.55	0.50	1.00	2.00
REG_PROC	Régimen de la escuela de procedencia	1373	1.36	0.48	1.00	2.00
TIPO_SEC	Tipo de secundaria	1373	1.89	0.45	1.00	5.00
PROM_SEC	¿Qué promedio tienes actualmente en secundaria?	1373	8.65	0.80	6.00	10.00
INST_ING	¿A qué institución de educación superior te gustaría ingresar al terminar tus estudios de educación media superior?	1373	5.23	2.85	1.00	11.00
INGR_FAM	¿Cuál es el ingreso familiar mensual?	1373	8.21	4.78	1.00	15.00
ESCO_PAD	Escolaridad del padre	1373	7.74	2.68	1.00	11.00
ESCO_MAD	Escolaridad de la madre	1373	7.17	2.68	1.00	11.00
OCU_PADR	Ocupación del padre	1373	9.32	3.22	1.00	14.00
OCU_MADR	Ocupación de la madre	1373	6.29	4.18	1.00	14.00
SER_TELE	¿Cuentas en tu casa con teléfono?	1373	0.90	0.30	0.00	1.00
SER_PC	¿Cuentas en tu casa con computadora?	1373	0.80	0.40	0.00	1.00
SER_INTE	¿Cuentas en tu casa con conexión a internet?	1373	0.55	0.50	0.00	1.00
DOM_ING	Dominio del Inglés	1373	3.00	1.23	1.00	5.00
ICNE_1	Índice Ceneval global	1373	1053.96	97.57	780.00	1258.00
PERCEN_1	Percentil	1373	52.43	28.55	0.81	100.00
NCNE_1	Número de aciertos global	1373	75.52	20.82	17.00	119.00
PCNE_1	Porcentaje de aciertos global	1373	58.99	16.27	13.28	92.96
IHV_1	Habilidad verbal (índice)	1373	1099.87	121.61	700.00	1300.00
IESP_1	Español (índice)	1373	1025.78	123.84	700.00	1300.00
IHIS_1	Historia (índice)	1373	1038.49	116.16	700.00	1300.00
IGEO_1	Geografía (índice)	1373	1033.61	109.30	700.00	1300.00
IFCE_1	Formación cívica y ética (índice)	1373	1089.44	127.48	700.00	1300.00
IHM_1	Habilidad matemática (índice)	1373	1069.06	140.45	738.00	1300.00
IMAT_1	Matemáticas (índice)	1373	1075.38	143.45	700.00	1300.00
IFIS_1	Física (índice)	1373	1018.14	133.82	700.00	1300.00

a) Estadísticas descriptivas del nivel 1

Nombre de variable	Descripción	N	Media	Desviación estándar	Mínimo	Máximo
IQUI_1	Química (índice)	1373	1020.79	134.44	700.00	1300.00
IBIO_1	Biología (índice)	1373	1048.62	142.58	700.00	1300.00
IHIB_1	Habilidades intelectuales básicas (índice)	1373	1084.52	114.96	719.00	1300.00
ICL_1	Comunicación y lenguaje (índice)	1373	1068.11	108.40	743.00	1300.00
ILM_1	Lenguaje matemático (índice)	1373	1071.75	130.49	721.00	1300.00
ICD_1	Conocimientos disciplinarios (índice)	1373	1043.69	97.13	769.00	1256.00
PHV_1	Habilidad verbal (% aciertos)	1373	66.67	20.30	0.00	100.00
PESP_1	Español (% aciertos)	1373	54.29	20.64	0.00	100.00
PHIS_1	Historia (% aciertos)	1373	56.41	19.36	0.00	100.00
PGEO_1	Geografía (% aciertos)	1373	55.60	18.22	0.00	100.00
PFCE_1	Formación cívica y ética (% aciertos)	1373	64.90	21.25	0.00	100.00
PHM_1	Habilidad matemática (% aciertos)	1373	61.53	23.44	6.25	100.00
PMAT_1	Matemáticas (% aciertos)	1373	62.56	23.91	0.00	100.00
PFIS_1	Física (% aciertos)	1373	53.02	22.30	0.00	100.00
PQUL_1	Química (% aciertos)	1373	53.46	22.41	0.00	100.00
PBIO_1	Biología (% aciertos)	1373	58.10	23.76	0.00	100.00
PHIB_1	Habilidades intelectuales básicas (% aciertos)	1373	64.10	19.18	3.12	100.00
PCL_1	Comunicación y lenguaje (% aciertos)	1373	61.36	18.08	7.14	100.00
PIM_1	Lenguaje matemático (% aciertos)	1373	61.97	21.76	3.57	100.00
PCD_1	Conocimientos disciplinarios (% aciertos)	1373	57.29	16.20	11.45	92.70
P_SEM1	Promedio semestre 1	1373	7.87	1.13	5.00	10.00
P_SEM2	Promedio semestre 2	1373	7.70	1.55	0.00	10.00
P_ANO1	Promedio año 1	1373	7.84	1.13	5.00	10.00
P_SEM3	Promedio semestre 3	1373	4.57	4.15	0.00	10.00
P_SEM4	Promedio semestre 4	1373	4.38	4.15	0.00	10.25
P_SEM5	Promedio semestre 5	1373	2.68	4.50	0.00	86.00
P_SEM6	Promedio semestre 6	1373	1.76	4.10	0.00	86.00
P_BACH	Promedio bachillerato	1373	1.71	3.38	0.00	9.98
ZCNE	Índice Ceneval estandarizado	1373	0.02	1.00	-4.00	2.58
ZPROMSEC	Promedio secundario estandarizado	1373	0.01	0.99	-3.67	2.32
Z_2	Índice Ceneval discretizado	1373	0.01	0.84	-3.33	2.17

b) Estadísticas descriptivas del nivel 2

Nombre de variable	Descripción	N	Media	Desviación estándar	Mínimo	Máximo
SEXO	Sexo	20	1.50	0.51	1.00	2.00
REG_PROC	Régimen de la escuela de procedencia	20	1.40	0.50	1.00	2.00
TIPO_SEC	Tipo de secundaria	20	2.10	0.72	1.00	5.00
PROM_SEC	¿Qué promedio tienes actualmente en secundaria?	20	8.70	0.76	7.10	9.80
INST_ING	¿A qué institución de educación superior te gustaría ingresar al terminar tus estudios de educación media superior?	20	4.15	2.80	1.00	10.00
INGR_FAM	¿Cuál es el ingreso familiar mensual?	20	9.05	4.98	2.00	15.00
ESCO_PAD	Escolaridad del padre	20	7.75	2.97	3.00	11.00
ESCO_MAD	Escolaridad de la madre	20	7.65	3.10	3.00	11.00
OCU_PADR	Ocupación del padre	20	9.70	3.57	2.00	14.00
OCU_MADR	Ocupación de la madre	20	6.75	4.24	3.00	14.00
SER_TELE	¿Cuentas en tu casa con teléfono?	20	0.90	0.31	0.00	1.00
SER_PC	¿Cuentas en tu casa con computadora?	20	0.80	0.41	0.00	1.00
SER_INTE	¿Cuentas en tu casa con conexión a internet?	20	0.65	0.49	0.00	1.00
DOM_ING	Dominio del Inglés	20	2.75	1.41	1.00	5.00
ICNE_1	Índice Ceneval global	20	1066.25	98.34	948.00	1248.00
PERCEN_1	Percentil	20	61.55	27.38	17.48	100.00
NCNE_1	Número de aciertos global	20	78.15	21.00	53.00	117.00
PCNE_1	Porcentaje de aciertos global	20	61.05	16.40	41.40	91.40
IHV_1	Habilidad verbal (índice)	20	1089.85	140.97	775.00	1300.00
IESP_1	Español (índice)	20	1035.00	133.87	850.00	1300.00
IHIS_1	Historia (índice)	20	1075.00	86.60	850.00	1200.00
IGEO_1	Geografía (índice)	20	1050.00	131.79	850.00	1250.00
IFCE_1	Formación cívica y ética (índice)	20	1120.00	121.83	850.00	1250.00
IHM_1	Habilidad matemática (índice)	20	1067.35	144.12	738.00	1225.00
IMAT_1	Matemáticas (índice)	20	1075.00	150.00	800.00	1300.00
IFIS_1	Física (índice)	20	1002.50	149.10	750.00	1200.00
IQUI_1	Química (índice)	20	1030.00	146.36	750.00	1300.00
IBIO_1	Biología (índice)	20	1110.00	139.17	850.00	1300.00
IHIB_1	Habilidades intelectuales básicas (índice)	20	1078.75	108.79	888.00	1262.00
ICL_1	Comunicación y lenguaje (índice)	20	1066.35	118.05	850.00	1278.00

b) Estadísticas descriptivas del nivel 2

Nombre de variable	Descripción	N	Media	Desviación estándar	Mínimo	Máximo
IUM_1	Lenguaje matemático (índice)	20	1070.60	138.55	807.00	1257.00
ICD_1	Conocimientos disciplinares (índice)	20	1062.10	98.68	931.00	1250.00
PHV_1	Habilidad verbal (% aciertos)	20	65.00	23.51	12.50	100.00
PESP_1	Español (% aciertos)	20	55.83	22.31	25.00	100.00
PHIS_1	Historia (% aciertos)	20	62.50	14.43	25.00	83.33
PGEO_1	Geografía (% aciertos)	20	58.33	21.96	25.00	91.66
PFCE_1	Formación cívica y ética (% aciertos)	20	70.00	20.30	25.00	91.66
PHM_1	Habilidad matemática (% aciertos)	20	61.25	24.04	6.25	87.50
PMAT_1	Matemáticas (% aciertos)	20	62.50	25.00	16.66	100.00
PFIS_1	Física (% aciertos)	20	50.41	24.85	8.33	83.33
PQUL_1	Química (% aciertos)	20	55.00	24.39	8.33	100.00
PBIO_1	Biología (% aciertos)	20	68.33	23.19	25.00	100.00
PHIB_1	Habilidades intelectuales básicas (% aciertos)	20	63.12	18.15	31.25	93.75
PCL_1	Comunicación y lenguaje (% aciertos)	20	61.07	19.69	25.00	96.42
PUM_1	Lenguaje matemático (% aciertos)	20	61.78	23.09	17.85	92.85
PCD_1	Conocimientos disciplinares (% aciertos)	20	60.36	16.45	38.54	91.66
P_SEM1	Promedio semestre 1	20	7.86	1.23	5.00	10.00
P_SEM2	Promedio semestre2	20	7.44	2.20	0.00	9.71
P_ANO1	Promedio año 1	20	7.84	1.23	5.00	9.86
P_SEM3	Promedio semestre 3	20	4.17	4.37	0.00	10.00
P_SEM4	Promedio semestre 4	20	4.02	4.24	0.00	9.88
P_SEM5	Promedio semestre 5	20	2.14	3.82	0.00	9.40
P_SEM6	Promedio semestre 6	20	1.78	3.66	0.00	9.56
P_BACH	Promedio bachillerato	20	1.66	3.43	0.00	9.40
ZCNE	Índice Ceneval estandarizado	20	0.30	0.93	-1.10	2.05
ZPROMSEC	Promedio secundario estandarizado	20	0.17	1.02	-2.08	1.75
Z_2	Índice Ceneval discretizado	20	0.25	0.79	-0.92	1.75

A2. Descripción de la base de datos para EXANI-I, Generación 1 (2005-2007)

La base de datos está organizada en variables descritas en dos tablas, como se presenta a continuación. Se indican los datos disponibles en la base en formato de EXCEL y los que se encuentran en la base de SPSS.

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
	D_R		X	Campo en blanco		
	Tipo_exa		X	Tipo de examen		
	Gen		X	Generación	1-3	Gen. 1=2005-2007, 2=2006-2008, 3=2007-2009
	ENTIDAD	X	X	Entidad donde se ubica la institución		
	CVE_INST	X	X	Clave de la institución en donde se llevó a cabo la aplicación		
	IDENTIFICA	X	X	Identificador asignado en calificación		
	FOLIO	X	X	Folio		
	APLIC_1	X	X	¿Presentó el examen cuando ingresó a la institución?	SI - NO	
	APLIC_2	X	X	¿Presentó el examen al inicio del tercer semestre?		
	APLIC_3	X	X	¿Presentó el examen al inicio del quinto semestre?		
	APLIC_4	X	X	¿Presentó el examen al final del sexto semestre?		
	ID_APLIC	X				
2	NOMBRE	X		Nombre	ABIERTA	
3.1	DIA_NAC	X	X	Día de nacimiento	01 - 31	
3.2	MES_NAC	X	X	Mes de nacimiento	01 a 12	
3.3	ANO_NAC	X	X	Año de nacimiento	1900 - 1999	1900 - 1999
4	EDO_NACI	X	X	Lugar de nacimiento	01 - 32, 99	Estado de la República, extranjero
5	SEXO	X	X	Sexo	1,2	Masculino,Femenino
6	LEN_IND	X	X	¿Tu lengua materna es indígena?	1,2	Sí,No
9	REG_PROC	X	X	Régimen de la escuela de procedencia	1,2	Pública,Privada
10	TIPO_SEC	X	X	Tipo de secundaria	1 - 5	Técn.,Gral.,Telesec., P.trabajadores,Abierta
11	PROM_SEC	X	X	¿Qué promedio tienes actualmente en secundaria?	6.0 - 10.0	6.0 - 10.0
13	INST_ING	X	X	¿A qué institución de educación superior te gustaría ingresar al terminar tus estudios de educación media superior?	01 - 11	Normal,I.Tecn.,U.privada, U.Tecn.,U.Estado,UPN, IPN,UAM,UANM, Otra, ninguna

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
16	VIVE_CON	X	X	¿Con quién vives actualmente?	01 - 12	PMH,PM,P,M,H,PH,MH, Otros,Cónyuge, Compañeros, Solo, Otra situación
19.1	EDAD_PAD	X	X	Edad del padre	01 - 10	<=30,31-35,36-40, 41-45,46-50,51-55, 56-60,>60, Fallecido,No lo sé
19.2	EDAD_MAD	X	X	Edad de la madre		
20	TRAB_ACT	X	X	¿En la actualidad desarrollas algún trabajo por el cual recibes un sueldo?	1,2	Sí,No
21	INGR_FAM	X	X	¿Cuál es el ingreso familiar mensual?	01 - 15	<=1,000,segmentos de 1000, >20000
22.1	ESCO_PAD	X	X	Escolaridad del padre	01 - 11	Ninguna, Lee/ escribe, Prim, Capac>P, Sec, Capa>Sec, Prof.téc., MediaSup, Normal, Lic., Posgrado
22.2	ESCO_MAD	X	X	Escolaridad de la madre		
23.1	OCU_PADR	X	X	Ocupación del padre	01 - 14	Ninguna, Jubil, Hogar, Campo/ pesca, Constr, Obrero, ..., empl. profes., empl. técn/ admv, profesión, otra
23.2	OCU_MADR	X	X	Ocupación de la madre		
24.10	SER_TELE	X	X	¿Cuentas en tu casa con teléfono?	1	Teléfono
24.17	SER_PC	X	X	¿Cuentas en tu casa con computadora?	1	Computadora
24.20	SER_INTE	X	X	¿Cuentas en tu casa con conexión a internet?	1	Conexión a internet
31.3	PAD_FELI	X	X	Te felicitan o premian cuando te va bien en la escuela	1 - 4	Casi nunca-casi siempre
31.4	PAD_RESP	X	X	Respetan tus opiniones sobre lo que ocurre en la escuela		
31.5	PAD_DECI	X	X	Promueven que tomes tus propias decisiones sobre lo que pasa en la escuela		
33	IDI_HABL	X	X	¿Cuál es el idioma que hablas en tu casa la mayor parte del tiempo?	1 - 3	Español, lenguas indígenas, otros
34.1	DOM_ING	X	X	Inglés	1 - 5	Nulo-avanzado
35.1	PRE_EXAM	X	X	Me reúno con mis compañeros para preparar un examen	1 - 4	Casi nunca-casi siempre
35.2	COM_TARE	X	X	Me reúno con mis compañeros para elaborar una tarea o trabajo en grupo		
35.3	PLAN_TRA	X	X	Al iniciar, identifico lo que necesito estudiar y elaboro un plan de trabajo		
35.4	REC_ESTU	X	X	Reviso qué es lo que recuerdo de lo que estudié		
35.5	IDE_CONC	X	X	Identifico los conceptos que aún no he comprendido		
35.6	BUS_INFO	X	X	Cuando no entiendo algo busco información que me aclare lo que estoy estudiando		
35.7	EST_MONO	X	X	Estudio principalmente con monografías		
35.8	APUN_CLA	X	X	Estudio principalmente con mis apuntes de clase		

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
35.9	EST_TEXT	X	X	Estudio principalmente con el libro de texto de la asignatura		
35.10	EST_APUN	X	X	Estudio principalmente con los apuntes de mis compañeros		
35.11	UTI_ENC	X	X	Utilizo enciclopedias, diccionarios o atlas		
35.12	UTI_PC	X	X	Utilizo computadora o internet para estudiar, hacer una tarea o resolver un examen		
36.1	APRE_MAS	X	X	Aprendo más cuando trabajo con otros compañeros		
36.2	APO_IDEA	X	X	Es de gran ayuda que todos aporten ideas cuando trabajo en grupo		
36.3	EST_ASEG	X	X	Estudio para asegurarme económicamente mi futuro		
36.4	BUEN_TRA	X	X	Estudio para obtener un buen trabajo		
36.5	EST_APRE	X	X	Estudio para aprender más		
36.6	VIV_MEJO	X	X	Estudio para vivir mejor		
36.7	ENT_ESTU	X	X	Tengo confianza que puedo entender lo que estudio, inclusive textos más difíciles		
36.8	EXC_TRAB	X	X	Tengo confianza que puedo realizar excelente trabajo en mis tareas y exámenes		
36.9	DOM_HABI	X	X	Tengo seguridad en que domino las habilidades que me enseñaron		
36.10	APRE_RAP	X	X	Aprendo rápidamente en la mayoría de las asignaturas	1 - 4	Totalmente en desacuerdo Totalmente de acuerdo
36.11	SOY_COMP	X	X	Soy competente en la mayoría de las asignaturas		
36.12	RES_EXAM	X	X	Resuelvo bien los exámenes en la mayoría de las asignaturas		
36.13	GUS_TRAB	X	X	Me gusta trabajar con otros compañeros		
36.14	LEO_OBLI	X	X	Solamente leo cuando tengo la obligación de hacerlo		
36.15	LEC_PASA	X	X	La lectura es uno de mis pasatiempos favoritos		
36.16	COM_LIBR	X	X	Me gusta comentar los libros con otras personas		
36.17	CUE_LEER	X	X	Me cuesta trabajo terminar de leer un libro		
36.18	REG_LIBR	X	X	Me gusta que me regalen libros		
36.19	LEC_PERD	X	X	La lectura me parece una pérdida de tiempo		
36.20	DIS_LIBR	X	X	Disfruto el visitar librerías o bibliotecas		
36.21	LEO_INFO	X	X	Solamente leo para obtener la información que necesito		
36.22	TRA_LEER	X	X	Me cuesta trabajo sentarme a leer por mucho tiempo	1 - 4	Totalmente en desacuerdo Totalmente de acuerdo

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
37.1	HORA_EST	X	X	Estudias o haces tarea fuera del horario escolar	00 - 11	0 a Más de 10
37.2	TIEM_LEE	X	X	Te dedicas a leer sobre lo que te gusta o interesa		
38.00	LEE_TEX	X	X	Cuántos libros completos has leído en los últimos 12 meses (no libros de texto)		
40	EST_ALCA	X	X	¿Cuál es el nivel de estudios que te gustaría alcanzar?	1 - 5	Bach., Prof.téc.,Téc.Sup. Univ.,licenc.,Posgrado
	VERSION_1	X	X	Versión del examen aplicado	01 - 99	
	POS_SUS_1	X		Lugar alcanzado por el sustentante	1 - 999999	
	ICNE_1	X	X	Índice Ceneval global	700 - 1300	
	PERCEN_1	X	X	Percentil	0.00 - 100.00	
	NCNE_1	X	X	Número de aciertos global	0 - 128	
	PCNE_1	X	X	Porcentaje de aciertos global	0.00 - 100.00	
Tema 1	IHV_	X	X	Habilidad verbal (índice)	700 - 1300	
Tema 2	IESP_1	X	X	Español (índice)		
Tema 3	IHIS_1	X	X	Historia (índice)		
Tema 4	IGEO_1	X	X	Geografía (índice)		
Tema 5	IFCE_1	X	X	Formación cívica y ética (índice)		
Tema 6	IHM_1	X	X	Habilidad matemática (índice)		
Tema 7	IMAT_1	X	X	Matemáticas (índice)		
Tema 8	IFIS_1	X	X	Física (índice)		
Tema 9	IQUI_1	X	X	Química (índice)}		
Tema 10	IBIO_1	X	X	Biología (índice)		
	IHIB_1	X	X	Habilidades intelectuales básicas (índice)		
	ICL_1	X	X	Comunicación y lenguaje (índice)		
	ILM_1	X	X	Lenguaje matemático (índice)		
	ICD_1	X	X	Conocimientos disciplinarios (índice)		
	NHV_1	X		Habilidad verbal (# aciertos)	0 - 16	
	NESP_1	X		Español (# aciertos)	0 - 12	
	NHIS_1	X		Historia (# aciertos)		
	NGEO_1	X		Geografía (# aciertos)		
	NFCE_1	X		Formación cívica y ética (# aciertos)		
	NHM_1	X		Habilidad matemática (# aciertos)	0 - 16	
	NMAT_1	X		Matemáticas (# aciertos)	0 - 12	
	NFIS_1	X		Física (# aciertos)		
	NQUI_1	X		Química (# aciertos)		
	NBIO_1	X		Biología (# aciertos)		

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
	PHV_1	X	X	Habilidad verbal (% aciertos)	0.00 - 100.00	
	PESP_1	X	X	Español (% aciertos)		
	PHIS_1	X	X	Historia (% aciertos)		
	PGEO_1	X	X	Geografía (% aciertos)		
	PFCE_1	X	X	Formación cívica y ética (% aciertos)		
	PHM_1	X	X	Habilidad matemática (% aciertos)		
	PMAT_1	X	X	Matemáticas (% aciertos)		
	PFIS_1	X	X	Física (% aciertos)		
	PQUI_1	X	X	Química (% aciertos)		
	PBIO_1	X	X	Biología (% aciertos)		
	PHIB_1	X	X	Habilidades intelectuales básicas (% aciertos)		
	PCL_1	X	X	Comunicación y lenguaje (% aciertos)		
	PLM_1	X	X	Lenguaje matemático (% aciertos)		
	PCD_1	X	X	Conocimientos disciplinarios (% aciertos)		
	VERSION_2	X		Versión del examen aplicado		
	POS_SUS_2	X		Lugar alcanzado por el sustentante	1 - 999999	
	ICNE_2	X		Índice Ceneval global	700 - 1300	
	PERCEN_2	X		Percentil	0.00 - 100.00	
	NCNE_2	X		Número de aciertos global	0 - 128	
	PCNE_2	X		Porcentaje de aciertos global	0.00 - 100.00	
Tema 1	IHV_	X		Habilidad verbal (índice)	700 - 1300	
Tema 2	IESP_2	X		Español (índice)		
Tema 3	IHIS_2	X		Historia (índice)		
Tema 4	IGEO_2	X		Geografía (índice)		
Tema 5	IFCE_2	X		Formación cívica y ética (índice)		
Tema 6	IHM_2	X		Habilidad matemática (índice)		
Tema 7	IMAT_2	X		Matemáticas (índice)		
Tema 8	IFIS_2	X		Física (índice)		
Tema 9	IQUI_2	X		Química (índice)		
Tema 10	IBIO_2	X		Biología (índice)		
	IHIB_2	X		Habilidades intelectuales básicas (índice)		
	ICL_2	X		Comunicación y lenguaje (índice)		
	ILM_2	X		Lenguaje matemático (índice)		
	ICD_2	X		Conocimientos disciplinarios (índice)		

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
	NHV_2	X		Habilidad verbal (# aciertos)	0 - 16	
	NESP_2	X		Español (# aciertos)	0 - 12	
	NHIS_2	X		Historia (# aciertos)		
	NGEO_2	X		Geografía (# aciertos)		
	NFCE_2	X		Formación cívica y ética (# aciertos)		
	NHM_2	X		Habilidad matemática (# aciertos)	0 - 16	
	NMAT_2	X		Matemáticas (# aciertos)	0 - 12	
	NFIS_2	X		Física (# aciertos)		
	NQUI_2	X		Química (# aciertos)		
	NBIO_2	X		Biología (# aciertos)		
	PHV_2	X		Habilidad verbal (% aciertos)	0.00 - 100.00	
	PESP_2	X		Español (% aciertos)		
	PHIS_2	X		Historia (% aciertos)		
	PGEO_2	X		Geografía (% aciertos)		
	PFCE_2	X		Formación cívica y ética (% aciertos)		
	PHM_2	X		Habilidad matemática (% aciertos)		
	PMAT_2	X		Matemáticas (% aciertos)		
	PFIS_2	X		Física (% aciertos)		
	PQUI_2	X		Química (% aciertos)		
	PBIO_2	X		Biología (% aciertos)		
	PHIB_2	X		Habilidades intelectuales básicas (% aciertos)		
	PCL_2	X		Comunicación y lenguaje (% aciertos)		
	PLM_2	X		Lenguaje matemático (% aciertos)		
	PCD_2	X		Conocimientos disciplinarios (% aciertos)		
	VERSION_3	X		Versión del examen aplicado	01 - 99	
	POS_SUS_3	X		Lugar alcanzado por el sustentante	1 - 999999	
	ICNE_3	X		Índice Ceneval global	700 - 1300	
	PERCEN_3	X		Percentil	0.00 - 100.00	
	NCNE_3	X		Número de aciertos global	0 - 128	
	PCNE_3	X		Porcentaje de aciertos global	0.00 - 100.00	
Tema 1	IHV_	X		Habilidad verbal (índice)		
Tema 2	IESP_3	X		Español (índice)		
Tema 3	IHIS_3	X		Historia (índice)		
Tema 4	IGEO_3	X		Geografía (índice)		
Tema 5	IFCE_3	X		Formación cívica y ética (índice)		

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
Tema 6	IHM_3	X		Habilidad matemática (índice)	700 - 1300	
Tema 7	IMAT_3	X		Matemáticas (índice)		
Tema 8	IFIS_3	X		Física (índice)		
Tema 9	IQUI_3	X		Química (índice)		
Tema 10	IBIO_3	X		Biología (índice)		
	IHIB_3	X		Habilidades intelectuales básicas (índice)		
	ICL_3	X		Comunicación y lenguaje (índice)		
	ILM_3	X		Lenguaje matemático (índice)		
	ICD_3	X		Conocimientos disciplinarios (índice)		
	NHV_3	X		Habilidad verbal (# aciertos)		
	NESP_3	X		Español (# aciertos)	0 - 12	
	NHIS_3	X		Historia (# aciertos)		
	NGEO_3	X		Geografía (# aciertos)		
	NFCE_3	X		Formación cívica y ética (# aciertos)		
	NHM_3	X		Habilidad matemática (# aciertos)	0 - 16	
	NMAT_3	X		Matemáticas (# aciertos)	0 - 12	
	NFIS_3	X		Física (# aciertos)		
	NQUI_3	X		Química (# aciertos)		
	NBIO_3	X		Biología (# aciertos)		
	PHV_3	X		Habilidad verbal (% aciertos)		
	PESP_3	X		Español (% aciertos)	0.00 - 100.00	
	PHIS_3	X		Historia (% aciertos)		
	PGEO_3	X		Geografía (% aciertos)		
	PFCE_3	X		Formación cívica y ética (% aciertos)		
	PHM_3	X		Habilidad matemática (% aciertos)		
	PMAT_3	X		Matemáticas (% aciertos)		
	PFIS_3	X		Física (% aciertos)		
	PQUI_3	X		Química (% aciertos)		
	PBIO_3	X		Biología (% aciertos)		
	PHIB_3	X		Habilidades intelectuales básicas (% aciertos)		
	PCL_3	X		Comunicación y lenguaje (% aciertos)		
	PLM_3	X		Lenguaje matemático (% aciertos)		
	PCD_3	X		Conocimientos disciplinarios (% aciertos)		
	VERSION_4	X		Versión del examen aplicado		
	POS_SUS_4	X		Lugar alcanzado por el sustentante	1 - 999999	

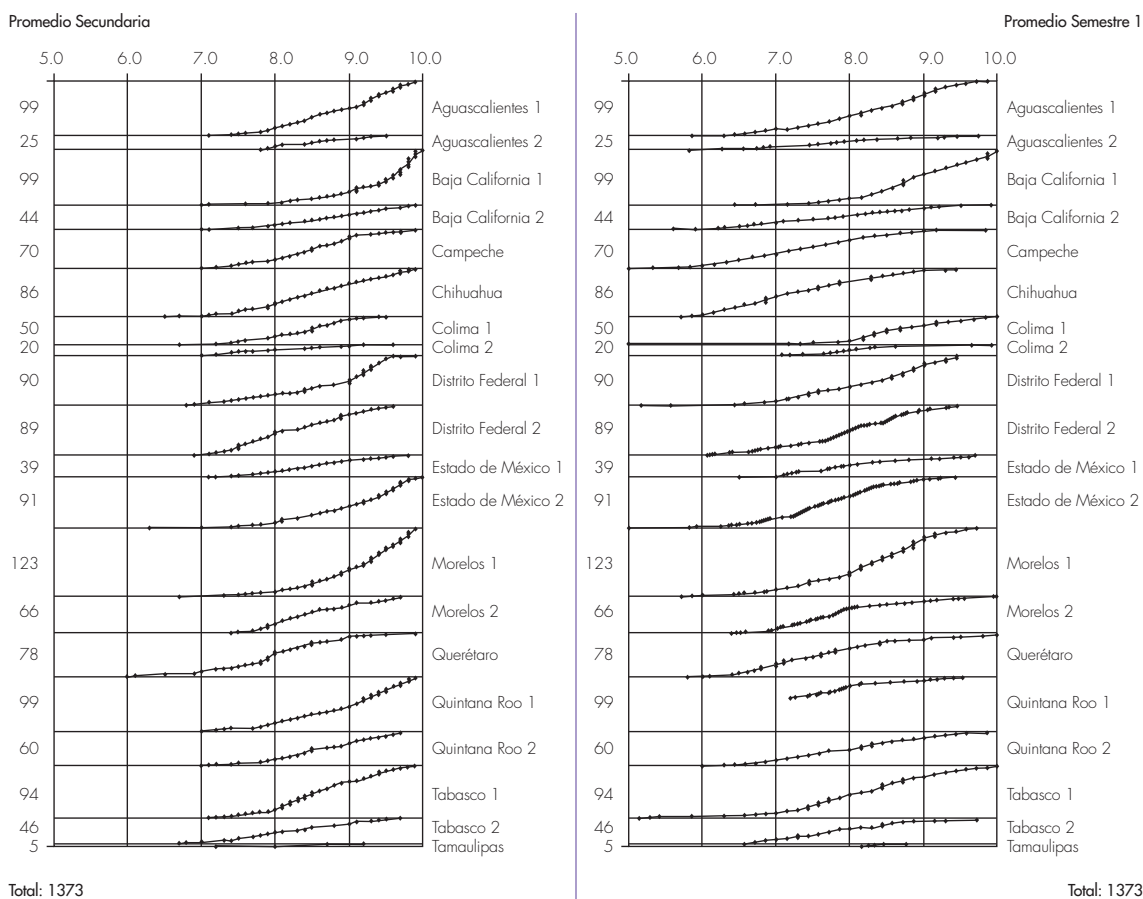
No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
	ICNE_4	X		Índice Ceneval global	700 - 1300	
	PERCEN_4	X		Percentil	0.00 - 100.00	
	NCNE_4	X		Número de aciertos global	0 - 128	
	PCNE_4	X		Porcentaje de aciertos global	0.00 - 100.00	
Tema 1	IHV_	X		Habilidad verbal (índice)	700 - 1300	
Tema 2	IESP_4	X		Español (índice)		
Tema 3	IHIS_4	X		Historia (índice)		
Tema 4	IGEO_4	X		Geografía (índice)		
Tema 5	IFCE_4	X		Formación cívica y ética (índice)		
Tema 6	IHM_4	X		Habilidad matemática (índice)		
Tema 7	IMAT_4	X		Matemáticas (índice)		
Tema 8	IFIS_4	X		Física (índice)		
Tema 9	IQUI_4	X		Química (índice)		
Tema 10	IBIO_4	X		Biología (índice)		
	IHIB_4	X		Habilidades intelectuales básicas (índice)		
	ICL_4	X		Comunicación y lenguaje (índice)		
	IUM_4	X		Lenguaje matemático (índice)		
	ICD_4	X		Conocimientos disciplinarios (índice)		
	NHV_4	X		Habilidad verbal (# aciertos)	0 - 16	
	NESP_4	X		Español (# aciertos)	0 - 12	
	NHIS_4	X		Historia (# aciertos)		
	NGEO_4	X		Geografía (# aciertos)		
	NFCE_4	X		Formación cívica y ética (# aciertos)		
	NHM_4	X		Habilidad matemática (# aciertos)	0 - 16	
	NMAT_4	X		Matemáticas (# aciertos)	0 - 12	
	NFIS_4	X		Física (# aciertos)		
	NQUI_4	X		Química (# aciertos)		
	NBIO_4	X		Biología (# aciertos)		
	PHV_4	X		Habilidad verbal (% aciertos)	0.00 - 100.00	
	PESP_4	X		Español (% aciertos)		
	PHIS_4	X		Historia (% aciertos)		
	PGEO_4	X		Geografía (% aciertos)		
	PFCE_4	X		Formación cívica y ética (% aciertos)		
	PHM_4	X		Habilidad matemática (% aciertos)		
	PMAT_4	X		Matemáticas (% aciertos)		
	PFIS_4	X		Física (% aciertos)		

No.	Variable	EXCEL	SPSS	Descripción	Valores	Etiquetas
	PQUI_4	X		Química (% aciertos)		
	PBIO_4	X		Biología (% aciertos)		
	PHIB_4	X		Habilidades intelectuales básicas (% aciertos)		
	PCL_4	X		Comunicación y lenguaje (% aciertos)		
	PLM_4	X		Lenguaje matemático (% aciertos)		
	PCD_4	X		Conocimientos disciplinarios (% aciertos)		
	ICNE	X		Promedio de Índice Ceneval	700 - 1300	
	NCNE	X		Promedio de número de aciertos (sólo temas)	0 - 120	
	PCNE	X		Promedio de porcentaje de aciertos (sólo temas)	0.00 - 100.00	
	p_sem1		X	Promedio del semestre 1		
	p_sem2		X	Promedio del año 1		
	p_ano1		X	Promedio del semestre 1		
	p_sem3		X	Promedio del semestre 3		
	p_sem4		X	Promedio del semestre 4		
	p_sem5		X	Promedio del semestre 5		
	p_sem6		X	Promedio del semestre 6		
	p_bach		X	Promedio general del bachillerato		
	P_SECmod		X			
	P_S1mod		X			
	P_S2mod		X			
	P_A1mod		X			
	P_S3mod		X			
	P_S4mod		X			
	P_S5mod		X			
	P_S6mod		X			
	P_Bacmod		X			

A3. Constataciones sobre las escalas empleadas en el análisis de predictividad.

Se comentó en el cuerpo del estudio (al presentar el tercer modelo predictivo de los resultados escolares) que los resultados exigen dimensionar racionalmente la participación del promedio de secundaria y del Índice Ceneval como predictores del resultado en el primer semestre de la educación media superior. Justamente se ha reportado en la literatura la mayor predictividad del promedio del sistema educativo anterior respecto de una prueba de selección, para lo cual se trabaja con regresiones lineales y la correlación de Pearson. El uso de un modelo multinivel permite interpretar de manera diferente la asociación entre variables.

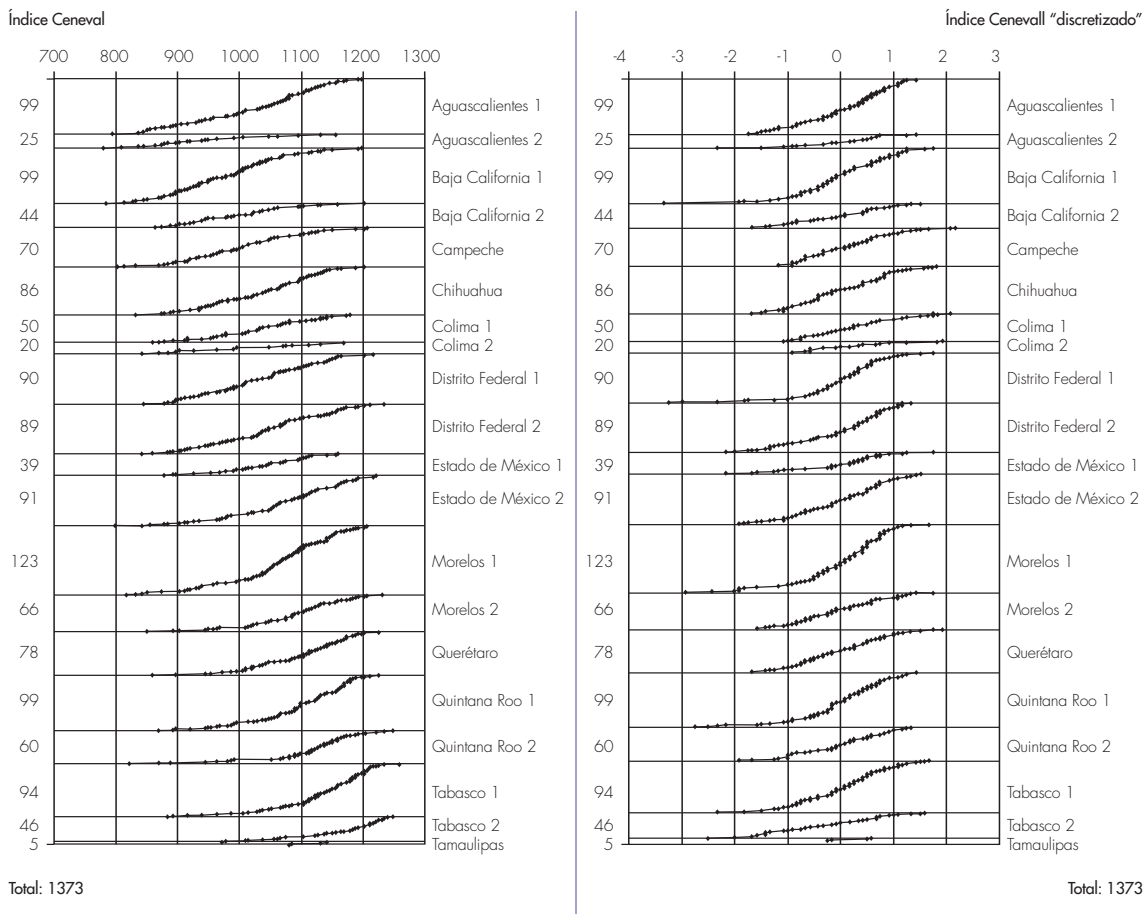
Figura A3a. Calificaciones escolares



Para comparar las escalas y su densidad, resulta interesante graficar las distribuciones acumuladas de las puntuaciones de los sujetos, como se presenta en las figuras a y b.

Estas distribuciones son elocuentes para explicar por qué hay una mayor participación entre escalas de calificaciones que con el puntaje procedente del Índice Ceneval, específicamente debido a la equivalencia de la densidad de puntos en la escala de calificaciones escolares. Asimismo se puede ver la diferencia de la escala una vez normalizada y “discretizada”, lo cual permite reportar mejores coeficientes del modelo para la prueba objetiva, recordando que no se trata de coeficientes de correlación lo que se está reportando. El cambio de escala (origen, rango y densidad) produce cambios suficientes por sí solos para modificar el modelo multinivel

Figura A3b. Puntaje Ceneval.



vel, en lugar de hacer la corrección por restricción de rango (Kobrin *et al.*, 2008) que encubre el comportamiento de las variables, en tanto que el cambio de escala hace mucho más visibles las implicaciones que se tienen en la distribución de los datos.

En consecuencia, se pueden plantear estas propiedades de interés para los estudios multinivel:

- a) Una variable que participe en el modelo multinivel no ve afectada su significancia estadística por las transformaciones que pueda experimentar en sus valores. La significancia es la misma, independientemente de que la variable se presente en forma bruta o estandarizada. Esta propiedad es similar a la que se tiene cuando se hace la correlación simple entre variables en valores brutos o estandarizados.
- b) El coeficiente de la variable en la ecuación multinivel sí se ve alterado por las transformaciones de valores en la variable y queda dimensionado a la escala de los datos utilizados en dicha variable.
- c) La calidad del modelo medida por el coeficiente de desajuste (Deviance) no se ve afectado significativamente por los cambios de las variables.
- d) Se recomienda que las variables que intervengan en el modelo multinivel tengan escalas similares, en origen, rango y densidad. La similitud de densidad en la escala de la variable influye en una mejora de la predictividad que se reporta en el coeficiente del modelo multinivel.
- e) Puede sugerirse emplear variables normalizadas o estandarizadas (media 0 y desviación estándar 1), para comparar fácilmente la participación de las variables.
- f) Es preferible juzgar la predictividad de las variables por la combinación de sus coeficientes en el modelo multinivel y la significancia estadística proporcionada por el mismo modelo, en lugar de comparar coeficientes de correlación.

