



2010 - Volumen 3, Número 2

**Procesos y Pruebas de Admisión a la
Educación Superior**



<http://www.rinace.net/riee/numeros/vol3-num2.html>

Revista Iberoamericana de Evaluación Educativa

La Revista Iberoamericana de Evaluación Educativa es una publicación oficial de la Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar.

Dirección:

F. Javier Murillo
Marcela Román

Editora:

Verónica González

Comité Directivo:

Marcela Gajardo (PREAL - Chile)
Sergio Martinic V. (PUC - Chile)
Carlos Pardo (ICFES - Colombia)
Margarita Poggi (IIPE/UNESCO - Argentina)
Francisco Soares (UMG - Brasil)
Héctor Valdés (ICCP - Cuba)

Comité Científico:

Juan Manuel Álvarez Méndez (UCM - España)
Patricia Arregui (GRADE - Perú)
Daniel Bogoya (UPN - Colombia)
Nigel Brooke (UMG - Brasil)
Leonor Cariola (MINEDUC - Chile)
María do Carmo Clímaco (Portugal)
Cristian Cox (PUC - Chile)
Santiago Cueto (Grade - Perú)
Tabaré Fernández (ANEP - Uruguay)
Juan Enrique Froemel (Min. Ed. – Qatar)
Rubén Klein (Fund. Cesgranrio – Brasil)
Luis Lizasoain (UPV - España)
Jorge Manzi (PUC - Chile)
Felipe Martínez Rizo (INEE - México)
Joan Mateo (UB - España)
Liliana Miranda (MINEDU - Perú)
Carlos Muñoz-Izquierdo (UIA – México)
Margarita Peña (ICFES - Colombia)
Dagmar Raczynski (Asesorías para el desarrollo - Chile)
Héctor Rizo (UAO - Colombia)
Guadalupe Ruiz (INEE - México)
Mario Rueda (UNAM - México)
Ernesto Schiefelbein (CIDE – Chile)
Alejandra Schulmeyer (INEP - Brasil)
Javier Tejedor (USAL – España)
Flavia Terigi (UBA/UNGS - Argentina)
Alexandre Ventura (IGE - Portugal)



Presentación

- Los Procesos y Pruebas de Admisión a la Educación Superior** 3
Jorge Manzi y Marcela Román

Sección temática: Procesos y Pruebas de Admisión a la Educación Superior

- El Balance entre Excelencia y Equidad en Pruebas de Admisión: Contribuciones de Experiencias en Sudáfrica y Costa Rica** 7
Alan Cliff y Eiliana Montero
- Validez Diferencial y Sesgo en la Predictividad de las Pruebas de Admisión a las Universidades Chilenas (PSU)** 29
Jorge Manzi, Angélica Bosch, David Bravo, Guido del Pino, Graciela Donoso, Manuel Martínez y Raúl Pizarro
- Medición de Atributos no Cognitivos para el Sistema de Admisión a la Educación Superior en Chile** 49
M. Verónica Santelices, Juan J. Ugarte, M. Paulina Flotts, Darinka Radovic, Ximena Catalán y Patrick Kyllonen
- Un Sistema Universitario sin Limitaciones de Acceso: el Caso de Uruguay** 76
Carlos Romero
- Selección y Permanencia en la Educación Superior: el Caso de la Universidad Autónoma de Yucatán** 90
Ileana López, Carlos Echazarreta, Silvia Pech y Brenda Gómez

Sección de Temática Libre

- Equiparación, Alineamiento y Predicción de Puntuaciones en Medición Educativa** 103
René Gemp
- Evaluar el Rendimiento Interno y Académico: un Desafío para la Macro y la Micro Política. Lecciones a Partir de un Estudio de Caso** 127
Silvina Gvirtz y Ángela Oría
- Modelo Explicativo del Bajo Rendimiento Escolar: Un Estudio con Adolescentes Mexicanos** 145
Joaquín Caso y Laura Hernández
- El Uso de Información Evaluativa Externa con Fines Formativos: El Caso de Establecimientos Educativos Chilenos Participantes de SEPA** 160
Andreas Hein y Sandy Taut
- El Tránsito a la Vida Post Secundaria. La Experiencia de la Primera Generación de Egresados de un Liceo de una Zona Popular en Santiago** 182
Nelson Paulus, Esteban Geoffroy y Manuela Mendoza
- Pasado, Presente y Futuro. Procesos y Procedimientos en la Evaluación de los Directores y Directoras de los Centros Escolares de Andalucía** 200
Esther Lorenzo y Antonia Ramírez



ISSN: 1989-0397

LOS PROCESOS Y PRUEBAS DE ADMISIÓN A LA EDUCACIÓN SUPERIOR

Paola Marisol Reyes Guevara y Mario Rueda Beltrán

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/editorial.pdf>

Los *Procesos y Pruebas de Admisión a la Educación Superior*, es el foco temático del presente número de la Revista Iberoamericana de Evaluación Educativa. Estrategias y mecanismos que, en los últimos años y, para un importante número de los países de la región, han adquirido gran relevancia puesto que las presiones hacia la ampliación de oportunidades educativas de nivel terciario, ha hecho necesario contar con sistemas de medición eficientes, confiables y equitativos para seleccionar candidatos a la educación superior. De hecho, la admisión a la educación superior es uno de los ámbitos más significativos y polémicos de la medición educacional. Se trata de procesos y pruebas de altas consecuencias, que afectan de manera muy significativa a examinados y al sistema educativo: tanto al nivel del que egresan los examinados como al que aspiran a ingresar.

En la mayor parte de los países la discusión pública en torno a estos sistemas de medición ha estado marcada por la preocupación en torno a la desigual representación en la educación superior de jóvenes provenientes de diversos grupos sociales y étnicos. En algunos casos, esta preocupación ha conducido al desarrollo de estudios que evalúen el potencial sesgo de medición de los sistemas de selección. En otros casos ha derivado en el desarrollo de mecanismos de selección especiales para grupos subrepresentados, y en otros, a la inclusión de criterios de selección potencialmente menos correlacionados con el origen socioeconómico de los postulantes a la educación superior (especialmente criterios no cognitivos). El efecto práctico de lo anterior es que los sistemas de admisión se están tornando cada vez más complejos y variados, incluyendo una gran diversidad de criterios, indicadores, instrumentos y procesos de selección de los estudiantes. Lamentablemente, en muchos casos no se ha estudiado ni demostrado la capacidad predictiva de estos dispositivos con respecto a los resultados académicos de los alumnos en el corto y largo plazo.

Para este número especial invitamos la presentación de manuscritos que emplearan información producida por los sistemas de admisión a universidades y otras instituciones de educación superior. Se invitó especialmente trabajos empíricos que analizaran características métricas de estas mediciones: confiabilidad, validez (especialmente predictiva), predicción diferencial y sesgo de medición. Este tipo de trabajos permite evaluar la calidad de estas mediciones para los propósitos selectivos que se persiguen. También se invitó a presentar trabajos que apuntaran a explicar las diferencias en los puntajes de estas pruebas según variables sociodemográficas y educacionales, así como a estudios y trabajos conceptuales que permitieran ampliar el espectro de criterios considerados para la admisión a la enseñanza superior. Así, estructuran la parte temática de esta revista cinco trabajos, los que en su conjunto abordan los diversos ámbitos planteados en la convocatoria, con antecedentes o evidencias de distintos países.

En el primero de ellos, Alan Cliff y Eiliana Montero describen dos experiencias de universidades que buscan integrar excelencia y equidad en sus procesos de admisión. Abordan así, los desafíos, hallazgos y aportes al ingreso de estudiantes que provienen de contextos o ambientes en desventaja educativa y social en Costa Rica y Sudáfrica.

La validez diferencial de las Pruebas de Admisión a las Universidades Chilenas, según género del estudiante y dependencia de los establecimientos educacionales, son el foco de interés de un equipo de académicos que conforman el Comité Técnico de las Pruebas de Selección Universitaria en Chile¹. Los

¹ El equipo de investigadores/autores está compuesto por: Jorge Manzi, Angélica Bosch, David Bravo, Guido del Pino, Graciela Donoso, Manuel Martínez y Raúl Pizarro.

autores utilizan información de las Pruebas de Selección de Matemática y de Lenguaje y Comunicación, además de las Notas de Educación Media (como predictores) y el rendimiento académico de los estudiantes en el primer año de sus estudios universitarios (como criterio), tratando de establecer si la capacidad predictiva de los factores de selección difiere entre hombres y mujeres o entre jóvenes que asistieron a distinto tipo de establecimientos educacionales.

El siguiente trabajo presenta evidencia acerca del proceso de validación de un conjunto de instrumentos desarrollados en la Universidad Católica de Chile², como parte de una iniciativa destinada a complementar el sistema de selección convencional actualmente existente en ese país. En esta iniciativa se considera la medición del pensamiento crítico y ciertos atributos no cognitivos de los estudiantes. El artículo comparte el proceso de validación piloto de instrumentos desarrollados para dichos fines.

Carlos Romero, docente de la Universidad Católica del Uruguay, discute en su artículo los pro y contra de contar con un sistema universitario de libre admisión, en una realidad y sistema educativo como el uruguayo.

Finalmente, Ileana López, Carlos Echazarreta, Silvia Pech y Brenda Gómez, académicos de la Universidad Autónoma de Yucatán de México, buscan determinar en qué medida el examen de selección utilizado por dicha universidad se asocia a la trayectoria escolar de los estudiantes en los primeros años en distintas carreras. Para ello analizan los resultados obtenidos por los estudiantes en la prueba EXANI II, en relación con el desempeño y riesgo escolar.

Dejando ya el sistema educativo superior, René Gempp en un interesante artículo, revisa los métodos de enlace de puntuaciones utilizados en medición educativa y pruebas estandarizadas para efectos de predicción, alineamiento y equiparación. A través de este recorrido, discute sus diferencias y ofrece sugerencias de aplicaciones prácticas de los mismos.

La mirada sobre el rendimiento escolar, en tanto antesala del abandono, es el propósito del texto de Silvina Gvirtz y Angela Oria, investigadoras argentinas. El texto comparte la experiencia y resultados de 'Escuelas del Bicentenario', proyecto que tienen entre sus objetivos, evitar la exclusión social y disminuir la repitencia.

Joaquín Caso y Laura Hernández, analizan el efecto de factores actitudinales y socioemocionales en el rendimiento académico de estudiantes de bachillerato en México. Para ello y mediante el modelamiento de ecuaciones estructurales, se incorporan las variables autoestima, establecimiento de metas, habilidades de estudio, adaptación escolar, asertividad y consumo de sustancias. Ofrecen así, un modelo explicativo del bajo rendimiento en adolescentes mexicanos y discuten posibles intervenciones preventivas a partir de dicho modelo.

Las diferentes formas de usar la información que genera la evaluación externa de los aprendizajes escolares, se constituye en eje de atención y discusión del artículo de Andreas Hein y Sandy Taut. Mediante el recorrido por la experiencia de un grupo de establecimientos educacionales que participan del Sistema de Evaluación de Progresos del Aprendizaje (SEPA), desarrollado e implementado por el Centro de Medición MIDE UC de la Universidad Católica de Chile, los autores comparten las dificultades y complejidades que enfrentan los establecimientos y sus actores, para una adecuada comprensión de los indicadores implicados en la medición y por ende, en la promoción de procesos o acciones de

² Los autores son María Verónica Santelices, Juan José Ugarte, María Paulina Flotts, Darinka Radovic, Ximena Catalán y Patrick C. Kyllonen

mejoramiento.

Nelson Paulus, Esteban Geoffroy y Manuela Mendoza, exponen y reflexionan sobre las percepciones, interpretaciones y atribuciones de sentido sobre el tránsito a la vida post secundaria, que hacen los jóvenes adolescentes que egresan del liceo en Chile. La investigación combina técnicas e instrumentos cuantitativos y cualitativos para identificar y analizar diversos factores implicados en la decisión de seguir estudiando, el tipo de trayectoria educativa de los jóvenes y sus eventuales implicancias en la deserción universitaria.

Cierra la revista el artículo de Esther Lorenzo y Antonia Ramírez, que ofrece al lector ideas e interpretaciones interesantes que agregan valor a la discusión sobre los desafíos de la dirección escolar. Mediante un completo recorrido histórico, ellas revisan las funciones y desafíos del director de centros educativos; analizan y discuten los procesos, procedimientos y criterios de evaluación de los directores y directoras de la Comunidad Autónoma de Andalucía.

Como siempre, los invitamos a conocerlos, discutirlos y difundirlos.



EL BALANCE ENTRE EXCELENCIA Y EQUIDAD EN PRUEBAS DE ADMISIÓN: CONTRIBUCIONES DE EXPERIENCIAS EN SUDÁFRICA Y COSTA RICA

THE BALANCE BETWEEN EXCELLENCE AND EQUITY ON ADMISSION TEST: CONTRIBUTIONS OF EXPERIENCES IN SOUTH AFRICA AND COSTA RICA

Alan Cliff y Eiliana Montero

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art2.pdf>

Fecha de recepción: 15 de marzo de 2010
Fecha de dictaminación: 19 de mayo de 2010
Fecha de segundo envío: 25 de mayo de 2010
Fecha de aceptación: 26 de mayo de 2010



1. ANTECEDENTES

Anivel internacional, las instituciones de educación superior en general, y las universidades en particular, enfrentan desafíos cada vez más complejos en lo que respecta a la selección, retención y conclusión de estudios de los estudiantes que se consideran aspirantes deseables (Terenzini et al, 1996; Tinto, 1998; Pascarella & Terenzini, 1998; Scott, Yeld & Hendry, 2007). El desafío parece constar de varias dimensiones importantes.

Primero está el desafío de valorar y comprender el alcance hasta el cual los aspirantes a la educación superior han recibido **preparación adecuada** durante su educación primaria y secundaria para **hacer frente a las demandas académicas** a las cuales se enfrentarán, tanto al ingresar como durante su curso de estudios en la educación superior (Griesel, 2006; Yeld, 2006; Bohlmann & Braun, 2006; Cliff & Yeld, 2006; Frith & Prince, 2006). En muchos países, existe un examen formal que se administra al terminar la secundaria y que se considera como una fuente importante para recopilar información sobre la preparación básica de los aspirantes. Por lo general, este examen lo administra a nivel nacional el ministerio responsable de las políticas y la práctica de educación secundaria. Los niveles aceptables de desempeño en este examen se determinan y certifican de modo que la educación superior pueda, por lo menos en teoría, considerar los niveles particulares de desempeño como los mínimos aceptables para ingresar a sus programas de estudio. Históricamente se ha asumido en este sistema, que niveles mínimos de desempeño implican que el conocimiento y los procesos de aprendizaje -que son prerrequisito- se han enseñado y aprendido y que el aspirante está 'preparado' para ingresar a los programas convencionales de estudio que son similares a las asignaturas para las cuales este aspirante tomó el examen de conclusión de secundaria.

Cuando el examen de conclusión de secundaria no existe, otro enfoque internacional común para abordar el desafío de evaluar el desempeño al concluir la secundaria y la preparación para ingresar a la educación superior de los estudiantes, es el uso de un examen o un conjunto de exámenes básicos de "lápiz y papel" (por ejemplo, AARP, 2009; ETS-SAT, 2009; ETS-TOEFL, 2009). Las metas de estas pruebas son también evaluar los que se consideran los conocimientos y procesos básicos que los aspirantes requieren para hacer frente a los estudios de educación superior. Sin embargo, estos exámenes a menudo se desarrollan de manera independiente de cualquier ministerio nacional de educación y con frecuencia se diseñan para evaluar tanto las destrezas académicas genéricas de lectura, escritura y capacidad de razonamiento, como para evaluar el conocimiento y los conceptos fundamentales disciplinares que se requieren para ingresar a la educación superior.

Las valoraciones en tales pruebas toman dos enfoques comunes: (1) tratan de evaluar el conocimiento aprendido por el individuo en el sistema de educación secundaria, y las tareas en las pruebas con frecuencia son objetivas y referenciadas, o sea, los exámenes evalúan lo que se ha aprendido con base en un nivel de aprendizaje anticipado que se considera el mínimo (ver por ejemplo, Foxcroft, 2006); y (2) pretenden asegurar si un individuo tiene el 'potencial' para enfrentar los estudios de educación superior, aunque ciertos niveles mínimos de conocimiento no se le han enseñado o no los haya aprendido adecuadamente (ver por ejemplo, Cliff, 2002; Cliff, Yeld & Hanslo, 2003; Cliff, Ramaboa & Pearce, 2007; Badenhorst, Cliff & Kidson, 2009, bajo revisión).

En el primer enfoque, los exámenes se usan para verificar o medir los niveles de conocimiento y los procesos que se supone deben haber sido aprendidos en la escuela secundaria. Tales pruebas se implementan con el fin de validar qué se ha aprendido en la escuela o con el fin de facultar a las instituciones de enseñanza superior para adoptar un enfoque estandarizado y comparar el conocimiento de los aspirantes a partir de una gran variedad de experiencias educativas, antecedentes o niveles de certificación (ver, por ejemplo, Pitoniak, Cliff & Yeld, 2008; Yeld et al, 2009).

El último enfoque, el uso de exámenes para evaluar el 'potencial' de los aspirantes, asume que, incluso en casos en los cuales los aspirantes no hayan desarrollado el conocimiento y los procesos requeridos esperados en un estudiante de primer ingreso, las pruebas pueden ayudar a comprender si el aspirante podrá tener éxito a pesar de la falta de conocimiento o procesos formales (Cliff & Hanslo, 2005; Cliff & Hanslo, 2009; Yeld, 2001). Tales pruebas se implementan para determinar la presencia de la capacidad o las capacidades que se consideran innatas o latentes, o sea, que no están relacionadas necesariamente de manera directa con el aprendizaje formal (Sternberg & Grigorenko, 2002). Con alguna controversia, las pruebas de potencial en todo el mundo tratan de hacer aflorar la capacidad de los aspirantes para aprender o su grado de reacción a una oportunidad de aprender, la cual se asume es inherente a la prueba misma. (Budoff, 1987; Campione & Brown, 1987; Feuerstein et al, 1991; Hessels & Hamers, 1993).

El uso de exámenes de primer ingreso se está volviendo cada vez más común a nivel internacional. En todo el mundo, las instituciones de educación superior creen que los graduados de secundaria no están preparados de manera adecuada para hacer frente a las demandas típicas de los estudios en la educación superior: las tasas de fracaso y no conclusión de estudios están aumentando, además, concluir los estudios está tomando mucho más tiempo que el mínimo establecido en los programas de estudio a nivel de pregrado. Los exámenes se usan para tratar de estimar los niveles mínimos de preparación, a pesar de que la certificación de conclusión de secundaria de los aspirantes confirma una preparación supuestamente adecuada. También, se usan para valorar la aprestamiento para aprender (o el potencial) de los aspirantes quienes no han estado expuestos necesariamente a una preparación adecuada o cuyos resultados de conclusión de secundaria no existen o no se pueden interpretar en los contextos en los cuales buscan acceso. Este último enfoque es especialmente relevante para los números crecientes de casos a nivel internacional en los cuales los estudiantes entran a la educación superior con antecedentes y niveles de preparación muy variados.

En este contexto, se resalta una segunda dimensión del desafío que enfrentan las instituciones de educación superior con respecto a la admisión, permanencia y graduación de los estudiantes. Es cierto que, en muchas partes del mundo (ciertos países asiáticos con bajos niveles de poblaciones migrantes pueden ser la excepción), las instituciones de educación superior enfrentan cada vez más el desafío de que los aspirantes que buscan ingresar a ellas poseen muy **diversos antecedentes educativos y vivenciales** (Terenzini et al, 1996; Pascarella & Terenzini, 1998; Hill, 2002). Varios factores universales parecen haber contribuido a esta diversidad:

- la masificación de la educación superior en los últimos 30 años;
- coyunturas económicas, que han llevado a índices de participación en la educación superior cada vez mayores;
- la migración a nivel mundial y el 'mercadeo' y la portabilidad de las ofertas de educación superior;
- la globalización tecnológica y la universidad 'virtual';

- la creciente participación en la educación superior de estudiantes adultos y estudiantes 'no tradicionales';
- la participación en el aprendizaje para toda la vida;
- la necesidad de justicia y equidad social.

En resumen, esta diversidad se puede ver en términos de características demográficas tales como grupo étnico, clase social, idioma, origen geográfico, experiencia e historia educativa, incluyendo la exposición previa a la educación superior. La dimensión de la diversidad presenta desafíos particulares para las instituciones de educación superior en términos de crear una misión y política para responder a esta diversidad. En términos sencillos, las universidades deben decidir si consideran importante reflejar la diversidad de las poblaciones nacionales (y regionales) en las composiciones demográficas de sus cuerpos estudiantiles. Y también tienen que decidir sobre el alcance hasta el cual desean ampliar el acceso a la diversidad de estudiantes que plantean retos para sus políticas y prácticas de enseñanza, aprendizaje y evaluación. Si deciden encarar esta diversidad, esto crea desafíos específicos en las instituciones para la admisión de grupos estudiantiles que pueden no estar preparados desde el punto de vista educativo para enfrentar las demandas de la educación superior, comparados con aquellos estudiantes que han sido educados en contextos de clase media con buenos recursos y bien desarrollados.

Está claro, a partir de la experiencia de países donde se ha promovido activamente la ampliación del acceso, como por ejemplo, Sudáfrica, que los estudiantes de contextos educativos con pocos recursos no pueden considerarse, en términos de preparación, en el mismo nivel que aquellos que vienen de contextos con recursos apropiados (Cliff, 2007; Cliff & Hanslo, 2009). Los factores siguientes caracterizan el perfil del estudiantado que presenta antecedentes con desventajas educativas. Es importante notar que los antecedentes de tales estudiantes no reflejan necesariamente todas estas características, pero a menudo incluyen por lo menos tres o cuatro de ellas. Asimismo, se debe notar que en los estudiantes con 'desventajas educativas' influyen tanto la presencia como el alcance de estas características en sus contextos. Ningún factor por sí mismo significa necesariamente que los estudiantes tengan desventajas educativas; además, debe existir evidencia sistemática y sustancial del impacto pernicioso considerable de uno o más de estos factores para que los antecedentes de los estudiantes se clasifiquen como con "desventajas educativas".

A continuación se presentan algunos factores asociados con la desventaja educativa en países como Sudáfrica y Costa Rica:

- Educación primaria y secundaria con pocos recursos (el estudiante viene de una escuela donde los recursos físicos y educativos han sido limitados, han estado ausentes o, cuando están presentes, son subutilizados; las escuelas tienen docentes mal preparados; los niveles de motivación y aspiración son bajos; la cultura de aprendizaje está casi ausente);
- Padres y madres de familia con bajos niveles de ingresos y educación;
- Los estudiantes provienen de minorías étnicas, sociales, lingüísticas o culturales cuando se les compara contra las normas prevalecientes en las instituciones a las que desean tener acceso,
- Los estudiantes provienen de comunidades rurales donde los niveles de desarrollo y educación varían de manera considerable con respecto al contexto al cual desean ingresar;
- Existen altos niveles de subdesarrollo socio-económico (por ejemplo, pobreza) y tensión psico-social (por ejemplo, criminalidad y violencia) en las comunidades donde viven estos estudiantes;

- Los estudiantes a menudo son la primera generación en sus familias que ingresan a la educación superior.

Dadas las características anteriores, debe considerarse como admirable y prometedor que muchos estudiantes con tales antecedentes tengan buen rendimiento académico, altos niveles de motivación y aspiración de ingresar a la educación superior y busquen contribuir a su propio desarrollo y mejorar así el de las comunidades de las cuales provienen. También está claro (con base en las pruebas que buscan identificar la presencia de talento académico) que muchos de estos estudiantes tienen la capacidad de enfrentarse a las demandas de la educación superior (Zaaiman et al, 2000; Bryson et al, 2002). Por razones que se relacionan con las necesidades de desarrollo social, equidad y justicia social y la necesidad de reflejar las características demográficas de la población nacional en las características demográficas de las universidades, tales estudiantes deben ingresar a la educación superior (Scott, Yeld & Hendry, 2007). Si se asume que el talento académico está distribuido normalmente en poblaciones nacionales, también se debe asumir que el desempeño educativo de los estudiantes con desventajas educativas se ha visto seriamente comprometido, en ciertos casos, por razones que no tienen nada que ver con sus habilidades innatas.

Una tercera dimensión del desafío de seleccionar, retener y graduar a los estudiantes de la educación superior se relaciona con la **necesidad de asegurar formas apropiadas y adecuadas para impartir el currículo** con el fin de apoyar la enseñanza, el aprendizaje y la evaluación. Dados los niveles de sub-preparación de muchos estudiantes, las dificultades que yacen con la interpretación del desempeño formal al terminar la secundaria; la identificación, mediante pruebas, del 'potencial' académico, las necesidades de aprendizaje de los estudiantes y la diversidad y amplia variedad de antecedentes educativos de los estudiantes que solicitan ingreso a la educación superior, es imperativo que el currículo sea dinámico y responda a dichos factores (Burch et al, 2006). Asimismo, los niveles altos y crecientes de 'deserción' estudiantil de la educación superior en muchos países sugieren, por lo menos, que el apoyo a la enseñanza y el aprendizaje debe aumentarse y mejorarse con el fin de que las naciones aprovechen al máximo el capital educativo latente o desarrollado en sus estudiantes de educación superior (Astin, 1993; Grayson, 2003; Cliff et al, 2003).

Dados los altos niveles de 'deserción' entre los estudiantes de educación superior, simplemente no tiene sentido suponer que solo es responsabilidad de los estudiantes mejorar sus propias deficiencias aparentes y también está claro que muchos estudiantes muy buenos no se están graduando necesariamente en el tiempo estándar (mínimo) en los programas de grado. Esto se da especialmente en los estudiantes con antecedentes de desventajas educativas, pero no solo en ellos. Para tales estudiantes, crear oportunidades mediante, por ejemplo, pruebas de 'potencial' para que tengan acceso a la educación superior, es solo un aspecto de un continuum de respuestas, pues se requiere que sean:

- seleccionados con criterios apropiados y defendibles;
- ubicados en programas de estudio adecuados y mejorados; desarrollados bajo la suposición de que los niveles de estos estudiantes no reflejan necesariamente su capacidad de enfrentarse a los estudios en la educación superior;
- apoyados integralmente mediante un currículo que brinde toda la gama de sistemas académicos, psico-sociales, de pares, de mentores, de tecnología de información y de apoyo residencial y financiero;

- rastreados y monitoreados durante todos sus estudios de grado, de modo que se les pueda apoyar cuando sus estudios flaqueen por razones académicas u otras, bajo la suposición de que tales estudiantes pueden requerir la duración total de sus estudios para convertirse en alumnos autocontrolados y autónomos.

Casi no hace falta decir que las formas anteriores de prestación curricular pueden ser apropiadas para todos los estudiantes, no solo para los estudiantes con antecedentes de desventajas educativas. No obstante, la prestación de un currículo adecuado no es un asunto de elegir entre unos y otros: no es para estudiantes con antecedentes de 'ventajas' educativas 'convencionales' o para estudiantes con antecedentes de 'desventajas educativas'. En un ambiente de educación superior con recursos limitados, esto supuestamente significa conseguir financiamiento adicional, con la justificación apropiada para ello, con el fin de brindar un currículo para estudiantes con antecedentes educativos con desventajas. Típicamente, tales recursos de financiamiento se pueden conseguir mediante individuos o entidades financieras donantes internacionales o nacionales, la industria, las corporaciones, las agencias de desarrollo social, las agencias gubernamentales, las organizaciones no gubernamentales, etc. Una vez que se han desarrollado, implementado y evaluado como exitosos los programas para estudiantes con antecedentes de desventajas educativas, se pueden desarrollar motivaciones dentro de las instituciones – o fuera de ellas – para el financiamiento adicional y la integración de tales programas en la vida académica de las instituciones.

Pasando ahora al caso específico de la situación en Costa Rica, hay cada vez más evidencias de inequidades crecientes, puesto que, exceptuando a la dimensión de género, en años recientes Costa Rica ha experimentado descensos continuos en sus indicadores de equidad. Se revela una desigualdad cada vez mayor entre diferentes grupos socioeconómicos y geográficos. El coeficiente de Gini es empleado con el fin de medir desigualdad en los ingresos, oscila entre 0 y 1, y cuanto más cercano a 0 más igualdad existe, por el contrario, cuanto más se acerque a 1 mayor desigualdad habrá. El índice de Gini es el cálculo del coeficiente multiplicado por 100. El Cuadro 1 presenta este indicador para algunos países, se puede observar el alto índice de Costa Rica, que supera a México, Estados Unidos y España.

Cuadro 1. Coeficientes de Gini para países seleccionados

País	Año de la encuesta	Índice
Sudáfrica ^b	2000	57.8
Brasil ^a	2004	57.0
Costa Rica ^a	2003	49.8
México ^b	2004	46.1
EEUU ^a	2000	40.8
España ^a	2000	34.7

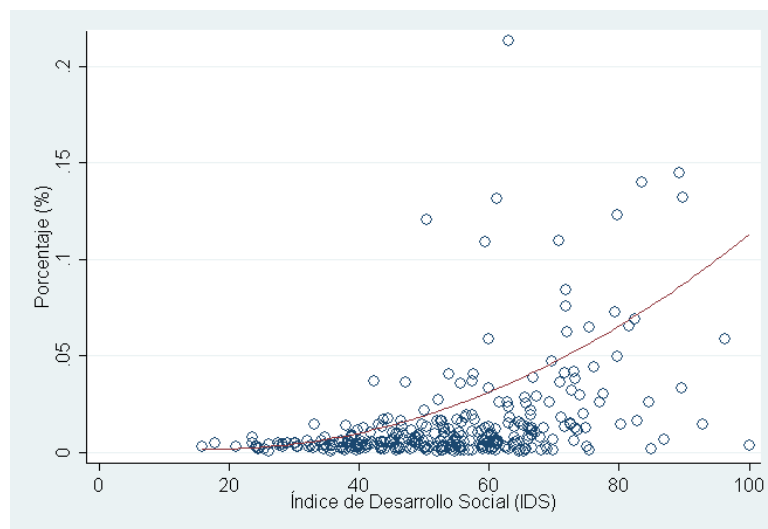
^a Los datos se refieren a la distribución del ingreso por percentiles de la población, ordenados por el ingreso per cápita.
^b Los datos se refieren a la distribución del gasto por percentiles de la población, ordenados por el gasto per cápita.
Fuente: Programa de las Naciones Unidas para el Desarrollo, Informe sobre Desarrollo Humano 2007-2008.

Según datos tomados de la Encuesta de Hogares de Propósitos Múltiples (Instituto Nacional de Estadística y Censos, 2008) la población masculina que asiste a la universidad es de 5.61% y la población femenina es de 6.83%. En un promedio realizado para los años del 2003, 2005 y 2006, el Estado de la

Educación reporta que el 78% de estudiantes que se encuentran en la educación superior provienen de los dos quintiles de ingreso superior (Programa del Estado de la Nación, 2008). A este indicador se le une otro aspecto relevante sobre el acceso a la educación universitaria, el informe indica, con respecto al nivel educativo del hogar "cerca del 40% de los alumnos procede de hogares cuyos jefes poseen formación universitaria". (p.94).

Para el caso específico de la Universidad de Costa Rica (UCR) y discutiendo el tema de la representatividad de la población costarricense en dicha institución, se presentan el gráfico 1 y el cuadro 2. En ellos se analiza la relación entre la participación relativa de estudiantes en la UCR y el Índice de Desarrollo Social del distrito al cual pertenece el colegio de procedencia de esos estudiantes. El Índice de Desarrollo Social (IDS), calculado por el Ministerio de Planificación costarricense (MIDEPLAN) "Es un índice resumen que se construye a partir de un conjunto de indicadores socioeconómicos (MIDEPLAN, 2007). Toma en cuenta cuatro dimensiones: económica, participación social, salud y educación. Se calcula a nivel distrital, siendo los distritos las unidades geográfico-administrativas en que se subdividen los cantones o municipios del país. Se analiza el porcentaje de población entre 17 y 24 años que está admitida en la UCR en el año 2009 según distrito del colegio de procedencia y según el valor del IDS.

Gráfico 1. Porcentaje de estudiantes que fueron admitidos en la Universidad de Costa Rica versus IDS, según distrito. Año 2009



Fuente: Elaboración propia con las Proyecciones de Población del Centro Centroamericano de Población de la Universidad de Costa Rica y datos suministrados por la Oficina de Registro de la UCR.

Cuadro 2. Población admitida en la UCR por cada 10000 personas con edades entre 17 y 24 años, según rangos del Índice de Desarrollo Social. 2009

Categoría	Población meta (por cada 10000) admitida en la UCR
Distritos con IDS menor o igual a 50%	55,14
Distritos con IDS mayor a 50%	126,72
Distritos con IDS menor o igual a 30%	32,76
Distritos con IDS mayor o igual a 70%	118,46
Número de distritos con IDS menor o igual a 30 = 17	
Número de distritos con IDS mayor o igual a 70 = 194	
Número total de distritos = 463	
Fuente: Elaboración propia con datos suministrados por la Oficina de Registro de la UCR y las Proyecciones de Población del Centro Centroamericano de Población.	

Esta información sugiere la necesidad y pertinencia de un programa para mejorar los niveles de representatividad de la población nacional en la UCR, pues la participación de la población entre 17 y 24 años en la Universidad de Costa Rica difiere grandemente según el nivel de desarrollo social de los diversos distritos.

Por ejemplo, resulta evidente la disparidad en términos de representación en la UCR para los jóvenes que residen en los distritos con menor desarrollo social, pues se estima que un joven entre 17 y 24 años proveniente de un distrito con IDS mayor a 70% tiene una probabilidad casi cuatro veces (3.61) mayor a la probabilidad asociada a un joven que proviene de un distrito con IDS menor a 30%.

2. LAS PROPUESTAS

Ante la necesidad de lograr que la población estudiantil en educación superior represente de mejor manera la diversidad de la población general de cada país y poder identificar de forma más precisa estudiantes con potencial intelectual para la educación superior que provienen de ambientes con desventajas educativas, tanto la Universidad de Costa Rica como la Universidad de Ciudad del Cabo en Sudáfrica han generado propuestas que intentan conciliar dimensiones de excelencia y equidad. Ambas se caracterizan, igualmente, por tratar de enfrentar científicamente la problemática. Mientras que en la Universidad de Costa Rica (UCR) se trata de un proyecto de reciente creación y que está aún en su fase de investigación, en la Universidad de Ciudad del Cabo (UCT) se han venido desarrollando y aplicando estos procedimientos por más de 20 años como parte de un programa alternativo de Admisión.

Ambas iniciativas buscan contribuir de manera efectiva a mejorar la representatividad de diversos sectores de población en sus Universidades y coadyuvar así en los procesos de desarrollo de sus respectivos países, sobre todo, porque pretenden lograr posibilidades más equitativas en cuanto a ingreso y permanencia. La meta es detectar a estos estudiantes que no han tenido contacto con la Academia y que por su situación socioeconómica o cultural, ni siquiera contemplan el ingreso a la Universidad como una posibilidad en sus vidas, esta es la población meta de estos proyectos. A continuación se presenta una descripción de ambas propuestas.

2.1. Sudáfrica: Programa de Admisión Alternativo de la Universidad de Ciudad del Cabo

La Universidad de Ciudad del Cabo en Sudáfrica cuenta con una experiencia de más de 20 años en el desarrollo e implementación de un programa alternativo de Admisión para estudiantes con desventajas educativas, cuyo coordinador actual y co-fundador es el Dr. Alan Cliff.

Dicha universidad ha tomado un camino bastante novedoso al no concentrarse en la "pureza psicométrica" de las pruebas, sino más bien enfocarse en sus propósitos educativos. Han tomado como referencia a Vygotsky (1978) y sus conceptos de zona de desarrollo próximo y "andamiaje" para construir pruebas que cumplen simultáneamente propósitos de enseñanza, selección y diagnóstico. El modelo fundamental subyacente es uno de evaluación dinámica (Sternberg & Grigorenko, 2002).

Se utilizan actualmente tres pruebas: *Alfabetismo Académico*, *Razonamiento Científico* y *Comprensión Matemática*. Estos instrumentos tienen la característica de que los conceptos más complejos se van "enseñando" a lo largo del desarrollo de la prueba. Además, los ejercicios que debe realizar el estudiante son bastante auténticos en el sentido de que se trata de simular, en la medida de lo posible, el tipo de contextos y tareas académicas que serán demandados del estudiante en educación superior. Otra característica interesante es que estas pruebas dependen, en buena medida, de habilidades de comprensión lectora, en este caso en idioma inglés o afrikáans, por ser estos los medios de instrucción de la educación superior en Sudáfrica. Por otra parte, el enfoque sudafricano, aunque requiere el conocimiento del lenguaje para realizar los exámenes, tiene características muy deseables desde una perspectiva educativa: ser también instrumentos de enseñanza, de diagnóstico, y "simular" demandas académicas a las que se enfrentarán los estudiantes. Estas características hacen de estas pruebas particularmente atractivas en términos de su posible valor agregado, especialmente refiriéndonos al caso de estudiantes con desventajas educativas.

El objetivo principal del Proyecto de Investigación de Acceso Alternativo (Alternative Access Research Project, AARP) de la Universidad de Cape Town en Sudáfrica (UCT) es crear tests de alfabetismo académico, matemática y razonamiento científico (Cliff & Hanslo, 2005; Cliff & Yeld, 2008). Se da un fuerte énfasis en identificar a los estudiantes de ingreso alternativo o en desventaja educativa que tienen el potencial de tener éxito en la UCT, brindando apoyos académico y estructuras de curriculares para cubrir las necesidades de aprendizaje de dichos estudiantes.

Dentro de un paradigma de pruebas dinámicas (Sternberg & Grigorenko, 2002) se busca desarrollar tests de razonamiento científico, matemático y del idioma, que sean predictores válidos del desempeño en contextos educativos, especialmente para estudiantes que provienen de ambientes con desventajas educativas.

Los objetivos del programa se pueden resumir así:

- Identificar, en particular, estudiantes talentosos pero que provienen de ambientes con desventajas educativas y cuya educación puede que no les haya permitido demostrar su verdadera capacidad.
- Garantizar que los estudiantes que se identifican con potencial, de acuerdo con los resultados de los tests, también reciban ofertas de apoyo y ubicación académica (tales como ayudas financieras, residencia, orientación social y psicológica, entre otras).

- Contribuir al diagnóstico de las fortalezas y debilidades académicas de todos los estudiantes que ingresan a la UCT, para poder nutrir el diseño de apoyos de aprendizaje y enseñanza y el desarrollo curricular.

2.1.1. Descripción de los tests de AARP

El test de alfabetismo académico o PTEEP por sus siglas en inglés (Cliff, 2002; Cliff, Yeld & Hanslo, 2003), tiene como objetivo evaluar las capacidades de los estudiantes en términos de dimensiones tales como: dar importancia a los textos que es probable que se encuentren en sus estudios; entender palabras y giros de discusión en sus contextos; identificar y seguirle la pista al razonamiento académico; entender y evaluar la base de evidencia del razonamiento; extrapolar y sacar evidencias y conclusiones a partir de lo que está indicado o dado; identificar ideas principales de las secundarias en la organización general de un texto; comprender información presentada visualmente (v.g. gráficos, tablas, etc.); y entender conceptos numéricos e información básica utilizados en el texto, incluyendo manipulaciones numéricas básicas.

El test de comprensión matemática (MCOM) está diseñado para proporcionar información acerca del potencial de un candidato de aprender nuevos conocimientos y habilidades matemáticas. Está diseñado con el principio de test dinámico (Sternberg & Grigorenko, 2002). Los siguientes criterios se han incorporado como parte de su diseño:

- Los temas, no incluidos en los programas de estudio de la secundaria, se seleccionan de forma que los candidatos que hayan tenido una educación deficiente no estén en desventaja.
- Los temas incorporan una variedad de hechos, habilidades, conceptos y principios que pueden ser evaluados en un rango de niveles cognitivos.
- El texto es usado como en el ambiente de enseñanza.
- Los temas minimizan la necesidad de destrezas matemáticas adquiridas en la educación formal, de manera que los candidatos que provienen de ambientes con menos oportunidades educativas no estén en desventaja.
- Las preguntas están clasificadas cuidadosamente y en una secuencia de acuerdo al grado de complejidad.
- El lenguaje utilizado en el texto es lo suficientemente simple para no generar desventajas en los candidatos con inglés como segunda y tercera lengua.

El test de razonamiento científico (SRT por sus siglas en inglés): Este test recientemente desarrollado – utilizado por primera vez en el 2003 – tiene el propósito de evaluar la capacidad de los estudiantes de involucrarse con el pensamiento lógico requerido típicamente por los estudiantes de Educación Superior. Esta prueba de razonamiento científico, evalúa la capacidad de los estudiantes de identificar la evidencia apropiada para apoyar una afirmación o un argumento; criticar suposiciones y pensamiento que subyacen al argumento; comprender la naturaleza del razonamiento tentativo y decisivo; ser capaz de ver las relaciones entre fenómenos; y entender el concepto de “azar”.

Los tests del AARP han contribuido a la equidad de acceso al asegurar que a los estudiantes con talento que provienen de ambientes con desventajas educativas se les dé la oportunidad de demostrar su potencial para la educación superior, aún cuando los resultados de sus exámenes de conclusión de estudios secundarios no necesariamente son un buen reflejo de ese potencial.

El desarrollo de estas pruebas ha contribuido a identificar el potencial académico del estudiante en dimensiones como razonamiento verbal, conocimientos básicos de aritmética y pensamiento matemático

y científico genéricos que se requieren en un estudiante a nivel de ingreso, en un ambiente de instrucción de un contexto particular de educación superior.

El constructo de alfabetismo académico que se articuló por primera vez por el AARP (y, en particular, por el profesor asociado Nan Yeld (Yeld, Cliff & Hanslo, 2002), con la colaboración de miembros del Grupo de Desarrollo del Lenguaje del Programa de Desarrollo Académico de la UCT, es una contribución substantiva a la Educación Superior en Sudáfrica, una contribución que se ha extendido a nivel nacional y que está generando creciente atención internacionalmente.

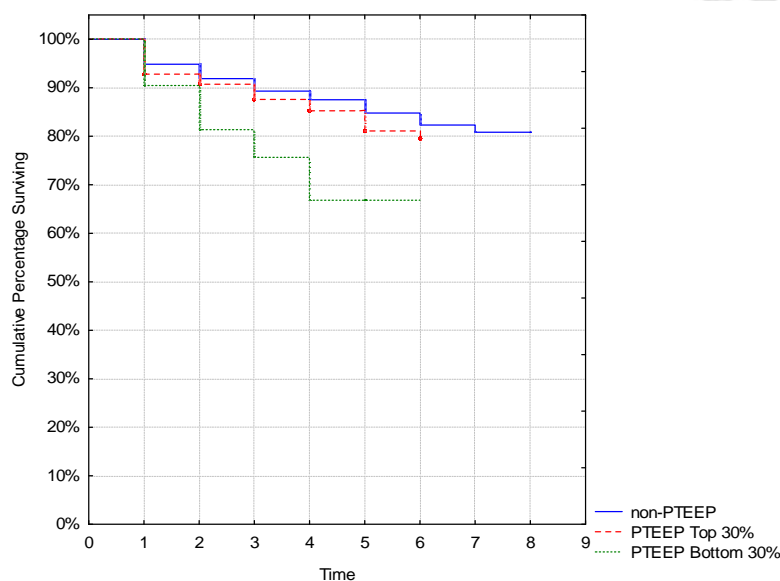
La contextualización e interpretación de las pruebas del AARP en relación con los antecedentes educativos y orígenes demográficos del estudiante, representan otra contribución poderosa del AARP a la identificación de talento educativo en Sudáfrica. El reporte de resultados en los tests del AARP ha disminuido las posibilidades que los estudiantes talentosos con pocas y pobres oportunidades educativas se "pierdan" por la competencia injusta con estudiantes de orígenes bien provistos de recursos educativos. (Cliff, Ramaboa & Pearce, 2007)

Para todos los estudiantes que buscan ingresar a la Educación Superior, los tests del AARP han contribuido a entender cómo se expresan sus niveles de preparación en constructos tan cruciales como razonamiento genérico científico, matemático y del lenguaje en los aspirantes a ingresar a estudios superiores.

A lo largo de los años se han realizado estudios predictivos relacionando el desempeño de los estudiantes en los tests del AARP y su desempeño académico subsecuente (Cliff & Yeld, 2008). Estas evaluaciones han mostrado, por medio del uso de métodos estadísticos tales como análisis de sobrevivencia, que se puede predecir la probabilidad de la progresión exitosa de los estudiantes a través de sus estudios académicos. Los gráficos 2 y 3 muestran funciones de sobrevivencia para estudiantes con ventajas y desventajas educativas respectivamente. A nivel individual del estudiante, las valoraciones que provee el AARP también permiten calcular el grado de "riesgo" académico adjunto al desempeño en un test para un estudiante particular, y dar recomendaciones con respecto a las disposiciones de currículum que se requieren para disminuir ese riesgo.

Esta figura muestra el tiempo de exclusión en términos académicos para estudiantes que provienen de ambientes escolares de buenos recursos.

Gráfico 2. Funciones de sobrevivencia de estudiantes con ventajas educativas para diferentes grupos de la prueba de alfabetismo académico (PTEEP por sus siglas en inglés)

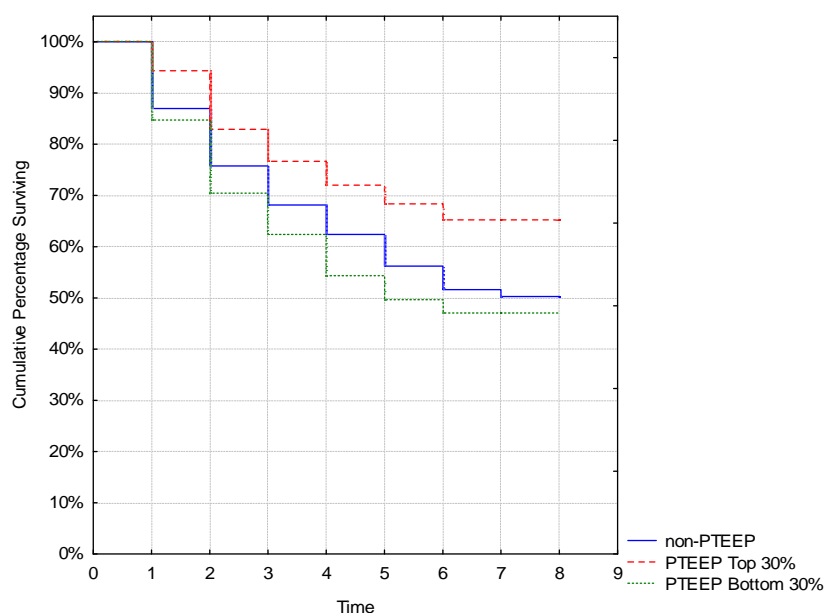


Notas:

- (1) El tiempo está dado en años.
- (2) "Non-PTEEP" se refiere a aquellos estudiantes que entraron a la Universidad sin haber tomado las pruebas del programa alternativo de admisión, "PTEEP Top 30%" se refiere a aquellos estudiantes que obtuvieron puntajes en los 3 deciles superiores de la distribución de puntajes en la prueba de Alfabetismo Académico, en el grupo que proviene de ambientes escolares de buenos recursos; "PTEEP Bottom 30%" se refiere a aquellos estudiantes que obtuvieron puntajes en los 3 deciles inferiores de la distribución de puntajes en la prueba de Alfabetismo Académico, en el grupo que proviene de ambientes escolares con buenos recursos.

La siguiente figura muestra un patrón un tanto diferente para estudiantes que provienen de ambientes educativos con escasos recursos.

Gráfico 3. Funciones de sobrevivencia de estudiantes con desventajas educativas para diferentes grupos de la prueba de alfabetismo académico (PTEEP por sus siglas en inglés)



Lo que parece mostrar el gráfico anterior es que, para estudiantes que provienen de ambientes con desventajas educativas, un buen desempeño en la prueba de Alfabetismo Académico se correlaciona mejor con el progreso académico, comparado con un pobre desempeño en esta misma prueba o en las pruebas de salida de la secundaria. La implicación es que un buen desempeño en esta prueba es una evaluación más confiable que las pruebas de salida de secundaria en términos de las probabilidades de progreso académico para los estudiantes que provienen de ambientes con desventajas educativas.

Estos análisis de sobrevivencia han sido replicados ya dos veces y los patrones que se muestran para la prueba de Alfabetismo Académico son similares a los que presenta la prueba de Comprensión Matemática.

Hay un beneficio claro para los estudiantes prometedores cuyo potencial puede identificarse mejor con base en su desempeño en los tests del AARP. La evaluación del AARP hace una contribución a la ampliación del acceso, particularmente para estudiantes que provienen de orígenes con desventajas educativas, pero también para estudiantes con orígenes educativos más "tradicionales" o de clase media.

2.2. Costa Rica: Proyecto de Pruebas Específicas de la Universidad de Costa Rica

La Universidad de Costa Rica, en su Estatuto Orgánico (1974), establece como propósito "obtener transformaciones que la sociedad necesita para el logro del bien común, mediante una política dirigida a la consecución de una verdadera justicia social, del desarrollo integral, de la libertad plena y de la total independencia de nuestro pueblo". (Artículo 3, Estatuto Orgánico).

De acuerdo con el modelo de admisión establecido por la Universidad de Costa Rica (UCR), el propósito de las pruebas de ingreso a la Universidad es medir el potencial académico de los estudiantes para la educación superior. Como universidad pública, la más antigua y prestigiosa del país, la UCR recibe continuamente presiones de diferentes sectores de la sociedad en torno a las dimensiones de excelencia y equidad en la admisión. De igual forma, hay una preocupación intrínseca en nuestra Institución, ante las crecientes desigualdades económicas y sociales, por mejorar la representatividad de diversas categorías de población estudiantil, logrando así contribuir de manera óptima con las metas de integración y movilidad social que son parte de su misión.

En este contexto, se justifica la creación de propuestas de investigación para diseñar modelos e instrumentos de admisión que, sin perder su carácter científico, tomen, de partida, a la equidad como uno de sus pilares básicos, conjuntamente con la excelencia. Enfrentar este reto es esencial, en los tiempos que corren, para que la Universidad pueda, efectivamente, ser capaz de contribuir a los procesos de movilidad e integración social que son parte de su misión fundamental.

Entonces en el marco de la equidad y de las políticas de integración y movilidad social, se buscan propuestas que permitan llevar a cabo una medición basada en instrumentos que posean altos índices de confiabilidad y validez, instrumentos que estén influenciados, en la menor medida posible, por el contexto socio cultural del estudiante y que permitan a la Universidad identificar a estudiantes con potencial académico, pero que no han tenido oportunidades educativas óptimas.

Es importante validar estos instrumentos de medición para probar científicamente que las diferencias de oportunidades pueden minimizarse, y de esta manera lograr un mayor acercamiento al potencial académico como constructo clave de la medición. De este modo se podrían identificar estudiantes con potencial, pero que por presentar desventajas educativas posiblemente se "perderían" al utilizar instrumentos de medición más tradicionales.

Así, es importante para la institución rescatar a estos estudiantes y brindarles la posibilidad de marcar una diferencia en sus familias y en sus comunidades, esto, acorde con el Estatuto Orgánico, que afirma: "...alcanzar niveles de excelencia en la formación de profesionales, que a su vez actúan como difusores y agentes de cambio en la comunidad en general". En la medida en que estos estudiantes sean estudiantes exitosos, serán ejemplo para otros jóvenes que se propongan como meta acceder a la educación superior. Cabe resaltar, que para estos estudiantes es necesario dar apoyo no sólo económico, también hay que velar por darles un acompañamiento integral que les permita continuar en la Universidad y, finalmente, graduarse.

Este acompañamiento con apoyo integral debe existir para que los estudiantes puedan pasar exitosamente por el proceso de adaptación a la Universidad. Aunque el apoyo económico es indispensable, no es suficiente, ellos requieren de tutores que los respalden y les ayuden a suplir sus carencias en cuanto a conocimientos formales (técnicas de estudio, características de la presentación de

un trabajo, técnicas básicas de investigación, cómo transitar y aprovechar los recursos de la Universidad, etc.) para evitar que se sientan frustrados(as) y deserten.

El proyecto de Pruebas Específicas del Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica (Montero, 2009) nace en el año 2005 respondiendo a un acuerdo del Consejo Universitario que data del año 2003. Desde su inicio el proyecto se ha ido enriqueciendo gracias a los aportes de la comunidad universitaria, incluyendo los diferentes niveles de la Administración, el Consejo Universitario, el Instituto de Investigaciones Psicológicas y las mismas unidades académicas.

A los objetivos planteados inicialmente y que se referían a construir y validar instrumentos para la selección de estudiantes en áreas y carreras específicas, se han añadido otros dos, uno de ellos tiene que ver con la necesidad de que las pruebas cumplan también un propósito diagnóstico, permitiéndole al estudiante (y a la Universidad) realizar una valoración cualitativa de su desempeño. Así, se busca brindarle a cada examinado no solo un puntaje numérico, sino una descripción de sus áreas fuertes y de sus áreas de oportunidad en términos de los contenidos medidos en las pruebas.

Esta misma información diagnóstica resulta también de relevancia para la institución, tanto a nivel agregado para toda la población de aspirantes, como para comparar los desempeños relativos de diversas categorías de población (v.g. estudiantes de colegios públicos y de colegios privados), y permite contar con una visión más detallada y descriptiva de las habilidades y conocimientos, presentes en las pruebas, que muestra la población de aspirantes a ingresar a las carreras participantes en el proyecto.

El otro objetivo principal que se ha añadido al proyecto, gracias a las recomendaciones de la comunidad universitaria, se relaciona con el principio de equidad. Por un lado especifica que las pruebas, en la medida de lo científicamente posible, deben contribuir a disminuir las brechas de oportunidades educativas existentes entre diferentes grupos de población, al ser capaces de identificar el potencial académico en jóvenes que provienen de ambientes con desventajas económicas, sociales y culturales. Por otro lado, también se reconoce que, para realizar un exitoso balance en ambas dimensiones (excelencia y equidad), no solo es necesario construir instrumentos de selección y de diagnóstico que permitan medir el potencial académico, sino también, en el contexto actual, se requiere el diseño de programas complementarios al proceso regular de admisión que permitan ofrecer opciones académicas a jóvenes con potencial académico, pero que provienen de ambientes con desventajas educativas.

De esta manera se ha incorporado la equidad como uno de los pilares fundamentales del proyecto, junto con la excelencia.

Estos dos objetivos del proyecto de Pruebas Específicas se pueden resumir así: (1) Lograr que las pruebas cumplan una labor diagnóstica permitiéndole al estudiante y a la universidad realizar una valoración cualitativa de su desempeño y (2) Contribuir a los procesos de movilidad e integración social que son parte de la misión fundamental de la universidad diseñando estrategias e instrumentos que permitan identificar estudiantes con potencial académico para la universidad, pero que han tenido menos oportunidades educativas en sus experiencias previas.

Entre los impactos deseados del proyecto a mediano plazo se incluyen una disminución en los porcentajes de deserción estudiantil y mejoramientos en el rendimiento académico y en la promoción, y también contribuir a los procesos de movilidad e integración social que son parte de la misión fundamental de la UCR.

Los referentes teóricos para la investigación parten de la psicometría y la psicología cognitiva (Nunnally, 1991; De Juan Espinoza, 1997; Bond & Fox, 2001; Martínez, 2005; Sternberg & Pretz, 2005; Montero, 2001; Montero, 2008a). Aunque puede decirse que tradicionalmente estos dos paradigmas se han desarrollado separadamente, en la actualidad hay propuestas que plantean posibilidades para su integración (Embretson & Gorin, 2001; Junker & Sijtsma, 2001; Cortada de Kohan, 2003; Leighton & Gierl, 2007). En nuestro caso particular, creemos que dicho enfoque nos ayudará mucho en el proceso de brindar interpretaciones diagnósticas cada vez más detalladas y precisas. Se trata de estudiar, analizar y probar empíricamente hipótesis sobre los diferentes procesos cognitivos involucrados en la resolución de los ítems, así como identificar características de los ítems que permitan predecir sus niveles de dificultad.

El análisis de validez predictiva, como ya se mencionó, es también fundamental si se desea que estas pruebas puedan usarse como instrumentos de selección en diversas áreas y carreras. Se busca establecer el grado de asociación entre el puntaje de la prueba y una medida apropiada de rendimiento académico. La validez predictiva implica que altos puntajes en las pruebas estarán asociados a altos niveles de rendimiento. En la actualidad se utilizan modelos de regresión y modelos de ecuaciones estructurales para estimar el grado de validez predictiva de los instrumentos.

A la fecha se han construido y recolectado diversas evidencias de validez para tres pruebas: Expresión Escrita (EE), Habilidades Cuantitativas (HC), y Razonamiento con Figuras (PRF). Además se realizó una aplicación operacional en población meta, en octubre del 2008, con propósitos diagnósticos.

La prueba de Expresión Escrita (EE) mide, con ítems de escogencia única, el uso correcto del lenguaje en textos escritos, desde el enfoque normativo. Los conocimientos básicos requeridos para esta prueba se clasifican en tres áreas: léxico, ortografía y morfosintaxis. Su tiempo de realización es de una hora.

La prueba de Habilidades Cuantitativas (HC) mide razonamiento basado en relaciones y propiedades matemáticas. Los procesos de razonamiento pueden ser inductivos, deductivos o cualquier combinación de ambos. Se presenta en cuatro áreas de contenido: Geometría, Álgebra, Aritmética y Análisis de Datos y Probabilidad. Este examen se considera de importancia en aquellas carreras en donde se debe utilizar la matemática como herramienta de trabajo. Es de escogencia única y el tipo de ejercicios utilizado no requiere el uso de la calculadora. Los conocimientos requeridos para resolver los ítems no van más allá de los contenidos del programa oficial de estudios de noveno año. Su duración es de una hora y cuarenta y cinco minutos.

La Prueba de Razonamiento con Figuras (PRF) mide habilidades generales de razonamiento frente a ejercicios novedosos que presentan una serie de figuras geométricas o líneas, para identificar reglas de clasificación, cumplimiento de condiciones, completar matrices y secuencias en series de eventos. Su duración es alrededor de 40 minutos. Con ella se intenta medir el constructo denominado inteligencia fluida (Cattell, 1971). Este constructo involucra procesos básicos de razonamiento y otras actividades mentales que dependen, en menor medida, de los conocimientos escolares, las oportunidades educativas y la cultura.

La creación de la PRF responde precisamente a la dimensión de equidad, preocupación fundamental de la Universidad de Costa Rica, puesto que busca medir de manera más precisa el potencial de aprendizaje del estudiante. Su uso nos ayudaría a identificar jóvenes que poseen potencial académico, pero que han tenido menos oportunidades educativas, por provenir de ambientes con desventajas económicas y sociales. Su empleo como parte del proceso de selección para ingreso a la Universidad, también podría justificar, científicamente, el hecho de ajustar las puntuaciones de las otras pruebas según los resultados

de inteligencia fluida y, con ello, no excluir a jóvenes que tienen potencial pero que han tenido menos oportunidades de aprender. Esto porque se reconoce que puede haber personas que tengan similar potencial para aprender y cursar apropiadamente una carrera universitaria, pero que han tenido muy diferentes oportunidades educativas. La implementación de una propuesta en esta dirección también implicaría una inversión importante de la Universidad en cursos o actividades llamados de "nivelación" o de "preparación".

La inteligencia fluida se puede definir como una energía o eficacia neuronal que fluye a través de diversas actividades mentales. Involucra procesos básicos de razonamiento y otras actividades intelectuales que dependen, en menor medida, del aprendizaje y la cultura. Es decir, que dependen menos de la educación formal y dependen más de influencias fisiológicas. Ejemplos de este constructo incluyen el razonamiento inductivo y deductivo y otras destrezas asociadas con la capacidad de resolver problemas novedosos, creando nuevos conceptos (De Juan Espinoza, 1997; Carroll, 1993). Las medidas más conocidas de inteligencia fluida son la Prueba de Matrices Progresivas de Raven y la prueba de Factor g de Cattell. Estas son pruebas específicamente diseñadas para minimizar el impacto del conocimiento adquirido, incluyendo destrezas lingüísticas. (Cattell & Cattell, 2001)

Por su parte, el concepto de inteligencia cristalizada se refiere a procesos mentales que reflejan no solo operaciones de inteligencia fluida, sino también los efectos del entorno (experiencia, aprendizaje y cultura). La comprensión de lectura y el conocimiento de vocabulario tienen una fuerte carga cristalizada; igualmente las pruebas tradicionales de logro o rendimiento académico. Así, la inteligencia cristalizada está relacionada a la extensión y profundidad del conocimiento del lenguaje, el conocimiento acerca de la información y conceptos de una cultura y también el conocimiento declarativo. Se adquiere a través de la educación y la experiencia. (Carroll, 1993; De Juan Espinoza, 1997; Kvist & Gustafsson, 2007). Factores de personalidad como motivación y oportunidades educativas y culturales son clave para su desarrollo.

Obviamente también reconocemos que la distinción fluida-cristalizada no es una dicotomía sino más bien un continuum, y, en general, diferentes medidas empíricas de habilidades intelectuales y conocimientos se ubican en diferentes rangos de éste. (Plucker, 2003).

La hipótesis fundamental subyacente en el Proyecto de Pruebas Específicas es que por medio de la PRF se podrán identificar de manera más precisa estudiantes con potencial académico para la educación superior, pero que provienen de ambientes con desventajas educativas. Así, de acuerdo con este marco de referencia, es posible predecir que las diferencias en medidas de inteligencia fluida van a ser más pequeñas que las diferencias en medidas más cargadas en aspectos cristalizados, cuando se comparan grupos socioeconómicos y culturales que difieren, en promedio, en sus oportunidades educativas.

Actualmente hay dos conjuntos de evidencias empíricas que apoyan la hipótesis básica subyacente en la medición del constructo en la PRF, bajo el modelo de inteligencia fluida y cristalizada de Cattell (1971). De acuerdo con este modelo, las diferencias entre categorías de población que difieren en sus oportunidades educativas serán menores en una medición de inteligencia fluida que en una medición de inteligencia cristalizada.

La primera evidencia es brindada por la aplicación operacional de una primera versión de la PRF (Prueba de Razonamiento con Figuras) que se realizó en octubre del año 2008 con estudiantes de secundaria que deseaban estudiar alguna de 9 carreras universitarias. De igual forma también se aplicó la prueba de prueba de Expresión Escrita (EE) en formato de escogencia única ("multiple choice") que mide uso correcto del lenguaje en términos normativos.

Dado que esta última es una prueba claramente cargada en conocimientos (inteligencia cristalizada) la hipótesis era que las diferencias entre estudiantes de colegios públicos y privados, deberían ser menores en la prueba de inteligencia fluida (PRF). Los hallazgos confirmaron esta hipótesis, pues las diferencias estandarizadas medidas por el estadístico "t de Student", mostraron efectivamente valores sustancialmente más bajos en el caso de la PRF. Estos resultados se muestran en el Cuadro 3.

Cuadro 3. Resultados obtenidos por una muestra de estudiantes aspirantes a ingresar a nueve carreras. Pruebas de inteligencia fluida (PRF) y cristalizada (EE). Comparación por dependencia del colegio. Proyecto de Pruebas Específicas. Universidad de Costa Rica. 2008

	Tamaño de muestra		Media (Puntaje Rasch*)		Desviación Estándar (Puntaje Rasch*)		Estadístico t**
	Privado	Público	Privado	Público	Privado	Público	
Prueba de inteligencia Fluida (PRF)	447	1054	0,68	0,29	0,71	0,78	-9,03
Prueba de Expresión Escrita (EE)	354	867	1,10	0,49	0,90	0,78	-11,81

* El puntaje de Rasch es cero en la dificultad promedio de los ítems.
** Se usa la prueba t de Student para variancias iguales dado que la prueba de Levene de igualdad de variancias resultó no significativa.

Otra evidencia a favor de la hipótesis básica subyacente en la medición del constructo en la PRF viene dada por los resultados obtenidos en un reciente estudio realizado en estudiantes de 13 colegios de secundaria costarricenses (Castelain, 2009; Morales & Fallas, 2009).

Se seleccionó una muestra aleatoria, no proporcional, tratando de sobre-representar centros educativos pertenecientes a regiones de menor desarrollo social, considerando cinco zonas de Costa Rica: zona central (más desarrollada), zona norte, zona sur, zona caribe y zona pacífica. A los estudiantes de décimo año de cada colegio se administraron dos pruebas, la PRF y la prueba de expresión escrita antes mencionada, que tiene una fuerte carga en inteligencia cristalizada. La hipótesis nuevamente era que las diferencias entre estudiantes de colegios públicos y privados, y también por zonas deberían ser menores en la PRF. En los cuadros 4 y 5 se presentan los resultados.

En el cuadro 4 la hipótesis se confirma al ser la t de Student en la prueba de Expresión Escrita casi el doble de la t correspondiente a la PRF comparando estudiantes de colegios públicos y privados (Castelain, 2009; Morales & Fallas, 2009). Igualmente en el cuadro 5 se muestra que, mientras que las diferencias a nivel muestral en la PRF no se pueden generalizar estadísticamente a toda la población, en el caso de la prueba de Expresión Escrita, sí se presentan diferencias generalizables estadísticamente en la comparación por zonas, conformándose tres grupos de promedios, brindando nuevamente evidencia a favor de la hipótesis subyacente, aún cuando esta muestra desagregada es reducida y con muy poco poder estadístico.

Cuadro 4. Resultados obtenidos por una muestra de estudiantes de décimo año de colegios costarricenses en pruebas de inteligencia fluida (PRF) y cristalizada (EE). Comparación por dependencia del colegio. Proyecto de Pruebas Específicas. Universidad de Costa Rica. 2009

	Tamaño de muestra		Media (% de correctas)		Desviación Estándar (% de correctas)		Estadístico t*
	Privado	Público	Privado	Público	Privado	Público	
Prueba de inteligencia Fluida (PRF)	100	197	55,38	48,28	12,70	13,36	-4,41
Prueba de Expresión Escrita (EE)	134	220	55,48	44,62	12,88	11,53	-8,22

* Se usa la prueba t de Student para variancias iguales dado que la prueba de Levene de igualdad de variancias resultó no significativa.

Cuadro 5. Resultados obtenidos por una muestra de estudiantes de décimo año de colegios costarricenses en pruebas de inteligencia fluida y cristalizada. Conjuntos de promedios homogéneos según la prueba de SNK Comparación por zonas de desarrollo Proyecto de Pruebas Específicas Universidad de Costa Rica. 2009

	EE (Inteligencia Cristalizada)			PRF (Inteligencia Fluida)	
	n	Grupo 1 (% de correctas)	Grupo 2 (% de correctas)	n	Grupo 1 (% de correctas)
Sur	113	44,5401		111	47,98
Pacífico	72	46,9444	46,9444	72	49,25
Central	23		51,2253	23	56,18
Norte	70		52,0519	21	53,33
Caribe	78		52,4709	70	54,82
Sig.		,330	,114		0,53

Así, estos resultados son alentadores e indican la necesidad de continuar trabajando en esta línea, con el objeto de dar una respuesta científica al reto de maximizar excelencia y equidad en el ingreso a la educación superior.

A manera de conclusión es importante reiterar que ambos enfoques, el de Costa Rica y el de Sudáfrica, partiendo de diferentes marcos de referencia y considerando las particularidades de sus respectivos contextos, tratan de abordar la problemática de equidad en la admisión para la educación superior, desde una perspectiva científica y no solamente política. Consideramos que éste es el mayor valor agregado de ambas propuestas.

El artículo no pretende hacer un análisis comparativo para concluir cual de las dos aproximaciones es más válida, sino que busca ilustrar dos formas posibles de enfrentar el difícil reto de maximizar excelencia y equidad en la admisión a la Universidad, partiendo de un enfoque de investigación. Cada una de las propuestas ha generado y está generando diversas evidencias de su valor y ciertamente también posee limitaciones.

REFERENCIAS BIBLIOGRÁFICAS

- AARP (2009) The Alternative Admissions Research Project Tests. Alternative Admissions Research Project, University of Cape Town. Available at <http://www.aarp.ac.za/hsc/index.html> . Accessed 17 August 2009.
- Astin, A.W. (1993) *What matters in college? Four critical years revisited*. San Francisco: Jossey-Bass.
- Badenhorst, E., Cliff, A. & Kidson, S.H. (2009, bajo revision). The effects of a mediated learning programme for underprepared students in a South African Medical School. *Medical Education*.
- Bohlmann, C. & Braun, M. (2006) Domain 2 – Cognitive Academic Mathematical Proficiency. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 35-41. Pretoria: Higher Education South Africa.
- Bond, T. and Fox, C. (2001). *Applying the Rasch model. Fundamental measurement in the human sciences*. United States: Lawrence Erlbaum Associates, Inc.
- Bryson, S., Smith, R. & Vineyard, G. (2002) Relationship of race, academic and non-academic information in predicting the first-year success of selected admissions first-year students. *Journal of the First-Year Experience and Students in Transition*, 14, pp. 65–80.
- Budoff, M. (1987) Measures for assessing learning potential. In C.S. Lidz (Ed) *Dynamic assessment: an interactional approach to evaluating learning potential*, pp. 173-195. New York: The Guilford Press.
- Burch, V.C., Sikakana, C.N.T., Yeld, N., Seggie, J.L. & Schmidt, H.G. (2006). Performance of academically at-risk medical students in a problem-based learning programme: a preliminary report. *Advances in Health Science Education*, 12(3), pp.345-358.
- Campione, J.C. & Brown, A.L. (1987) Linking dynamic testing with school achievement. In C.S. Lidz (Ed) *Dynamic assessment: an interactional approach to evaluating learning potential*, pp. 82-115. New York: The Guilford Press
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Castelain, T. (2009). *Prueba de Razonamiento con Figuras: Informe de labores 2009*. San José, Costa Rica: Proyecto de Pruebas Específicas, Instituto de Investigaciones Psicológicas. Universidad de Costa Rica.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York: Houghton Mifflin.
- Cattell, R.B. & Cattell, A.K.S. (2001). *Factor "g", escalas 2 y 3*. Madrid: TEA Ediciones.
- Cliff, A. (2002). *The Placement Test in English for Educational Purposes (PTEEP): A construct analysis*. Paper presented at the annual international conference of the Association for the Study of Evaluation in Education in Southern Africa, Johannesburg, South Africa
- Cliff, A. (2007). *Possible contributions of standardized assessment tests to the development of foundation programmes*. Workshop presented at the Conversations about Foundation conference, Cape Town, South Africa.
- Cliff, A. & Hanslo, M. (2005). *The use of 'alternate' assessments as contributors to processes for selecting applicants to Health Sciences Faculties*. Invited symposium paper presented at the annual conference of the Association for Medical Education in Europe (AMEE), Amsterdam, August.

- Cliff, A. & Hanslo, M. (2009). The design and use of 'alternate' assessments of academic literacy as selection mechanisms in Higher Education. *Southern African Linguistics and Applied Language Studies*, 27 (3).
- Cliff, A., Ramaboa, K. & Pearce, C. (2007) The assessment of entry-level students' academic literacy: does it matter? *Ensovoort*, 11, pp. 33-48.
- Cliff, A. & Yeld, N. (2006) Domain 1 – Academic Literacy. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 19-27. Pretoria: Higher Education South Africa.
- Cliff, A. & Yeld, N. (2008). *The design and use of 'alternate' assessments of academic literacy as selections mechanisms in Higher Education*. Symposium paper presented at the annual AILA Conference, Essen, Germany, August.
- Cliff, A., Yeld, N. & Hanslo, M. (2003). *Assessing the academic literacy skills of entry-level students, using the Placement Test in English for Educational Purposes (PTEEP)*. Paper presented at the biennial international conference of the European Association for Research in Learning and Instruction, Padova, Italy and at the annual national conference of the South Africa Academic Development Association, Cape Town
- Cortada de Kohan, N. (2003). *Posibilidad de integración de las teorías cognitivas y la psicometría moderna*. *Revista Argentina de Neuropsicología*, 1, pp.8 –23. Extraído el 23 de marzo de 2009 desde http://www.revneuropsi.com.ar/pdf/Posibilidad_de_integracion.pdf
- De Juan Espinoza, M. (1997). *Geografía de la Inteligencia Humana: Las aptitudes cognitivas*. Madrid: Pirámide.
- Embretson, S. and Gorin, J. (2001, Winter). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement*, 38(4), pp.343-368.
- ETS-SAT. (2009) *The SAT*. Princeton, New Jersey, USA: Educational Testing Service.
- ETS-TOEFL. (2009) *Test of English as a Foreign Language (TOEFL)*. Princeton, New Jersey, USA: Educational Testing Service.
- Feuerstein, R., Klein, P. & Tannenbaum, A.J. (Eds) (1991) *Mediated learning experience (MLE): theoretical, psychosocial and learning implications*. Tel Aviv and London: Freund Publishing House
- Foxcroft, C. (2006) The nature of benchmark tests. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 7-16. Pretoria: Higher Education South Africa.
- Frith, V. & Prince, R. (2006) Domain 1 – Quantitative Literacy. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 28-34. Pretoria: Higher Education South Africa
- Grayson, J.P. (2003) The consequences of early adjustment to university. *Higher Education*, 46, pp. 411-429.
- Griesel, H. (2006) The context of the National Benchmark Tests Project. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 1-6. Pretoria: Higher Education South Africa.
- Hessels, M.G.P. & Hamers, J.H.M. (1993) The learning potential test for ethnic minorities. In J.H.M. Hamers et al (Eds) *Learning potential testing*, pp. 285-311. Amsterdam: Swets and Zeitlinger.

- Hill, C. (2002). *Linguistic and cultural diversity – a growing challenge to American Higher Education*. CSE Technical Report 556. UCLA: National Center for Research on Evaluation, Standards and Student Testing (CRESST).
- Instituto Nacional de Estadística y Censos (INEC). (2008). *Encuesta de hogares de propósitos múltiples*. <http://www.inec.go.cr/>
- Junker B. and Sijtsma K. (2001). Cognitive Assessment Models with Few Assumptions, and Connections with Nonparametric Item Response Theory. *Applied Psychological Measurement*, 25(3), pp. 258-272
- Kvist, A.V. & Gustafsson, J. (2007). *The relation between fluid intelligence and the general factor as a function of cultural background: a test of Cattell's investment theory*. WORKING PAPER 2007:23. Uppsala, Sweden: The Institute for Labour Market Policy Evaluation (IFAU).
- Leighton, J. and Gierl, M. (2007). *Cognitive diagnostic assessment. Theory and applications*. United States: Cambridge University Press.
- Martínez, R. (2005). *Psicometría: Teoría de los Tests Psicológicos y Educativos*. Madrid: Editorial Síntesis, S.A.
- Montero, E. (2001). La teoría de respuesta a los ítemes: una moderna alternativa para el análisis psicométrico de instrumentos de medición. *Revista de Matemática: Teoría y Aplicaciones*, 7(1,2)págs., pp. 217-228.
- Montero, E (2008a). Escalas o Índices para la medición de constructos: El dilema del analista de datos. *Avances en Medición*, 6, pp.15–24. Bogotá, Colombia: Laboratorio de Psicometría, Facultad de Ciencias Humanas de la Universidad Nacional de Colombia.
- Montero, E. (2008b). *The balance between excellence and equity in admission tests: Contributions of cognitive psychology in an emergent experience from the University of Costa Rica*. World Linguistic Conference, Essen, Germany.
- Montero, E. (2009). Informe ejecutivo del Proyecto No. 723-A5-192: "Construcción de Pruebas Específicas para ingreso a carrera" entregado a la Rectoría el 20 de enero, 2009. Universidad de Costa Rica.
- Morales, A. & Fallas, S. (2009). *Prueba de Expresión Escrita: Informe de labores 2009*. San José, Costa Rica: Proyecto de Pruebas Específicas, Instituto de Investigaciones Psicológicas. Universidad de Costa Rica
- Nunnally, J. (1991). *Teoría Psicométrica*. México : Trillas
- Pascarella, E.T. & Terenzini, P.T. (1998) Studying college students in the 21st century: meeting new challenges. *The Review of Higher Education*, 21, pp.151-165.
- Pitoniak, M.J., Cliff, A. & Yeld, N. (2008) Technical report on the Standard Setting process for Higher Education South Africa's National Benchmark Tests – Academic Literacy Test. Report submitted to Higher Education South Africa, Pretoria
- Plucker, J. A. (Ed.). (2003). Human intelligence: Historical influences, current controversies, teaching resources. Retrieved May 12, 2008, from <http://www.indiana.edu/~intell>.
- Programa de las Naciones Unidas para el Desarrollo (PNUD). (2007). *Informe sobre desarrollo humano 2007 – 2008. La lucha contra el cambio climático: Solidaridad frente a un mundo dividido*. Publicado por el Programa de las Naciones Unidas para el Desarrollo (PNUD). http://hdr.undp.org/en/media/HDR_20072008_SP_Complete.pdf

- Programa del Estado de la Nación (2008). *Estado de la Educación. Capítulo 2: La evolución de la educación superior*. <http://www.estadonacion.or.cr/Educacion/educacion.html>.
- Scott, I., Yeld, N. & Hendry, J. (2007) A case for improving teaching and learning in South African Higher Education. *Higher Education Monitor*. Pretoria: The Council on Higher Education.
- Sternberg, R. and Pretz, J. (2005). *Cognition and Intelligence. Identifying the mechanisms of the mind*. United States: Cambridge University Press.
- Sternberg, R.J. & Grigorenko, E.L. (2002). *Dynamic testing: the nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Terenzini, P., Springer, L., Yeager, P., Pascarella, E. & Nora, A. (1996) First-generation college students: characteristics, experiences, and cognitive development. *Research in Higher Education*, 37, pp.1-22.
- Tinto, V. (1998) Colleges as communities: taking research on student persistence seriously. *Review of Higher Education*, 21, pp.167-177.
- Universidad de Costa Rica (1974). Estatuto Orgánico. San Pedro de Montes de Oca, San José, Costa Rica.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Yeld N. (2001) Equity, assessment and language of learning: key issues for Higher Education selection and access in South Africa. Unpublished PhD Thesis, Cape Town, University of Cape Town.
- Yeld, N. (2006) Test domains and constructs: overview. In H.Griesel (Ed) *Access and Entry-level Benchmarks: The National Benchmark Tests Project*, pp. 17-18. Pretoria: Higher Education South Africa.
- Yeld, N. & Cliff, A. & Hanslo, M. (2002). *Broadening access to Higher Education through assessing the language abilities and thinking approaches of first-year university applicants*. Invited paper presented at an international symposium: International Perspectives: Global Voices for Gender Equity, Washington, U.S.A.
- Yeld, N., Prince, R., Cliff, A. & Bohlmann, C. (2009). The National Benchmark Tests Project. Summary report March 2009. Report submitted to Higher Education South Africa, Pretoria.
- Zaaiman, H., Van der Flier, H. & Thijs, G.D. (2000) Selection as contract to teach at the student's level. Experiences from a South African mathematics and science foundation year. *Higher Education* 40, pp.1-21.

VALIDEZ DIFERENCIAL Y SESGO EN LA PREDICTIVIDAD DE LAS PRUEBAS DE ADMISIÓN A LAS UNIVERSIDADES CHILENAS (PSU)

DIFFERENTIAL VALIDITY AND SLANT ON THE PREDICTABILITY OF THE ADMISSION TESTS TO THE CHILEAN UNIVERSITIES (PSU)

Jorge Manzi, Angélica Bosch, David Bravo, Guido del Pino, Graciela Donoso y Raúl Pizarro

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art2.pdf>

Fecha de recepción: 4 de abril de 2010.

Fecha de dictaminación: 30 de mayo de 2010.

Fecha de aceptación: 30 de mayo de 2010.

La sostenida brecha en los resultados entre establecimientos públicos y privados en Chile ha generado un permanente debate acerca de la equidad y del posible sesgo socioeconómico de las Pruebas de Selección Universitaria (PSU). Sin embargo, este debate no es exclusivo de la experiencia chilena ni del contexto de admisión universitaria. Las pruebas estandarizadas a gran escala suelen revelar diferencias en los puntajes obtenidos por ciertas minorías étnicas, socioeconómicas o de género. No obstante, Zwick (2007) plantea que las brechas en los resultados no representan necesariamente un sesgo de las pruebas y sostiene, por otro lado, que “la equidad de una prueba está inextricablemente ligada a su validez”.

En este sentido, la inequidad podría manifestarse a través del sesgo de medición del instrumento (que ocurre cuando aparecen diferencias sistemáticas en los puntajes obtenidos por algunos grupos, que no son atribuibles a diferencias en la habilidad que se desea medir), o bien, a través de un uso de los puntajes que perjudica de manera consistente a ciertos grupos, en relación al objetivo para el cual fue diseñada la prueba.

En el caso de las pruebas de selección universitaria, la evidencia más pertinente acerca de su validez es el grado en que los puntajes permiten predecir el éxito de los estudiantes en la enseñanza universitaria. En rigor, se espera que el ordenamiento de los postulantes según los puntajes obtenidos en una prueba de selección coincida con el ordenamiento que tendrían respecto de su potencial desempeño universitario. Una condición necesaria para que esto se cumpla es que la capacidad predictiva de la prueba sea invariante respecto de las características sociodemográficas de los examinados. Ello no ocurre si, para un mismo resultado en la prueba, algún subgrupo de la población obtiene sistemáticamente mejores o peores resultados en su desempeño académico. De ser así, los puntajes de las pruebas de selección estarían siendo usados de un modo que resulta inequitativo para ciertos grupos de la población examinada (Young, 2004).

La constante preocupación por la equidad en medición educacional ha llevado a desarrollar diversos procedimientos para evaluar los riesgos de sesgo asociados a las características sociodemográficas de los examinados. Dos técnicas que se utilizan para establecer si la capacidad predictiva de un instrumento de medición es equivalente entre distintos grupos de examinados son los análisis de *Validez Diferencial* y de *Predicción Diferencial*. Los estudios de Validez Diferencial se basan en comparar la magnitud de la relación entre el resultado de la prueba de selección y el rendimiento en la universidad, para diferentes grupos. Cuando el coeficiente de correlación estimado para cada subgrupo es sustancialmente distinto en alguno de ellos, entonces se habla de Validez Diferencial. La evaluación de Predicción Diferencial, por otra parte, consiste en analizar si la relación entre los factores de selección y el rendimiento académico puede describirse a través de un modelo de regresión único para todos los subgrupos examinados. Si los residuos de la ecuación de regresión empleada para la predicción son desiguales entre grupos, entonces se produce una subestimación (o sobreestimación) del rendimiento académico de alguno de éstos. En otras palabras, los alumnos de dicho grupo tienden a tener un desempeño que es mejor (o peor) al que el modelo de regresión común predice.

Para el caso norteamericano, Young (2004) realizó una revisión de los estudios llevados a cabo desde 1974 sobre validez y predicción diferencial en diferentes pruebas de admisión universitaria. La mayor parte de los 49 estudios incluidos en este trabajo analizan la relación entre puntajes de pruebas de selección y notas de primer año en la universidad, para diferentes grupos étnicos o raciales y/o entre mujeres y hombres. Aún cuando los resultados varían entre estudios, en general se observan correlaciones más bajas para los grupos minoritarios conformados por negros e hispanos, en comparación

con los estudiantes blancos, mientras que para las comparaciones de género, estas correlaciones suelen ser mayores en el caso de las mujeres. En términos de predicción diferencial, se observa una sobrepredicción de las notas de primer año para los grupos que conforman las minorías étnicas y raciales, y una subpredicción del rendimiento de las mujeres.

De las pruebas de admisión usadas actualmente en Estados Unidos, el SAT es una de las de mayor cobertura en términos del número de examinados anual (Young, 2004). El SAT es una prueba de razonamiento que se compone de tres secciones: redacción (SAT-W), lectura crítica (SAT-CR) y matemáticas (SAT-M), en las cuales la mayoría de las preguntas son de selección múltiple. Respecto a esta prueba, el estudio más reciente sobre validez y predicción diferencial fue llevado a cabo por Mattern et al. (2008) en base a una muestra de 155.316 alumnos, provenientes de 110 instituciones. En su trabajo, los autores encuentran un patrón de resultados que es consistente con el conocido hasta entonces:

En los análisis de validez diferencial, para todos los factores de selección obtienen correlaciones con el rendimiento académico que son superiores en el grupo de estudiantes blancos versus los grupos de negros e hispanos. En las comparaciones por género, obtienen correlaciones entre notas en la universidad y el SAT que son más altas para mujeres que para hombres, mientras que las correlaciones con las notas de colegio resultan ser más parejas. (Los principales resultados de este análisis se resumen en la Tabla 1 que se presenta a continuación).

TABLA 1. VALIDEZ DIFERENCIAL DEL SAT

Grupo Sociodemográfico	SAT-CR	SAT-M	SAT-W	Notas Colegio
Hombres	.44	.45	.47	.52
Mujeres	.52	.53	.54	.54
Blancos	.48	.45	.51	.56
Negros	.40	.40	.43	.44
Hispanos	.43	.41	.46	.46

(Fuente: Mattern et al, 2008)

En cuanto a la predicción diferencial, los autores concluyen que existe una sobrepredicción del rendimiento académico de los hombres y una subpredicción del rendimiento académico de las mujeres. En el caso de estudiantes blancos versus minorías raciales y étnicas, observan un patrón contrario: una leve subpredicción del rendimiento para alumnos blancos, una sobrepredicción del rendimiento para alumnos negros y una sobrepredicción algo menor para alumnos hispanos. (Los resultados, que se presentan en la Tabla 2 se interpretan de la siguiente manera: valores negativos indican una sobrepredicción del rendimiento y valores positivos, una subpredicción.)

TABLA 2. PREDICCIÓN DIFERENCIAL DEL SAT

Grupo Sociodemográfico	SAT-CR	SAT-M	SAT-W	NEM
Hombres	-.14	-.20	-.11	-.08
Mujeres	.12	.17	.10	.07
Blancos	.04	.05	.04	.06
Negros	-.30	-.26	-.26	-.32
Hispanos	-.17	-.16	-.16	-.27

(Fuente: Mattern et al, 2008)

El presente estudio tiene como objetivo analizar la validez diferencial y el sesgo de predicción en los factores de selección utilizados por las 25 universidades que conforman el Consejo de Rectores de las Universidades Chilenas (CRUCH). Estos factores de selección incluyen dos pruebas obligatorias (Lenguaje

y Comunicación y Matemáticas) y dos pruebas opcionales (Ciencias e Historia y Ciencias Sociales), además de las Notas de Enseñanza Media (NEM).¹

La batería de pruebas, denominada Prueba de Selección Universitaria (PSU), reemplazó a partir de la admisión 2004, a las pruebas precedentes o Pruebas de Aptitud Académica (PAA). En este contexto de cambio, el Consejo de Rectores ha propiciado la realización de un conjunto de estudios técnicos que permitan contar con antecedentes sobre el funcionamiento de las PSU. A la fecha, esta tarea ha sido encomendada al Comité Técnico Asesor para las Pruebas de Selección y Actividades de Admisión del Consejo de Rectores de las Universidades Chilenas (CTA). El año 2006 el Comité Técnico Asesor entregó el primer estudio respecto de la validez predictiva de la PSU. En dicha oportunidad se consideró importante comparar la capacidad predictiva de las nuevas pruebas con la de aquéllas que habían regido hasta la admisión del año anterior. Por lo tanto, el estudio contempló el análisis de la validez predictiva de las baterías de admisión para los alumnos ingresados los años 2003 y 2004, considerando en el rendimiento de estos estudiantes al término de su primer año de estudios en la universidad. Las técnicas de análisis incluyeron correlaciones simples para cada uno de los factores de selección y regresiones lineales para explicar el aporte conjunto de los diferentes factores de selección. El año 2008 se publicó una actualización del estudio anterior, incorporando en esta versión a los alumnos de las admisiones 2005 y 2006. Los resultados reportados en ambos informes revelan que las pruebas con mayor capacidad predictiva promedio en las PSU corresponden a las pruebas de Matemática y Ciencias, las que muestran valores predictivos individuales semejantes o superiores a la validez predictiva de las Notas de Enseñanza Media. Por su parte, las pruebas de Lenguaje y Ciencias Sociales presentan valores promedio positivos, pero marcadamente más bajos. Finalmente, ambos estudios concluyen que la comparación entre las dos baterías de selección es favorable a las nuevas pruebas en la gran mayoría de las universidades, carreras y áreas de estudio. (Manzi, et al., 2006, 2008)

El trabajo que se presenta a continuación constituye un complemento a los estudios recién citados, al incorporar los análisis de sesgo de validez y predictividad de los factores de selección universitaria. Las características de agrupación aquí consideradas son el género del estudiante y la Dependencia² del establecimiento de egreso de la Educación Media.

1. ASPECTOS METODOLÓGICOS

1.1. Bases de Datos

Los estudios de validez predictiva se basan en analizar la relación que existe entre ciertas variables predictoras (los factores de selección) y alguna variable de criterio, normalmente relacionada con el desempeño académico. En este trabajo, la variable criterio utilizada corresponde al Promedio Ponderado

¹ El Sistema de Selección se basa en la asignación de un puntaje que combina estos factores con ponderadores diferentes dependiendo de cada carrera.

² En Chile, los establecimientos escolares se clasifican en 3 Dependencias, de acuerdo al tipo de administración y financiamiento. Los Particulares Pagados son de financiamiento privado, mientras que los Particulares Subvencionados y Municipales reciben subvención del Estado. Estos últimos se diferencian en que los Particulares Subvencionados son de administración privada mientras que los Municipales son establecimientos públicos. En Chile, dada la alta segregación educativa que existe, la Dependencia suele considerarse como proxy del nivel socioeconómico del estudiante.

Acumulado (PPA) de los estudiantes al término del primer año universitario. Las variables predictoras consideradas incluyen las pruebas de selección obligatorias de Matemática y Lenguaje y Comunicación (PSUM/PAAM y PSUL/PAAV), el puntaje por notas de educación media (NEM), y el Puntaje de Selección (PSEL) con el cual cada alumno es admitido a su carrera.

El dato de rendimiento académico fue obtenido a partir de las bases entregadas por las 25 universidades del Consejo de Rectores para los estudios de Validez Predictiva llevados a cabo por el CTA. Esta información fue complementada con las bases de datos del DEMRE³, para obtener los antecedentes relativos a los resultados en los factores de selección (puntajes de pruebas y NEM) y a las características de género y dependencia del establecimiento de egreso de estos estudiantes. Los Puntajes de Selección fueron computados a partir de los puntajes y de los ponderadores de los factores de selección, según la carrera a la que es admitido cada alumno.

En total, se contó con 214.259 datos válidos de notas y puntajes de selección para las admisiones 2003 (última PAA) a 2007. La distribución de alumnos por año de ingreso según género y dependencia se presenta en la Tabla 3.

TABLA 3. DISTRIBUCIÓN DE ALUMNOS CONSIDERADOS EN EL ESTUDIO

Grupo	Año admisión					Total
	2003	2004	2005	2006	2007	
Hombre	22.955 53.7%	20.554 51.2%	21.576 52.2%	22.445 51.5%	23.645 50.9%	111.175
Mujer	19.770 46.3%	19.606 48.8%	19.758 47.8%	21.130 48.5%	22.820 49.1%	103.084
Particular Pagado	8.745 20.5%	7.824 19.5%	7.733 18.7%	7.868 18.1%	7.999 17.3%	40.169
Particular Subvencionado	16.555 38.9%	16.421 41.0%	17.564 42.6%	18.900 43.5%	2.1125 45.6%	90.565
Municipal	17.297 40.6%	15.805 38.5%	15.944 38.7%	16.707 38.4%	17.223 37.1%	82.976
Total	42.725	40.160	41.334	43.575	46.465	214.259

Nota: En las bases de datos empleadas, el dato de género estaba presente para todos los casos, no así el de dependencia del establecimiento de egreso. Es por ello que el número de casos por dependencia es algo menor que el número de casos totales. El porcentaje es sobre el total de casos con dato de dependencia.

2. ASPECTOS ANALÍTICOS

2.1. Validez Diferencial

Para determinar la relación entre cada una de las variables predictoras y el rendimiento académico, se calcularon correlaciones simples (correlación Producto-Momento de Pearson) separadamente para cada subgrupo de estudiantes. Con esta metodología, la validez diferencial de los factores de selección se manifiesta cuando los coeficientes de correlación son sustancialmente diferentes entre ciertos subgrupos.

³ El DEMRE es un organismo dependiente de la U. de Chile que está encargado del desarrollo y construcción de las PSU y de la administración del Sistema de Selección a la Educación Superior.

Es importante destacar que, de modo de asegurar condiciones de homogeneidad para la variable criterio, se decidió emplear a cada carrera como unidad de análisis básica. Con ello, valores diferentes en los coeficientes de correlación calculados revelarían la existencia de validez diferencial y no otras diferencias en las experiencias educacionales que podrían darse entre carreras.

Por otra parte, dado que el número de estudiantes válidos por carrera es en algunos casos muy pequeño, se establecieron límites inferiores de manera de evitar que las estimaciones de la capacidad predictiva de los factores de selección fueran imprecisas. Así, para las estimaciones de los coeficientes de correlación se exigieron al menos 10 casos válidos por subgrupo en cada carrera. En consecuencia, el número de carreras sobre el cual se calcula el sesgo es inferior al total de carreras disponibles para el análisis.

Los resultados que se presentan corresponden a promedios de estas correlaciones a nivel de carrera. La metodología para promediar correlaciones se basa en el método de Hedges-Olkin, que consiste en calcular el promedio ponderado (en función del tamaño de cada muestra) de las transformaciones Z para las correlaciones individuales. Este resultado es reconvertido en un índice de correlaciones mediante la transformación inversa tangente hiperbólica ($\bar{r} = \tanh(\bar{Z})$).

Finalmente, es importante advertir que, dado que los factores de selección son usados para tomar decisiones de admisión, se produce una inevitable restricción en el rango de los valores de ellos, en la medida que sólo se dispone de información para los postulantes que fueron aceptados a las universidades. Esta condición lleva a una subestimación de la capacidad predictiva de estos factores, pues solo se analiza la porción de su variabilidad que queda disponible una vez hecha la selección. No obstante lo anterior, los coeficientes de correlación pueden ser utilizados como base para realizar inferencias comparativas entre poblaciones distintas (Wainer, Saka & Donoghue, 1993).

Alternativamente, es posible usar algún método estadístico que permita corregir la restricción de rango en las correlaciones. El método más usado en estudios de Validez Predictiva corresponde a la corrección multivariada de Pearson-Lawley, conocida también como la corrección de Gulliksen (Gulliksen, 1950). Éste consiste en estimar las correlaciones en la población restringida (la muestra de estudiantes universitarios) y corregir posteriormente estas correlaciones, considerando las desviaciones estándar de los factores de selección para la población completa de examinados. La fórmula utilizada en este método se presenta a continuación:

$$r_{corr} = \frac{\sigma_x r_{xy}}{\sqrt{\sigma_x^2 (1 - r_{xy}^2) + \sigma_x^2 r_{xy}^2}}$$

Donde:

r_{corr} = Correlación corregida por restricción de rango.

r_{xy} = Correlación original.

σ_x = Desviación estándar de puntajes en población restringida (universitarios).

σ_x = Desviación estándar en población no restringida (todos los examinados).

Las correlaciones presentadas en el trabajo de Mattern et. al. (2008) incluyen este tipo de corrección. Por lo tanto, y con el objeto de poder comparar los resultados para las PSU con la evidencia internacional, en el presente estudio se incluyen ambos resultados: coeficientes de correlación con y sin correcciones por restricción de rango.

2.2. Predicción diferencial

Para analizar el sesgo predictivo, se ajustó una regresión lineal común para todos los estudiantes de una misma carrera, y se consideró una ecuación para cada predictor en forma separada: Prueba de Matemática, Prueba de Lenguaje y Comunicación, NEM y Puntaje de Selección. Posteriormente se comparó el promedio de los residuos de estas regresiones (es decir, la diferencia entre las notas obtenidas y las predichas por el modelo) para cada subgrupo. Sólo se consideraron para el promedio, carreras con al menos 10 casos en cada grupo.

Para facilitar la combinación de resultados, previo al ajuste de los modelos, las notas de primer año se estandarizaron con media 0 y desviación estándar 1. De esta forma, los residuos de cada regresión estimada también se encuentran en esta escala estandarizada. Calculados de esta manera, la interpretación de los resultados es bastante sencilla: el promedio de los residuos para la muestra completa es igual a cero, por lo tanto, cuando el promedio de los residuos para algún subgrupo es distinto de cero, el instrumento presenta predicción diferencial. Así por ejemplo, si el promedio de los residuos es positivo para alguno de los grupos, entonces la prueba tiende a subpredecir el rendimiento académico de ese grupo. En otras palabras, los alumnos de dicho grupo tienden a tener un desempeño mejor al que el modelo de regresión común predice. Análogamente, si el promedio de los residuos es negativo para alguno de los grupos, entonces la prueba tiende a sobrepredecir el rendimiento académico de dicho grupo. Por otra parte, el sesgo predictivo, calculado como la diferencia entre los promedios de los residuos para cada grupo, tiene valores que están en términos de las desviaciones estándar de las notas de primer año. Por lo tanto, este método tiene una interpretación directa en cuanto a evaluar el sesgo en la capacidad predictiva del instrumento, lo que representa una ventaja respecto a la comparación de las correlaciones entre desempeño y factores de selección.

3. RESULTADOS

3.1. Descriptivos

Las tablas siguientes muestran los promedios por género y dependencia del establecimiento de egreso para los puntajes obtenidos en los diferentes factores de selección

3.1.1. Puntajes según Género

En la prueba de Lenguaje se observan resultados muy parejos entre hombres y mujeres para todas las admisiones. Distinto es el caso de la Prueba de Matemáticas donde aparece sistemáticamente una brecha en los resultados a favor de los hombres. Es importante advertir que las diferencias de puntajes entre años no son comparables directamente ya que sólo a partir de la admisión 2005 comenzaron a estandarizarse los puntajes de las PSU⁴. En términos de las desviaciones estándar de cada prueba, las diferencias varían relativamente poco, entre un 25% y un 31% y además, estas diferencias son muy similares a las que se observan en la población completa de examinados en cada admisión. Por otra parte, en el puntaje por NEM, son las mujeres las que obtienen mejores resultados, con una ventaja de

⁴ A partir del año 2005 los puntajes son convertidos a escalas normalizadas, con media 500 y desviación estándar 110. Los años 2003 y 2004 las desviaciones estándar para la prueba de matemática fueron de 140 y 109, respectivamente, y para la de lenguaje, de 122 en ambos años de admisión.

entre 30 y 40 puntos promedio. Pese a estas diferencias en algunos factores, una vez combinados, los promedios en puntajes de selección son bastante homogéneos para todas las admisiones (con una diferencia a favor de los hombres de 4 puntos como máximo).

TABLA 4. PROMEDIOS EN FACTORES DE SELECCIÓN SEGÚN GÉNERO. ADMISIONES 2003 A 2007

Admisión	N carreras	Género	N alumnos	Lenguaje		Matemática		NEM		PSEL	
				Media	D.E.	Media	D.E.	Media	D.E.	Media	D.E.
2003	776	Hombres	22955	611	90	665	96	601	96	629	67
		Mujeres	19770	612	85	630	97	645	85	625	64
2004	816	Hombres	20554	598	99	611	88	612	91	611	72
		Mujeres	19606	603	93	577	80	648	81	609	65
2005	850	Hombres	21576	591	85	615	84	616	92	613	71
		Mujeres	19758	592	81	583	78	649	83	609	65
2006	890	Hombres	22445	590	85	616	85	615	91	612	72
		Mujeres	21130	594	81	583	78	646	82	609	65
2007	941	Hombres	23645	597	83	619	82	615	91	616	70
		Mujeres	22820	599	80	590	75	644	83	612	65

3.1.2. Puntajes según Dependencia

Al comparar los resultados según dependencia de egreso, se observa que los alumnos provenientes de establecimientos Particulares Pagados obtienen, en promedio, mejores puntajes en todos los factores de selección.

TABLA 5. PROMEDIOS EN FACTORES DE SELECCIÓN SEGÚN DEPENDENCIA. ADMISIONES 2003 A 2007.

Admisión	N carreras	Dependencia	N alumnos	Lenguaje		Matemática		NEM		PSEL	
				Media	D.E.	Media	D.E.	Media	D.E.	Media	D.E.
2003	178	Particular Pagado	8745	661	78	714	83	648	90	674	62
		Particular Subvencionado	16555	606	83	643	91	621	92	623	60
		Municipal	17297	592	87	621	96	608	94	608	62
2004	149	Particular Pagado	7824	659	82	660	86	657	85	661	67
		Particular Subvencionado	16421	597	90	586	75	628	88	604	62
		Municipal	15805	575	97	571	80	618	87	591	63
2005	145	Particular Pagado	7733	648	80	665	82	663	84	664	67
		Particular Subvencionado	17564	586	76	591	73	627	89	604	62
		Municipal	15944	570	79	577	77	621	88	593	62
2006	154	Particular Pagado	7868	653	80	671	82	662	85	668	69
		Particular Subvencionado	18900	586	76	591	72	626	88	603	62
		Municipal	16707	570	78	577	77	620	87	592	62
2007	148	Particular Pagado	7999	660	76	674	78	661	84	671	66
		Particular Subvencionado	21125	593	75	598	71	625	87	608	61
		Municipal	17223	576	78	580	74	619	87	595	61

Las mayores diferencias se dan entre establecimientos Particulares Pagados y Municipales, en la Prueba de Matemática (alrededor de 90 puntos, poco más de 80% de una desviación estándar de las PSU) y en la de Lenguaje y Comunicación (alrededor de 80 puntos, 75% de una desviación estándar). Esta brecha disminuye en el caso de las NEM. Las diferencias entre Particulares Subvencionados y Municipales, si bien son estables en el tiempo, son bastante menores: poco menos de 20 puntos en el caso de las pruebas obligatorias y de 10 puntos en el caso de las NEM.

En ambas comparaciones por dependencia, las brechas presentan las mismas tendencias que las observadas en la población completa de examinados que rinden las PSU. Sin embargo, las magnitudes de estas brechas son algo menores⁵ en el caso de los alumnos admitidos a universidades del Consejo de Rectores.

Para el caso del Puntaje de Selección, se mantienen los mismos patrones anteriores (con una diferencia de 70 puntos entre Particulares Pagados y Municipales y una diferencia de 15 puntos entre Particulares Subvencionados y Municipales).

3.2. Validez Diferencial

3.2.1. Correlaciones Simples con Notas de Primer Año (PPA)

La siguiente tabla muestra las correlaciones promedio entre factores de selección y notas de primer año (PPA). Los resultados, que coinciden con los presentados en Estudios de Validez Predictiva de las PSU, muestran una mayor capacidad predictiva promedio para la PSU de Matemática, la que presenta correlaciones semejantes a la validez predictiva de las notas de enseñanza media. La Prueba de Lenguaje también presenta valores promedio positivos, pero marcadamente más bajos. En ambos subsectores las PSU presentan correlaciones con el PPA superiores a las obtenidas para la última PAA (Manzi et al, 2006).

TABLA 6. CORRELACIONES PROMEDIO DE LOS FACTORES DE SELECCIÓN CON PPA. ADMISIONES 2003 A 2007.

Año Admisión	N carreras	N alumnos	Lenguaje	Matemática	NEM	PSEL
2003	766	42725	0.02	0.13	0.25	0.30
2004	777	40160	0.08	0.29	0.27	0.39
2005	824	41334	0.11	0.25	0.26	0.38
2006	851	43575	0.10	0.24	0.24	0.36
2007	918	46465	0.10	0.25	0.25	0.36

3.2.2. Validez Diferencial según Género

En el caso de las comparaciones por género, no se observan diferencias importantes en las correlaciones calculadas para ninguno de los factores. Los resultados son similares a los obtenidos al considerar la muestra completa en cada carrera (tabla 4).

⁵ Las brechas entre Particulares Pagados y Municipales se reducen en alrededor de 20 a 35 puntos, (lo que se traduce en reducciones que varían entre un 20 y 30% dependiendo del año y prueba), mientras que las brechas entre Particulares Subvencionados y Municipales se reducen en alrededor de 10 puntos, (lo que se traduce en reducciones que varían entre un 10 y 40% dependiendo del año y prueba)

TABLA 7. CORRELACIONES PROMEDIO DE LOS FACTORES DE SELECCIÓN CON PPA SEGÚN GÉNERO. ADMISIONES 2003 A 2007.

Año admisión	N carreras	Género	N alumnos	Lenguaje	Matemática	NEM	PSEL
2003	519	Hombres	17825	0.02	0.15	0.23	0.31
		Mujeres	15093	0.03	0.14	0.26	0.31
2004	469	Hombres	15447	0.09	0.30	0.26	0.40
		Mujeres	14131	0.08	0.29	0.28	0.39
2005	483	Hombres	15560	0.12	0.27	0.26	0.39
		Mujeres	14438	0.12	0.25	0.26	0.37
2006	534	Hombres	16883	0.10	0.25	0.23	0.36
		Mujeres	15938	0.10	0.27	0.24	0.37
2007	572	Hombres	17836	0.10	0.26	0.24	0.36
		Mujeres	17422	0.11	0.26	0.24	0.36

3.2.3. Validez Diferencial según Dependencia

En el caso de la PSU de Lenguaje y Comunicación, no pareciera haber importantes diferencias en las magnitudes de las correlaciones con el PPA entre grupos de estudiantes egresados de distintas dependencias.

En Matemática, en cambio, se observan correlaciones relativamente más altas para Particulares Subvencionados y Municipales que para Particulares Pagados, en los dos primeros años de las PSU (2004 y 2005). Las diferencias se atenúan los siguientes dos años de admisión. Para la última PAA de Matemática (2003) tampoco se aprecian diferencias en las correlaciones aunque éstas son marcadamente más bajas que las observadas para las PSU.

TABLA 8. CORRELACIONES PROMEDIOS DE LOS FACTORES DE SELECCIÓN CON PPA SEGÚN DEPENDENCIA. ADMISIONES 2003 A 2007.

Año admisión	N carreras	Dependencia	Lenguaje	Matemática	NEM	PSEL
2003	178	Particular Pagado	0.06	0.13	0.31	0.37
		Particular Subvencionado	0.00	0.14	0.30	0.32
		Municipal	0.05	0.13	0.20	0.31
2004	149	Particular Pagado	0.09	0.23	0.36	0.41
		Particular Subvencionado	0.06	0.28	0.28	0.39
		Municipal	0.07	0.27	0.23	0.38
2005	145	Particular Pagado	0.10	0.18	0.37	0.38
		Particular Subvencionado	0.10	0.26	0.28	0.39
		Municipal	0.10	0.24	0.19	0.35
2006	154	Particular Pagado	0.11	0.20	0.33	0.40
		Particular Subvencionado	0.07	0.22	0.23	0.34
		Municipal	0.07	0.24	0.18	0.35
2007	148	Particular Pagado	0.10	0.22	0.36	0.41
		Particular Subvencionado	0.08	0.22	0.23	0.34
		Municipal	0.06	0.24	0.15	0.32

En cuanto a las NEM, se observa un patrón diferente. En la última admisión de las PAA, las correlaciones para el grupo de Particulares Pagados y Particulares Subvencionados son del orden de 0.3, mientras que para los Municipales la correlación promedio con el PPA resulta bastante más baja (de 0.2). Durante las admisiones siguientes (baterías PSU), se observa una diferencia marcada entre los tres grupos: las correlaciones más altas se dan para Particulares Pagados, seguidas por las del grupo de Particulares Subvencionados y, finalmente, los Municipales que presentan correlaciones bastante más bajas.

En el Puntaje de Selección, se observa una tendencia de correlaciones levemente más altas para el grupo de Particulares Pagados que para Particulares Subvencionados y Municipales.

3.2.4. Correlaciones Corregidas por Restricción de Rango

Las siguientes tablas muestran las correlaciones de las pruebas obligatorias y NEM, corregidas por la restricción de rango. En el caso del Puntaje de Selección no es posible llevar a cabo esta corrección ya que sólo es posible determinar su varianza para los alumnos que fueron efectivamente admitidos a alguna carrera.

En términos generales, al comparar con las correlaciones sin corrección (tabla 9), notamos que el efecto de la restricción de rango es más importante en las pruebas de Lenguaje y Matemática que en el caso de las NEM.

TABLA 9. CORRELACIONES PROMEDIOS DE LOS FACTORES DE SELECCIÓN CON PPA, CORREGIDAS POR RESTRICCIÓN DE RANGO. ADMISIONES 2003 A 2007.

Año Admisión	N carreras	N alumnos	Lenguaje	Matemática	NEM
2003	766	42725	0.04	0.33	0.37
2004	777	40160	0.15	0.60	0.42
2005	824	41334	0.23	0.55	0.41
2006	851	43575	0.21	0.55	0.39
2007	918	46465	0.22	0.57	0.40

3.2.5. Validez Diferencial según Género, con correcciones por Restricción de Rango

Una vez corregidas por la restricción de rango, las diferencias en las correlaciones entre hombres y mujeres aumentan un poco en algunos años, pero las NEM son el único factor que presenta una tendencia más clara en todas las admisiones, siendo las correlaciones con las notas de primer año algo mayores para las mujeres. Este efecto tiende a disminuir para los últimos años.

Estos resultados son distintos a los presentados por Mattern et. al. (2008), donde se obtienen correlaciones con las secciones del SAT que son más altas para las mujeres, mientras que las correlaciones con las notas del colegio son relativamente similares (ver tabla 1)

TABLA 10. CORRELACIONES PROMEDIOS DE LOS FACTORES DE SELECCIÓN CON PPA SEGÚN GÉNERO, CORREGIDAS POR RESTRICCIÓN DE RANGO. ADMISIONES 2003 A 2007

Año admisión	N carreras	Género	N alumnos	Lenguaje	Matemática	NEM
2003	519	Hombres	17825	0.06	0.39	0.35
		Mujeres	15093	0.09	0.33	0.44
2004	469	Hombres	15447	0.19	0.61	0.41
		Mujeres	14131	0.17	0.61	0.48
2005	483	Hombres	15560	0.26	0.57	0.40
		Mujeres	14438	0.26	0.54	0.44
2006	534	Hombres	16883	0.23	0.56	0.38
		Mujeres	15938	0.23	0.59	0.42
2007	572	Hombres	17836	0.23	0.59	0.40
		Mujeres	17422	0.26	0.58	0.42

3.2.6. Validez Diferencial según Dependencia, con correcciones por Restricción de Rango

Al comparar por dependencia también se obtiene un patrón sistemático en las diferencias en las correlaciones entre notas de primer año y las NEM. En este caso, las diferencias son más importantes que en el caso de las comparaciones por género y mantienen la misma tendencia que en el caso de las correlaciones sin corrección por restricción de rango: las más altas se dan en el grupo de Particulares Pagados, seguido por el de Particulares Subvencionados y finalmente, las correlaciones más bajas se dan para el grupo proveniente de establecimientos Municipales. Este resultado es concordante con el patrón encontrado por Mattern et al. (2008) para las minorías raciales y étnicas (en todos los factores de selección). Sin embargo, en el caso de las NEM, las diferencias son algo mayores que las reportadas en el estudio de Mattern et al. (2008) para las notas de colegio: correlación para estudiantes blancos de 0.56, para Hispanos de 0.46 y para negros de 0.44 (ver tabla 1).

En la Prueba de Matemática, el año 2003 (PAA) se observa una mayor correlación para Particulares Pagados y correlaciones similares para los otros dos grupos de dependencia. En cambio, para las PSU de Matemática se observa un efecto contrario: mayores correlaciones para Particulares Subvencionados y Municipales que para Particulares Pagados, aunque este efecto tiende a atenuarse los últimos dos años.

Respecto a la Prueba de Lenguaje y Comunicación, los resultados son menos concluyentes y sólo pareciera darse una tendencia (que sería similar a la de los SAT y PAAM) en las últimas dos admisiones: correlaciones algo mayores para Particulares Pagados que para Particulares Subvencionados y Municipales.

TABLA 11. CORRELACIONES PROMEDIOS DE LOS FACTORES DE SELECCIÓN CON PPA SEGÚN DEPENDENCIA, CORREGIDAS POR RESTRICCIÓN DE RANGO. ADMISIONES 2003 A 2007

Año admisión	N carreras	Dependencia	Lenguaje	Matemática	NEM
2003	178	Particular Pagado	0.14	0.42	0.53
		Particular Subvencionado	0.02	0.32	0.45
		Municipal	0.11	0.31	0.29
2004	149	Particular Pagado	0.20	0.51	0.62
		Particular Subvencionado	0.12	0.57	0.45
		Municipal	0.16	0.55	0.36
2005	145	Particular Pagado	0.21	0.40	0.64
		Particular Subvencionado	0.19	0.54	0.46
		Municipal	0.21	0.51	0.28
2006	154	Particular Pagado	0.25	0.47	0.60
		Particular Subvencionado	0.16	0.49	0.39
		Municipal	0.16	0.53	0.29
2007	148	Particular Pagado	0.22	0.50	0.65
		Particular Subvencionado	0.18	0.50	0.41
		Municipal	0.17	0.54	0.26

3.3. Predicción diferencial

3.3.1. Predicción Diferencial según Género

A partir de una regresión común para cada carrera, se estimaron los residuos estandarizados para cada alumno, como la diferencia entre las notas obtenidas y las predichas por cada modelo de regresión. En la tabla 12 se presentan los promedios de estos residuos para cada uno de los factores, según género.

Se observan, para las pruebas obligatorias, promedios negativos para los estudiantes hombres y promedios positivos para las mujeres. Esto se traduce en un sesgo predictivo en favor del género masculino, que en el caso de la prueba de Matemática es de 0.2 desviaciones estándar y en el caso de Lenguaje y Comunicación es de al menos 0.12 desviaciones estándar. En ambos subsectores los resultados son bastante estables entre admisiones, excepto para la prueba de lenguaje del año 2003 (última versión de la PAA Verbal). En dicha admisión se observa un sesgo predictivo que es levemente superior a los obtenidos posteriormente con las PSU de Lenguaje y Comunicación (0.17 vs. 0.12 a 0.14 desviaciones estándar).

Estos resultados son consistentes con los patrones encontrados en el estudio de Mattern et al. (2008), aunque las magnitudes en los sesgos en el caso de las PSU son un poco inferiores a los reportados en dicho trabajo para el SAT (en la tabla 2, la diferencia absoluta entre las predicciones diferenciales de hombres y mujeres es de 0.37 desviaciones estándar para el SAT Matemáticas, de 0.26 para el SAT-CR y de 0.22 para el SAT-W; todos estos sesgos son a favor de los estudiantes hombres).

En el caso de las Notas de Enseñanza Media también se observa un sesgo predictivo a favor de los hombres, sin embargo, en este caso, la magnitud es sustancialmente inferior y no supera las 0.04 desviaciones estándar.

Finalmente, el Puntaje de Selección presenta, en promedio, un sesgo predictivo a favor de los hombres con un patrón muy similar a la prueba de Lenguaje y Comunicación.

TABLA 12. PROMEDIOS DE RESIDUOS ESTANDARIZADOS POR GÉNERO. ADMISIONES 2003 A 2007

Año admisión	N carreras	Género	N alumnos	Lenguaje	Matemática	NEM	PSEL
2003	519	Hombres	17825	-0.08	-0.09	-0.02	-0.08
		Mujeres	15093	0.09	0.11	0.02	0.09
2004	469	Hombres	15447	-0.06	-0.09	0.00	-0.06
		Mujeres	14131	0.06	0.10	0.00	0.06
2005	483	Hombres	15560	-0.06	-0.09	-0.01	-0.06
		Mujeres	14438	0.07	0.10	0.01	0.06
2006	534	Hombres	16883	-0.06	-0.10	-0.01	-0.06
		Mujeres	15938	0.07	0.10	0.02	0.06
2007	572	Hombres	17836	-0.07	-0.10	-0.02	-0.07
		Mujeres	17422	0.07	0.10	0.02	0.07

3.3.2. Predicción Diferencial según Género, por área del conocimiento

Las siguientes tablas presentan la magnitud total de la predicción diferencial entre hombres y mujeres, por área de conocimiento, para las admisiones de las baterías PSU. Los resultados, aunque varían un poco en algunos años, presentan tendencias bastante estables.

En el caso de la prueba de Matemática, son las áreas de Ciencias Sociales, Economía y Administración y Educación las que presentan mayor sesgo de predicción a favor de los hombres. Las áreas con menor sesgo son las de Arte, Ingeniería y Ciencias.

En la Prueba de Lenguaje y Comunicación, también son las áreas de Ciencias Sociales y Economía y Administración las que presentan mayor sesgo de predicción a favor de los hombres, mientras que las de Arte, Ingeniería y Ciencias, son las que tienen menores sesgos.

TABLA 13. DIFERENCIAS EN LOS PROMEDIOS DE LOS RESIDUOS ESTANDARIZADOS ENTRE HOMBRES Y MUJERES, POR ÁREA. ADMISIONES 2003 A 2007. CASO PSU MATEMÁTICA.

Área del Conocimiento	Año de Admisión			
	2004	2005	2006	2007
Agronomía-Forestal	-0.24	-0.18	-0.21	-0.20
Arte	-0.12	-0.04	-0.09	-0.12
Ciencias	-0.13	-0.12	-0.21	-0.16
Ciencias Sociales	-0.31	-0.30	-0.27	-0.34
Economía-Administración	-0.26	-0.35	-0.36	-0.30
Educación	-0.23	-0.27	-0.29	-0.24
Humanidades	-0.27	-0.16	-0.22	-0.21
Ingeniería	-0.14	-0.11	-0.13	-0.16
Salud	-0.20	-0.22	-0.16	-0.23
Tecnología	-0.14	-0.18	-0.17	-0.24

TABLA 14. DIFERENCIAS EN LOS PROMEDIOS DE LOS RESIDUOS ESTANDARIZADOS ENTRE HOMBRES Y MUJERES, POR ÁREA. ADMISIONES 2003 A 2007. CASO PSU LENGUAJE.

Área del Conocimiento	Año de Admisión			
	2004	2005	2006	2007
Agronomía-Forestal	-0.10	-0.07	-0.12	-0.07
Arte	-0.10	-0.01	-0.04	-0.09
Ciencias	0.00	-0.02	-0.12	-0.06
Ciencias Sociales	-0.28	-0.28	-0.23	-0.31
Economía-Administración	-0.20	-0.30	-0.30	-0.24
Educación	-0.16	-0.22	-0.23	-0.20
Humanidades	-0.23	-0.13	-0.19	-0.19
Ingeniería	-0.02	-0.01	0.00	-0.06
Salud	-0.17	-0.20	-0.12	-0.21
Tecnología	-0.13	-0.11	-0.06	-0.14

En el caso de las NEM los sesgos predictivos por área son más bajos que en el caso de las PSU, y no todas son a favor de los hombres. Las áreas de Ciencias Sociales y de Economía y Administración son las que presentan mayor sesgo a favor de los hombres mientras que el área de las Ingenierías presenta un leve sesgo a favor de las mujeres.

TABLA 15. DIFERENCIAS EN LOS PROMEDIOS DE LOS RESIDUOS ESTANDARIZADOS ENTRE HOMBRES Y MUJERES, POR ÁREA. ADMISIONES 2003 A 2007. CASO NEM.

Área del Conocimiento	Año de Admisión			
	2004	2005	2006	2007
Agronomía-Forestal	0.06	0.07	0.00	0.00
Arte	0.04	0.06	0.06	0.00
Ciencias	0.12	0.10	-0.02	0.06
Ciencias Sociales	-0.13	-0.15	-0.12	-0.19
Economía-Administración	-0.06	-0.15	-0.14	-0.09
Educación	-0.02	-0.10	-0.10	-0.11
Humanidades	-0.06	0.04	-0.04	-0.05
Ingeniería	0.08	0.07	0.07	0.03
Salud	-0.05	-0.10	-0.05	-0.12
Tecnología	0.04	-0.01	-0.02	-0.10

Finalmente, respecto de los Puntajes de Selección, se observa un patrón de resultados similar al de la PSU de Lenguaje, donde las áreas de Ciencias Sociales y de Economía y Administración son las que presentan

mayor sesgo de predicción a favor de los hombres, mientras que las de Arte, Ingeniería y Ciencias, son las que tienen menores sesgos.

TABLA 16. DIFERENCIAS EN LOS PROMEDIOS DE LOS RESIDUOS ESTANDARIZADOS ENTRE HOMBRE Y MUJERES, POR ÁREA. ADMISIONES 2003 A 2007. CASO PUNTAJE DE SELECCIÓN.

Área del Conocimiento	Año de Admisión			
	2004	2005	2006	2007
Agronomía-Forestal	-0.11	-0.09	-0.11	-0.10
Arte	-0.06	0.00	-0.05	-0.08
Ciencias	0.00	-0.02	-0.12	-0.07
Ciencias Sociales	-0.28	-0.30	-0.22	-0.28
Economía-Administración	-0.18	-0.26	-0.26	-0.20
Educación	-0.16	-0.21	-0.21	-0.20
Humanidades	-0.23	-0.10	-0.18	-0.18
Ingeniería	0.00	0.01	-0.02	-0.05
Salud	-0.16	-0.18	-0.12	-0.20
Tecnología	-0.08	-0.08	-0.03	-0.10

3.3.3. Predicción Diferencial según Dependencia

Los promedios por dependencia de los residuos estandarizados revelan una subpredicción en el rendimiento de estudiantes provenientes de establecimientos particulares pagados y una sobrepredicción del rendimiento para alumnos de colegios municipales en la Prueba de Lenguaje, aunque con un sesgo predictivo relativamente bajo a favor de los establecimientos municipales (de 0.08 desviaciones estándar como máximo).

Los sesgos predictivos son aún más bajos en la Prueba de Matemáticas y en el Puntaje de Selección (diferencia que no supera las 0.04 desviaciones estándar), y no se observa una tendencia clara a favor o en contra de alguno de los grupos de dependencia.

TABLA 17. PROMEDIOS DE RESIDUOS ESTANDARIZADOS POR DEPENDENCIA. ADMISIONES 2003 A 2007

Año admisión	N carreras	Dependencia	Lenguaje	Matemática	NEM	PSEL
2003	178	Particular Pagado	0.01	-0.01	0.04	-0.02
		Particular Subvencionado	0.01	0.01	0.01	0.02
		Municipal	-0.02	-0.01	-0.02	0.00
2004	149	Particular Pagado	0.05	0.00	0.07	0.01
		Particular Subvencionado	0.01	0.01	0.01	0.01
		Municipal	-0.03	-0.01	-0.04	-0.02
2005	145	Particular Pagado	0.04	0.00	0.07	0.01
		Particular Subvencionado	0.01	0.01	0.02	0.02
		Municipal	-0.03	-0.01	-0.05	-0.02
2006	154	Particular Pagado	0.03	-0.02	0.07	0.00
		Particular Subvencionado	0.00	0.01	0.01	0.01
		Municipal	-0.02	0.00	-0.04	-0.01
2007	148	Particular Pagado	0.04	-0.02	0.08	0.00
		Particular Subvencionado	0.01	0.01	0.02	0.02
		Municipal	-0.03	-0.01	-0.05	-0.02

Las NEM constituyen el factor que presenta mayor sesgo predictivo (de hasta 0.13 desviaciones estándar), con un sesgo positivo a favor de los establecimientos municipales. Esto significa que, en promedio, un alumno egresado de colegio municipal tendrá un menor desempeño relativo en la universidad en relación con un alumno de colegio Particular Pagado con igual NEM. Estas diferencias podrían explicarse por mecanismos diferentes de asignación de notas entre colegios de distinta dependencia. Sin embargo, en términos de equidad, estos resultados se traducen en una ventaja a favor de los alumnos de establecimientos municipales, ya que incrementa en forma artificial sus puntajes de postulación a la universidad.

4. CONCLUSIONES

En un sistema de admisión universitario es primordial contar con un mecanismo de selección que sea equitativo para todos los postulantes. En Chile, la selección de estudiantes a las universidades del Consejo de Rectores se realiza en base a los resultados obtenidos en una batería compuesta de pruebas estandarizadas (PSU, desde el año 2004) y las notas de la Enseñanza Media.

Para los años de admisión considerados en este estudio (2003 a 2007), las brechas en resultados en las pruebas de selección, entre los estudiantes admitidos a las universidades del Consejo de Rectores, siguen las mismas tendencias que las que se observan en el universo de postulantes que las rinden. Esto es, un mejor desempeño relativo de hombres versus mujeres en la Prueba de Matemáticas y mejores desempeños relativos de alumnos provenientes de establecimientos Particulares Pagados versus Particulares Subvencionados y Municipales, en ambas pruebas obligatorias.

Sin embargo, y más allá de las diferencias en resultados, en términos de equidad del sistema de admisión, lo que se espera es que los factores de selección sean capaces de discriminar correctamente a aquellos alumnos que tienen una mayor probabilidad de éxito en su desempeño académico universitario. Para que ello ocurra, la capacidad predictiva de los instrumentos de selección debe ser insesgada respecto a características sociodemográficas de los postulantes. De no ser así, se corre el riesgo de que los puntajes de selección estén siendo usados de un modo que, en definitiva, resulta consistentemente inequitativo para ciertos grupos.

El objetivo de este estudio era evaluar posibles sesgos en la capacidad predictiva de los factores de selección obligatorios a las universidades del Consejo de Rectores (Prueba de Matemática, Prueba de Lenguaje y Comunicación (Verbal) y NEM), además del Puntaje de Selección de cada carrera. Se emplearon el género y dependencia de los establecimientos educacionales como variables de agrupamiento de los estudiantes.

En los resultados obtenidos de los análisis de Validez Diferencial, que se basaron en la comparación de las correlaciones entre los factores de selección y el rendimiento académico universitario de primer año, **no se evidencian diferencias importantes en las comparaciones por género para ninguno de los factores.** Al corregir por restricción de rango, las diferencias en las correlaciones entre hombres y mujeres aumentan un poco en algunos años, pero es el puntaje por NEM el único factor que presenta una tendencia más clara en todas las admisiones, con correlaciones un poco mayores para las mujeres. Este efecto, sin embargo, tiende a disminuir para los últimos años de admisión.

En relación a las comparaciones por dependencia, aunque se observan diferencias en las magnitudes de las correlaciones con las pruebas obligatorias en algunos años, **no se encuentra ningún patrón**

sistemático que evidencie la existencia de validez diferencial. En el caso de las NEM ocurre algo diferente. En la última admisión del Sistema PAA, las correlaciones para el grupo de Particulares Pagados y Particulares Subvencionados son similares y algo mayores que para el grupo de Municipales. Sin embargo, durante las admisiones siguientes (baterías PSU), se observa una diferencia marcada entre los tres grupos: las correlaciones más altas se dan para Particulares Pagados, seguidas por las del grupo de Particulares Subvencionados, y finalmente, los Municipales, que presentan correlaciones bastante más bajas. Estos patrones se mantienen bastante similares después de corregir por la restricción de rango.

En términos de la predicción diferencial de las pruebas obligatorias, los resultados obtenidos aportan evidencia equivalente a los patrones observados en estudios norteamericanos. Respecto a las diferencias predictivas por dependencia, se observa una leve sobrepredicción del rendimiento de estudiantes provenientes de establecimientos municipales en la Prueba de Lenguaje, lo que se traduce en un sesgo de 0,07 desviaciones estándar de las notas de primer año a favor de este grupo, en relación a los estudiantes de establecimientos Particulares Pagados. Por otra parte, los sesgos predictivos son muy bajos en la Prueba de Matemáticas y no se observa una tendencia clara a favor de alguno de los grupos de dependencia. En cuanto a las comparaciones por género, existe una subpredicción del rendimiento de las mujeres, que es del orden de un 13% de la desviación estándar para la prueba de Lenguaje y de un 20% para la prueba de Matemática. Estos resultados entre hombres y mujeres son algo menores a los encontrados por Mattern et al. en el último estudio de validez diferencial del SAT (de 26% para el SAT-CR y 37% para el SAT-M), sin embargo, confirman las tendencias observadas en este tipo de pruebas. Al realizar los análisis de género por área de conocimiento, también se confirman estos resultados, aunque las magnitudes varían un poco: los mayores sesgos predictivos a favor de los hombres están en las áreas de Ciencias Sociales y Economía y Administración (sobre 30% en Matemáticas y sobre 25% en Lenguaje para la mayoría de las admisiones) y los menores sesgos en las áreas de Arte, Ingeniería y Ciencias (poco más de 10% en Matemáticas y bajo 5% en Lenguaje para la mayoría de las admisiones),

Por otra parte, las Notas de Enseñanza Media presentan una tendencia similar a la de las pruebas pero tienen comportamientos diferentes en relación a la magnitud relativa de los sesgos. En las comparaciones por género también se observa un sesgo predictivo a favor de los hombres pero, en este caso, la magnitud es sustancialmente inferior a la de las pruebas y no supera las 0.04 desviaciones estándar. En cambio, en las comparaciones por dependencia, son las NEM las que constituyen el factor que presenta mayor sesgo predictivo (de hasta 0.13 desviaciones estándar). En términos de equidad, estos resultados se traducen en una ventaja a favor de los establecimientos municipales.

Por último, cabe señalar que los Puntajes de Selección, que se calculan como una combinación ponderada de los factores de selección, tienen comportamientos coherentes con los resultados de validez y predicción diferencial observadas para los factores que lo constituyen.

Entre la literatura dedicada a este tema, diferentes hipótesis se han planteado para explicar los fenómenos de sesgo en las capacidades predictivas de los factores de selección.

Una explicación que se ha dado, especialmente en el contexto norteamericano, hace referencia a la validez en el uso de las notas de la universidad como criterio de análisis, especialmente cuando los alumnos tienen la posibilidad de elegir sus cursos. Si las asignaturas escogidas por algún grupo son sistemáticamente más fáciles que las escogidas por los demás, entonces es posible que estos alumnos obtengan relativamente mejores notas, y por tanto, los análisis realizados en estas condiciones podrían sugerir, en forma errónea, predicción diferencial de los factores de selección. Para minimizar este riesgo,

en el presente trabajo se decidió utilizar como criterio de comparación sólo las notas de primer año (donde existe poca libertad para la elección de ramos) y, como unidad de análisis básica cada carrera (de modo de controlar por las diferencias en las experiencias educacionales que podrían darse entre carreras). Con ello, la variabilidad de ramos cursados por alumnos en una misma carrera es bastante baja, haciendo más comparable los rendimientos académicos entre estudiantes. En consecuencia, la hipótesis anterior resulta poco plausible para explicar los sesgos de validez y predicción el caso abordado en este estudio.

Por otra parte, y suponiendo que se acepta como válido el uso de las notas universitarias como variable de criterio, los estudios de predicción diferencial pueden servir para detectar la presencia de sesgo de medición en las pruebas. En este sentido, predicciones diferenciales del rendimiento podrían deberse a diferencias sistemáticas en los puntajes obtenidos por ciertos grupos, las que no serían atribuibles a diferencias en el constructo (habilidad, capacidad) que la prueba está tratando de medir. No obstante, la mayoría de las pruebas estandarizadas de altas consecuencias, como la PSU, realizan análisis del funcionamiento diferencial de sus ítems (DIF) como forma de detectar la existencia de sesgo de medición en la prueba. Por tanto, aún en ausencia de sesgo de medición, es posible encontrar evidencia de validez diferencial en las pruebas de selección universitaria.

El caso de los puntajes asociados a las NEM es un poco distinto ya que este es el único factor de selección que no constituye un instrumento estandarizado de medición y que, por tanto, no representa una escala común de aprendizajes. Por el contrario, las NEM dependen directamente del sistema de evaluación que posea el establecimiento educacional de origen. Este hecho podría explicar por qué, a pesar de que este factor tiene una alta capacidad predictiva promedio (de 0.25, sin corrección de rango y de 0.4, con corrección), es que se observan diferencias importantes en las magnitudes de estas correlaciones según la dependencia de origen. Por otra parte, el sesgo de predicción a favor de los establecimientos municipales revelaría en qué forma operan estos mecanismos diferenciales de asignación de notas, ya que los resultados indican que, en promedio, un alumno egresado de un colegio municipal tendría un menor desempeño relativo en la universidad comparado con un alumno de un colegio Particular Pagado con igual NEM.

Para concluir, es importante mencionar que existen hipótesis alternativas que intentan explicar el sesgo en las capacidades predictivas de los factores de selección. Estas suponen entre otras cosas, que existen expectativas, propias o del entorno, que pueden afectar el desempeño académico de ciertas minorías, o que, en el contexto universitario, ciertos grupos enfrentan opciones o poseen hábitos o actitudes diferentes a las de sus pares, los que se traducen en un rendimiento académico diferencial.

Dentro de estas hipótesis está la que se conoce como "amenaza de estereotipo", que señala que, estereotipos culturalmente compartidos sobre el desempeño de cierto grupo, y que son evidenciados en un contexto que apela a dicho estereotipo, pueden afectar el desempeño de un individuo que se identifica con ese grupo (Steele y Aronson, 1995, 1997). Dicho en otras palabras, cuando un individuo está consciente de las bajas expectativas de rendimiento asociadas a su grupo, la tensión o ansiedad que dicha conciencia produce, puede llevar a rendir por debajo de su potencial, confirmando las expectativas.

Si bien esta hipótesis podrían explicar el comportamiento diferencial de ciertos grupos socioeconómicos, como minorías étnicas y raciales en EEUU y dependencia educacional de origen, en Chile, la evidencia respecto a los resultados diferenciales en predicción del rendimiento entre hombres y mujeres no son consistentes con dicha hipótesis (pues en este caso, las mujeres demuestran un mejor rendimiento que el previsto según sus puntajes en las pruebas de admisión).

En definitiva, no contamos aun con buenas explicaciones para los resultados detectados en los estudios de validez diferencial (ver Zwick, 2006). El hecho que los resultados de este estudio sean consistentes con los norteamericanos (sobrepredicción de grupos socialmente desventajados y subpredicción del rendimiento de las mujeres, sugiere una cierta consistencia en distintos contextos culturales, lo que debiera motivar esfuerzos para comprender los mecanismos que explican tales resultados. Los sesgos detectados no son importantes en su magnitud (especialmente los asociados a la posición social –como grupo étnico en EEUU y dependencia educacional en Chile–), pero su consistencia en el tiempo y entre culturas los debiera constituir en aspectos a ser monitoreados sistemáticamente. Asimismo, se debieran establecer programas de investigación que permitan establecer sus orígenes o al menos sus correlatos.

REFERENCIAS BIBLIOGRÁFICAS

- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (1999). *The Standards for Educational and Psychological Testing*. Washington, D.C.: Autor
- Bravo, D., G. Del Pino, G. Donoso, G. Hawes, J. Manzi y M. Martínez (2008), "Resultados de la Aplicación de Pruebas de Selección Universitaria Admisión 2006 a 2008", Comité Técnico Asesor, Honorable Consejo de Rectores de las Universidades Chilenas, Serie Documentos Técnicos, Septiembre. Disponible en www.cta-psu.cl
- Burton, N. and Ramist, L.. Predicting success in college: SAT studies of classes graduating since 1980. New York: College Entrance Examination Board.
- Camara, Wayne, and Gary Echternacht. 2000. "The SAT I and High School Grades: Utility in Predicting Success in College." The College Board Research Notes RN-10:1-10.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley and Sons.
- Manzi, Jorge, David Bravo, Guido Del Pino, Graciela Donoso, Manuel Martínez, y Raúl Pizarro. 2006. "Estudio acerca de la Validez Predictiva de los Factores de Selección a las Universidades del Consejo de Rectores." Comité Técnico Asesor del CRUCH.
- Manzi, Jorge, David Bravo, Guido Del Pino, Graciela Donoso, Manuel Martínez, y Raúl Pizarro. 2008. "Estudio acerca de la Validez Predictiva de los Factores de Selección a las Universidades del Consejo de Rectores." Comité Técnico Asesor del CRUCH.
- Mattern, Krista D., Brian F. Patterson, Emily J. Shaw, Jennifer L. Kobrin, and Sandra M. Barbuti. "Differential Validity and Prediction of the SAT" (College Board Research Report No. 2008-4). New York: The College Board.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811.
- Wainer, H., & Steinberg, L. S. (1992). Sex differences in performance on the mathematics section of the Scholastic Aptitude Test: A bidirectional validity study, *Harvard Educational Review*, 62, 323–36.

- Young, J. W. (2001). Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis (College Board Research Report No. 2001-6). New York: The College Board.
- Zwick, R. (2006). Higher Education Admissions Testing, en R. Brennan (Ed.), Educational measurement (4th ed., pp. 647-679). National Council on Measurement in Education Greenwood Press., West Port, CT.



MEDICIÓN DE NUEVOS ATRIBUTOS PARA EL SISTEMA DE ADMISIÓN A LA EDUCACIÓN SUPERIOR EN CHILE

MEASUREMENT OF NEW ATTRIBUTES FOR THE ADMISSION SYSTEM TO HIGHER EDUCATION IN CHILE

María Verónica Santelices, Juan José Ugarte, María Paulina Flotts, Darinka Radović, Ximena Catalán, y Patrick C. Kyllonen

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art3.pdf>

Fecha de recepción: 15 de marzo de 2010
Fecha de dictaminación: 27 de abril de 2010
Fecha de segundo envío: 10 de mayo de 2010
Fecha de aceptación: 11 de mayo de 2010

1. ANTECEDENTES

Este documento describe el desarrollo y proceso de validación inicial de nuevos instrumentos creados para medir atributos que complementan la prueba cognitiva estandarizada usada actualmente para la admisión de pregrado en Chile. Específicamente, los resultados de esta investigación influirán en el proceso de toma de decisiones en cuatro universidades que en la actualidad están considerando el potencial uso de estos instrumentos en sus procesos de admisión de alumnos de pregrado.

Los instrumentos desarrollados son:

- (a) un Formulario Ampliado de Postulación, que solicita información académica en profundidad, información sobre participación en actividades extracurriculares, honores y premios recibidos, y puestos de liderazgo
- (b) una Reflexión Personal
- (c) un Ensayo de Pensamiento Crítico
- (d) tres cuestionarios estandarizados, que evalúan estrategias metacognitivas, habilidades sociales y proactividad

Estas autoevaluaciones, instrumentos de recolección de información biográfica y ensayos fueron desarrollados con el propósito de medir los siguientes constructos: (1) motivación para el desarrollo personal, social y académico, y (2) pensamiento crítico. Todos estos son atributos considerados relevantes por un panel de autoridades de la universidad y representantes de las facultades y de los estudiantes, quienes se reunieron durante el año 2007 para decidir cuáles eran las características de interés más importantes en un postulante.

Se espera que el desarrollo de estos nuevos instrumentos, y su futuro uso en admisión, disminuyan algunas de las bien conocidas ventajas dadas por los indicadores académicos tradicionales a los estudiantes de mayor nivel socioeconómico. Además, estos instrumentos enviarán una señal clara a los establecimientos educacionales sobre los atributos que éstos deberían ayudar a desarrollar, expandiendo así el currículum anual de los estudiantes de último año de enseñanza secundaria, el cual actualmente está enfocado en la preparación de la prueba estandarizada de admisión a las universidades. El desarrollo de los instrumentos descritos en este documento da la oportunidad de considerar las múltiples dimensiones de inteligencia que son relevantes para el logro del éxito universitario y profesional. Los resultados además servirán de referencia para esfuerzos similares que se están realizando en otras partes del mundo.

En Chile existen actualmente 61 universidades, 25 de las cuales reciben directamente financiamiento público, en tanto que las otras 36 son instituciones privadas. Las instituciones que reciben fondos del Estado conforman el Consejo de Rectores de las Universidades Chilenas (CRUCH), creado en el año 1954, y entre cuyas principales funciones está la definición de los criterios y procedimientos de postulación y selección a estos planteles de educación superior.

A partir del año 1967, el CRUCH estableció un mecanismo de admisión común, basado en atributos de carácter cognitivo: las calificaciones obtenidas por el alumno durante su enseñanza secundaria y su desempeño en pruebas estandarizadas diseñadas para el proceso de selección¹. Desde el año 1967 hasta el año 2002, las pruebas estandarizadas de admisión fueron la Prueba de Aptitud Académica (PAA), y la Prueba de Conocimientos Específicos (PCE). La PAA consistía de dos instrumentos (PAA Verbal y PAA Matemáticas), y fue diseñada para medir la capacidad académica de los egresados de la educación secundaria, centrándose en las habilidades del estudiante. Por otro lado, en las PCE, los contenidos evaluados provenían del currículo oficial de la enseñanza secundaria.

En el año 2003, la PAA y las PCE fueron sustituidas por la Prueba de Selección Universitaria (PSU), un conjunto de pruebas desarrollada y administrada por el Departamento de Evaluación, Medición y Registro Educativo (DEMRE)². La PSU consiste en cuatro instrumentos de evaluación cognitiva que miden la capacidad de razonamiento de los postulantes egresados de la enseñanza secundaria, teniendo como referente el currículo común de primero a cuarto año de enseñanza secundaria. Estas pruebas son: Lenguaje y Comunicación, Matemática, Historia y Ciencias Sociales, y Ciencias. Esta última prueba incluye un módulo común y otro a elección entre Biología, Física y Química. Los candidatos deben rendir en forma obligatoria la prueba de Lenguaje y Comunicación y la de Matemática, y pueden elegir entre la prueba de Historia y Ciencias Sociales y la de Ciencias (para mayor información revisar el sitio www.cruch.cl).

En la actualidad, la Pontificia Universidad Católica de Chile (PUC) está liderando un estudio que busca complementar los criterios utilizados actualmente en la admisión a la educación superior, evaluando la posibilidad de incluir la medición de atributos no cognitivos y la capacidad de pensamiento crítico de los postulantes. La PUC es una de las dos instituciones de educación superior más prestigiosas del país, siendo reconocida tanto por la calidad de sus programas de pregrado y postgrado, como por su investigación. Cada año esta universidad recibe alrededor de 21.000 postulaciones, seleccionando alrededor de 3.300 estudiantes.

Durante el año 2008, se sumaron a esta investigación otros planteles de educación superior pertenecientes al CRUCH: la Universidad de Concepción, la Universidad de Santiago y la Universidad Técnica Federico Santa María. La participación de estas universidades en el estudio, junto a la de 20 establecimientos de educación secundaria, aumentó de manera significativa la validez de los resultados obtenidos.

2. REVISIÓN BIBLIOGRÁFICA

La literatura estadounidense sobre el proceso de admisión a las instituciones de educación superior da cuenta de que un gran número de colleges y universidades han decidido complementar los indicadores académicos tradicionales con indicadores de atributos personales relevantes para cada institución. Esta incorporación se ha hecho a la luz de las misiones institucionales y como una manera de disminuir las

¹ El desarrollo del test y su administración son centralizados y el proceso a nivel nacional fue delegado por el CRUCH a la Universidad de Chile, una de las universidades miembros.

² Organismo encargado de la creación de pruebas y la administración del proceso de admisión y selección de alumnos para las Universidades del Consejo de Rectores desde el año 1996.

diferencias que han existido entre grupos socioeconómicos al utilizarse criterios académicos tradicionales (Breland, Maxey, Gernand, Cumming, & Trapani, 2001; Geiser & Santelices, 2007; Rigol, 2003; Zwick & Grief Greene, 2007).

Aunque ha sido un ejercicio común entre las instituciones privadas desde hace mucho tiempo, las cifras de la Encuesta de Prácticas de Admisión (Breland et al., 2001; 2000) sugieren que también se ha extendido a las instituciones públicas. Las decisiones de admisión consideran diferentes dimensiones del postulante dependiendo de la misión y filosofía institucional (Perfetto, 1999). Por ejemplo, algunas instituciones ponderan en un mayor porcentaje los indicadores académicos, mientras otras le dan un mayor valor al potencial que tiene el postulante para contribuir a la institución. Este alineamiento entre la misión institucional y los criterios de admisión tiene una importancia fundamental y su consistencia debería ser monitoreada constantemente (College Board, 2002).

Desde un punto de vista teórico, además, la medición y consideración de múltiples atributos en los postulantes a una institución de educación superior se fundamenta en al menos dos paradigmas de la psicología de la inteligencia: (1) el paradigma de las inteligencias múltiples (Gardner, 1983) y (2) el paradigma de la inteligencia emocional (Goleman, 1996). Gardner enfatiza que cada inteligencia es un sistema en sí mismo, que éstas son independientes entre sí pero que a la vez interactúan. Goleman por su parte considera que la inteligencia emocional es una metahabilidad que determina el éxito con que podemos utilizar otras habilidades (Rosas, Boetto, & Jordan, 2004).

La evidencia ha confirmado la importancia de las habilidades no cognitivas para lograr el éxito tanto en las escuelas de postgrado como en la vida laboral (Heckman & Rubinstein, 2001; Kyllonen, 2005). Dada esta importancia, el College Board, ETS y ACT han estado realizando varios proyectos enfocados en el desarrollo de evaluaciones no cognitivas, con una variedad de propósitos y en diversos contextos educacionales (Camara, 2007; Kyllonen, 2005; Robbins, Lauver, Le, Davis, Langley, & Carlstrom, 2004).

El College Board junto con investigadores de la Universidad del Estado de Michigan han estado desarrollando instrumentos que combinan preguntas de juicio situacional y de información biográfica para evaluar los constructos que hay detrás de la declaración de misión de 35 instituciones de educación superior de Estados Unidos. Hace un par de años se terminó un primer estudio, realizado con aproximadamente 2.000 estudiantes y 10 instituciones, y posteriormente se inició otro de mayor envergadura (8.000 estudiantes y 15 instituciones). Esta iniciativa ha abordado, entre otras cosas, el falseamiento de la información, especialmente de la información biográfica (P. Kyllonen 2007), y los costos y políticas de implementación. Para mayor información ver: <http://iopsych.msu.edu/cbstudy/index.htm>.

Por otro lado, ETS estrenó en julio de 2009 la carta de recomendación estandarizada. El Personality Potential Index (PPI) es un sistema de calificación en línea en el cual los evaluadores califican los futuros estudiantes de postgrado, complementando así el puntaje obtenido por éstos en el GRE. El PPI fue desarrollado a partir de investigación en profundidad sobre los atributos más importantes para lograr el éxito y ha sido piloteado con practicantes de verano del ETS (Liu, Minsky, Ling, & Kyllonen, 2009). Este índice proporciona una imagen del candidato que va más allá de las calificaciones y de los puntajes en las pruebas, enriqueciendo el proceso de evaluación y permitiendo a los estudiantes dar evidencia de un amplio rango de capacidades (<http://www.ets.org/ppi>).

Robert Sternberg también ha dirigido importantes esfuerzos en el desarrollo e implementación de evaluaciones no cognitivas, basados en décadas de investigación acerca de su "teoría de la inteligencia triárquica" (práctica, creativa y analítica) (Sternberg, 1999, 2003). Su primer proyecto tuvo lugar en la Escuela de Negocios de la Universidad de Michigan, en donde experimentó con mediciones de inteligencia práctica, incluyendo ítems de juicio situacional y problemas de caso. El Proyecto Arcoiris (Rainbow Project), patrocinado por el College Board, fue su segundo intento e involucró admisiones de pregrado en un grupo de universidades. Los instrumentos de evaluación no cognitiva fueron de tres tipos: (1) Creativo – Interpretando caricaturas; escribiendo ensayos, (2) Analítico- Mediciones tradicionales de habilidad, y (3) Práctico - Pruebas de juicio situacional. La iniciativa más reciente de Sternberg, el Proyecto Caleidoscopio (Kaleidoscope Project), también trabajó con procesos de admisión de pregrado y se desarrolló en la Universidad Tufts. Las mediciones incluyeron ensayos optativos, evaluados por su creatividad ("qué pasa si"), capacidad práctica y sabiduría, usando una rúbrica de corrección.

También en Estados Unidos, Robbins, Allen & Sawyer (2007) estudiaron la capacidad del Inventario de Preparación del Estudiante (STI) para predecir el promedio de notas y la persistencia en primer año de universidad. El STI fue desarrollado por investigadores de ACT (Le, Casillas, Robbins, & Langley, 2005) y tuvo como objetivo medir motivación, hábitos de estudio, auto control y participación social. Oswald et al. (2004) estudiaron la validez de un instrumento de medición de datos biográficos y de un inventario situacional en una muestra de 600 estudiantes de college. Los autores basaron el desarrollo de los instrumentos en doce constructos relevantes para la admisión, agrupándolos en tres dimensiones (cognitiva, personal e interpersonal) e indagaron en la capacidad de estas mediciones para predecir la asistencia a clases durante el primer año de universidad y el promedio de notas del alumno. Estos dos estudios en particular sirvieron como modelo para el diseño y la metodología de la investigación presentada en este artículo.

Durante los últimos años, en Europa han comenzado una serie de proyectos de evaluación no cognitiva. Algunas universidades alemanas y austriacas están implementando autoevaluaciones para ayudar a los estudiantes a elegir apropiadamente sus carreras (ver EAPA Symposium, 2007), reduciendo así la tasa de deserción. Las nuevas mediciones incluyen evaluación de la personalidad, inventarios de interés y test cognitivos.

Los estudios que han explorado la validez de estos instrumentos en las universidades han arrojado que sólo contribuyen moderadamente a predecir los resultados académicos de corto y/o largo plazo como las calificaciones, pero que pueden mejorar la predicción de definiciones más amplias de éxito universitario (incluyendo liderazgo y participación cívica), así como también la predicción de la persistencia y de las tasas de graduación. La consideración de estos atributos también tiene como consecuencia menores diferencias en el desempeño de estudiantes de distintos grupos socio-demográficos (Camara, 2005; Cliffordson & Askling, 2006; Willingham, 1985).

2.1. Preguntas de investigación

Este artículo tiene las siguientes preguntas de investigación:

- 1) ¿Cuál es la validez y confiabilidad de los nuevos instrumentos?
- 2) ¿Cuál es la relación entre el desempeño en los nuevos instrumentos y las características sociodemográficas de los estudiantes?

3) ¿Cuál es la relación entre el desempeño en los nuevos instrumentos y el desempeño de los estudiantes en los indicadores de admisión actual?

3. INSTRUMENTOS, MUESTRA, CORRECCIÓN Y METODOLOGÍA

La siguiente sección presenta detalles sobre los instrumentos y su proceso de desarrollo, la muestra de estudio y las pautas de corrección. Por su parte, la sección de metodología explica el análisis realizado.

3.1. Instrumentos y su proceso de desarrollo

Los cuatro tipos de instrumentos de esta investigación fueron desarrollados en base a un perfil de postulante, el cual fue definido por un comité formado tanto por autoridades de la Pontificia Universidad Católica de Chile, como por estudiantes y representantes de las facultades. Este perfil incluyó una serie de atributos personales, interpersonales y académicos de los cuales se escogieron aquellos más factibles de ser evaluados de forma estandarizada y que pudiesen ser interesantes para otras universidades. En base a la literatura de medición de estas habilidades, se decidió poner foco en la evaluación de pensamiento crítico y de la motivación para el desarrollo académico, personal y social³. Las evaluaciones fueron desarrolladas en base a la definición de constructos encontrada en la literatura y el proceso incluyó el desarrollo de ítems por el equipo del proyecto, los aportes de los colegios participantes, la asesoría de miembros de las facultades y de un experto internacional del ETS, entrevistas en profundidad con estudiantes universitarios y una aplicación de pre-pilotaje. La tabla 1 muestra los constructos de interés y sus instrumentos.

TABLA 1. DIMENSIONES E INSTRUMENTOS

Dimensión	Instrumentos
Pensamiento Crítico	Ensayo de Pensamiento Crítico
Motivación para el desarrollo personal, académica y social	Formulario Ampliado de Postulación
	Reflexión Personal
	Cuestionarios Estandarizados (estrategias metacognitivas, habilidades sociales, proactividad)

3.1.1. Ensayo de Pensamiento Crítico

Se desarrolló un ensayo para medir pensamiento crítico, entendido como una habilidad de pensamiento reflexivo y cuestionamiento que se expresa a través de la capacidad de argumentación. Los estudiantes recibieron aleatoriamente uno de dos temas disponibles, los cuales presentaban dos posiciones opuestas en relación a un tópico general. Se pidió a los estudiantes analizar los argumentos que fundamentaban cada una de las posiciones, elegir una posición y explicar su decisión. La pauta de corrección del Ensayo de Pensamiento Crítico constó de diez sub-dimensiones teóricamente divididas en dos grupos: (1) el contenido o calidad de la argumentación y (2) aspectos formales. La dimensión de contenido incluyó la calidad del análisis de los argumentos dados, la presencia y calidad de la tesis, la formulación de argumentos, la coherencia de la argumentación en contra de uno o más contra-argumentos, la existencia

³ El perfil del postulante incluyó las siguientes características: (1) persistencia/superación de metas/resiliencia, (2) desarrollo personal /estudiantes reflexivos, (3) tolerancia, (4) creatividad, (6) motivación, (7) solidaridad, (8) conocimiento del área de interés y (9) concordancia con la misión de la institución.

y calidad de la conclusión, y la reflexión acerca de su propio punto de vista. Por otro lado, la subdimensión de aspectos formales incluyó la evaluación de la ortografía, el vocabulario, la cohesión textual, el uso de párrafos y una evaluación general.

3.1.2. Formulario Ampliado de Postulación

Se desarrollaron dos instrumentos con el objetivo de evaluar el nivel de motivación para el desarrollo personal, la participación social y cívica, y el desarrollo académico. El primero fue el Formulario Ampliado de Postulación, el cual se enfocó en la manifestación conductual de la motivación, recogiendo datos biográficos a través de preguntas cerradas agrupadas en tres sub-secciones: (1) participación en actividades extracurriculares, (2) honores y premios y (3) liderazgo. Además, el formulario tenía una pregunta final que solicitaba dar una explicación a los estudiantes que no habían participado en ninguna actividad durante su enseñanza secundaria. Cada estudiante podía reportar hasta un máximo de cinco actividades extracurriculares, tres premios y un cargo de liderazgo. Además, se solicitó la información de contacto de las personas o instituciones relacionadas con el desarrollo de esas actividades, con el fin de evaluar, en una muestra aleatoria, la veracidad de la información entregada por los estudiantes y disuadirlos de entregar información falsa⁴.

3.1.3. Reflexión Personal

El segundo instrumento desarrollado para medir motivación fue el ejercicio de Reflexión Personal, en el cual se solicitó a los estudiantes utilizar hasta tres actividades que fueran interesantes para ellos con el fin de ejemplificar (1) cómo encaran la dificultad, (2) cuánto perseveran, (3) cuál es su percepción de la autoeficacia, (4) describir su locus de control y (5) su capacidad para aprender desde la experiencia y proyectarse hacia el futuro. Aunque se trató de un ejercicio escrito, no se revisaron aspectos como ortografía y vocabulario.

3.1.4. Cuestionarios estandarizados

Se administraron tres cuestionarios cortos de alternativas múltiples, con el fin de evaluar proactividad, estrategias metacognitivas y habilidades sociales. Estos cuestionarios son mucho más fáciles de administrar y corregir que los ensayos, y por lo tanto su desempeño tuvo un especial interés para procesos de admisión de pregrado que podrían incluir entre 5.000 y 150.000 postulaciones. Estos cuestionarios ya habían sido desarrollados, y las alternativas de respuesta consistían en es escalas Likert cuyas opciones varían desde "totalmente de acuerdo" a "totalmente en desacuerdo".

3.1.5. Otros instrumentos

Además de los cuatro tipos de instrumentos descritos anteriormente, se administraron dos instrumentos adicionales, sólo como parte del estudio y sin intención de usarlos en el proceso de admisión. Estos instrumentos fueron el Cuestionario del Profesor y el Cuestionario de Deseabilidad Social (que es la traducción al español de la escala de Marlowe-Crowne). En el caso de este último instrumento, su incorporación tuvo como objetivo monitorear el grado de influencia de la deseabilidad social en las respuestas dadas por los alumnos, un problema recurrente en las mediciones no cognitivas. Este

⁴ Se realizó un proceso aleatorio de verificación como parte de esta investigación y también está contemplado como parte fundamental de la implementación operacional de estos instrumentos. La incorporación del "método de advertencia" ha sido exitosa en investigaciones anteriores (P. Kyllonen, Walters & Kaufman, 2005).

cuestionario consistía en ítems con respuestas de opción múltiple y fue administrado a los estudiantes junto con los instrumentos descritos arriba.

Por otro lado, el Cuestionario del Profesor fue desarrollado especialmente para este estudio, siendo distribuido y completado durante las dos semanas previas a la aplicación de instrumentos en las escuelas. Los profesores llenaron una página con preguntas cerradas por cada estudiante del curso. Se pidió a los profesores calificar a cada estudiante en constructos no cognitivos tales como: motivación por aprender, autoeficacia, visión de futuro, liderazgo, responsabilidad social, participación en actividades extracurriculares, entre otros. La escala de respuesta fue de cuatro puntos y la evaluación del alumno siempre estuvo referida a su relación con el resto de la clase. El Cuestionario del Profesor también incluyó preguntas relacionadas con el tiempo que el profesor conocía al estudiante y la profundidad de ese conocimiento, de manera de excluir del análisis aquellos casos en que el profesor no tenía suficiente información para caracterizar adecuadamente a los alumnos. Este instrumento fue administrado con el objetivo de recolectar información de una tercera persona que permitiese la triangulación de la información auto-reportada por el alumno, lo cual permite analizar la validez de las mediciones que se han estado desarrollando.

3.2. Muestra

La recolección de datos fue realizada durante el año 2008, sin consecuencias para los estudiantes que participaron del estudio. La batería de instrumentos fue administrada a una muestra de aproximadamente 1.400 estudiantes de primer año de varias carreras en cuatro instituciones de educación superior (ver tabla 2), todas miembros del CRUCH, y aproximadamente 1.600 estudiantes de último año de enseñanza secundaria de colegios particulares pagados, particulares subvencionados y municipalizados (los tres tipos de dependencia administrativa predominantes en el país). Se incluyeron estos dos grupos con el objetivo de obtener datos lo más parecidos posibles a los de un postulante universitario. Si bien el escenario ideal habría sido evaluar a los postulantes al mismo tiempo que rendían la prueba estandarizada para la admisión a la universidad (PSU), en esas circunstancias los estudiantes habrían tenido pocos incentivos para participar en el ejercicio, y aquellos que hubiesen participado podrían no ser necesariamente representativos de la población estudiantil.

En vez de esto, se decidió evaluar a los estudiantes en ambientes educacionales formales en los que podrían ser encontrados más fácilmente y evaluados de forma estandarizada. Cabe señalar al respecto que los estudiantes de último año de enseñanza secundaria se asemejan mucho a los postulantes universitarios, especialmente a finales de su último año, y los estudiantes de primer año de universidad son también parecidos a los postulantes, especialmente a comienzos del primer año de universidad, cuando aún no han sido modificados por la experiencia universitaria⁵. En base a estas consideraciones, la información de los estudiantes de primer año de universidad fue recogida entre abril y julio de 2008⁶ y para la muestra de escolares de último año fue recolectada entre septiembre y octubre de 2008. Todos los participantes firmaron un consentimiento formal para participar. Adicionalmente, los estudiantes de último año de enseñanza secundaria que eran menores de edad al momento de contestar los instrumentos, contaron con un consentimiento firmado por sus apoderados legales.

⁵ Es importante recordar que el año académico en Chile comienza en marzo y termina en diciembre.

⁶ La idea original era recolectar la información de las universidades entre abril y mayo de 2008, pero fue imposible debido a la existencia de protestas por temas legales y financieros en las escuelas y universidades.

La muestra de universitarios incluyó estudiantes de tres universidades no públicas: la Pontificia Universidad Católica de Chile (n=586), la Universidad de Concepción (n=283) y la Universidad Federico Santa María (n=176), además de una universidad pública, la Universidad de Santiago (n=398). Respecto a su ubicación, dos de estas instituciones están localizadas en la capital del país y dos están fuera del área metropolitana de Santiago, pero en zonas que también son densamente pobladas. En relación a sus procesos de admisión, estas instituciones presentan diferencias en sus grados de selectividad: mientras las tres primeras están dentro de las más selectivas del país, la cuarta admite a casi todos sus postulantes⁷.

Las carreras seleccionadas para participar en la investigación fueron elegidas con el propósito de tener una adecuada representación de (1) áreas y disciplinas, (2) selectividad y (3) características de los estudiantes, tales como género, rendimiento académico y perfil socioeconómico (ver tabla 2). Los estudiantes incluidos en la muestra presentaron diferencias tanto en sus características socioeconómicas como académicas.

Las escuelas secundarias participantes en el estudio representaron a los tres tipos existentes en el sistema educacional chileno (particular, particular subvencionado y municipal), y la mayoría es representativa de las escuelas cuyos estudiantes han estado tradicionalmente en la lista de espera de la Pontificia Universidad Católica de Chile (en otras palabras, los puntajes de sus egresados han estado tradicionalmente cerca del punto de corte en sus respectivas carreras)^{8,9,10}. La inclusión de tres colegios municipales y dos particulares subvencionados, de especial interés para algunas de las universidades participantes del estudio, amplían significativamente las características académicas de la muestra de estudiantes.

Es importante destacar que la participación de las instituciones en el estudio fue voluntaria, pero motivada por miembros de cada una de las escuelas y universidades. Para motivar la participación, el equipo de investigación hizo presentaciones acerca del estudio en algunas de las escuelas y universidades, reuniéndose con los estudiantes y/o profesores. Además de esto, tanto los estudiantes de último año de enseñanza secundaria como los universitarios de primer año que contestaron todos los instrumentos participaron en el sorteo de 300 reproductores portátiles de música. Además, en el caso de los universitarios, se ofreció un incentivo alimenticio.

⁷ En el año 2007 la PUC aceptó a un 39% de sus postulantes, la Universidad de Concepción a un 63%, la Universidad Federico Santa María a un 68% y la Universidad de Santiago a un 95%.

⁸ Además fueron invitados a participar estudiantes del programa PENTA UC, de la PUC, y de 3 escuelas municipales de bajo rendimiento, con las cuales la Universidad de Santiago trabaja de cerca. El programa PENTAUC está diseñado para reforzar la enseñanza y aprendizaje de estudiantes talentosos provenientes de entornos desfavorecidos.

⁹ Aunque el Ministerio de Educación chileno fija los lineamientos pedagógicos, y define y provee los recursos para las escuelas del país, la autoridad municipal administra las escuelas públicas localmente. Actores privados, sean individuos o instituciones, administran tanto las escuelas particulares pagadas como las particulares subvencionadas. En el año 2007 había aproximadamente 11,800 escuelas trabajando en el sistema, 50% de las cuales eran municipales, 43% particulares subvencionadas y 6% particulares pagadas (Ministerio de Educación 2009). Aproximadamente un 45% de los estudiantes en edad escolar asisten a escuelas municipales, 46% a particulares subvencionadas y sólo un 7% asiste a escuelas particulares pagadas. En general, los estudiantes de nivel socioeconómico más bajo asisten a escuelas públicas, mientras que los de mayor nivel socioeconómico asisten a escuelas particulares.

¹⁰ Esta selección fue hecha en base al análisis de datos de tres años (2004-2007) y considerando una distancia de 10 y 20 puestos por arriba y por abajo del puntaje de corte. La PSU de matemática y la de lenguaje y comunicación tiene un rango de puntajes de 200 a 850 puntos. Un total de 243,575 estudiantes rindieron la PSU de Lenguaje y Comunicación, y 242,483 la PSU de Matemáticas. Para ambas pruebas la media es ajustada para estar en los 500 puntos y la desviación estándar es de 109 puntos.

La administración de las pruebas se realizó durante el horario de clases, con una duración promedio de 3 horas y con una tasa de participación del 50% de los alumnos, aproximadamente. Todos los instrumentos fueron administrados por un equipo capacitado, siguiendo un protocolo estandarizado de administración. Sólo fueron considerados en la muestra final los estudiantes que se graduaron de la enseñanza secundaria en el año 2002 o posteriormente, y que comenzaron la universidad en el año 2007 ó 2008. Los estudiantes de la muestra final tenían en promedio 18,5 años de edad, 52% eran hombres y aproximadamente un tercio provenía de escuelas públicas (34%), un tercio de escuelas particulares (32%) y un tercio de escuelas particulares subvencionadas (31%).

TABLA 2. CARACTERÍSTICAS DEMOGRÁFICAS DE LA MUESTRA DE UNIVERSITARIOS POR UNIVERSIDAD Y CARRERA¹¹

Universidad	Carrera	Edad promedio	Ingreso familiar mensual (2009 US\$)	% Mujeres	% Padres con educación superior	Promedio PSU Matemática (2008)	Promedio PSU Lenguaje (2008)	N Total
P. Universidad Católica de Chile	Agronomía	19	1,015-1,269	52%	57%	672	640	79
P. Universidad Católica de Chile	Arquitectura	19	1,523-2,030	41%	85%	734	717	27
P. Universidad Católica de Chile	Bachillerato	19	1,523-2,030	65%	84%	670	695	124
P. Universidad Católica de Chile	Educación	20	1,015-1,269	92%	54%	615	636	179
P. Universidad Católica de Chile	Enfermería	19	761-1,015	75%	63%	675	678	55
P. Universidad Católica de Chile	Ingeniería	19	1,523-2,030	14%	93%	812	714	122
Universidad de Concepción	Arquitectura	20	508-761	51%	44%	623	603	73
Universidad de Concepción	Educación	20	254-508	96%	21%	542	557	70
Universidad de Concepción	Enfermería	19	508-761	68%	44%	645	647	101
Universidad de Concepción	Ingeniería	19	761-1,015	13%	45%	656	590	39
Universidad de Santiago	Arquitectura	20	761-1,015	57%	38%	605	632	53
Universidad de Santiago	Bachillerato	19	508-761	45%	37%	609	596	38
Universidad de Santiago	Enfermería	20	508-761	76%	34%	648	626	70
Universidad de Santiago	Ingeniería	19	761-1,015	20%	47%	687	636	237
Univ Federico Santa María	Ingeniería	19	1,015-1,269	21%	54%	682	616	176

TABLA 3. CARACTERÍSTICAS DEMOGRÁFICAS DE LA MUESTRA DE ALUMNOS DE ÚLTIMO AÑO DE ENSEÑANZA SECUNDARIA POR TIPO DE ESCUELA

Dependencia administrativa del establecimiento	Edad promedio	Ingreso familiar mensual (2009 US\$)	% Mujeres	% Padres con educación superior	Promedio PSU Matemática (2009)	Promedio PSU Lenguaje (2009)	N total
Municipal/Público	18	812-1,083	48%	27%	610	611	651
Particular subvencionado	18	1,083-1,353	40%	46%	564	561	398
Particular pagado	18	2,707-2,977	51%	95%	687	660	491
Programa Penta UC	18	541-812	54%	37%	649	641	28

¹¹ Fuente: archivos de datos 2008 y 2009 de los estudiantes inscritos para tomar la PSU, del DEPARTAMENTO DE EVALUACIÓN, MEDICIÓN Y REGISTRO EDUCACIONAL (DEMRE).

3.3. Pauta y proceso de corrección

Se desarrollaron pautas de corrección para cada uno de los instrumentos de respuesta abierta (Ensayo de Pensamiento Crítico, Reflexión Personal y Formulario Ampliado de Postulación), a través de un proceso multi-etápico. Un primer bosquejo de la rúbrica se hizo desde una perspectiva puramente teórica; este bosquejo fue posteriormente aplicado a una muestra de 20 evaluaciones por un corrector experimentado, proponiéndose cambios en base a la capacidad de las pautas de corrección de discriminar y a la factibilidad de su aplicación. En base a este proceso, se desarrollaron pautas específicas para cada instrumento siguiendo una aproximación analítica, es decir, fueron diseñadas para evaluar sub-dimensiones específicas dentro de cada instrumento, en oposición a una evaluación holística.

La corrección se realizó en dos procesos: los instrumentos aplicados en la Pontificia Universidad Católica de Chile, al estar disponibles antes, fueron evaluados en junio de 2008 y el resto de la muestra, administrada más tarde, fue evaluados en octubre de 2008. El Ensayo de Pensamiento Crítico y la Reflexión Personal fueron corregidos por correctores entrenados. En el caso del Formulario Ampliado de Postulación, las preguntas abiertas también fueron evaluadas, mientras que las preguntas cerradas fueron registradas directamente en una base de datos especialmente diseñada para ello. Las respuestas de los cuestionarios estandarizados, el Cuestionario de Deseabilidad Social y el Cuestionario del Profesor también fueron digitados directamente.

En el primer proceso de corrección trabajaron dos equipos diferentes de correctores a tiempo completo, reunidos en un mismo lugar físico durante diez días de trabajo. El primer equipo estuvo formado por 9 psicólogos y evaluó la Reflexión Personal y el Formulario Ampliado de Postulación. El segundo equipo, compuesto de 4 profesores de lingüística con experiencia en la evaluación de instrumentos similares, trabajó en el Ensayo de Pensamiento Crítico. Cada equipo contó con un supervisor y tuvo un día completo de entrenamiento.

El segundo proceso de corrección también tomó dos semanas, desarrollándose en el mismo espacio físico, y tuvo un equipo de correctores a tiempo completo por cada instrumento (tres equipos en total): 17 evaluadores corrigieron el Ensayo de Pensamiento Crítico, 8 corrigieron el Formulario Ampliado de Postulación y 14 evaluaron la Reflexión Personal. Cada uno de estos equipos tuvo un supervisor, el perfil de los correctores fue similar al descrito para el primer proceso de corrección, y todos recibieron una capacitación similar. Además, todos los correctores firmaron un compromiso de confidencialidad.

En ambos casos el entrenamiento consistió en una revisión detallada de la rúbrica de evaluación, presentación y discusión de respuestas tipo para cada nivel de la rúbrica y discusión grupal de un primer caso evaluado.

El nivel de acuerdo entre evaluadores fue monitoreado a través de la doble corrección de todos los ensayos de pensamiento crítico y reflexiones personales, y de un 20% de las preguntas abiertas del Formulario Ampliado de Postulación. Diferencias de puntaje de más de un punto se consideraron "inconsistentes", siendo necesaria una tercera evaluación, otorgada por el equipo supervisor, en tanto que los evaluadores discrepantes tuvieron un re-entrenamiento sobre la pauta de corrección (las escalas de cada instrumento se muestran en la tabla 5).

3.4. Metodología

La validez y confiabilidad de los instrumentos fue medida a través de indicadores de consistencia interna, dificultad relativa de los ítems, confiabilidad entre jueces y análisis factorial. Se estudió la correlación

entre el auto-reporte del estudiante y el reporte del profesor jefe, así como la correlación de puntajes de los nuevos instrumentos con el Cuestionario de Deseabilidad Social, las notas de enseñanza secundaria y los puntajes en pruebas estandarizadas. La relación entre desempeño y variables socioeconómicas fue analizada observando la diferencia de medias de los puntajes estandarizados y la proporción de la varianza explicada en el puntaje estandarizado.

4. RESULTADOS

Esta sección presenta un resumen de los resultados por tipo de instrumento, y la relación entre los nuevos instrumentos y variables de interés tales como el puntaje en la prueba cognitiva estandarizada usada actualmente para la admisión (PSU), el Cuestionario de Deseabilidad Social y el Cuestionario del Profesor. Se analizaron también las diferencias en el impacto de las nuevas medidas.

4.1. Descriptivos por instrumento

La tabla 4 muestra el rendimiento promedio en cada ítem de cada instrumento. En el Ensayo de Pensamiento Crítico, los estudiantes obtuvieron altos puntajes en la calidad de los argumentos dados y en la presentación de su propia tesis, y menores puntajes en la evaluación de su propio punto de vista y en ítems como vocabulario y ortografía. Estos resultados son consistentes con los observados en la aplicación de pruebas similares en estudiantes chilenos en edad escolar (Manzi & Flotts, 2006). El análisis factorial de los ítems confirma la presencia de dos sub-dimensiones: una que evalúa contenido y otra que evalúa la forma en que escriben los estudiantes (ver Apéndice 1 para detalles). El puntaje final ha sido propuesto como un promedio ponderado en el que la dimensión de contenido pesa un 70% y los aspectos formales un 30%.¹²

Los resultados presentados en la tabla 4 indican que la Reflexión Personal fue el instrumento más difícil de todos. En particular, los alumnos tuvieron dificultades para presentar evidencia que les permitiera sustentar sus proposiciones. Junto con esto, pudo observarse la existencia de un mejor desempeño en los ítems que aparecieron primero en la página de instrucciones (Ej. perseverancia y esfuerzo). El análisis factorial exploratorio realizado confirmó la presencia de 1 factor (ver el Apéndice 2). El puntaje final para este instrumento ha sido propuesto como un promedio simple de todas las sub-dimensiones.

En el Formulario Ampliado de Postulación, los estudiantes reportaron un promedio de 2.4 actividades, pero sólo 1.4 de esas actividades fueron motivadas por el deseo de lograr desarrollo personal, académico o social. Los estudiantes de colegios particulares mostraron una mayor participación en actividades extracurriculares que los provenientes de colegios municipales o particulares subvencionados. Para evitar la existencia de algún sesgo derivado del tipo de colegio (municipal, particular subvencionado o particular pagado) en estos resultados, la participación en actividades extracurriculares fue ajustada por el número de actividades disponible en cada tipo de escuela.¹³

¹² Esta ponderación se sugiere como una forma de considerar aspectos formales en el puntaje final preservando la mayor parte de la ponderación para aspectos de contenido. Los ponderadores podrían ser revisados en una eventual implementación o reanálisis de los datos.

¹³ Si bien este ajuste tiene por objetivo corregir por diferencias en la disponibilidad de actividades entre establecimientos de distinta dependencia, pueden aun existir diferencias por lugar de residencia y/o nivel socioeconómico de los alumnos.

TABLA 4. DESEMPEÑO EN LOS ÍTEMS DE LOS INSTRUMENTOS

Variable	N	Media	Desviación estándar	Mínimo	Máximo
ENSAYO DE PENSAMIENTO CRÍTICO*					
Análisis de argumentos	2592	3.05	0.69	1	4
Tesis	2592	3.44	0.56	1	4
Argumentos	2592	2.98	0.60	1	4
Contrargumento	2592	1.67	0.78	1	4
Conclusión	2592	3.12	0.76	1	4
Autoreflexión	2592	1.43	0.73	1	4
Ortografía	2592	1.72	0.75	1	4
Vocabulario	2592	1.58	0.62	1	4
Uso de oraciones	2592	2.85	0.79	1	4
Uso de párrafos	2592	2.82	0.72	1	4
Evaluación general	2592	2.20	0.57	1	4
FORMULARIO AMPLIADO DE POSTULACIÓN**					
Actividades extracurriculares	3003	0.23	0.22	0	1
Diversidad de la motivación	3003	0.34	0.29	0	1
Nº de años de la participación	3003	0.50	0.27	0	1
Perseverancia entre actividades	2756	0.14	0.34	0	1
Nivel máximo de premios	3003	0.21	0.27	0	1
Tipo de liderazgo	1443	0.33	0.33	0	1
Efectividad del liderazgo	1380	0.24	0.32	0	1
REFLEXIÓN PERSONAL***					
Significado de la dificultad	2132	0.82	0.79	0	2
Perseverancia	2132	0.78	0.72	0	2
Autoeficacia	2132	0.52	0.60	0	2
Locus de control interno	2132	0.71	0.78	0	2
Experiencia & visión de futuro	2132	0.44	0.54	0	2
Evaluación global	2132	1.85	0.70	1	4
CUESTIONARIOS ESTANDARIZADOS****					
Habilidades sociales	2743	61.90	6.42	38	79
Estrategias metacognitivas	2740	52.23	7.85	20	72
Proactividad	2735	28.83	2.52	8	32

*Los ensayos de pensamiento crítico extremadamente cortos no fueron considerados en el análisis. Los ensayos más largos fueron considerados pues se asemejan mejor al comportamiento de estudiantes en una situación de prueba con altas consecuencias.

** Todas las variables fueron estandarizadas para facilitar la comparación.

***Sólo las reflexiones personales en las cuales los alumnos entregaron evidencia apoyando su motivación para el desarrollo personal, social o académico fueron consideradas.

****Sólo los cuestionarios que tuvieron más del 50% de las preguntas completadas fueron considerados como casos válidos pues se asemejan mejor al comportamiento de estudiantes en una situación de prueba con altas consecuencias.

La mayoría de los estudiantes reportaron en el formulario la participación en actividades deportivas (31%) y académicas (17%). En relación a las razones esgrimidas para realizar estas actividades, un 22% señaló una motivación académica, un 21% una motivación personal, un 18% una motivación social y un

38% reportó otros tipos de motivación¹⁴. Sólo una pequeña proporción de estudiantes que no reportó participación en actividades extracurriculares mencionó variables de contexto (es decir, que están fuera de su control) como razón para no realizar actividades, no existiendo cambios significativos en la proporción según el tipo de escuela (20 de 100 en escuelas públicas versus 8 de 54 en escuelas privadas).

Los estudiantes podían reportar un máximo de tres premios cada uno, pero en promedio reportaron sólo uno. En relación al tipo de premio, cerca de un 80% del total estuvieron relacionados con deportes o actividades académicas; respecto a su procedencia, un 75% de los premios fueron entregados por escuelas, especialmente escuelas privadas (42%). El puntaje dado en este ítem fue asignado en base a cuál fue el premio reportado de mayor nivel (Ej. si el premio fue a nivel de escuela, a nivel municipal, a nivel regional o internacional), y no en base al número de premios.

Cerca de la mitad de los estudiantes que participaron en el estudio reportó haber tenido un puesto de liderazgo durante la enseñanza secundaria (48,9%). De ellos, 42% describió actividades tales como la conducción o monitoreo de una o más personas orientadas a la misma tarea (el nivel más básico de liderazgo evaluado), mientras que un 12% refirió a niveles más complejos de liderazgo¹⁵. El puntaje general de liderazgo consideró tanto el tipo de liderazgo como la efectividad en el logro de objetivos del rol de liderazgo, según lo descrito por el estudiante. El puntaje final del Formulario Ampliado de Postulación dio un 40% de peso a la participación en actividades extracurriculares (consideradas como un indicador de motivación), 20% a la perseverancia en esas actividades, 20% a los premios y reconocimientos y 20% al puntaje de liderazgo.¹⁶

En relación a la confiabilidad de los instrumentos, existen algunos matices. La Reflexión Personal arrojó un coeficiente Alfa de Cronbach de 0.8, mientras que el Ensayo de Pensamiento Crítico mostró un menor Alfa de Cronbach, alcanzando un valor de 0,57. En tanto, los cuestionarios estandarizados arrojaron valores Alpha de Cronbach de 0.76 para el Cuestionario de Habilidades Sociales, 0.83 para el Cuestionario de Habilidades Metacognitivas y 0.66 para el Cuestionario de Proactividad.

TABLA 5. TASA DE ACUERDO ENTRE JUECES: NÚMERO DE CASOS CON DIFERENCIAS DE PUNTAJE DE 1 Ó 0.5 PUNTOS POR INSTRUMENTO

Proceso de corrección	Instrumento	Puntaje máximo	N doblemente corregido	Diferencia de 0.5 puntos o menos	Diferencia de 1 punto o menos
Primero	Ensayo de Pensamiento Crítico	4	383	98%	99%
	Formulario Ampliado de Postulación(Liderazgo)	3	116	98%	100%
Segundo	Reflexión Personal	3	516	86%	99%
	Ensayo de ensamiento crítico	4	2,306	87%	99%
	Formulario Ampliado de Postulación(Liderazgo)	3	489	96%	99%
	Reflexión Personal	3	2,424	77%	98%

¹⁴ 31% de estos estudiantes reportó hacer actividades "para entretenerse".

¹⁵ Estos niveles más altos de liderazgo incluyen la coordinación de personas trabajando en distintas tareas, responsabilidad de toma de decisiones, la supervisión y evaluación del trabajo de otros y/o la representación de los intereses u opiniones de un grupo.

¹⁶ Los ponderadores asignados tienen por objetivo relevar la información que es objetiva y más fácilmente verificable (participación en actividades extracurriculares) entregando igual ponderación a los ítems restantes.

La tabla 5 muestra el acuerdo entre jueces por instrumento y el proceso de corrección. Se puede observar que más del 85% de los instrumentos doblemente corregidos obtuvieron puntajes totales con una diferencia de 0.5 o menos. La única excepción es el segundo proceso de corrección de la Reflexión Personal, en donde un 77% del total de puntajes en los instrumentos presentaron esta diferencia.

4.2. Relación entre instrumentos

Esta sección presenta la relación de los instrumentos del Estudio Piloto de Admisión entre ellos mismos y con otras variables de interés, a saber, los instrumentos actualmente usados para admisión (pruebas PSU y notas de enseñanza secundaria), el Cuestionario de Deseabilidad Social y la Encuesta del Profesor.

4.2.1. Correlación entre los nuevos instrumentos

En general hay bajas correlaciones entre las sub-dimensiones del Ensayo de Pensamiento Crítico, la Reflexión Personal y el Formulario Ampliado de Postulación, lo que puede interpretarse como indicador de que los instrumentos miden atributos diferentes. La mayor correlación es de 0.17 ($p < 0.001$) observada entre la Reflexión Personal y la subsección de contenido del Ensayo de Pensamiento Crítico. Excepto por esto, y aunque son estadísticamente significativas, la mayoría de las otras correlaciones no excede el 0.10.

TABLA 6. CORRELACIÓN ENTRE INSTRUMENTOS

Instrumento	Estadísticos	Pensamiento Crítico (Contenido)	Pensamiento Crítico (Aspectos Formales)	Formulario Ampliado de Postulación	Reflexión Personal	C.Habilidades Sociales	C. Estrategias Metacognitivas	C. Proactividad
Pensamiento Crítico (Contenido)	r	1	0.21	0.10	0.17	0.02	0.12	0.05
	Valor- P		<.0001	<.0001	<.0001	0.4204	<.0001	0.006
	N	2714	2714	2707	1985	2711	2709	2704
Pensamiento Crítico (Aspectos Formales)	r	0.21	1	0.05	0.05	0.00	0.10	0.04
	Valor- P	<.0001		0.0181	0.0242	0.8039	<.0001	0.0568
	N	2714	2714	2707	1985	2711	2709	2704
Formulario Ampliado de Postulación	r	0.10	0.05	1.00	0.08	0.16	0.16	0.11
	Valor- P	<.0001	0.0181		0.0001	<.0001	<.0001	<.0001
	N	2707	2707	3003	2129	2737	2734	2729
Reflexión Personal	r	0.17	0.05	0.08	1.00	0.07	0.12	0.09
	Valor- P	<.0001	0.0242	0.0001		0.0031	<.0001	<.0001
	N	1985	1985	2129	2132	2007	2005	2003
C.Habilidades Sociales	r	0.02	0.00	0.16	0.07	1.00	0.42	0.36
	Valor- P	0.4204	0.8039	<.0001	0.0031		<.0001	<.0001
	N	2711	2711	2737	2007	2743	2740	2735
C. Estrategias Metacognitivas	r	0.12	0.10	0.16	0.12	0.42	1.00	0.36
	Valor- P	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
	N	2709	2709	2734	2005	2740	2740	2734
C. Proactividad	r	0.05	0.04	0.11	0.09	0.36	0.36	1.00
	Valor- P	0.006	0.0568	<.0001	<.0001	<.0001	<.0001	
	N	2704	2704	2729	2003	2735	2734	2735

Se observan mayores correlaciones entre los puntajes de los Cuestionarios Estandarizados, las que persisten después de controlar por el puntaje del Cuestionario de Deseabilidad Social¹⁷. La correlación entre el Cuestionario de Proactividad y el Cuestionario de Estrategias Metacognitivas es de 0.36

¹⁷ Las correlaciones parciales controlando por los puntajes en el Cuestionario de Deseabilidad Social pueden revisarse en el Apéndice 3.

($p < 0.001$), al igual que la correlación entre el Cuestionario de Proactividad y el Cuestionario de Habilidades Sociales ($r = 0.36$, $p < 0.001$). Por otra parte, la relación entre el Cuestionario de Proactividad y el Cuestionario de Habilidades Sociales es de 0.42 ($p < 0.001$). Estos resultados parecen indicar cierto grado de redundancia entre los tres cuestionarios estandarizados¹⁸.

4.2.2. Relación con los indicadores actuales de admisión

La correlación entre los nuevos instrumentos y los indicadores actuales de admisión es baja, aunque la mayoría es estadísticamente significativa. La actual prueba estandarizada de matemática se correlaciona débilmente con las sub-dimensiones del Ensayo de Pensamiento Crítico ($r = 0.21$, $p < 0.001$ y $r = 0.21$, $p < 0.001$), así como también lo hacen las notas de enseñanza secundaria ($r = 0.21$, $p < 0.001$ y $r = 0.22$, $p < 0.001$). Las notas de enseñanza secundaria también se correlacionan débilmente con el puntaje en el Cuestionario de Estrategias Metacognitivas ($r = 0.26$, $p < 0.001$). Correlaciones algo mayores se observan entre los nuevos instrumentos y la prueba estandarizada de Lenguaje y Comunicación, especialmente con las dos sub-dimensiones del Ensayo de Pensamiento Crítico ($r = 0.28$, $p < 0.001$ y $r = 0.28$, $p < 0.001$) y el Formulario Ampliado de Postulación ($r = 0.20$, $p < 0.001$), pero estas correlaciones siguen siendo pequeñas en magnitud.

TABLA 7. CORRELACIÓN CON LOS INSTRUMENTOS ACTUALES DE ADMISIÓN

Instrumentos	Estadísticos	PSU Matemática	PSU Verbal	Notas de enseñanza secundaria
Pensamiento Crítico (Contenido)	r	0.21	0.28	0.21
	Valor- P	<.0001	<.0001	<.0001
	N	2452	2452	2451
Pensamiento Crítico (Aspectos Formales)	r	0.21	0.28	0.22
	Valor- P	<.0001	<.0001	<.0001
	N	2452	2452	2451
Formulario Ampliado de Postulación	r	0.16	0.20	0.18
	Valor- P	<.0001	<.0001	<.0001
	N	2721	2722	2721
Reflexión Personal	r	0.08	0.12	0.03
	Valor- P	0.0004	<.0001	0.2154
	N	1969	1969	1968
C.Habilidades Sociales	R	-0.05	0.00	0.00
	Valor- P	0.0072	0.8136	0.9513
	N	2480	2481	2480
C. Estrategias Metacognitivas	R	0.09	0.17	0.26
	Valor- P	<.0001	<.0001	<.0001
	N	2479	2480	2479
C. Proactividad	R	0.00	0.05	0.08
	Valor- P	0.9491	0.0182	<.0001
	N	2473	2474	2473

Cabe destacar que existen algunas diferencias en estos resultados si es que distinguimos entre la muestra de escolares y la de universitarios. Para el caso de la muestra de escolares, las correlaciones presentes entre los instrumentos de admisión tradicionales y los nuevos instrumentos alcanzan valores mayores a los presentados para la muestra de universitarios. La razón que podría explicar la menor asociación entre

¹⁸ Se exploró la hipótesis de que esta redundancia podría deberse al hecho de que los tres instrumentos tuvieran un formato de respuesta múltiple, pero no se encontró evidencia para sustentarla (ver Apéndice 4).

ambos tipos de instrumentos en la muestra de universitarios es la menor varianza existente en los puntajes de estos alumnos en las pruebas de selección actual, debido a que al momento de participar en el estudio, ya habían sido parte de un proceso de selección basado en esas pruebas.

Considerando esto, es posible hipotetizar que en una eventual aplicación real de los instrumentos, la asociación con las pruebas de selección tradicional muestre un comportamiento más similar al observado en la muestra de escolares, al tratarse de individuos que no han pasado por un proceso de selección universitaria. Tomando en cuenta estos datos, de todas formas es posible concluir que los nuevos instrumentos propuestos evalúan mayoritariamente atributos distintos a los incluidos en los procesos de admisión a las universidades chilenas.

4.2.3. Relación con el Cuestionario de Deseabilidad Social

Los resultados arrojan correlaciones moderadas y estadísticamente significativas entre el Cuestionario de Deseabilidad Social y el puntaje en el Cuestionario de Habilidades Sociales ($p=0.39$, $p<0.001$), el puntaje en el Cuestionario de Estrategias Metacognitivas ($p=0.33$, $p<0.001$) y el puntaje en el Cuestionario de Proactividad ($p=0.25$, $p<0.001$). Las correlaciones del Cuestionario de Habilidades Sociales con el resto de los instrumentos incluidos en el estudio son menores que 0.1 y la mayoría de ellas no son significativas en un nivel de confianza del 1%.

TABLA 8. CORRELACIÓN CON EL CUESTIONARIO DE DESEABILIDAD SOCIAL

	<i>Estadísticos</i>	Puntaje Cuestionario de Deseabilidad Social
Pensamiento Crítico (Contenido)	r	0.01
	Valor- P	0.6928
	N	2705
Pensamiento Crítico (Aspectos Formales)	r	0.00
	Valor- P	0.857
	N	2705
Formulario Ampliado de Postulación	r	0.06
	Valor- P	0.0012
	N	2730
Reflexión Personal	r	0.05
	Valor- P	0.0245
	N	2005
C.Habilidades Sociales	r	0.39
	Valor- P	<.0001
	N	2736
C. Estrategias Metacognitivas	r	0.33
	Valor- P	<.0001
	N	2735
C. Proactividad	r	0.25
	Valor- P	<.0001
	N	2733

4.2.4. Relación con el Cuestionario del Profesor

La correlación del promedio por clase (z de Fisher) fue calculado para analizar la relación entre las evaluaciones de los profesores y los puntajes de los estudiantes en los cuatro tipos de instrumentos incluidos en el estudio, así como entre las evaluaciones de los profesores y los puntajes en las pruebas estandarizadas y notas de enseñanza secundaria. Los profesores que declararon haber conocido a los estudiantes hace menos de 6 meses o "muy poco" fueron excluidos del análisis.

En general, los profesores evaluaron a sus estudiantes muy positivamente (menos de 10% de los estudiantes fueron evaluados “bajo el promedio” en cada ítem) y por lo tanto los puntajes discriminaron poco entre ellos. Los resultados muestran que las evaluaciones de los profesores están muy relacionadas con las notas de enseñanza secundaria y los puntajes en las pruebas estandarizadas, lo cual puede esperarse de un profesor jefe de último año de enseñanza secundaria, el cual tiene pocas interrelaciones con los estudiantes y que conoce sobre todo el desempeño académico de éstos. En términos de los nuevos instrumentos, se observan correlaciones cercanas o sobre el 0.3 entre la calificación del profesor y los puntajes en el Cuestionario de Estrategias Metacognitivas y en el Formulario Ampliado de Postulación.

El Cuestionario del Profesor muestra dos factores (ver Apéndice 5). El primero, que hemos llamado “académico”, debido a su mayor correlación con puntajes estandarizados y notas de enseñanza secundaria, incluye los siguientes ítems: motivación por aprender, motivación académica, superación de los obstáculos, visión de futuro perseverancia, responsabilidad personal, autoconfianza, reflexión personal, argumentación y habilidad escrita.

El segundo factor fue llamado “interés social y motivación” e incluyó los siguientes ítems del cuestionario: interés por los otros, responsabilidad social, liderazgo, participación en actividades extracurriculares. Mientras el factor “académico” mostró correlaciones de cerca de un 0.3 con el puntaje del Cuestionario de Estrategias Metacognitivas, el ítem acerca de participación en actividades extracurriculares se correlaciona moderadamente con el puntaje en el Formulario Ampliado de Postulación (0.27).

La correlación con los otros instrumentos del Estudio Piloto de Admisión no superó el 0.25 (las subdimensiones de aspectos formales y contenido del Ensayo de Pensamiento Crítico, el Cuestionario de Proactividad y el de Habilidades Sociales) y por lo tanto no han sido incluidas en la tabla. Se hipotetiza que el profesor jefe no conoce en profundidad las dimensiones evaluadas por el cuestionario en cada uno de sus alumnos y el puntaje asignado está influido seriamente por el desempeño académico del alumno.

TABLA 9. CORRELACIÓN CON EL CUESTIONARIO DEL PROFESOR

Subdimensiones del Cuestionario del Profesor	C. Estrategias Metacognitivas	Formulario Ampliado de Postulación	PSU de Matemática	PSU de Verbal	Notas de Enseñanza secundaria
Motivación por aprender	0.32	0.18	0.49	0.46	0.67
Motivación académica	0.31	0.17	0.49	0.44	0.68
Superación de obstáculos	0.30	0.17	0.44	0.40	0.62
Visión de futuro	0.26	0.17	0.43	0.40	0.60
Interés en los otros	0.16	0.14	0.17	0.19	0.27
Responsabilidad social	0.16	0.18	0.15	0.20	0.25
Liderazgo	0.16	0.18	0.15	0.20	0.25
Perseverancia	0.29	0.17	0.42	0.39	0.59
Responsabilidad personal	0.28	0.17	0.44	0.39	0.60
Autoconfianza	0.24	0.21	0.38	0.35	0.54
Reflexión personal	0.25	0.19	0.33	0.39	0.50
Argumentación	0.24	0.19	0.32	0.41	0.50
Habilidad de escritura	0.26	0.18	0.36	0.45	0.52
Participación en Actividades Extracurriculares	0.17	0.27	0.21	0.21	0.32
Motivación General	0.29	0.19	0.43	0.43	0.61

4.2.5. Diferencias entre grupos

Se estudió también la relación entre el desempeño en los nuevos instrumentos y características sociodemográficas de los examinados. En concreto, las diferencias estandarizadas de rendimiento entre hombres y mujeres, así como entre estudiantes de escuelas privadas y municipales, fueron analizadas y comparadas con las diferencias estandarizadas de desempeño observadas en la prueba de admisión actual y en las calificaciones de enseñanza secundaria¹⁹. La intención es observar las diferencias en el desempeño de estudiantes de diferentes tipos de escuelas como forma de examinar diferencias entre estudiantes de diferentes grupos socioeconómicos.

La tabla 10 muestra que los estudiantes de escuelas particulares pagadas tienen un mejor desempeño que los de escuelas municipales en las dos sub-dimensiones del Ensayo de Pensamiento Crítico (especialmente en la sub-dimensión de contenido), así como también en el Formulario Ampliado de Postulación. En ambos casos, sin embargo, las diferencias son de entre un 30% y un 50% de las diferencias observadas en las pruebas cognitivas actuales y en las notas de enseñanza secundaria. La diferencia en la sub-dimensión de aspectos formales del Ensayo de Pensamiento Crítico no es estadísticamente significativa. Es interesante notar que las diferencias a favor de los estudiantes de escuelas municipales tanto en la Reflexión Personal y en el Cuestionario de Estrategias Metacognitivas son estadísticamente significativas.

En la tabla 11 se puede observar que todos los instrumentos piloteados en este estudio favorecen a las mujeres, con las mayores (y estadísticamente significativas) diferencias observadas en los cuestionarios estandarizados. El desempeño de las mujeres superan al de los hombres en el Cuestionario de Habilidades Sociales y en el de Estrategias Metacognitivas en entre un 25% y 40% desviación estándar.

TABLA 10. DIFERENCIA EN LOS PUNTAJES ESTANDARIZADOS ENTRE ESCUELAS MUNICIPALES Y PARTICULARES PAGADAS EN PRUEBAS COGNITIVAS Y NO COGNITIVAS

Variable	N Part. Pagado	N Municipal	Media Part. Pagado	Media Municipal	Diferencia de medias (P. Pagado-Municipal)	Valor t	Prob t
C. Habilidades sociales	830	972	0.00	0.02	-0.02	-0.39	0.69
C. Estrategias Metacognitivas	830	971	-0.09	0.07	-0.16	-3.45	0.00
C. Proactividad	829	969	-0.06	-0.02	-0.05	-0.94	0.34
Ensayo de Pensamiento Crítico (contenido)	799	909	0.20	-0.06	0.26	5.35	0.00
Ensayo de Pensamiento crítico (asp. formales)	799	909	0.09	0.07	0.02	0.51	0.61
Formulario Ampliado de Postulación	957	1016	0.15	-0.18	0.33	7.78	0.00
Reflexión Personal	694	710	-0.04	0.19	-0.23	-4.17	0.00
PSU Matemática	955	987	0.56	-0.25	0.82	20.10	0.00
PSU Verbal	955	988	0.43	-0.16	0.59	13.86	0.00
Notas de Ens. Secundaria	955	987	0.28	-0.30	0.57	13.52	0.00

¹⁹ Estos análisis fueron realizados estandarizando el desempeño de los estudiantes en cada muestra por separado (universitarios y muestra de estudiantes) y los resultados obtenidos eran similares a los presentados arriba.

TABLA 11. DIFERENCIA EN LOS PUNTAJES ESTANDARIZADOS ENTRE HOMBRES Y MUJERES EN PRUEBAS COGNITIVAS Y NO COGNITIVAS

Variable	N Hombres	N Mujeres	Media Hombres	Media Mujeres	Diferencia de medias (hombres-mujeres)	Valor t	Prob t
C. Habilidades sociales	1421	1286	-0.13	0.14	-0.26	-6.93	0.00
C. Estrategias Metacognitivas	1418	1286	-0.19	0.21	-0.41	-10.85	0.00
C. Proactividad	1416	1283	-0.07	0.07	-0.14	-3.63	0.00
Ensayo de Pensamiento Crítico (contenido)	1315	1241	-0.02	0.02	-0.05	-1.19	0.23
Ensayo de Pensamiento crítico (asp. formales)	1315	1241	-0.02	0.02	-0.03	-0.86	0.38
Formulario Ampliado de Postulación	1542	1418	-0.03	0.03	-0.05	-1.44	0.15
Reflexión Personal	1000	1097	-0.04	0.04	-0.08	-1.78	0.07
PSU Matemática	1491	1371	0.20	-0.21	0.41	11.32	0.00
PSU Verbal	1491	1372	0.07	-0.07	0.14	3.71	0.00
Notas de Ens. Secundaria	1490	1371	-0.08	0.08	-0.16	-4.32	0.00

5. DISCUSIÓN

Los resultados revisados sugieren que los cuatro tipos de instrumentos piloteados en el estudio están midiendo diferentes constructos y que esos constructos son diferentes a los evaluados por los criterios actuales de admisión. Sin embargo, es posible observar cierta superposición entre los tres cuestionarios estandarizados, lo cual pareciera no estar relacionado con la deseabilidad social o con el hecho de que los tres cuestionarios presentan un formato de respuesta múltiple. Estos resultados, además del hecho de que solamente el Cuestionario de Estrategias Metacognitivas está correlacionado con el Cuestionario del Profesor, sugieren que sólo el último podría ser recomendado para un uso operativo.

Los resultados también sugieren que existe una moderada relación entre la Prueba de Lenguaje y Comunicación y el Ensayo de Pensamiento Crítico, la Reflexión Personal y el Formulario Ampliado de Postulación. Ya que dos de estos nuevos instrumentos son ensayos y el otro incluye preguntas abiertas, este resultado no es tan sorprendente. La habilidad verbal, la lectura y escritura son aptitudes importantes que se correlacionarán en general con la mayoría de las autoevaluaciones. En este caso, la correlación es sólo moderada y la inclusión de este nuevo instrumento se justifica ya que la prueba actual de Lenguaje y Comunicación utilizada para admisión es una prueba de respuesta múltiple (Sincoff & Sternberg, 1987).

Lamentablemente, la evaluación de los profesores jefes de último año de educación secundaria no es tan informativa como se esperaba. Sin embargo, el hecho de que haya una correlación moderada entre la evaluación del profesor de la participación en actividades extracurriculares y los puntajes de los estudiante en el Formulario Ampliado de Postulación da apoyo a la validez del último instrumento. Además, la correlaciones observadas están en línea con los resultados reportados por Oswald et al (2004) entre la evaluación de los pares y mediciones de información biográfica.

Entre los nuevos instrumentos, el Ensayo de Pensamiento Crítico muestra la mayor diferencia de desempeño entre estudiantes de bajo y alto nivel socioeconómico (NSE); estas diferencias, sin embargo,

son de magnitud menor (entre un tercio y la mitad) a las diferencias observadas en los instrumentos actuales de admisión). Por otro lado, la Reflexión Personal es la que muestra los resultados más positivos en términos de favorecer a los estudiantes de bajo nivel socioeconómico. Aunque estos resultados son alentadores, es importante recordar que la participación en el estudio no tuvo consecuencias en los estudiantes. Una vez que se asocien consecuencias a las nuevas pruebas, es probable que aumente la preparación para rendir estas pruebas, incrementándose por tanto la diferencia entre grupos socioeconómicos. Los resultados de experiencias en otros países, sin embargo, sugieren que las diferencias entre grupos en el desempeño en instrumentos no cognitivos, como la Reflexión Personal y el Formulario Ampliado de Postulación, no llegarían a ser tan grandes como las observadas en evaluaciones cognitivas como las usadas hoy día en Chile para la admisión a la educación superior.

La mayoría de los instrumentos muestra coeficientes de confiabilidad de nivel aceptable. El Ensayo de Pensamiento Crítico es la excepción más importante, mostrando el Alpha de Cronbach más bajo (0.57). Es importante destacar que los resultados de confiabilidad para este instrumento no aumentan si son calculados separadamente para las dos sub-dimensiones (0,58 para la sub-dimensión de contenido y 0.51 para la sub-dimensión de aspectos formales). Una hipótesis para este fenómeno es que cada subsección del Ensayo de Pensamiento Crítico tiene pocos ítems (la dimensión de contenido tiene 6 y la de aspectos formales, 4) y cada una de estas mide aspectos muy distintos del constructo. Además, la literatura en general muestra bajos coeficientes de confiabilidad para evaluaciones escritas basadas en sólo un ensayo. El reducido número de tareas muestreado por la evaluación es señalado como el origen de este fenómeno (Breeland, Bridgeman, & Fowles, 1999).

Durante el año 2009 se revisaron todos los instrumentos y pautas de corrección en base a los resultados de la aplicación realizada en el año 2008. El bajo desempeño de los estudiantes en el instrumento de Reflexión Personal sugiere que deberían revisarse especialmente el lenguaje y el orden en el cual se dan las instrucciones para este instrumento. Junto con esto, la menor tasa de acuerdo entre correctores para este instrumento estaría indicando que los evaluadores necesitan pautas más claras de corrección, instrucciones más precisas y un proceso de supervisión más estricta. En tanto, el nivel de acuerdo entre jueces observado en los otros dos instrumentos (Ensayo de Pensamiento Crítico y Formulario Ampliado de Postulación), son similares a los observados en otros estudios.

6. CONCLUSIÓN

Los resultados preliminares del primer año de investigación son promisorios. Los nuevos instrumentos desarrollados para este estudio son capaces de medir atributos que podrían complementar los indicadores que actualmente son usados para los procesos de admisión de pregrado en Chile. Aunque la validez y confiabilidad varía dependiendo del instrumento, en general se muestran coeficientes aceptables de consistencia interna y de acuerdo entre jueces. Junto con esto, el análisis factorial confirma la estructura teórica del Ensayo de Pensamiento Crítico y de la Reflexión Personal. Además de lo anterior, los nuevos instrumentos muestran menores diferencias en el desempeño de las mujeres y de los estudiantes provenientes de entornos socioeconómicos más bajos, en comparación a los observados en los instrumentos de admisión actuales.

Es importante tener en cuenta, sin embargo, el nivel de generabilidad de estos resultados, considerando que no hubo consecuencias asociadas a la participación de los estudiantes o a sus resultados en los

instrumentos. Una vez que existan consecuencias asociadas, estos instrumentos serán susceptibles de ser preparados y/o falseados. Como una forma de disuadir a los estudiantes del falseamiento, se ha considerado un proceso aleatorio de verificación, tal como se describía en la sección que explica el Formulario Ampliado de Postulación. A los estudiantes que han sido admitidos en la universidad y que no han sido capaces de demostrar su participación en una actividad reportada o un premio recibido podría cancelárseles su matrícula. Los efectos de la preparación en los puntajes del Ensayo de Pensamiento Crítico y en la Reflexión Personal serán cuidadosamente monitoreados. Se espera, sin embargo, que existan algunos beneficios educacionales asociados a la preparación de las pruebas si esto se traduce en mejores habilidades de escritura en los estudiantes que postulan a la educación superior.

En el futuro, los resultados presentados en este documento serán complementados con información relativa a la capacidad de los nuevos instrumentos de predecir el éxito académico y no académico de estos alumnos. La información relativa al éxito no académico fue recogida a través de una encuesta telefónica durante el primer semestre de 2009 (encuesta de seguimiento) y refiere principalmente a la satisfacción con la actividad que están realizando, la transición hacia la vida universitaria, la fidelidad a disciplina de estudio liderazgo, actividades extraprogramáticas y expectativas.

Junto con esto, la información sobre las motivaciones de los estudiantes, su autoeficacia, su perseverancia y sus habilidades de liderazgo, entre otras, fueron obtenidos de una tercera persona también durante el primer semestre de 2009, a través de otra encuesta telefónica para la muestra de estudiantes universitarios. Una vez que la información recolectada esté disponible, será usada para triangular el autoreporte de los estudiantes de la misma forma que el Cuestionario del Profesor fue usado en la muestra de escolares. Esta información, así como la encuesta de seguimiento a los estudiantes descrita más arriba y los registros administrativos, incluyendo las calificaciones de los estudiantes y su persistencia en la universidad, estarán disponible durante el primer semestre del año 2010.

Durante el año 2009 además se hizo un nuevo pilotaje de los mismos instrumentos piloteados durante el año 2008, incorporándose algunas pequeñas modificaciones. El Piloto 2009 se centró en colegios públicos de la Región Metropolitana y en colegios de diversas dependencias en varias regiones del país. Los resultados obtenidos a través de estas mediciones complementarán la evidencia aquí presentada.

REFERENCIAS BIBLIOGRÁFICAS

- Breeland, H., Bridgeman, B., & Fowles, M. (1999). *Writing Assessment in Admission to Higher Education: Review and Framework* (No. 99-3). New York: College Board.
- Breeland, H., Maxey, J., Gernand, R., Cumming, T., & Trapani, C. (2001). *Trends in College Admission 2000. A Report of a Survey of Undergraduate Admissions Policies, Practices, and Procedures*. ACT, Inc., Association for Institutional Research, The College Board, Educational Testing Service, The National Association for College Admission Counseling.
- Camara, W. J. (2005). Broadening Criteria of College Success and the Impact of Cognitive Predictors. In W. J. Camara & E. W. Kimmel (Eds.), *Choosing Students: Higher Education Admissions Tools for the 21st Century*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Camara, W. J. (2007). *Thinking Outside the Box: New Admissions Tools*. *Non-Cognitive Forum*. Paper presented at the Presentation at the National Council of Measurement in Education, Chicago.

- Cliffordson, C., & Askling, B. (2006). Different Grounds for Admission: Its Effects on Recruitment and Achievement in Medical Education. *Scandinavian Journal of Educational Research*, 50(1), 45-62.
- College Board. (2002). *Best Practices in Admissions Decisions. A Report on the Third College Board Conference on Admission Models.*, New York.
- Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic.
- Geiser, S., & Santelices, M. V. (2007). Validity Of High-School Grades In Predicting Student Success Beyond The Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes. *CSHE Research and Occasional Paper Series*(CSHE.9.07).
- Goleman. (1996). *La Inteligencia Emocional: ¿por qué es más importante que el CI?* Buenos Aires: Javier Vergara Eds.
- Heckman, J., & Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *The American Economic Review*, 91(2), 145-149.
- Kyllonen, P. C. (2005). *The Case for Noncognitive Assessments.* Princeton: ETS.
- Le, H., Casillas, A., Robbins, S. B., & Langley, R. (2005). Motivational and Skills, Social, and Self-management Predictors of College Outcomes: Constructing the Student Readiness Inventory. *Educational and Psychological Measurement*, 65(3).
- Liu, L., Minsky, J., Ling, G., & Kyllonen, P. (2009). Using the Standardized Letters of Recommendation in Admissions: Results from a Multidimensional Rasch Model. *Educational and Psychological Measurement*, 69(3), 475-492.
- Rigol, G. W. (2003). *Admissions-Making Model. How US Institutions of Higher Education Select Undergraduate Students.* New York: College Board.
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. *Psychological Bulletin*, 130(2), 261-288.
- Rosas, R., Boetto, C., & Jordan, V. (2004). *Introducción a la Psicología de la Inteligencia.* Santiago: Ediciones Universidad Católica de Chile.
- Sternberg, R. J. (1999). A Triarchic Approach to the Understanding and Assessment of Intelligence in Multicultural Populations. *Journal OF School Psychology*, 37(2), 145-159.
- Sternberg, R. J. (2003). Our Research Program Validating the Triarchic Theory of Successful Intelligence: Reply to Gotfredson. *Intelligence*, 31, 399-413.
- Willingham. (1985). *Success in College. The Role of Personal Qualities and Academic Ability.* New York: College Board.
- Zwick, R., & Grief Greene, J. (2007). New Perspectives on the Correlation of SAT Scores, High School Grades and Socioeconomic Factors. *Journal of Educational Measurement*, 44(3), 23-45.

APÉNDICES

A1. Análisis factorial del Ensayo de Pensamiento Crítico

El análisis factorial arroja una solución de dos factores (autovalores mayores que uno), que explica el 24% de la varianza y coincide con la teoría del constructo. El primer factor incluye ítems que miden "contenido" y el segundo factor "aspectos formales" de la escritura. El análisis fue realizado usando el método de estimación de Mínimos Cuadrados Ordinarios, ya que este se comporta mejor cuando las variables no tienen una distribución normal (variables ordinales). El ítem de uso de párrafos tuvo un importante peso en ambos factores mientras que el ítem de Autorreflexión no presentó peso en ninguno de los dos factores.

TABLA 1. PESOS FACTORIALES

	1	2
Análisis de argumento	0.519	0.069
Tesis	0.561	0.154
Argumentos	0.613	0.147
Contra Argumentos	0.409	0.041
Conclusión	0.580	0.093
Autorreflexión	0.037	0.057
Ortografía	0.086	0.583
Vocabulario	0.035	0.511
Cohesión	0.119	0.466
Uso de párrafos	0.230	0.289

Método de extracción: Mínimos Cuadrados Generalizados.

Método de rotación: Normalización Promax con Kaiser.

A2. Análisis factorial del instrumento de Reflexión Personal. Componentes Principales del Método con Rotación Varimax

Basado en el gráfico de sedimentación, los autovalores sobre el valor uno y sobre la proporción de la varianza explicada por cada factor se decidió mantener un factor. Los resultados han sido preliminarmente confirmados por la implementación del análisis paralelo con 50, 100 y 200 iteraciones.

TABLA 2. AUTOVALORES DE LA MATRIZ DE CORRELACIÓN. TOTAL=5 PROMEDIO=1

	Autovalor	Diferencia	Proporción	Acumulativo
1	2.31000010	1.36110619	0.4620	0.4620
2	0.94889391	0.22768097	0.1898	0.6518
3	0.72121295	0.15920875	0.1442	0.7960
4	0.56200420	0.10411536	0.1124	0.9084
5	0.45788884		0.0916	1.0000

TABLA 3. MATRIZ FACTORIAL ROTADA

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
<i>Significado de la dificultad</i>	0.16126	0.21654	0.07567	0.24466	0.92818
<i>Perseverancia</i>	0.05578	0.05525	0.99437	0.03278	0.06310
<i>Autoeficacia</i>	0.96339	0.11766	0.06095	0.18081	0.14707
<i>Locus de control</i>	0.19887	0.18879	0.03774	0.92951	0.24372
<i>Experiencia</i>	0.12052	0.95386	0.06170	0.17679	0.20141

Sólo fueron considerados los ensayos de Reflexión Personal en los cuales los estudiantes dieron evidencia apoyando su motivación para el desarrollo personal, social o académico

TABLA 4. VARIANZA EXPLICADA POR CADA FACTOR

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1.0113090	1.0092738	1.0034528	0.9888625	0.9871019

A3.

TABLA 5. COEFICIENTES DE CORRELACIÓN PARCIAL DE PEARSON ENTRE LOS CUESTIONARIOS ESTANDARIZADOS, CONTROLANDO POR EL CUESTIONARIO DE DESEABILIDAD SOCIAL (N = 2696, PROB > |R| BAJO H0: RHO PARCIAL=0)

	Habilidades Sociales	Estrategias Metacognitivas	Proactividad
Habilidades Sociales	1	0.34 <.0001	0.30 <.0001
Estrategias Metacognitivas	0.34 <.0001	1	0.30 <.0001
Proactividad	0.30 <.0001	0.30 <.0001	1

A4. Análisis factorial de todos los ítem de los cuestionarios estandarizados combinados

En total se encontraron 46 ítems (20 del Cuestionario de Habilidades Sociales, 18 del Cuestionario de Estrategias Metacognitivas y 8 del Cuestionario de Proactividad). Un análisis de Componente Principal con Rotación Varimax arrojó 11 factores usando la regla de decisión de autovalores sobre 1. Esos 11 factores explicaron un 49% de la varianza total. El hecho de que los ítems de los diferentes cuestionarios se agruparan en diferentes factores (ver tabla 6) demostró que el formato de los ítems (de respuesta múltiple) no fue el factor principal detrás de la varianza de puntaje observada.

FIGURA 1. GRÁFICO DE SEDIMENTACIÓN

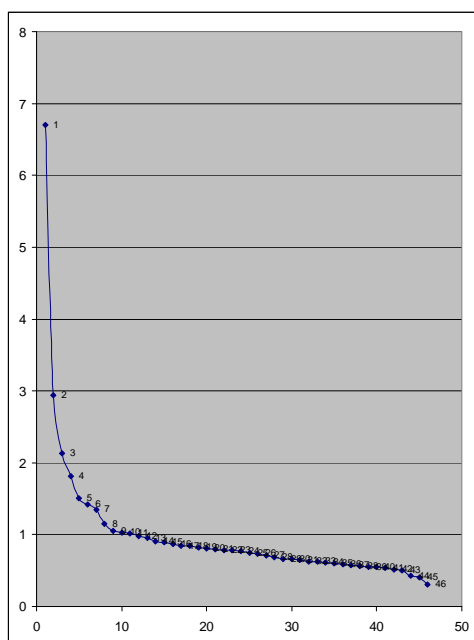


TABLA 6. ÍTEMS CON PESOS SOBRE 0.35 POR FACTOR

Factor	Ítems
Factor 1	H7, H8, H12, H13, H15, H19
Factor 2	M1, M2, M5, M8, M12, P1
Factor 3	M3, M13, M15, M16, M17
Factor 4	H2, H3, H4, H7, H15, H16
Factor 5	P1, P2, P3, P4, P5, P6, P7, P8
Factor 6	H5, H8, H9, H11, H20
Factor 7	M2, M4, M14, M18
Factor 8	M6, M7, M9, M13
Factor 9	H10, H17, H18
Factor 10	H6, H14
Factor 11	H1, M10, M11

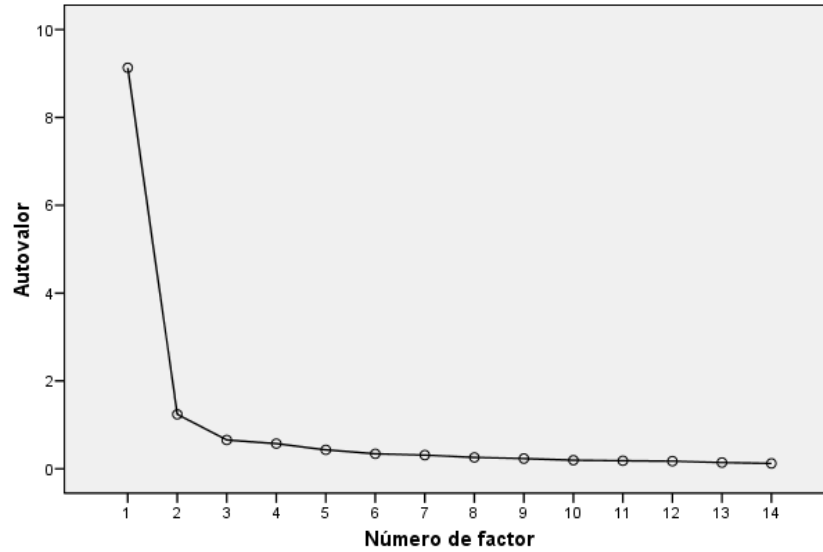
A5. Análisis factorial del Cuestionario del Profesor

El Cuestionario del Profesor fue administrado en 82 cursos, correspondientes a 21 colegios, con un total de 1,517 estudiantes evaluados (97% de la muestra de escolares). La información fue excluida del análisis en 52 casos, debido a que los profesores declararon no conocer a sus alumnos lo suficientemente bien (por menos de un semestre o "muy poco"). La muestra final tuvo 1,475 casos y los puntajes se concentraron en el "promedio" y "sobre el promedio" para la mayoría de los ítems (cerca del 75% de los casos).

El análisis factorial, desarrollado usando el método de Mínimos Cuadrados Ordinarios, arrojó dos factores (basado en autovalores sobre 1 y a partir del gráfico de sedimentación de la figura 2 con una alta correlación entre ellos (0.695). Esos dos factores fueron llamados (1) interés/desempeño académico e (2) interés y motivación social, y el peso de los ítem en cada factor es presentado en la tabla 7.

Estos dos factores explican un 70% de la varianza total.

FIGURA 2. GRÁFICO DE SEDIMENTACIÓN



Habilidad de escritura	0.609	0.242
Participación extracurricular	0.248	0.457

UN SISTEMA UNIVERSITARIO SIN LIMITACIONES DE ACCESO: EL CASO DE URUGUAY

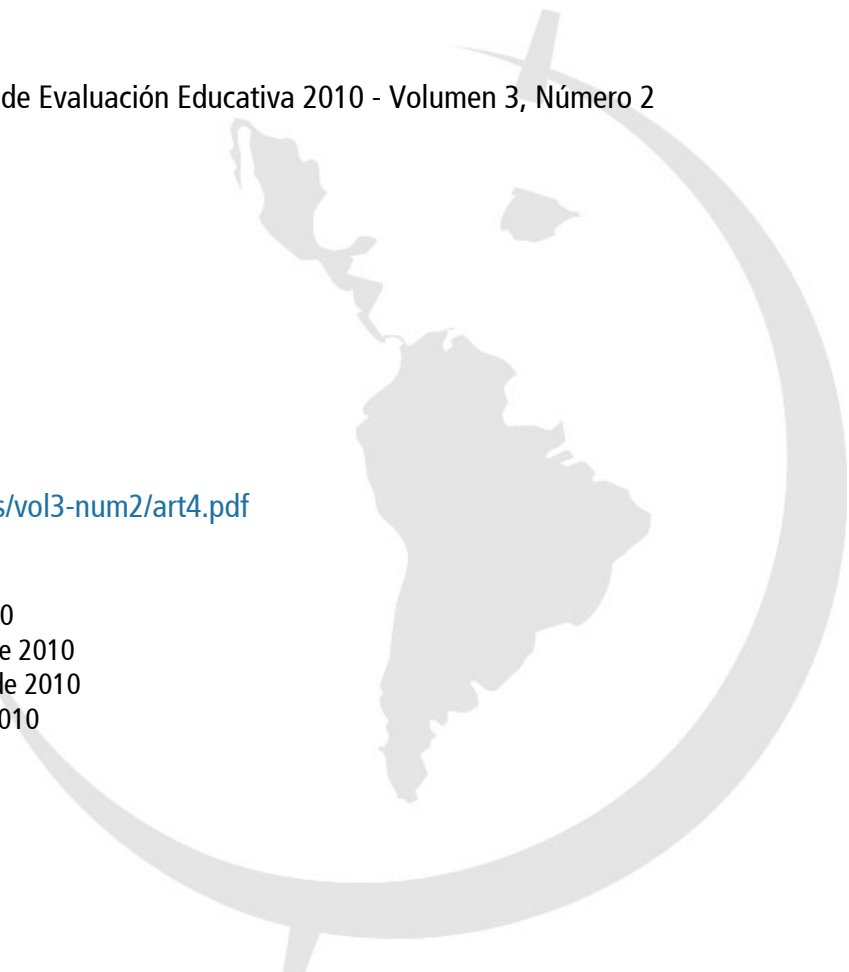
A UNIVERSITY SYSTEM WITHOUT RESTRICTIONS OF ACCESS: THE CASE OF
URUGUAY

Carlos Romero

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art4.pdf>

Fecha de recepción: 02 de abril de 2010
Fecha de dictaminación: 19 de mayo de 2010
Fecha de segundo envío: 22 de mayo de 2010
Fecha de aceptación: 23 de mayo de 2010



El sistema educativo uruguayo reúne varias particularidades que lo diferencian de la mayoría de los países de la región. Estas particularidades constituyen un desafío a la hora de pretender hacer un aporte que permita comparar lo que aquí sucede con las situaciones en otros contextos. El tema que centra el análisis de la presente edición de la Revista Iberoamericana de Evaluación Educativa es uno de ellos.

La situación uruguaya podría resumirse en pocas líneas, al señalar que en esencia, el sistema universitario¹ tiene libre admisión a ese nivel de estudios una vez que se hayan aprobado con éxito los niveles anteriores de la educación. Por lo señalado, el aporte de este artículo quedaría limitado a esa indicación, lo que al tiempo de ajustarse a la realidad no deja traslucir la historia que está detrás de esa práctica, las condicionantes que hacen posible la misma más allá de los problemas que representa, y las miradas encontradas de algunos de sus actores, entre aquellos que propugnan por no innovar y aquellos que piden que se limite o regule el acceso.

Este artículo intentará contribuir con una perspectiva conjunta sobre estos intereses contrapuestos y por dónde se encamina el tema en los años venideros.

1. UBICACIÓN EN EL CONTEXTO

Se podría comenzar con una afirmación que hemos sostenido en más de una ocasión. No existe en nuestro país un verdadero "sistema educativo", o al menos es cuestionable su existencia si por él consideramos aquel en el cual "las partes tienen sentido y responden al funcionamiento articulado del todo" (Romero y Landoni, 2001). Tomando en cuenta nuestra realidad es más razonable pensar en la existencia de subsistemas que funcionan con un cierto grado de articulación.

Sería una tarea inútil pretender argumentar la importancia de disponer de un verdadero sistema educativo, dado que su existencia se da como un dato de la realidad y como uno de los principios fundantes de las estructuras educativas de los países, incluido el nuestro. Pero una cosa es la afirmación de su existencia y otra el que la misma pueda verificarse en los hechos.

En el caso de nuestro país se ha impuesto, a lo largo de decenios, una visión que propone el funcionamiento autónomo e independiente de dos niveles educativos (el universitario por un lado y todo el resto de la educación, por otro), sin que exista un organismo conductor o articulador ejecutivo entre esos niveles. Es más, si se abundara en lo que se refiere a las instituciones educativas uruguayas, encontraríamos otro sector que tiene un conjunto de centros donde se brinda formación de diverso tipo, muchas de ellas del Estado (formación militar, policial, diplomática, etc) que no están regidos por ninguna organización que los aglutine.

Probablemente una de las peculiaridades que más llama la atención del 'sistema'² educativo uruguayo sea el escaso peso del Ministerio de Educación y Cultura dentro del mismo, reflejado en las pocas

¹ Entre los aspectos complejos del sistema uruguayo se encuentra las diferentes acepciones y alcances que se le da a "educación superior" y cuáles son las instituciones que deben integrarla. Para no introducir una polémica adicional y compleja de abordar, se optó a los efectos de este artículo asumir una de las concepciones mayoritarias, la que asocia el concepto de "educación superior" con el de "educación universitaria".

² Emplearemos la palabra sistema más como un sustantivo que de la idea de lo que queremos expresar, más que de una entidad real porque, como se indicó, sostenemos que no existe en cuanto tal.

responsabilidades asignadas en los temas educativos en comparación con las que tiene en otras áreas (registros, fiscalías, cultura, etc).

El Ministerio es el responsable de la política educativa en su sentido más amplio, pero la gestión y administración de enseñanza primaria, secundaria, técnica y formación docente dependen de un Ente Autónomo, la Administración Nacional de Educación Pública (ANEP) en tanto que la educación superior pública están en la órbita de otro Ente Autónomo, la Universidad de la República (UDELAR).” (Romero, 2005)

Esta distribución de responsabilidades entre variados órganos públicos ha sido uno de los principales escollos para articular los distintos niveles de la educación en una propuesta orgánica. Sustentada sobre la base de “autonomías” plenas, la falta de armonización resultaba un verdadero desafío que debía ser superado y estuvo en la base de los cambios impulsados en los últimos años por las autoridades nacionales de la educación.

Uno de los ejemplos más destacados de estos cambios es la Ley General de Educación, N° 18.437 de diciembre de 2008, por la cual se incorporan cambios sustanciales a la situación anterior que, según las autoridades impulsoras permitieron crear un ‘real sistema educativo’, que integre desde la educación formal hasta la no formal, pasando por todos los niveles educativos.

Así es que el art. 20 de esta Ley se establece la creación del “sistema nacional de educación como un conjunto de propuestas educativas integradas y articuladas para todos los habitantes y para toda la vida” (Ley 18.437)

Pero esa afirmación no se sustenta en los hechos y resulta en un buen propósito sin concretar. Es difícil encontrar a lo largo de la norma los ejemplos concretos que permitan afirmar que el sistema realmente existe y que la compartimentación anterior ha sido superada. Es verdad que se crean órganos de coordinación como la Comisión Nacional de Educación (COMINE) -art. 43-, el Congreso Nacional de Educación, -art. 45-, con funciones de asesoramiento y consulta, pero en todos los casos carecen potestades ejecutivas.

Es también cierto que el Ministerio de Educación y Cultura adquiere un protagonismo que no tenía hasta ese entonces, pero sus cometidos tampoco le confieren la posibilidad de conducir la educación. Sus funciones son las de “desarrollar los principios generales de la educación”, “facilitar la coordinación de políticas educativas”, “articular las políticas educativas” con las áreas de innovación, ciencia, desarrollo, entre otras, “promover la articulación”, “presidir los ámbitos de coordinación”, y otras actividades de similar naturaleza, que están indicadas en el art. 51 de la norma.

Pero lo que esta ley no modifica es la característica fundamental de lo que ha sido tradicional en nuestro país: el carácter autónomo de los dos Entes que rigen la educación. Estos son la Administración Nacional de Educación Pública (que atiende a la enseñanza Inicial y Primaria, Media Básica y Media Superior y Formación Técnico Profesional) y la Universidad de la República, que además no es incluida en esta norma sino que mantiene su Ley Orgánica de 1958.

La Ley General de Educación innova con la creación de dos nuevas instituciones: el Instituto Universitario de Educación (que incluirá por primera vez a la formación de docentes en el nivel universitario) y los institutos de Educación Terciaria.

Tanto por la confirmación de las autonomías de los entes, como por las características de las nuevas instituciones creadas, el actual sistema no hace más que confirmar la gestión en compartimentos

estancos y las estructuras en paralelo, lo que potencialmente lejos de superar la dispersión actual, tienda a aumentarla.

Esta dispersión no se supera siquiera con las disposiciones de la Ley que destacan la voluntad de unificar criterios y procesos, como por ejemplo cuando se indica que los Consejos Directivos Autónomos deberán “coordinar sus acciones con el fin de cumplir con los principios, orientaciones y los fines de la educación...” (art. 48). Sin una autoridad superior que regule las actividades de las instancias autónomas, la coordinación queda librada a la voluntad de acuerdo de los actores, tal como ha sido hasta ahora, y por tanto, con sobrados ejemplos de lo imposible que resulta hacerlo.

De todas maneras no es posible realizar afirmaciones categóricas sobre los efectos de la Ley en vigencia, pues aún se encuentra en la fase de puesta en funcionamiento de las nuevas estructuras creadas, y se continúa con la inercia anterior en las operaciones cotidianas. Lo que sí ha demostrado este año y medio de actividades orientadas a poner en acción las creaciones de la Ley, es que el ritmo en que se logra avanzar dista mucho del que requiere una conducción ejecutiva.

Los acuerdos son difíciles de alcanzar entre organismos autónomos y cogobernados, y esto se potencia porque usualmente los representantes de las instituciones autónomas que forman parte de los grupos de trabajo creados para hacer operativa la Ley, no pueden asumir una opinión definitiva y deben estar refrendando los acuerdos con las instancias superiores de gobierno de sus respectivas instituciones.

En concreto eso ha demostrado que los ‘acuerdos definitivos’ conseguidos en los grupos de trabajo, parecen no ser tales, y siempre puede existir la posibilidad de tener que renovar las discusiones.

2. LA ARTICULACIÓN EN UN SISTEMA DESARTICULADO

Lo señalado hasta aquí intentó contextualizar a nivel macro cómo se ha diseñado la estructura organizacional de la educación uruguaya. El acento puesto en lo que se espera comience a operar de aquí en más sirve para resaltar que la compartimentación ha sido la característica distintiva de la educación nacional.

En lo que se refiere al nivel de la educación superior, se puede afirmar que hasta 1985 ésta “era sinónimo de la Universidad de la República” (Martínez, 2003). Sin otra institución pública o privada con la cual debiera coordinar sus acciones o recibir de ella orientación, supervisión o control, se fue moldeando a sí misma y por consiguiente a lo que era aplicable en la educación superior.

En efecto, tal como lo afirma Bentancur “su porte relativo, tradición y naturaleza jurídica, en un contexto de incipiente pero reducido desarrollo del tramo privado de la educación superior y alienidad gubernamental, constituyen a “la” Universidad en un centro de gravedad tan pronunciado que prácticamente la convierten en un sistema en sí misma. De esta forma, sus políticas y problemas son en buena medida, los de la universidad toda”. (Bentancur, 2004)

Esa ‘universidad toda’ que señala el autor tiene su traducción en un sistema que se encuentra compuesto de una única universidad pública (la de la República -UDELAR- creada en 1833), y 4 de gestión privada: la Universidad Católica del Uruguay (1985), la Universidad ORT (1996), Universidad de Montevideo (1997) y Universidad de la Empresa (1998).

Existe también un conjunto de instituciones dentro del nivel universitario, compuesto por los “Institutos Universitarios” de gestión privada. La diferencia con las universidades es que se concentran en 1 o 2 áreas de conocimiento, en tanto que las universidades deben hacerlo en 3 o más.

Los institutos universitarios son mayoría en cuanto número de instituciones (12 en este momento), pero con una cantidad de alumnos muy exiguo que ronda en el 1% del sistema, con 1.177 matriculados en el año 2008. (MEC, 2009).

El sistema en su conjunto cuenta con 127.371 estudiantes que se matricularon en carreras de grado en el año 2008, de los cuales 112.707 corresponden a UDELAR y 13.487 universidades privadas (MEC, ob cit), lo que confirma que la importancia de la universidad pública en la captación de los estudiantes. Con las características de gratuidad total y de acceso libre, sumado a su trayectoria centenaria, es entendible su enorme peso relativo dentro del subsistema universitario.

Sumado al hecho que durante 152 años fue la única institución universitaria, se explican las razones por las cuales la universidad pública fue construyendo por sí misma el conjunto de sus requisitos y formas de acción, a prescindencia de instancias externas a ella.

El peso y la historia de la universidad pública se extienden aún en las privadas. Las normas legales y reglamentarias que se imponen a éstas, dentro de las cuales se inscriben los procedimientos que el Ministerio de Educación asume con estas instituciones, refleja consistentemente lo delineado originalmente para aquella. El conjunto de factores inciden también en el comportamiento que asumen las universidades privadas, quienes han demostrado con mayor o menor énfasis y con más o menos acciones planificadas en ese sentido, un comportamiento de “isomorfismo” con relación a la universidad pública. (Landoni, 2007)

En ese escenario los requisitos para acceder a la educación universitaria son también similares, aunque no hayan sido objeto de una política educativa expresa que buscara la homogeneización en el sistema.

2.1. Requisitos de ingreso a la educación superior

En este entorno, ¿cuáles son las exigencias que debe satisfacer un estudiante para ingresar a la educación superior? Para poder matricularse en ella alcanza con la aprobación completa del segundo ciclo de la educación secundaria, en cualquiera de las dos modalidades actuales, denominados como “bachillerato diversificado” y “bachillerato tecnológico”.

La comprobación de esta exigencia se materializa en la emisión de un certificado que es expedido por la institución en la que se cursaron los estudios. Esa certificación de egreso de la secundaria es condición suficiente para aspirar a la educación superior, sin más trámite ni instancias de otra organización ante las cuales se deban requerir supervisiones.³

No existe una prueba de egreso de la enseñanza secundaria, sino que el egreso se comprueba con la aprobación de todas y cada una de las materias o asignaturas que componen los diferentes planes de

³ Las instituciones de educación secundaria pueden ser de gestión pública, como de gestión privada, que en este caso tienen que ser “habilitadas” por la autoridad pública correspondiente, la Administración Nacional de Educación Pública.

estudio de las orientaciones de bachillerato, y por lo tanto las universidades públicas o privadas deben exigir el documento del bachillerato completo para poder inscribir a los estudiantes.

Una reciente Ordenanza de UDELAR flexibiliza mínimamente la exigencia de tener totalmente aprobado el Bachillerato previamente a la inscripción en la Universidad. Así es que se dispuso que se pueda "admitir el ingreso condicional a la UDELAR con una materia previa, la que deberá ser aprobada en el período especial a abril para que el ingreso quede firme". (UDELAR, 2009). Es de hacer notar que el año académico comienza en el mes de marzo, por lo cual la flexibilidad es realmente poca ya que se le da al estudiante un solo período adicional para que salve su materia previa.

En el caso de las instituciones privadas no hay una indicación legal o normativa expresa sobre la cantidad de materias previas que pueden tomar como condicionales a la hora de inscribir a los estudiantes para el primer año de las carreras. Pese a ello existe una recomendación del Ministerio de Educación que permite el ingreso hasta con tres materias previas, que deberán ser aprobadas por el estudiante antes de poder rendir ningún examen universitario.

En la práctica, la falta de controles efectivos al respecto hace imposible saber si el procedimiento sugerido se aplica a cabalidad, quedando librado a la responsabilidad de las autoridades institucionales hacer efectiva esta recomendación.

3. EL ESTRECHO VINCULO ENTRE EL BACHILLERATO Y LA UNIVERSIDAD

La historia de la vinculación entre el bachillerato y la universidad es entendible al analizar el proceso histórico en que ambas se fueron gestando. Como se indicó antes, la Universidad de la República tiene sus inicios en el año 1833 y en 1880 se instala en su seno la "sección secundaria" como un servicio que le permitiera brindar a los estudiantes una necesaria "educación preparatoria" claramente propedéutica y orientada al ingreso a las facultades.

Así lo señala la propia UDELAR en su historia. "En sus comienzos la enseñanza secundaria fue esencialmente "preparatoria" para los estudios universitarios: no se concebía cursar la misma si no se pensaba continuar los estudios. A principios de este siglo (s. XX) se planteó un enfoque algo distinto, vinculado a la creación de liceos en el Interior del país, cuya función sería dar una formación integral, que se extendería por cuatro años y una posterior, especializada, que duraría dos años. Los liceos del Interior se establecieron entre 1912 y 1913, a lo que debe sumarse la creación -en Montevideo- de un liceo "femenino" en 1912 y otro nocturno para trabajadores en 1919, además de otros liceos diurnos en la capital. Este crecimiento institucional llevó la matrícula de unos 500 alumnos a fines del siglo pasado a unos 6300 en 1923 y a los 11360 en 1931". (web RAU, 2010)

En diciembre de 1935, en medio de una crisis política, el gobierno separa la enseñanza secundaria de la órbita universitaria dándole carácter autónomo. (ANEP, 2008). Aún con la separación funcional ambas instituciones siguieron transitando un camino que suponía su vinculación, en especial cuando el conjunto de la oferta de carreras universitarias estaba concentrado en las 'carreras tradicionales' (medicina, derecho, ingeniería, arquitectura, agronomía, veterinaria, química, economía). Este número de carreras limitado permitían mantener un razonable control sobre la cantidad de cursos a vincular (unos como previos de otros), y lo más importante, con pocos centros de educación secundaria y una sola universidad. La escala del sistema permitía una armonización razonable sin demasiada articulación.

Lo cierto es que con los años las carreras se fueron multiplicando y los centros educativos también, pero tampoco se crearon las instancias de debate sobre la forma de optimizar los resultados vinculando formalmente ambos niveles. Asimismo estuvo ausente la reflexión sobre adoptar pruebas de acceso o admisión a la Universidad. Si bien no hubo debates sobre la conveniencia de su aplicación, si existió un período en donde la universidad pública las aplicó: durante la dictadura militar (1973 -1985).

Este requisito se efectivizó en un muy corto período dentro de los años indicados, sólo en dos años para el conjunto de la institución, aunque previamente se había aplicado parcialmente en algunas de sus facultades. La historia de la Universidad lo describe así, *"... en 1975 comenzaron a aplicarse mecanismos restrictivos de la admisión de estudiantes. Establecido el examen de ingreso primero parcialmente (Medicina y Odontología) se quiso justificarlo a partir de una insuficiencia locativa, que las autoridades interventoras agravaron al no construir prácticamente ningún edificio durante su gestión. En 1980 se generalizó el régimen a toda la Universidad; no obstante, un sistema de cupos prefijados que se impuso en 1982 constituyó el puente hacia la desaparición de todas las trabas en el último año de intervención. Nada se dijo si el sistema había fracasado, simplemente se modificó."* (web RAU.)

Es posible sospechar que existe una suerte de vinculación inconciente entre el examen de ingreso y el momento en que se dio (dictadura, con una universidad que había perdido su autonomía), que ponga fuera de cualquier discusión actual esta medida como una de las posibles a asumir, pero esta afirmación no resulta sencilla de probar.

La superación de este breve capítulo de la historia es considerada como la reconquista de un principio fundamental en esa institución, que pervive hasta el presente.

La restricción en el ingreso de estudiantes no es un tema que integre la agenda universitaria en el país. sin embargo el número de estudiantes que recibe año a año preocupa cada vez más a distintos niveles por las dificultades innegables que presenta para la gestión y las repercusiones sobre la calidad de enseñanza en una Universidad que, en complejidad y exigencias no es la misma de décadas atrás.

"El ingreso a la UdelaR ha sido tradicionalmente irrestricto para quienes completan la educación secundaria, con la excepción de un breve período durante la intervención en el que se aplicó un examen de ingreso. Esta es una política claramente orientada por el principio de equidad, pues otorga a todos la misma oportunidad de cursar estudios superiores. Sin embargo, en las condiciones actuales en que se imparten los cursos en la mayor parte de la Facultades, especialmente en los primeros años de las carreras, es evidente que esa amplitud impacta en la calidad de la enseñanza impartida." (Bentancur, ob. cit.). Hay algunas excepciones que se aplican en cursos o carreras de UDELAR, como son los exámenes previstos para el ingreso a Traductorado, a la Escuela Universitaria de Música, la Escuela de Tecnología Médica y la Escuela de Nutrición y Dietética.

Las instituciones privadas comparten la misma situación de ausencia de exámenes de admisión, con pocas salvedades, pero por razones bien diferentes. Al contrario de la Universidad estatal, ninguna de las instituciones privadas recibe apoyo económico del Estado para su funcionamiento, por lo cual se basa exclusivamente en el ingreso que generan las cuotas que pagan sus estudiantes. En otras palabras, lejos de preocuparse por limitar el número de estudiantes que deseen inscribirse en sus carreras, el interés de estas instituciones está en captar el mayor número posible de alumnos o el número que al menos satisfaga el punto de equilibrio que permita su funcionamiento.

La excepción recién indicada está en la Universidad de Montevideo y para una de sus Facultades, la de Humanidades. Allí se propone una prueba de nivel para sus carreras (Licenciatura en Humanidades, los profesorado, y las carreras de Traductorado y Magisterio Bilingüe, Licenciatura en Humanidades, Profesorado de Historia, Profesorado de Filosofía, Profesorado de Literatura, Profesorado de Inglés y Profesorado de Matemática).

4. LOS REQUISITOS DE INGRESO Y SU LECTURA POLÍTICA

Las excepciones indicadas en donde se aplican exámenes de ingreso no son sino las excepciones que confirman la regla de un sistema sin limitaciones.

Resulta claro que la ausencia de estos procedimientos en las instituciones privadas refleja un valor culturalmente aceptado, y en este caso además, una medida innecesaria por el volumen de los estudiantes atendidos y por las razones de sostenimiento económico. Por lo tanto el eje de la reflexión debe tornarse hacia las causas de fondo que llevan a UDELAR a sostener el acceso irrestricto, dado que allí es donde el volumen de estudiantes representa un problema y la no adopción de medidas parece un contrasentido a la calidad.

Es posible hacer más de una lectura sobre el hecho del acceso irrestricto, tomando en cuenta la incidencia de diferentes factores y también perspectivas diversas. En ese sentido consideremos la reflexión sobre este tema que realiza uno de los actores con activa participación dentro de la universidad pública y que conoce de primera mano las condiciones en que esa institución funciona.

"El panorama es conocido: dictado de clases en grandes salones o incluso en locales no diseñados específicamente para la enseñanza (cines, anfiteatros, etc.), que aún así suelen no ser lo suficientemente amplios para permitir el acceso a todos los estudiantes matriculados, bajísima relación cuantitativa docente-estudiante y, como consecuencia, metodología de exposiciones docentes de tipo magistral con casi nulas posibilidades de intercambio con los alumnos. De este modo la selectividad se produce de manera indirecta, por medio del abandono estudiantil en esas instancias liminares, dispositivo de ajuste claramente indeseable. El ingreso irrestricto y la consiguiente masificación son habitualmente ligados, además, con otros indicadores de eficiencia universitaria, que dan cuenta de una baja performance de la institución: altas tasas de deserción, tiempo de cursado de las carreras muy superiores a los previstos y desempleo de parte de los egresados universitarios, sobre todo en algunas profesiones." (Bentancur, ob cit)

Por otro lado consideramos algunas expresiones del Rector de una Universidad privada, la ORT, que reflexiona como un actor involucrado en la marcha del sistema universitario en general, y en particular sobre el efecto que las políticas de la UDELAR –y por ende, de la universidad pública sostenida por toda la sociedad- sobre los temas de equidad. "La conclusión que yo saco de esos datos es que el gasto público en educación superior en el Uruguay es altamente inequitativo, y que los tres instrumentos que estamos usando: la financiación directa de la Universidad de la República, la "gratuidad" y el acceso irrestricto son ineficientes en el sentido que no alcanzan el objetivo de asegurar equidad en el acceso a la educación superior. La Universidad de la República es financiada por todos pero la usan mayoritariamente los más ricos." (Grunberg, 2001)

De estos dos breves ejemplos surgen algunas de las constantes que son frecuentes de encontrar en las discusiones sobre el rol de la universidad pública: por un lado una posición crítica sobre la real

contribución de ésta a garantizar una educación para todos, sin exclusiones que vayan más allá de las capacidades personales (pues los que realmente logran ingresar y mantenerse, son los que provienen mayoritariamente de las clases económicamente pudientes), y por otro lado una suerte de aceptación irremediable de que aunque no se esté contribuyendo a la equidad con las condiciones actuales, es poco lo que se puede hacer para cambiar.

En otras palabras, el problema está bien diagnosticado, lo que no es sinónimo de que eso se traduzca en acciones para superar los problemas detectados.

La opción por limitar o no el acceso a la universidad excede en mucho los aspectos meramente pedagógicos y docentes que garanticen una correcta formación de profesionales y ciudadanos comprometidos. También surgen desde el entorno reclamos para que se integren otras exigencias en relación a lo que la sociedad impone, propone o aspira de las instituciones de educación superior, y de la universidad pública en particular. La rendición de cuentas está en la base de estas aspiraciones.

La situación se complejiza cuando se trata de una universidad autónoma que fija sus propias condiciones de funcionamiento, aparentemente lejos de una 'lógica de mercado'. Entre las principales críticas que recibe la universidad pública está el tiempo que se demora en generar profesionales, al considerar la eficiencia terminal.

Los datos numéricos de los estudiantes que ingresan cada año y los que egresan dentro del tiempo previsto para concluir sus respectivas carreras, están muy lejos de ofrecer buenos resultados. (Anexo 1).

Según el Anuario Estadístico del Ministerio de Educación y Cultura 2008 –último publicado- en los nueve años de esta década UDELAR tuvo un ingreso promedio de 17.344 estudiantes por año, frente a un egreso de 4.228 anual, también en promedio. En grandes números, por cada estudiante que egresa por año, entran otros tres que se van rezagando en la institución.

Hay dos características adicionales de la Universidad de la República que dificultan disponer de cifras precisas por cohorte para analizar la eficiencia terminal: por una parte la 'calidad de estudiante' nunca se pierde, y por otra parte como consecuencia de lo anterior, la universidad no realiza el seguimiento por generaciones y tampoco lleva la cantidad anual de estudiantes matriculados. Éstos surgen del Censo General de Estudiantes que se realiza cada cierto período de tiempo, el último en el año 2007.

En relación a lo indicado en primer lugar, "UDELAR no cuenta con restricciones al nivel de actividad académica del estudiante (no hay plazos para aprobar los niveles, no hay límite de repeticiones, ni hay límites de exámenes). En el caso de UDELAR la situación reglamentaria que deriva de su Ley Orgánica de 1958 y de modificaciones a su marco electoral de 1972, hace que la 'titularidad' en el 'orden estudiantil' - y también en el de egresado - sólo se extingue con la muerte del titular. Es decir que quien accede a la condición de estudiante universitario *siempre tendrá derecho a estudiar, a continuar sus estudios, o mejor dicho, a retomarlos indefinidamente hasta recibirse*. En cierto sentido, este complemento a la gratuidad, es un acceso ilimitado al tiempo necesario para los que tienen limitantes de contexto social externo a la Universidad. (Boado, 2005)

En posible tener una aproximación de la eficiencia terminal del sistema, no sólo la UDELAR, utilizando la misma información estadística que muestra a las privadas con un mejor desempeño relativo.

Mientras UDELAR representa el 78% de los ingresos, tiene el 71% de los egresos, en tanto que las privadas tienen el 22% de los ingresos y el 28% de los egresos. (Anexo 2), aunque el Anuario destaca

que “la brecha entre ingresos y egresos se incrementa en el sector privado y disminuye en la UdelaR” (MEC, 2008).

Ante la búsqueda de explicaciones, y por qué no, de soluciones a la situación que enfrenta a la universidad pública con el ingreso masivo de estudiantes, es posible encontrar argumentos que resultan bastante coincidentes entre los actores de esa institución.

En cierta forma podría relativizarse el problema de la “masividad”, visto desde el sistema en su conjunto, dado que los datos sobre la tasa de cobertura están por debajo del 30% de la población que debería estar atendida en educación superior (27,7% según ADUR, 2008), por lo cual ese dato de por sí no expresa un problema.

El problema si surge cuando se cruza este valor con la cantidad de instituciones incluidas, que se reduce por lo que hemos analizado hasta ahora, a una sola institución con predominio sobre las demás. Por lo tanto, la masividad como problema no puede considerarse tomando en cuenta al sistema (desde el cual aún habría mucho trabajo para hacer para ampliar la inclusión a este nivel educativo), sino desde lo que sucede en su universidad estatal.

“Una solución aparentemente efectiva sería, pues, la de limitar el ingreso de estudiantes en procura de obtener una Universidad más pequeña, eficiente, y menos costosa, tal como lo proponen organismos como el Banco Mundial (1993). No consideramos que esta sea una medida acertada en nuestro país, por varias razones.” (Bentancur, ob.cit.)

También se relativiza la importancia de hacer la medición de la eficiencia universitaria en términos de número de graduados y años de cursado, desde la mirada de UDELAR, ya que se propugna a que ésta tenga un rol de generación de ciudadanía y valores que van más allá del tiempo que demora ‘formar’ un egresado.

“En nuestro país, (la eficiencia terminal es tener) una mirada estrecha. En contextos de alta desocupación juvenil, la Universidad no cumple solamente la función de producir egresados: también es un valor público a preservar su rol socializador y de continentación de jóvenes que, aunque no obtengan su diploma o luego de hacerlo no puedan desempeñarse en su profesión específica, emplean mientras tanto productivamente su tiempo y probablemente, luego verán incrementado su nivel de ingreso por la formación recibida. Teniendo en cuenta las razones sociales y académicas que desaconsejan la aplicación de mecanismos de limitación al ingreso, y la dificultad política para la imposición de barreras de esa naturaleza, se han ensayado y propuesto varias alternativas para mantener niveles de enseñanza decorosos en un contexto de masificación.” (Bentancur, ob cit)

Este convencimiento sigue presente hasta el día de hoy dado que UDELAR se encuentra analizando la renovación de su Ley Orgánica, proceso en el que ha estado ocupada en los últimos tres años y para lo cual se han producido varios documentos de trabajo. Lo que resulta claro es que hay principios que no forman parte de la discusión para el cambio, entre los que se encuentra en primera línea los referidos a la gratuidad y el ingreso libre.

Algunos de los documentos del rectorado que han sido puestos a consideración de la comunidad universitaria, como el denominado “Notas para la Actualización de la Ley Orgánica”, el de “Plan de Trabajo del Equipo Rectoral para Impulsar la Reforma Universitaria” y el “Plan de Obras de Mediano y Largo Plazo en Perspectiva Académica” han dejado explícitos que existen consensos básicos de política universitaria, como son “Fortalecer la autonomía y cogobierno de la institución, manteniendo la Libertad

de ingreso y cursado, sin examen de ingreso ni cobro de matrícula, (como ha sido tradicional en nuestra Universidad) y la total vigencia del artículo 66 de la actual Ley Orgánica.” (CGU 2008)

Lo cierto es que dentro del mismo proceso de renovación institucional, no todas las visiones sobre este consenso tienen la misma valoración. Contrariamente a ello, es posible encontrar expresiones que asumen posturas formales, articuladas y explícitas pero desde la ironía, con una lectura alternativa a esos mismos consensos.

“No todo es negativo al interior de la UDELAR. Hay algunos progresos en la transparencia sobre la actuación universitaria y sus resultados. Hoy se sabe que se reciben muy pocos. Se conoce que los procesos formativos implementados son extremadamente igualitarios (de liberalismo a ultranza) y que perjudican a aquellos que tienen condiciones más desfavorables (vulnerabilidades que no son atendidas).

Que la deserción es el fenómeno más notorio durante los primeros años de estudio y que en consecuencia no es necesario poner un examen de ingreso porque el propio modelo funciona a la perfección para restringir el número de estudiantes (aunque sea tremendamente costoso e injusto). Que el promedio de calificaciones o rendimientos porcentual para la UR está por debajo de un nivel medio y más concretamente es de por sí mediocre y que la falta de calidad en la enseñanza campea por toda la institución sin determinar medidas de fondo. Que pese a que se diga que la mayoría de la investigación se realiza en la UR es una mayoría dentro de una comunidad científica muy pequeña que apenas alcanza a 1 investigador cada 1000 integrantes de la población económicamente activa (para alcanzar los niveles regionales al menos se debería duplicar ese guarismo), además las patentes que se registran y la innovación tecnológica se concreta mayoritariamente por fuera de ella. Estos datos salen de los estudios que la propia UR realiza, lo que torna más incomprensible la ausencia de planes de impacto que detenga este retroceso públicamente testimoniado.” (Alternativa Universitaria, 2009)

5. CONCLUSIONES

Ante la ausencia de definiciones políticas que provengan desde un “sistema educativo” articulado que acoja a todas las instituciones involucradas con los mismos derechos, posibilidades y obligaciones, se continuará con desequilibrios imposibles de salvar en el corto plazo. Por lo tanto seguirá primando una suerte de ley de ‘mercado imperfecto’ que continuará orientando (o desorientando) en dirección a las distintas propuestas de las instituciones.

En el equilibrio entre el Estado como orientador y el mercado como espacio de oportunidades, se da la ‘paradoja’ que existe una institución (UDELAR) que ocupa un lugar en ambos lados. Por una parte porque comparte el rol de ‘Estado’ con el Ministerio de Educación y Cultura, pues tiene la posibilidad de definir políticas sobre el conjunto del subsistema, que la convierten en juez y parte. Pero adicionalmente UDELAR realiza su oferta en el mercado junto con las instituciones privadas –aunque con un claro privilegio a su favor, al contar con el 100% de sus recursos financieros asegurados por los fondos públicos- cuando a sale a competir por ‘captar’ estudiantes.

Aún en este escenario de mercado ‘imperfecto’ existe un amplio rango de acciones que no están sujetas a reglas predeterminadas o lo están para una institución pero se aplican por analogía en las demás, aunque no siempre destinándose el tiempo necesario para analizar su aplicabilidad o conveniencia de utilizar procedimientos de un sector en otro (público y privado). La realidad ha demostrado que muchas cosas funcionan libremente o aún que son ‘reguladas indirectamente’, por la fuerza de los hechos y las

costumbres. Entre estos últimos se inscriben los tópicos analizados en este artículo en relación al ingreso de los estudiantes.

El sistema educativo uruguayo se encuentra en un dilema en relación a su necesidad de incrementar el número de estudiantes que optan por realizar estudios de enseñanza superior, la posibilidad que ellos se desenvuelvan en el sector público (gratuito) y el peso que su universidad mayor tiene en el sistema. Este reto debería dar lugar a que en el país se propusiera abordarlo desde una perspectiva conjunta, sistemática, en donde se integraran todas las instituciones involucradas.

Pero la agenda pública no da señales en ese sentido, ni se verifica en los planteos de los principales partidos políticos o sus dirigentes, avanzar hacia un sistema de enseñanza superior que vincule e integre a las instituciones públicas junto con las privadas. Lejos de ello, la recientemente aprobada Ley General de Educación prevé la creación del sistema nacional de educación pública, lo que resulta en una clara señal de mantener el doble estándar.

Eso lleva a que se pueda afirmar que las instituciones privadas seguirán optando los mismos mecanismos que utiliza actualmente para la incorporación de sus estudiantes, que lejos de propugnar criterios selectivos buscan la incorporación de todo estudiante con las condiciones de hacerlo y las posibilidades económicas que conlleva. Lo anterior sin menoscabar que existan criterios de selección para algunas carreras que requieran determinados niveles de conocimiento previo, tal como existe actualmente de manera casi testimonial, pero no como un principio general para toda su oferta.

En la Universidad estatal el debate central NO pasa por estructurar mecanismos de acceso restringido, sino todo lo contrario, propugnar por un incremento aún mayor del alumnado.

La apuesta por la creación de nuevas instituciones terciarias, como las dos creadas por la actual Ley de Educación General, el Instituto Terciario Superior y el Instituto Universitario de Educación, pueden ser en parte una solución al dilema de 'cantidad vs. calidad'. De cualquier manera la Universidad de la República seguirá teniendo un número de estudiantes que, en opinión de todos los actores, aún los más firmes defensores del modelo UDELAR, entienden como sobredimensionado.

Un debate adicional sobre el modelo ha sido puesto "sobre la mesa de discusión" por el Presidente de la República José Mujica, al impulsar un debate sobre la 'descentralización de la universidad al interior del país', buscando alternativas a la actual situación de la educación superior que se ubica fundamentalmente en la capital del país. Pero aún si esta iniciativa tiene éxito y según sea la forma en que se opte por resolverla (no está definido ni planteado el modelo a impulsar), el tema del acceso ilimitado no está entre los que forma parte de la discusión.

Tampoco parece ser la solución a estos problemas las propuestas por ofrecer una orientación vocacional temprana y efectiva a los jóvenes, como manera de canalizar las inclinaciones profesionales antes de que se inscriban en la Universidad. Sin otras instituciones de alternativa con el mismo prestigio y amplitud de oferta que las de la universidad estatal, de poco parece servir que se canalicen los intereses hacia espacios que no existen. En suma todo parece indicar que hay pocas probabilidades que se produzcan cambios radicales en los procedimientos de admisión actuales, y el "status quo" prevalecerá en estos temas por un largo tiempo.

REFERENCIAS BIBLIOGRÁFICAS

- ADUR, Asociación de Docentes de la Universidad de la República. (2008). *La transformación educativa en perspectiva comparada: tendencias y cambios en la escala internacional y principalmente latinoamericana*. Montevideo: UDELAR. Disponible en <http://www.adur.org.uy/trauniv/documento1.pdf>.
- Alternativa Universitaria (2009). *Mitos, Desencuentros y Compromisos de la Reforma Universitaria: por una Universidad Abierta a la gente*. Disponible en http://www.universidadur.edu.uy/claustro/docs/Doc_agosto2009_alternativa_universitariaVF1.pdf
- ANEP (2008). *Historia de Educación Secundaria 1935-2008*. Montevideo: Consejo de Educación Secundaria.
- Bentancur, N. (2004). Cinco dilemas universitarios de comienzos de siglo. *Revista Uruguaya de Ciencia Política*, 14. Páginas 85- 101 Montevideo: ICP.
- Boado, M. (2005). La deserción universitaria en Udelar, algunas tendencias y reflexiones. M. Boado, DS/FCS (2005); Disponible en <http://www4.iesalc.unesco.org.ve/programas/Deserci%C3%B3n/Informe%20Deserci%C3%B3n%20RESUMEN%20-%20Uruguay.pdf>.
- CGU, CORRIENTE GREMIAL UNIVERSITARIA, 2008: Disponible en http://www.cgu.edu.uy/reforma/AporteDeCGU_NuevaLeyOrganica_20080724.pdf.
- Grunberg, J. (2001). El sistema universitario uruguayo. Algunas consideraciones sobre calidad y equidad, Conferencia del Rector de la Universidad ORT en el British Schools de Montevideo. Junio de 2001.
- Landoni, P. (2007). Isomorfismo y calidad: redefiniendo los espacios públicos y privados en la educación superior uruguaya. *Revista Uruguaya de Ciencia Política* – 17, número 1 -. Páginas 183- 203 ICP – Montevideo.
- Ley 18.736 de 2008, General de Educación.
- Ley 2.549 de 1958, Orgánica de la Universidad de la República.
- Martínez Larrechea, E. (2003). *La educación superior de Uruguay, inercias y horizontes de cambio*. Montevideo: Fundación de Cultura Universitaria.
- RAU (web de UDELAR) <http://www.rau.edu.uy/uruguay/cultura/histoweb1.htm#6>
- Romero, C. y Landoni, P. (2001). Educación superior uruguaya: del monopolio a la consolidación del sistema. En *Economía Política de las Reformas Educativas en América Latina*. Páginas 259 - 286. Santiago de Chile: Preal.
- Romero, C. (2005). *Diagnóstico de los títulos de la educación superior en Latinoamérica y El Caribe; informe de Uruguay, Proyecto "Observatorio de la Educación Superior en América Latina y el Caribe*. Caracas: Instituto Internacional para la Educación Superior en América Latina y el Caribe (IESALC).
- UDELAR (2009). Resolución del Consejo Directivo Central de la Universidad sobre el ingreso con previas a la Universidad de la República, Sesión Ordinaria CDC - 22 de diciembre de 2009.

ANEXOS

Anexo 1.

INGRESOS Y EGRESOS DE LA UDELAR

(2000-2008)

	2000	2001	2002	2003	2004	2005	2006	2007	2008
Ingresos	14895	15648	16360	17346	17744	18497	18668	18569	18374
Egresos	3053	4182	3714	4247	3999	4235	4776	4808	5038

Fuente: Ministerio de Educación y Cultura (2009), Anuario Estadístico de la Educación.

Anexo 2

INGRESOS, MATRÍCULA Y EGRESOS SEGÚN FORMA DE ADMINISTRACIÓN EN CANTIDADES ABSOLUTAS Y PORCENTAJES AÑO 2008

FORMA DE ADMINISTRACIÓN	INGRESOS	MATRÍCULA	EGRESOS
VALORES ABSOLUTOS TOTAL	23.566	100.452	7.018
Universidad de la República	18.374	81.123	5.038
Instituciones Universitarias Privadas	5.192	19.329	1.980
PORCENTAJES TOTAL	100,0%	100,0%	100,0%
Universidad de la República	78,0%	80,8%	71,8%
Instituciones Universitarias Privadas	22,0%	19,2%	28,2%

Fuente: Ministerio de Educación y Cultura (2009), Anuario Estadístico de la Educación.

SELECCIÓN Y PERMANENCIA EN LA EDUCACIÓN SUPERIOR: EL CASO DE LA UNIVERSIDAD AUTÓNOMA DE YUCATÁN

SELECTION AND PERMANENCE AT HIGHER EDUCATION: THE CASE OF THE AUTONOMOUS UNIVERSITY OF YUCATÁN

*Ileana R. López López, Carlos M. Echazarreta González,
Silvia J. Pech Campos, y Brenda A. Gómez Ortegón*

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art5.pdf>

Fecha de recepción: 15 de marzo de 2010
Fecha de dictaminación: 5 de mayo de 2010
Fecha de segundo envío: 11 de mayo de 2010
Fecha de aceptación: 12 de mayo de 2010

La Universidad pública, en nuestro país, es la opción a la que más recurren los estudiantes que concluyen el bachillerato. Los motivos pueden atribuirse al nivel de la calidad en la formación que se imparte en ella, que refleja un cambio importante en la última década, a las condiciones de acceso, permanencia y los diversos servicios que ofrece a todo tipo de población, y a situaciones de índole socioeconómica. En consecuencia, este tipo de instituciones se ven sometidas a las fuertes presiones causadas por la demanda, que se incrementa de manera importante conforme avanzan los años. No obstante ésta se aminora con un relativo incremento de las plazas o lugares ofertados, este incremento aún no resulta suficiente para atender a toda la población demandante. Tal condición confirma el hecho de que “el acceso a la universidad pública sea selectivo” (García, 2005, p. 19), en el sentido de que las universidades se ven obligadas a implementar diversos mecanismos para la selección de los estudiantes de nuevo ingreso.

Desde el momento del ingreso, los sustentantes seleccionados, comienzan a generar un historial académico que conforma la trayectoria escolar, término que se refiere a la “cuantificación del comportamiento escolar de un conjunto de estudiantes (cohorte) durante su estancia educativa” (Altamira, 1997)

Los estudios acerca de la trayectoria escolar permiten a las instituciones educativas identificar las necesidades de diferentes grupos de estudiantes, lo cual posibilita planear e implementar diversas estrategias específicas para atender dichos grupos y mejorar el desempeño de los mismos. De esta manera, el alcance de estos estudios se logra debido a que “el análisis de las TE implica la observación continua de los movimientos de una población estudiantil a lo largo de los ciclos escolares especificados en una cohorte” (Barranco & Santacruz, 1995, citado en González, 2000, p.18). En otro sentido, las trayectorias escolares retroalimentan las decisiones referentes a la selección de estudiantes de nuevo ingreso, y los criterios e instrumentos utilizados para tal fin. La esencia u objetivo de la selección apunta a la elección de los sujetos con mayores probabilidades de delinear una trayectoria escolar con éxito, evitando la deserción y el rezago.

En cuanto a este sentido, los nuevos esquemas que buscan regular el acceso a la Universidad, la tarea de seleccionar apunta a “elegir” a los aspirantes con un mayor potencial para estudiar una carrera profesional, y cuyo dominio de habilidades y conocimientos establezcan mayores probabilidades de éxito académico.

En el caso de la Universidad Autónoma de Yucatán (UADY) la demanda total de sustentantes, en el período comprendido de 2001 a 2007 se ha incrementado a razón de un 10% anual; sin embargo, como se mencionó, el cupo disponible no ha crecido en la misma proporción. En caso de mantenerse este incremento, la demanda para ingresar a la UADY en el año 2010 será alrededor de 12,000 aspirantes y el cupo de 3500 (UADY, 2007); lo anterior, obliga a hacer más rigurosos, confiables, capaces de predecir el desempeño académico de los estudiantes y hacer más transparentes los mecanismos de selección.

Entre los instrumentos y criterios que se utilizan en los últimos tiempos para seleccionar a los estudiantes de nuevo ingreso en México, se encuentra el EXANI II del CENEVAL, el Examen de Habilidades y Conocimientos Básicos (EXHCOBA) de la Universidad Autónoma de Baja California y la PAA del College Board, por mencionar algunos. Por otro lado, existen algunas universidades que utilizan instrumentos diseñados *ex profeso*.

El EXANI II es un instrumento que mide habilidades y conocimientos indispensables para realizar estudios de nivel superior en forma exitosa, es de alto impacto, porque sus consecuencias afectan

directamente la vida académica de los aspirantes; tiene un diseño estandarizado, porque cuenta con reglas fijas de diseño, elaboración, aplicación y calificación, lo que permite ubicar las puntuaciones de individuos y grupos con respecto a las normas de calificación establecidas nacional y regionalmente, al igual que es objetivo porque emplea procesos de calificación confiables (CENEVAL, 2009).

En cuanto a la organización de la información que generan los estudiantes desde el momento en que ingresan a la universidad en indicadores que describan el tránsito de los mismos por el programa educativo, se toma, para el presente estudio, la propuesta de Chain y Jácome (2007) en el que conforma el desempeño escolar con tres indicadores: el Índice de Aprobación en Ordinario (IAO), el Índice de Promoción (IP) y el Promedio de calificaciones. El IAO mide "el nivel de aprobación de las asignaturas o créditos en la primera opción"; el IP se refiere a "la proporción de asignaturas o créditos que el estudiante logra promover, independientemente de si lo hace en ordinario o en cualquiera de las demás oportunidades que las reglas de escolaridad de la institución permiten"; y el promedio global de las calificaciones refleja el "aprovechamiento que la institución considera que ha obtenido un estudiante" (Chain & Jácome, 2007, p.188).

1. PROPÓSITO Y OBJETIVOS DEL ESTUDIO

Determinar la relación entre el desempeño en el examen de selección (EXANI II) y el Riesgo y Desempeño Escolar de los estudiantes de una cohorte, durante los primeros cuatro semestres en la Universidad.

1.1. Objetivos específicos

1. Describir los resultados del desempeño en el examen de selección,
2. Determinar el desempeño escolar (DE) con base en el Índice de Aprobación en Ordinario (IAO), Índice de promoción (IP) y Promedio de calificaciones (P),
3. Determinar la situación escolar (SE) con base en tres niveles (óptimo, irregular y rezagado)
4. Describir la trayectoria escolar y los niveles de riesgo en los estudiantes,
5. Determinar la relación entre el desempeño en el examen de selección y los niveles de riesgo y desempeño escolar.

2. METODOLOGÍA

El estudio se realizó con una muestra de 440 estudiantes que ingresaron en septiembre del 2005 y avanzaron en los cuatro primeros semestres de los programas educativos de los campus de Ciencias biológicas y agropecuarias (CBA) Ciencias exactas e ingenierías (CEI), Ciencias de la salud (CSalud) y Ciencias sociales (CSociales), en la Tabla 1 se muestra la distribución de la muestra por área de conocimiento, facultad y programa educativo.

La información de las puntuaciones obtenidas por los estudiantes en el EXANI II, se organizó en ocho áreas del conocimiento (razonamiento verbal, razonamiento matemático, español, matemáticas, ciencias naturales, ciencias sociales, mundo contemporáneo e inglés; y en una escala de desempeño con cuatro niveles (muy bajo, bajo, alto y muy alto)

TABLA 1. DISTRIBUCIÓN DE LA MUESTRA

Área de conocimiento	Facultades	Programas Educativos	f	%
CSociales	Economía	Economía y Comercio Internacional	91	20.7
CSalud	Enfermería	Enfermería	35	7.95
CBA	Medicina, Veterinaria y Zootecnia	Biología y Medicina Veterinaria y Zootecnia	119	27
CEI	Matemáticas	Actuaría, Ciencias de la Computación, Ingeniería de Software, Ingeniería en Computación, Enseñanza de las Matemáticas, Matemáticas	195	44.3

La información de las calificaciones, asignaturas, tipo de examen y número de créditos de los cuatro semestres estudiados fue organizada en una base de datos por áreas de conocimiento

Las variables de estudio fueron dos: la primera, es el desempeño escolar (DE) compuesta por tres indicadores, Índice de aprobación en ordinario (IAO), Índice de promoción (IP) y promedio de calificaciones (PC). La combinación de estos tres indicadores nos dan tres niveles de DE: alto, regular y bajo (Tabla 2).

La segunda variable es la situación escolar (SE) expresada en porcentaje de asignaturas o créditos cubiertos con relación al mínimo esperado al momento del análisis y se describe en tres niveles óptimo (100%), irregular (90-99%) y rezago (menor a 90%).

Los indicadores estudiados fueron los propuestos por Chain y Jácome 2007 y el programa para su análisis fue Trayectorias Escolares Ver 3.2.17.

La Trayectoria escolar descrita por las combinaciones posibles entre los tres niveles de DE y SE se expresan en nueve tipos de trayectoria y estos a su vez se traducen en tres niveles de riesgo (Alto, medio y bajo) como se expresa en la Tabla 3.

TABLA 2. ESQUEMA PARA LA CONSTRUCCIÓN DE LOS INDICADORES DE TRAYECTORIA ESCOLAR

Indicador	Fórmula	Niveles
Índice de aprobación en ordinario	(Número de asignaturas o créditos aprobados en ordinario/Total de asignaturas o créditos cursados) *100	Bajo, Medio, Alto
Índice de promoción	(Número de asignaturas o créditos promovidos/Total de asignaturas o créditos cursados) *100	Bajo, Medio, Alto
Promedio	(Calificaciones numéricas y aprobatorias/Total de asignaturas)*100	Bajo, Medio, Alto
Desempeño escolar	Resultado de la combinación de IAO, IP y Promedio	Alto, Medio, Bajo

Nota: los datos de la Tabla provienen del libro *Perfil de ingreso y trayectoria escolar en la Universidad* de Chain, y Jácome, 2007, impreso con autorización.

TABLA 3. TRAYECTORIA ESCOLAR Y NIVELES DE RIESGO

Desempeño escolar	Situación escolar	Tipos de trayectoria escolar	Niveles de Riesgo
Bajo	Rezago	1	Alto
Bajo	Irregular	2	Alto
Bajo	Óptimo	3	Medio
Regular	Rezago	4	Alto
Regular	Irregular	5	Medio
Regular	Óptimo	6	Bajo
Alto	Rezago	7	Medio
Alto	Irregular	8	Medio
Alto	Óptimo	9	Bajo

Nota: los datos de la Tabla provienen del libro *Perfil de ingreso y trayectoria escolar en la Universidad* de Chain, y Jácome, 2007, impreso con autorización.

Por último, se calcularon las correlaciones entre el desempeño en el EXANI II y el riesgo escolar y DE.

3. RESULTADOS

3.1. Desempeño en el EXANI II

El desempeño que los estudiantes tuvieron en el EXANI II, organizado por niveles, se presenta en la Tabla 4. Se puede observar una distribución casi homogénea en los niveles de desempeño; sin embargo predominan los porcentajes en el nivel Muy bajo, con aproximadamente entre el 27 y el 45 % de los estudiantes en las diferentes áreas del EXANI II.

TABLA 4. DESEMPEÑO EN EL EXANI II

Áreas de EXANI II	Niveles de desempeño							
	Muy bajo		Bajo		Alto		Muy alto	
	f	%	f	%	F	%	f	%
Razonamiento verbal	157	35.7	89	20.2	131	29.8	63	14.3
Razonamiento matemático	170	38.6	89	20.2	112	25.5	69	15.7
Español	165	37.5	72	16.4	123	28	80	18.2
Matemáticas	129	29.3	123	28	84	19.1	104	23.6
Ciencias naturales	171	38.9	65	14.8	133	30.2	71	16.1
Ciencias sociales	119	27	163	37	71	16.1	87	19.8
Mundo contemporáneo	201	45.7	104	23.6	76	17.3	59	13.4
Inglés	145	33	102	23.2	90	20.5	103	23.4
Global	119	27.0	113	25.7	105	23.9	103	23.4

3.1.1. Trayectorias escolares

Utilizando la propuesta de Chain y Jácome (2007) se calcularon los indicadores de trayectoria escolar: 1. El desempeño escolar, 2. Los nueve tipos de trayectorias escolares y 3. Los niveles de riesgo escolar (Tabla 3, 4 y 5)

3.2. Desempeño escolar

3.2.1 Índices de aprobación en ordinario

En la Tabla 5 se muestra la distribución en cuanto al Índice de Aprobación en Ordinario (IAO). El 43.9% de los estudiantes se ubica en el nivel Alto; un porcentaje cercano (42.2%) representa a los estudiantes con nivel Bajo y el 13.9% restante se caracteriza por tener nivel medio. Para el campus de CBA y CSalud la mayor proporción de estudiantes se ubica en el nivel alto de IAO (58% y 62.9%), esto significa que en esas dos áreas una mayor proporción de los estudiantes aprueban en el ordinario entre el 90% y el 100% de sus asignaturas cursadas. Lo contrario sucede con CSociales, CEI, en las que el mayor porcentaje de los estudiantes se concentra en el nivel Bajo (45.1% y 57.4%), lo que significa que el mayor porcentaje de los estudiantes de estas dos áreas se caracteriza por aprobar en ordinario menos del 80% de las asignaturas cursadas.

TABLA 5. ÍNDICE DE APROBACIÓN EN ORDINARIO

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	29	24.4	21	17.6	69	58
CSalud	35	4	11.4	9	25.7	22	62.9
CSociales	91	41	45.1	11	12.1	39	42.9
CEI	195	112	57.4	20	10.3	63	32.3

3.2.2 Índice de promoción

En el Índice de Promoción, el campus de CEI es la única área en la que el mayor porcentaje de estudiantes presenta un nivel Bajo, lo cual significa que éstos aprueban menos del 90% de las asignaturas cursadas. En todas las otras áreas, el mayor porcentaje de estudiantes se concentra en el nivel Alto, es decir, la mayor proporción de los estudiantes promueve el 100% de las asignaturas cursadas en cualquiera de las oportunidades que les otorgue la institución, incluyendo el ordinario. En el campus de CSociales, donde el mayor porcentaje de los estudiantes se caracterizaba por aprobar en ordinario menos del 80% de las asignaturas (45% de los estudiantes), se aprecia que aunque los estudiantes reprueben los exámenes ordinarios, se recuperan al aprobar en las otras opciones que el plan de estudios les permite (extraordinarios, especiales, etc.). En esta área el 41.8% ha aprobado el 100% de los créditos o asignaturas, el 28.6% promueve entre el 90 y menos del 100% de sus asignaturas o créditos, y únicamente el 29.7% ha promovido o aprobado menos del 90% de las asignaturas. En comparación con el porcentaje de asignaturas aprobadas en el ordinario este último indica que existe un 34.1% de sujetos que han aumentado su aprobación utilizando otras opciones (comparado con el 45.1% de estudiantes que aprueban en ordinario menos del 80% de las asignaturas cursadas), lo cual supone que los exámenes extraordinarios y especiales, constituyen una opción importante para aprobar las asignaturas. En la Tabla 6 se expresan las frecuencias y porcentajes en cuanto a este indicador.

TABLA 6. ÍNDICE DE PROMOCIÓN

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	25	21	6	5	88	73.9
CSalud	35	0	0	2	5.7	33	94.3
CSociales	91	27	29.7	26	28.6	38	41.8
CEI	195	110	56.4	26	13.3	59	30.3

3.2.3 Promedio de calificaciones

En cuanto al Promedio de calificaciones, predominan los niveles medios. En los campus de CBA y CEI, el mayor porcentaje de los estudiantes tiene un promedio de entre 75 y 85; en el caso de los campus de CSociales y de CSalud, el mayor porcentaje de los estudiantes tiene un promedio de entre 80 y 90. En el caso de CSalud ningún estudiante tiene promedio menor a 80. En cuanto a los estudiantes en su conjunto, predomina el promedio Medio en todas las áreas del conocimiento; es decir, que en el caso de los campus de CBA y CEI los estudiantes tienden a tener con mayor frecuencia promedios de entre 75 y menos de 85, y en los campus de CSociales y CSalud los estudiantes presentan mayor frecuencia en el rango de promedios de entre 80 y menos de 90 (Tabla 7).

TABLA 7. PROMEDIO DE CALIFICACIONES

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	28	23.5	68	23	19.3	
CSalud	35	0	0	23	65.7	12	34.3
CSociales	91	5	5.5	52	57.1	34	37.4
CEI	195	77	39.5	83	42.6	35	17.9

3.2.4 Desempeño escolar

Los indicadores: IAO, IP y Promedio, conforman el Desempeño Escolar. En éste, los campus de CBA y CSalud presentan el mayor porcentaje de los alumnos con un desempeño en el nivel alto (52.1% y 62.9%); por otra parte, los estudiantes del campus de CSociales y CEI se ubican en el nivel de desempeño escolar Bajo (47.3% y 61%). En su conjunto, el 44% de los estudiantes tiene un nivel de desempeño Bajo, esto indica que el mayor porcentaje de los estudiantes se caracteriza por tener niveles bajos o medios en cuanto a la aprobación en ordinario, promoción y promedios. No obstante, un porcentaje no muy lejano al anterior (39.5%) se ubica en el nivel de desempeño Alto (Tabla 8).

TABLA 8. DESEMPEÑO ESCOLAR

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	31	26.1	26	21.8	62	52.1
CSalud	35	1	2.9	12	34.3	22	62.9
CSociales	91	43	47.3	11	12.1	37	40.7
CEI	195	119	61	23	11.8	53	27.2

3.3. Situación escolar

En cuanto a la situación escolar, se expresa en la Tabla 9, en este indicador los campus de CEI y CByA, presentan el mayor porcentaje de sus estudiantes en estado de *Rezago*, predominando significativamente en el campus de CEI. Por otra parte, el 38% de los estudiantes presenta un estado *Óptimo*, es decir, son aquellos estudiantes que han promovido el 100% de las asignaturas o créditos requeridos por el programa de licenciatura. Destaca el área de CSalud con el 100% de sus alumnos con esta característica como se muestra en la Tabla 9.

TABLA 9. SITUACIÓN ESCOLAR

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	26	21.8	1	8	92	77.3
CSalud	35	0	0	0	0	35	100
CSociales	91	38	41.8	21	23.1	32	35.2
CEI	195	178	91.3	7	3.6	10	5.1

3.4. Situación escolar

Entre éstas sobresalen las trayectorias de estudiantes caracterizados por un desempeño Alto y un estado Óptimo con un 26.8%, y los estudiantes con un desempeño escolar Bajo y un estado de Rezago con un 38.6%. De igual manera, pero en menor proporción, destacan los estudiantes con desempeño Alto pero

Rezagados con un 11.6% y los estudiantes con un desempeño Regular y estado Óptimo con un 9.1%. El mayor porcentaje de los estudiantes se caracteriza por tener un desempeño Bajo y estar en Rezago (38.6%); de este porcentaje, el 47%, que representa a casi la mitad de los estudiantes, se encuentran en situación de baja.

3.5. Niveles de riesgo escolar

El mayor porcentaje de los estudiantes se encuentra en riesgo alto (46.4%), determinado en gran medida por los alumnos del campus de CEI quienes presentan el mayor porcentaje; por otra parte el 35.9% se encuentra sin riesgo, determinado en gran medida por los estudiantes de ciencias de la salud (Tabla 10).

TABLA 10. NIVELES DE RIESGO CALCULADOS EN EL CUARTO SEMESTRE

Campus	n	Bajo		Medio		Alto	
		f	%	F	%	f	%
N=440							
CBA	119	82	68.9	15	12.6	22	18.5
CSalud	35	34	97.1	1	2.9	0	0
CSociales	91	32	35.2	13	14.3	46	50.5
CEI	195	10	5.1	49	25.1	136	69.7

3.6. Niveles de riesgo escolar

3.6.1. Población y muestra

Al estudiar la relación entre el Desempeño Escolar de los estudiantes del campus de CSalud con el desempeño en el EXANI II, se encontró que no existe relación lineal entre dichas variables. Este mismo resultado fue obtenido al estudiar las relaciones entre el riesgo escolar con el resultado global del EXANI II y el resultado en cada una de sus áreas.

3.6.2. Campus de ciencias sociales

En el campus de CSociales se encontró una relación lineal inversa entre la calificación en el área de ciencias sociales del EXANI II y el riesgo escolar ($r = -.222, p < .05$) lo cual significa que conforme mejora el desempeño en ésta área de EXANI II, disminuye el riesgo escolar. En cuanto a la relación entre el desempeño en el EXANI II y el desempeño escolar, las áreas Global ($r = .219, p < .05$), matemáticas ($r = .212, p < .05$), ciencias naturales ($r = .212, p < .05$) y ciencias sociales ($r = .235, p < .05$) del EXANI II, demostraron que existe una relación lineal directa, lo cual significa que conforme aumenta el desempeño en estas áreas del examen, aumenta el desempeño escolar.

3.6.3. Campus de ciencias exactas e ingenierías

En cuanto al riesgo escolar, se encontró una relación inversa con las áreas Global ($r = -.254, p < .05$), razonamiento matemático ($r = -.291, p < .05$), español ($r = -.166, p < .05$), matemáticas ($r = -.323, p < .05$) e Inglés ($r = -.141, p < .05$) del EXANI II, lo cual significa que conforme aumenta el desempeño en éstas áreas, disminuye el riesgo escolar. Se encontraron relaciones significativas entre el desempeño escolar y las áreas global ($r = .265, p < .05$), razonamiento matemático ($r = .216, p < .05$), español ($r = .203, p < .05$) y matemáticas ($r = .316, p < .05$) del EXANI II, lo cual significa que conforme aumenta el desempeño en el examen (en estas áreas), aumenta el desempeño escolar.

3.6.4. Campus de ciencias biológicas y agropecuarias

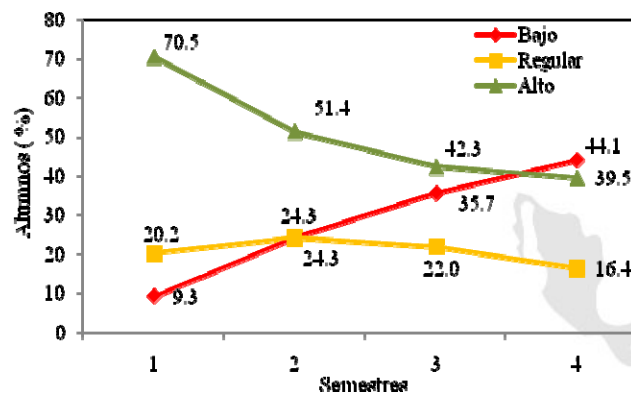
No se encontró relación lineal entre el desempeño en el EXANI II y el riesgo escolar, pero al analizar la relación con el desempeño escolar surge que existe relación lineal directa con las áreas global ($r=.265, p<.05$), razonamiento verbal ($r=.188, p<.05$), razonamiento matemático ($r=.199, p<.05$), matemáticas ($r=.250, p<.05$) y ciencias naturales ($r=.295, p<.05$) del EXANI II, lo cual significa que conforme aumenta el desempeño en el examen, aumenta el desempeño escolar.

3.7. Evolución de los indicadores de Desempeño escolar, Estado y Riesgo a través de los 4 semestres.

Para el caso del desempeño escolar de aquellos estudiantes que tenían un desempeño Bajo (9.3%), el 97.6% se mantuvo en ese nivel, mientras que únicamente un 2.4% pasó al desempeño Regular. De los estudiantes que en el primer semestre representan el desempeño Regular, el 73% cambió al desempeño Bajo en el cuarto semestre; el 19.1% se mantuvo en esa situación y únicamente el 7.9% mejoró su desempeño al ubicarse en el nivel Alto.

Por otra parte, de los estudiantes que en el primer semestre presentaban un desempeño escolar Alto, el 53.9% se mantuvo en ese nivel de desempeño en el cuarto semestre; el 16.4% bajó al nivel de desempeño Regular y el 44.1% pasó del desempeño Alto al Bajo en el cuarto semestre. De manera general, se aprecia en la Figura 1 el decremento significativo de la cantidad de alumnos en el nivel de desempeño Alto, y por otra parte, el aumento del número de alumnos con desempeño escolar Bajo.

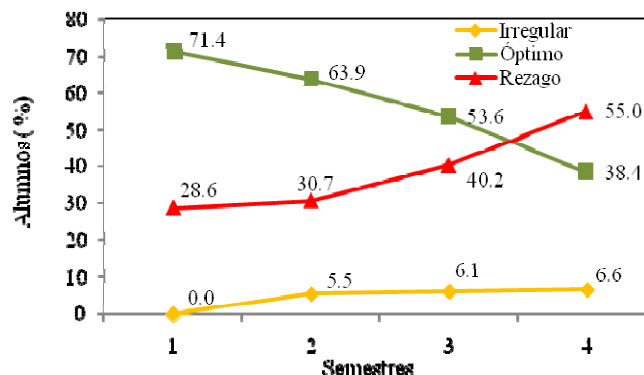
FIGURA1. EVOLUCIÓN DEL DESEMPEÑO ESCOLAR DE LOS ESTUDIANTES A TRAVÉS DE LOS CUATRO PRIMEROS SEMESTRES



En la Figura 2 se muestra la evolución de los porcentajes de estudiantes en cuanto a su Estado por cada semestre; se aprecia que la cantidad de estudiantes con situación escolar Óptima en el primer semestre (71.4%) disminuye en el cuarto semestre en un 53.8%. De manera contraria, el número de estudiantes que en el primer semestre tenían una situación escolar de Rezago aumentó en el cuarto semestre en un 92.1% (28.6% a 55%).

En la Figura 3 se muestra la evolución del número de estudiantes por cada semestre en cuanto a los niveles de riesgo. Se aprecia a simple vista que, al igual que en las figuras anteriores, la relación es inversa entre los niveles de Riesgo bajo y Riesgo alto, siendo que conforme una disminuye la otra aumenta, y es en el paso del tercer al cuarto semestre cuando los porcentajes se invierten quedando por debajo el nivel de Riesgo bajo.

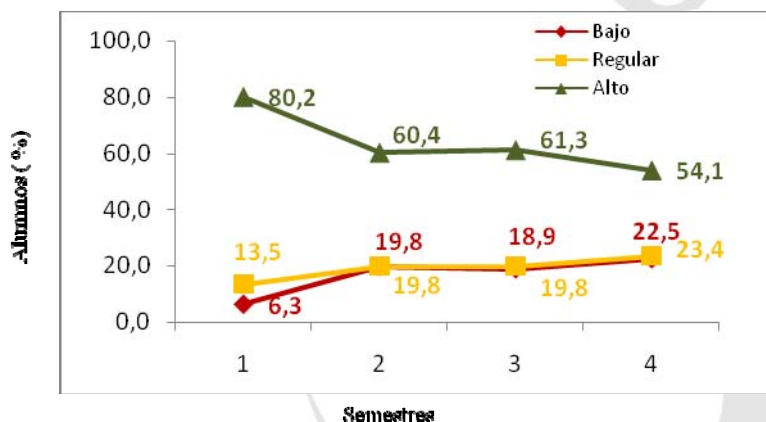
FIGURA 2. EVOLUCIÓN DEL ESTADO DE LOS ESTUDIANTES A TRAVÉS DE LOS CUATRO PRIMEROS SEMESTRES



Del porcentaje de estudiantes que en el primer semestre se encontraban Riesgo bajo, el 52.2% se mantiene en este nivel, pero el 22.2% pasa al nivel de Riesgo medio y el 25.6% restante forma parte de los estudiantes con Riesgo alto. En el caso de los estudiantes que en el primer semestre se encuentran en riesgo, el 4% mejora su situación al ubicarse en el nivel de Riesgo bajo; el 12% se mantiene con Riesgo medio mientras que el 84% agrava su situación al pasar al nivel de Riesgo alto. Por otra parte, de los estudiantes que en el primer semestre se encontraban con riesgo alto, el 95.6% se mantiene en el cuarto semestre en esa misma situación, mientras que únicamente 4.4% pasa al nivel de Riesgo medio y ninguno a Riesgo bajo. De igual manera, es entre los períodos tres y cuatro en los que los estudiantes, debido a su situación escolar, optan por darse de baja temporal de la licenciatura, lo cual implica atraso con respecto a la cohorte a la que pertenecen.

En todos los casos se puede identificar la transición del tercer al cuarto período como crítico. Sin embargo, es importante considerar que el análisis incluye a estudiantes de todas las áreas del conocimiento, por lo que la influencia de los datos del campus de CEI es en sentido opuesto al campus de CSalud, dado que una se caracteriza por tener índices de rezago altos y desempeños bajos, y la otra, por tener el mayor porcentaje de sus alumnos con desempeños Altos-Regulares y en estado Óptimo.

FIGURA 3. EVOLUCIÓN DE LOS NIVELES DE RIESGO DE LOS ESTUDIANTES A TRAVÉS DE LOS CUATRO PRIMEROS SEMESTRES



4. DISCUSIÓN

La UADY ha buscado en los últimos años propiciar que existan condiciones de igualdad en la selección de aspirantes, cambiando los instrumentos y junto con esto, el ajuste de los criterios de selección. Con el uso del resultado obtenido en un examen estandarizado (EXANI II), desarrollado por una instancia evaluadora externa (CENEVAL), se pretende lograr un proceso claro y justo, equitativo y libre de prejuicios o de criterios faltos de transparencia.

Este proceso trata de discriminar a aquellos sujetos que, dados sus conocimientos y habilidades, tendrán mayores probabilidades de cursar los estudios con éxito, con respecto a los que no, esto le da un carácter predictivo al examen, que se confirmaría con los estudios de la trayectoria escolar.

Diversos estudios han analizado las relaciones entre el EXANI II y las trayectorias escolares, como son: González & Velázquez, 2006, Carrasco, Franco, Herrera & Contreras, 2006; Casillas, Chain & Jácome, 2007; González, Morfin & Peña, 2005. Sin embargo, este estudio resulta el primero en su género en la UADY.

En este estudio, los resultados del EXANI II se relacionan de manera diversa con la desempeño escolar y el riesgo de los estudiantes en los primeros cuatro periodos en la universidad.

En algunos casos resultaron relacionadas de manera lineal y positiva las áreas del EXANI II con las áreas correspondientes a la temática de la disciplina, lo cual puede sugerir que los estudiantes, al elegir una carrera han orientado sus esfuerzos a un campo específico, fortalecido desde su formación anterior. También es posible que sea consecuencia de la naturaleza de las asignaturas optativas que elige en el bachillerato o bien la fortaleza y tiempo invertido en el momento de estudiar áreas preferidas al prepararse para sustentar el examen; luego, estos conocimientos adquiridos impactan en su tránsito por la universidad. Tal es el caso del campus de CSociales en el que resultó significativa la relación del desempeño escolar con el desempeño en el área de Ciencias sociales del EXANI II, en el campus de CEI resultó significativa la relación del desempeño escolar con el desempeño en las áreas relacionadas a las matemáticas (Razonamiento matemático y Matemáticas) y en el área de CBA el área de Ciencias naturales del mismo examen.

Existen diferencias en los resultados en cuanto la naturaleza de las licenciaturas. En el campus de Ciencias de la Salud, no se encontraron relaciones significativas entre el desempeño en el EXANI II con el desempeño y riesgo escolar; lo que significa que tanto los estudiantes con desempeño bajo, como los de desempeño alto en el EXANI II tienen desempeños similares en los cuatro primeros periodos escolares. Esta área se caracteriza por tener los mejores índices de desempeño y riesgo escolar bajo, así como la menor deserción y alta permanencia de sus estudiantes. Estos resultados parecen explicar la efectividad de los programas de apoyo estudiantil como la tutoría, así como un alto perfil de disciplina y esfuerzo que se fomenta en los estudiantes de esta área y se refleja en el trabajo de los profesionales en los hospitales.

Por el contrario el campus de CEI, se caracteriza por tener bajos índices de desempeño escolar y mayores porcentajes de estudiantes en riesgo alto, pasando de un 20 % de estudiantes en el primer periodo a 55% en el cuarto periodo escolar. Se puede apreciar diferencias en las trayectorias que muestran los estudiantes de estos dos campus, que en el caso del campus CEI los resultados de bajo desempeño y riesgo alto, parecen estar asociados a la complejidad y naturaleza abstracta de las ciencias duras y la cultura de la disciplina que se ilustra con ideas como "mientras más tarden en la carrera más aprenden" y "reprobar es normal".

Existen diferencias en cada campus con respecto a las relaciones con el desempeño en el EXANI II y sus áreas específicas. Por lo tanto, sería deseable atender esas características particulares, en cuanto a las necesidades que tienen con respecto a las habilidades y conocimientos que requiere un alumno de nuevo ingreso para transitar con éxito y mantener su permanencia en los estudios. Los resultados del EXANI II muestran un panorama de las habilidades y conocimientos que poseen los alumnos en las áreas del examen, por lo que también puede indicar posibles deficiencias o debilidades en su formación, herramienta que puede utilizarse para establecer mecanismos de apoyo para los estudiantes, en los primeros años de la carrera. Estos resultados arrojan datos importantes para el conocimiento de los estudiantes que ingresan en la universidad y sus necesidades, así como, constituyen datos preliminares en cuanto al estudio longitudinal de las trayectorias escolares universitarias. Estos resultados hacen surgir nuevas necesidades de profundizar en aspectos tales como la relación entre los perfiles de egreso de la educación media superior y superior: así como en la búsqueda de modelos de predicción del éxito en los estudios universitarios y vincularlos a su vez, con el éxito profesional y laboral.

El porcentaje de estudiantes con nivel de riesgo alto aumenta durante los primeros cuatro semestres, particularmente, el paso del tercer al cuarto periodo es crítico pues casi el 50 % del total de los estudiantes se ubican en esta situación. Esta situación puede deberse a que conforme avanzan los periodos, las asignaturas aumentan su complejidad y con esto las exigencias escolares. De igual manera es posible que exista un efecto de acumulación del número de asignaturas que los estudiantes aun no acreditan. Esta situación se agrava con las oportunidades adicionales para presentar exámenes que les permita acreditar las asignaturas, debido a que los estudiantes continúan con sus cargas académicas sin haber acreditado por completo las antecedentes, y en caso de reprobarlas, se invalidará el avance obtenido en sus otras asignaturas. Esta situación muy probablemente eleva los niveles de estrés de los estudiantes, lo cual podría afectar de manera importante su rendimiento escolar, contribuyendo al aumento del porcentaje de estudiantes en riesgo alto.

En términos de predictibilidad para un buen desempeño en la educación superior, Chain, Casillas & Cruz, 2006 encontraron que el promedio de bachillerato es un elemento que se considera como un indicador relevante del desempeño de los estudiantes, siendo evidente que cuando un sujeto presenta un promedio bajo en el bachillerato, se asocia a una trayectoria baja en la universidad o una escasa probabilidad de obtener una trayectoria escolar alta. Por el contrario, los casos de promedios de bachillerato alto, se asocian a trayectorias altas o con escasas probabilidades de presentar trayectorias regular o baja, por lo que se confirma que el desempeño en el bachillerato es la variable más relevante y la que mayor información aporta en la predicción del desempeño en la universidad.

Para la UADY el uso del EXANI II representa la oportunidad de tener un proceso transparente, equitativo y objetivo. El uso de criterios adicionales para seleccionar a los futuros estudiantes, podría favorecer la predictibilidad del éxito de los mismos en sus trayectorias escolares.

El promedio de bachillerato, es utilizado en otras instituciones públicas en las que ha demostrado que puede fortalecer las decisiones de selección y favorecer los resultados de las trayectorias escolares, tal como se encontró en el estudio mencionado. Esto representa un área de oportunidad que podría explorarse en la UADY para implementarse como mecanismo complementario en el proceso de selección de aspirantes.

Adicionalmente, tal y como sugiere la situación del Campus de Ciencias de la Salud, los mecanismos de apoyo a los estudiantes y de atención a sus necesidades académicas tales como los cursos de nivelación,

cursos inter semestrales en lo que se pueden cursar asignaturas de manera intensiva, sin necesidad de esperar los períodos regulares y los programas de tutorías, pueden contribuir a mejorar las trayectorias escolares de los estudiantes.

Para concluir, el presente estudio representa una primera aproximación al conocimiento de los estudiantes y sus trayectorias escolares en la UADY y sus resultados ilustran las diversas relaciones que se dan con el desempeño en el EXANI II y las áreas que lo conforman, por lo que es necesario realizar estudios enfocados a este tema, en períodos más amplios y explorando otras variables que afectan el desempeño escolar, como lo es la situación socioeconómica, trayectoria escolar previa, hábitos de estudio, entre otros.

REFERENCIAS BIBLIOGRÁFICAS

- Altamira, A. (1997). *El análisis de las trayectorias escolares como herramienta de evaluación de la actividad académica universitaria: Un modelo ad hoc para la Universidad Autónoma de Chiapas, el caso de la Escuela de Ingeniería Civil*, Tesis de Maestría en Educación México.
- Aguilar, A. (2002). *El cambio de la política de admisión de estudiantes en la Universidad de Guadalajara y su implantación 1995-1999*. México: ANUIES.
- Chain, R., Cruz, N., Martínez, M. & Jácome, N. (2003). Examen de selección y probabilidades de éxito escolar en estudios superiores. Estudio en una universidad pública estatal mexicana. *Revista Electrónica de Investigación Educativa*, 5. Recuperado el 3 de noviembre de 2007 de: <http://redie.uabc.mx/vol5no1/contenido-chain.html>
- Chain, R. & Jácome, N. (2007). *Perfil de ingreso y trayectoria escolar en la Universidad*. México: Instituto de Investigaciones en Educación-UV.
- García, C. (2005). *Financiamiento de la Educación Superior en América Latina, en GUNI-UNESCO, Educación Superior en el mundo 2006. El financiamiento de las universidades*. España: Mundi-Prensa.
- Vázquez, R. (1985). *La influencia de los estilos cognitivos en el rendimiento escolar, en la trayectoria escolar en la educación superior*. México: ANUIES.

EQUIPARACIÓN, ALINEAMIENTO Y PREDICCIÓN DE PUNTUACIONES EN MEDICIÓN EDUCATIVA

SCORE EQUATING, ALIGNING AND PREDICTION IN EDUCATIONAL
MEASUREMENT

René Gempp

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art6.pdf>

Fecha de recepción: 04 de abril de 2010

Fecha de dictaminación: 18 de mayo de 2010

Fecha de aceptación: 18 de mayo de 2010

Las pruebas educativas estandarizadas son dispositivos de evaluación masiva diseñados para proveer evidencia replicable (i.e. "fiable") a partir de la cual formular inferencias relevantes y fundamentadas (i.e. "válidas") sobre los conocimientos y habilidades de quienes las responden.

El desarrollo de una prueba estandarizada requiere resolver varios problemas técnicos de alta complejidad. Entre ellos, uno de los más cruciales es también el menos conocido fuera del círculo de especialistas en psicometría. Nos referimos al conjunto de problemas asociados a la *comparabilidad* de las mediciones que, paradójicamente, afecta directamente la interpretación y uso de las puntuaciones, que son el resultado más visible de prueba.

Desde el lado de los examinados, los problemas de *comparabilidad* aparecen cuando enfrentan pruebas estandarizadas cuyos resultados se expresan en métricas distintas para evaluaciones que se supone miden lo mismo. Por ejemplo, un estudiante podría obtener 30 puntos en una prueba de inglés contratada y aplicada por su escuela, 250 puntos en otra prueba de inglés rendida en el proceso de selección a un instituto y 400 puntos en una prueba aplicada por alguna agencia nacional o regional que esté realizando un diagnóstico sobre el nivel de inglés de los estudiantes. ¿Qué significan esos resultados dispares? ¿Hay alguna relación entre ellos? ¿En cuál de todas las pruebas el resultado es mejor? ¿Y por qué son diferentes las escalas de puntuación si todas las pruebas evalúan inglés?

Desde el lado de otros usuarios importantes, los investigadores en educación, la situación tampoco es fácil cuando deben enfrentarse a problemas de comparabilidad de mediciones. Para fines ilustrativos, sólo un ejemplo: ¿Cómo se explica que el resultado de un grupo de escuelas en una prueba estandarizada aplicada por el sistema nacional de medición sea de 240 puntos y que las mismas escuelas hayan obtenido 450 puntos en una prueba internacional en la que el país ha participado? ¿Qué significa esa diferencia? ¿Es posible llevar las puntuaciones de la prueba nacional a la escala de reporte de la prueba internacional?

Por otro lado, las instituciones responsables de estas pruebas (por ejemplo, las agencias nacionales de medición educativa o los organismos encargados de los procesos de selección universitaria) enfrentan una cara menos visible pero aún más delicada del problema de comparabilidad al tratar de resolver cuestiones del tipo ¿Cuánto podemos cambiar las especificaciones de la prueba de este año sin comprometer la comparabilidad con los resultados de años anteriores? ¿En qué medida un cambio curricular afecta la capacidad de nuestras pruebas para realizar un monitoreo escolar válido? ¿Cómo hacemos para distribuir los ítems de la prueba en varios cuadernillos distintos sin comprometer su validez ni sesgar los resultados? ¿Qué efecto podría tener en los resultados la técnica de *equiparación* que utilicemos?

Estos y otros muchos problemas asociados a la *comparabilidad* de las mediciones educativas son una fuente continua de dolores de cabeza para los especialistas. La mayoría de ellos puede resolverse mediante metodologías básicas de *equiparación* [del inglés *equating*]¹ pero otros requieren de métodos especiales cuyas limitaciones y alcances no siempre son bien comprendidos.

¹ La traducción preferida para el término *equating* está fuertemente condicionada por la geografía. Aunque la voz hispana *equiparación* es, sin duda, la correcta y cuenta, además, con plena aceptación entre los especialistas, en Latinoamérica se escuchan frecuentemente otras alternativas, como "igualación", "equivalencia" o "ecualización". En nuestra opinión, ninguna de ellas supera a "equiparación" en fidelidad con el original y pueden, además, inducir a serios malentendidos. Por ejemplo, la acepción castellana de "igualación" conlleva la idea equivocada de que es posible lograr que dos pruebas sean iguales, cuando de lo que se trata es apenas de lograr equivalencia entre sus puntuaciones. Por otro lado, "equivalencia" respeta el significado del concepto, pero complica la tarea de expresarlo como verbo (i.e. ¿cómo decir que queremos equiparar dos pruebas?). Finalmente, "ecualización" es la traducción correcta del término "equalization", cuya definición específica en psicometría moderna (Kolen, 1988) es más cercana a su significado en ingeniería acústica que a la idea de *equiparación*.

En nuestra experiencia, por ejemplo, hemos encontrado que un error muy común entre los investigadores es creer que las técnicas basadas en regresión múltiple son un método óptimo para establecer el grado de equivalencia entre dos mediciones. Otra creencia muy arraigada es suponer que cualquier par de mediciones pueden *equipararse* o *igualarse* a todo evento, mediante la técnica de estadística adecuada, ignorando que la *equiparación* es una meta que no siempre se puede lograr empíricamente. Un equívoco relacionado, también muy frecuente, es creer que las técnicas de *equiparación* permiten igualar dos pruebas diferentes, cuando la verdad es que su único propósito es ajustar diferencias de dificultad entre las puntuaciones de ambas pruebas.

Estos y otros malentendidos se originan en una comprensión equivocada de las condiciones que posibilitan que una *equiparación* sea posible y en el desconocimiento de otros procedimientos para establecer *comparabilidad* entre mediciones. Sin embargo, métodos alternativos se han venido aplicando desde principios del siglo XX para, por ejemplo, alinear puntuaciones de pruebas distintas o monitorear progreso educativo. En algunos casos se trata de técnicas relativamente populares (e.g. *Escalamiento Vertical*) y en otros, de procedimientos apenas conocidos por un puñado de especialistas (i.e. *Escalamiento de Baterías*).

Familiarizarse con estas metodologías es necesario por dos razones. Primero, porque facilita visualizar soluciones para problemas de medición educativa que a veces se consideran irresolubles o, peor aún, sobre los cuales no se ha tomado la debida conciencia. Segundo, porque el contraste de estas metodologías con las técnicas de *equiparación* tradicionales ayuda a comprender mejor las condiciones en las cuales una *equiparación* es o no válida.

Lamentablemente, los pocos trabajos que ofrecen exposiciones de metodologías de *comparabilidad* adicionales a la equiparación (Kolen, 2004; Kolen & Brennan, 2004; Holland & Dorans, 2006; Dorans, Pommerich & Holland, 2007) están dirigidos a un público angloparlante y altamente especializado en psicometría. Por ello, y con el propósito de contribuir a la divulgación en nuestro medio de algunas metodologías alternativas a la *equiparación*, en este artículo ofrecemos una revisión sintética de los métodos más importantes para lograr *comparabilidad* entre mediciones. En cada caso, se explica brevemente el propósito de método, se revisa un ejemplo prototípico y se sugieren algunas aplicaciones potenciales en nuestro contexto regional. De acuerdo al objetivo didáctico del trabajo, se optó por no describir los diseños de recogida de información o las técnicas estadísticas para cada procedimiento, aunque se sugieren referencias especializadas para el lector interesado.

La exposición que presentamos a continuación se basa en el marco conceptual propuesto por Holland y Dorans (2006), también descrito por Holland (2007) y por Dorans, Holland & Petersen (2007). Aunque a través de los años diferentes especialistas han desarrollado marcos teóricos para sistematizar los métodos de *comparabilidad* (ver, por ejemplo, los trabajos de Flanagan, 1951; Angoff, 1971; Mislevy, 1992; Linn, 1993; Feuer, Holland, Green, Bertenthal & Hemphill, 1999; Dorans, 2000, 2004), la taxonomía de Holland y Dorans (2006) es la más comprehensiva y mejor fundamentada en el presente.

1. CONCEPTOS BÁSICOS: ENLACE DE PUNTUACIONES

Para comenzar, es necesario circunscribir el problema de *comparabilidad entre mediciones* a aquellas situaciones que requieren construir algún tipo de regla de correspondencia que nos permita expresar el resultado de una prueba en la métrica de la otra.

Técnicamente hablando, la solución a estos problemas es aplicar alguna metodología de *enlace de puntuaciones* [del inglés *score linking*]² es decir, *una transformación estadística que permita encontrar a qué puntuaciones en una prueba equivalen cada una de las puntuaciones de la otra prueba* (Holland & Dorans, 2006).

Nótese que en adelante utilizaremos el término *enlace* en lugar de "*comparabilidad*" [del inglés *comparability*], porque este último no tiene una definición inequívoca en la literatura psicométrica. Por ejemplo, Flanagan (1951) lo utilizó en un sentido más restringido que *enlace*, para describir algunas de las metodologías que hoy en día se agrupan bajo la noción de *alineamiento de puntuaciones*, que explicaremos más adelante. Veinte años después, Angoff (1971) acotó aún más el concepto, utilizándolo sólo para aquellos *enlaces* entre pruebas que midieran constructos diferentes. En la siguiente década (APA, 1985) se lo utilizó para referir a todos los tipos de *enlace* que no alcanzaran el rigor técnico de una *equiparación* propiamente tal [*scaling to achieve comparability*]. Desde entonces se lo emplea de vez en cuando como sinónimo informal de *equiparación* o para referir a problemas de *enlace* en un sentido amplio, pero sin mediar una definición técnica precisa ni mucho menos completamente aceptada. Para complicar aún más las cosas, es cada vez más frecuente utilizar el término *comparabilidad* para denominar al área específica de estudio que se ocupa de evaluar la equivalencia psicométrica entre las versiones informatizada y tradicional de una misma prueba [*"test comparability studies"*; ver por ejemplo, ejemplo, Sireci & Zenisky, 2006)].

Por las razones anteriores, las publicaciones especializadas privilegian el uso del término *enlace*, cuya definición técnica es precisa, en lugar de *comparabilidad*, de uso coloquial y significado ambiguo.

La definición del *enlace de puntuaciones* es suficientemente amplia como para describir situaciones muy diversas, que involucran metodologías distintas. En términos generales, las metodologías de *enlace* se diferencian entre sí por la validez, la precisión y la reversibilidad con que se pueden transformar las puntuaciones de una prueba en otra, pudiéndose distinguir tres grandes grupos de métodos, desde los más laxos hasta los más rigurosos en el cumplimiento de esas condiciones: *predicción*, *alineamiento* y *equiparación* de puntuaciones.

2. PREDICCIÓN DE PUNTUACIONES

La *predicción* es la metodología de enlace más antigua que se conoce y una de las más utilizadas en investigación aplicada. Su propósito es predecir la puntuación esperada en una evaluación a partir de otra información relevante. Esa información puede consistir en los resultados de una o más pruebas pero también puede incluir uno o varios antecedentes sociodemográficos de importancia (e.g. género, etnia, nivel socioeconómico). Aunque no es una regla, es habitual que haya una diferencia temporal entre el momento en que se mide esa información relevante y el momento en que se obtiene la puntuación a ser predicha.

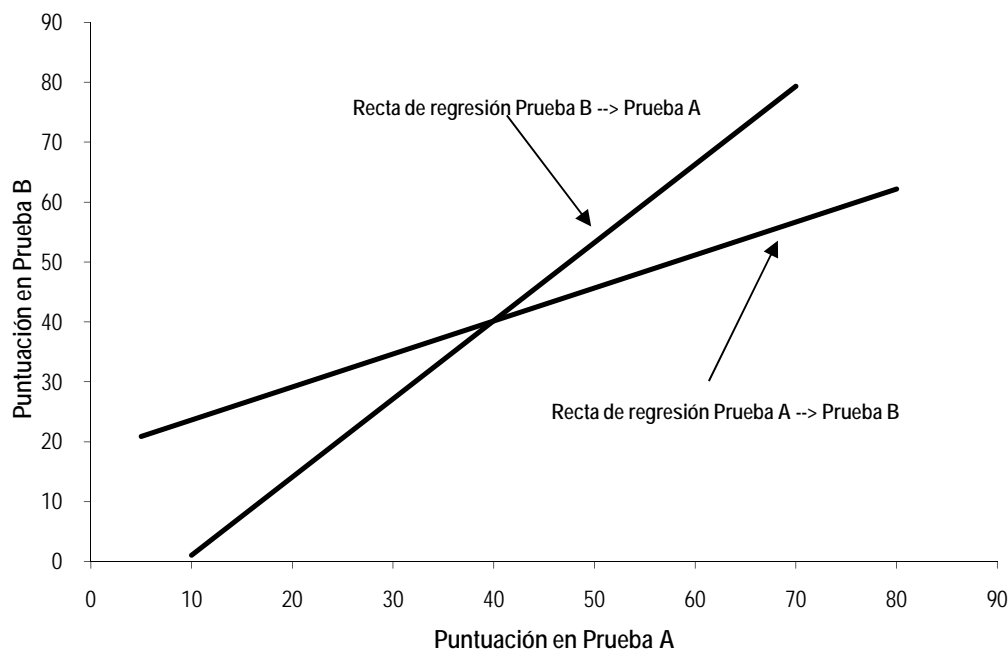
² En la psicometría hispanoparlante no existe un equivalente técnico universalmente aceptado para los conceptos psicométricos *link* y *linking*. Hemos optado por la traducción *enlace* porque nos parece más fiel al significado psicométrico del original, en comparación con las otras alternativas disponibles en castellano (i.e. *vínculo*, *conexión*, *unión*, *ligazón*).

El caso más común en que aplican técnicas de *predicción* corresponde a las pruebas de selección universitaria, cuya validez depende de su capacidad para predecir, con la mayor precisión posible, el rendimiento posterior en la Universidad (otra medición basada, por ejemplo, en el promedio de calificaciones del primer semestre). Otro ejemplo clásico es el intento de predecir el resultado en una prueba de selección universitaria rendida al término de la escolaridad (normalmente, al finalizar del 12° grado), a partir de los resultados obtenidos en pruebas de rendimiento educativo aplicadas durante los años previos de escolaridad (por ejemplo, en 10° grado).

En los *enlaces por predicción* de puntuaciones, el énfasis no suele estar puesto en predecir el resultado específico de cada examinado, sino en la puntuación esperada para el alumno "típico" de acuerdo a ciertas variables de contexto o explicativas. Por ejemplo, predecir el rendimiento más probable en la Universidad, a partir de su prueba de selección, para los alumnos de una determinada titulación, en función de que hayan estudiado en escuelas públicas o privadas. En este sentido, los *enlaces por predicción* generalmente se utilizan en estudios de *validez predictiva* y en la construcción de modelos de crecimiento escolar [*growth models*].

Una característica crucial de la *predicción* es que existe una asimetría fundamental entre la puntuación predicha y las variables que se utilizan como predictores. Hay razones lógicas y estadísticas para ello. Por una parte, los predictores pueden ser varios y anteceder temporalmente a la puntuación predicha (que es sólo una), por lo cual resulta lógico que la predicción opera en una dirección pero no al revés. Por ejemplo, tiene sentido predecir el rendimiento en el primer semestre de Universidad a partir de los resultados de una o varias pruebas de selección universitaria, pero no parece razonable predecir el rendimiento pasado en esas pruebas a partir del rendimiento presente en la universidad.

FIGURA 1. EJEMPLO DE ASIMETRÍA EN LA REGRESIÓN LINEAL SIMPLE



Existe, además, una poderosa razón estadística para esta asimetría. La técnica preferente para los estudios de predicción es la regresión lineal o alguna de sus variantes. Es bien sabido, desde la invención de esta técnica, a fines del siglo XIX (Galton, 1888), que la recta de regresión para predecir Y a partir de X no es exactamente la inversa de la recta para predecir X a partir de Y. Esta idea se ilustra en la Figura 1, en que se grafican las rectas de regresión entre las pruebas A y B, cuyas puntuaciones tiene una correlación a $r=0.85$. Podemos observar que si predecimos los resultados de B utilizando A y luego intentamos predecir la puntuación de A mediante el valor de B recién obtenido, llegaremos a un resultado distinto del original. Este fenómeno, bien conocido por los estadísticos, pero muchas veces olvidado por los investigadores aplicados, redundo en que las predicciones conseguidas mediante regresión lineal no son reversibles. Como veremos más adelante, la falta de reversibilidad o, lo que es lo mismo, la asimetría de la predicción, supone una distinción fundamental entre el *enlace* de puntuaciones por *predicción* y los *enlaces* por *alineamiento* o *equiparación*. Existen dos subtipos básicos de *predicción*: *predicción de puntuaciones individuales* y *proyección de distribuciones*.

2.1. Predicción de puntuaciones individuales

Este tipo de *enlace* es la predicción por antonomasia y se define por las mismas características antes mencionadas. El ejemplo prototípico, como hemos repetido varias veces, es el de las pruebas de selección, aunque otros usos interesantes también son posibles. Entre ellos, predecir el rendimiento en una prueba de término de grado (e.g. 8° grado) a partir de los resultados de una prueba anterior (e.g. 6° grado), o predecir el resultado de una prueba de selección universitaria, rendida al término de la escolaridad, a partir de pruebas estandarizadas rendidas en los años previos y/o a través de una combinación sensata de antecedentes académicos del estudiante.

Una vez obtenida la función de *enlace* entre ambas puntuaciones (i.e. una ecuación de regresión lineal), ésta puede aplicarse para predecir el rendimiento de otros evaluados que provengan de una población equivalente a aquella con la cual fue realizado el estudio.

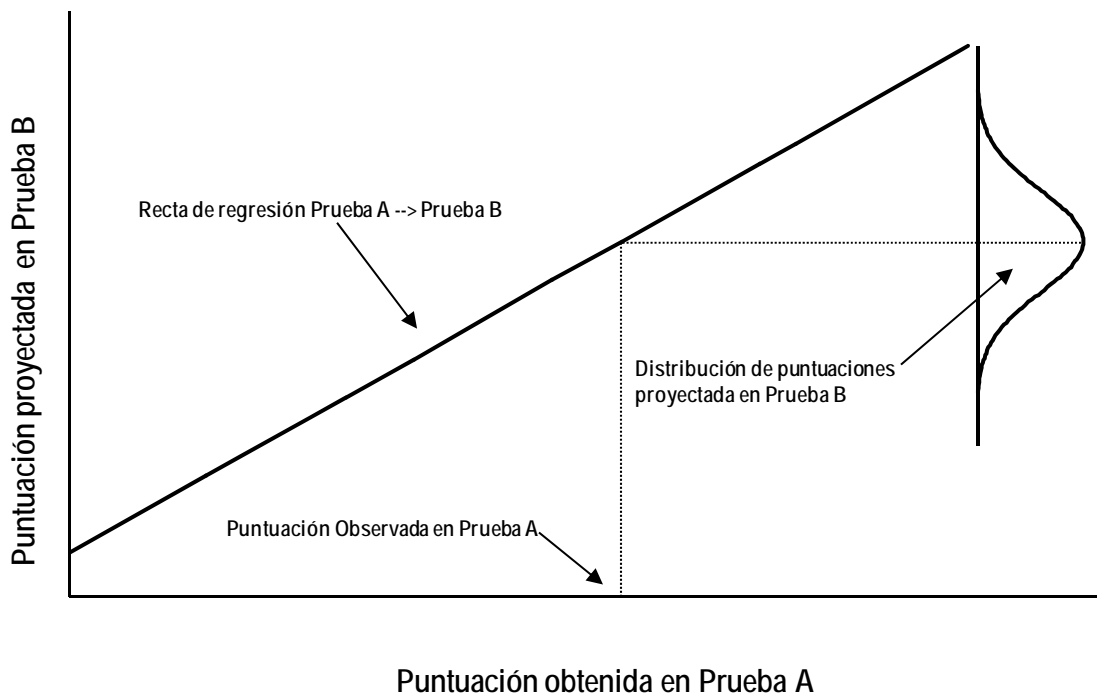
La *predicción de puntuaciones individuales* siempre es unidireccional y tiene un margen de error que viene dado por la fuerza de la asociación entre las variables y el respectivo coeficiente de determinación de la regresión, simple o múltiple (R^2), que permite cuantificar la proporción de varianza explicada por los predictores. Otro punto importante, es que la eficacia de predicción está sujeta al cumplimiento de los supuestos del modelo de regresión, pudiendo ésta ser de tipo lineal, simple o múltiple, aunque también es admisible trabajar con regresión no lineal o con regresión jerárquica (también llamada regresión multinivel o modelos HLM) si los predictores así lo ameritan.

Finalmente, hay que advertir que los *enlaces por predicción* suelen ser vulnerables a dos problemas metodológicos que requieren de un cuidadoso control. Primero, como típicamente la muestra final está sesgada selectivamente "hacia arriba" (sólo los alumnos que obtuvieron altos rendimientos en la prueba de selección ingresan efectivamente a la Universidad) la *predicción* puede parecer menos robusta de lo que realmente es (i.e. una correlación más baja de lo esperado), debido a problemas de restricción de rango (para una revisión ver, por ejemplo, Sackett & Yang, 2000). Segundo, es necesario determinar si la *predicción* está sujeta al llamado *sesgo externo* o *predicción diferencial* (Camilli, 2006), como se denomina al hecho de que la ecuación de regresión funcione diferencialmente para los examinados de un determinado grupo socioeconómico, por ejemplo. En el contexto de las metodologías de *enlace de puntuaciones*, la *predicción diferencial* atenta contra el requerimiento de *invarianza* de la función de *enlace* (ver sección 4.2).

2.2. Proyección de distribuciones

Un problema relacionado, aunque menos conocido, consiste en proyectar la distribución de puntuaciones en una prueba a partir de los resultados de otra. En este caso, a partir de los resultados obtenidos en una muestra de evaluados que haya respondido ambas pruebas, se intenta predecir la distribución condicional de la prueba B para cada puntuación observada de la prueba A, mediante ciertas variantes de los métodos de regresión lineal. Posteriormente, la función de *enlace* así obtenida se puede utilizar para proyectar la distribución de puntuaciones de la prueba B en otra muestra de estudiantes que sólo haya respondido la prueba A, pero que tenga características comparables a la muestra original (i.e. provenga de la misma población). Una representación gráfica de la idea de *proyección de distribuciones* se presenta en la Figura 2.

FIGURA 2. ENLACE POR PROYECCIÓN DE DISTRIBUCIONES



La característica más importante de la *proyección de distribuciones* es que cada puntuación de la prueba A no predice valores puntuales sino una distribución de puntuaciones en la prueba B; estas distribuciones son posteriormente agregadas para obtener un resultado total, como sucede con los *valores plausibles* utilizados en pruebas como NAEP o PISA. De hecho, las principales aplicaciones de la proyección pueden encontrarse en los esfuerzos para enlazar pruebas que utilizan esa metodología en mediciones a nivel agregado. Un ejemplo emblemático es el estudio realizado por Pashley y Phillips (1993) para proyectar las puntuaciones del *Internacional Assessment of Educational Progress* (IAEP) en la escala de *Nacional Assessment o Educational Progress* (NAEP) de Estados Unidos.

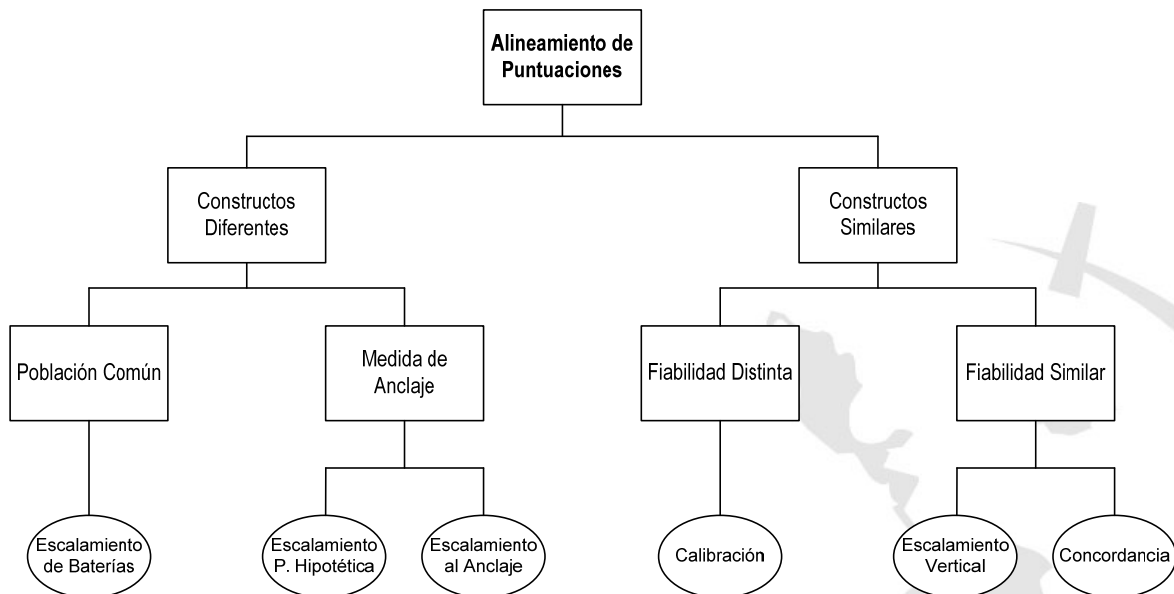
Cabe notar que un área de investigación potencialmente interesante y, hasta donde sabemos, inexplorada en Ibero América, es el enlace entre las mediciones nacionales de aprendizaje de cada país y los resultados obtenidos en las pruebas internacionales en las que se participe. El método de *proyección de distribuciones* es una alternativa metodológica en el caso que la prueba internacional entregue resultados basados en valores plausibles (e.g. PISA).

3. ALINEAMIENTO DE PUNTUACIONES

Bajo el concepto de alineamiento de puntuaciones [del inglés *score aligning*] se agrupan un conjunto de métodos que comparten entre sí el propósito de transformar los resultados de mediciones distintas a una escala común, para obtener así puntuaciones comparables. Mientras la meta de la predicción era lograr la predicción más exacta posible, el objetivo del alineamiento es lograr *puntuaciones comparables*. Ejemplos típicos de metodologías de *enlace* por *alineamiento* son los estudios de *Concordancia* de puntuaciones y la construcción de *Escalas Verticales*, que se explicarán más adelante.

El desarrollo de los métodos de *alineamiento*, iniciado a principios del siglo XX, ha sido bastante desordenado y no sigue una secuencia necesariamente lógica. En su taxonomía de métodos de *enlace* Holland y Dorans (2006) intentan clasificarlos atendiendo a cuestiones tales como el propósito de la evaluación, el diseño empleado y las características psicométricas de las pruebas. Para los fines de esta revisión no haremos una presentación exhaustiva sino que intentaremos describir brevemente los métodos más utilizados, de acuerdo a la organización conceptual presentada en la Figura 3, adaptada de Holland (2007). Como puede observarse, la primera distinción importante es si se pretenden *alinear* pruebas que evalúan constructos similares o distintos. Las restantes distinciones tienen que ver con el diseño de recogida de información y con las características técnicas de la pruebas.

FIGURA 3: TIPOS DE ALINEAMIENTO ENTRE PUNTUACIONES



3.1. Escalamiento de Baterías: Constructos diferentes evaluados en la misma población de examinados

Quando dos o más pruebas que miden constructos diferentes son administradas a la misma población de evaluados, los resultados de cada una de ellas pueden ser transformadas de manera que tengan una distribución común para esa población, por ejemplo, una distribución normalizada. En otras palabras, ambas pruebas son transformadas a una escala de referencia con características distribucionales conocidas.

La denominación *Escalamiento de Baterías* para este tipo de situaciones puede encontrarse en Kolen (2004), aunque tiene sus orígenes en los trabajos de Kelley (1914, 1923), Flanagan (1951) y Angoff (1971).

Sin entrar en los pormenores estadísticos, un ejemplo de este tipo de alineamiento es el estudio para re-centrar las escalas del SAT, en Estados Unidos (Dorans, 2002), uno de cuyos objetivos era que las escalas Verbal (SAT-V) y Matemática (SAT-M) tuvieran la misma distribución en la población de referencia, definida como la muestra de estudiantes evaluada en 1990. Gracias a este procedimiento un estudiante que obtiene una puntuación más alta en el SAT-V que en el SAT-M puede concluir, con toda propiedad, que su rendimiento fue mejor en el área Verbal que en el área Matemática de la prueba. En este ejemplo, queda muy claro como el Escalamiento de Baterías es una técnica para conseguir escalas con una distribución comparable (¡y no puntuaciones equivalentes!) en pruebas que evalúan constructos distintos en una población común de examinados. No hace falta decir que el enlace entre ambas pruebas no implica que las puntuaciones puedan utilizarse en forma intercambiable y que el error de este tipo de alineamiento es conceptualmente mayor que en los métodos que alinean pruebas dirigidas a evaluar constructos similares. Finalmente, es obvio que el error del alineamiento, en este caso, es inversamente proporcional a la magnitud de la correlación entre ambas pruebas.

3.2. Escalamiento mediante anclajes: Diferentes constructos, evaluados en poblaciones distintas

A veces es necesario alinear pruebas que miden constructos diferentes pero que no han sido respondidas por miembros de la misma población. En este caso, una solución es utilizar una *medida de anclaje* [del inglés *anchor measure*], es decir una tercera medición que sí ha sido administrada a miembros de ambas poblaciones. Esta *medida de anclaje* es utilizada como un puente para *enlazar* las puntuaciones originales.

Un ejemplo sería, por ejemplo, *enlazar* dos pruebas de idiomas, una de Francés y otra de Inglés, aplicadas a estudiantes de enseñanza secundaria, para establecer una métrica comparable. De este modo podríamos expresar en una misma escala ambos resultados y comparar el rendimiento de los estudiantes en una y otra prueba. El problema, es que normalmente cada estudiante responderá sólo aquella prueba que corresponda al idioma que cursa. Esto impide tener una muestra de examinados que hayan rendido ambas pruebas, lo cual imposibilita utilizar un *Escalamiento de Baterías*. La alternativa es emplear como *medida de anclaje* el resultado de una o varias pruebas de otras asignaturas que hayan sido administradas a ambos grupos de evaluados (o incluso variables sociodemográficas de interés) y que correlacionen con el resultado de las pruebas de idiomas. De este modo, es posible *enlazarlas* en una escala con una distribución común. En este sentido, el *escalamiento mediante anclajes* puede concebirse como una aproximación al *escalamiento de baterías* en caso de no contar con muestras equivalentes (i.e. pertenecientes a la misma población).

Respecto a los procedimientos estadísticos utilizados es posible diferenciar dos aproximaciones. El *Escalamiento a una Población Hipotética* es similar a la *Proyección* (sección 2.2) en tanto proyecta las distribuciones que corresponderían a cada prueba a partir de los resultados obtenidos en la *medida de anclaje*.

Algo más común es el *Escalamiento al Anclaje*, que algunos autores han denominado *moderación estadística* (McGaw, 1977; Mislevy, 1992; Keeves, 1998). Por ejemplo, Linn (1993) describe un caso en el cual este método fue utilizado para lograr la comparabilidad de puntuaciones en varias pruebas distintas que fueron aplicadas por algunas escuelas de un distrito escolar. El problema era cómo comparar

resultados que se habían obtenido con instrumentos diferentes, y que no evaluaban el mismo constructo. Utilizando como *medida de anclaje* una prueba estandarizada que se administró a todos los estudiantes de ese distrito, se consiguió poner en una escala comparable los resultados de las pruebas específicas de cada escuela.

Aunque no conocemos trabajos de este tipo de Latinoamérica, creemos que estudios de esta naturaleza serían muy interesantes de replicar para, por ejemplo, lograr resultados comparables en aquellos países en que exista alguna prueba estandarizada nacional (o regional) de logro escolar y se desee obtener una escala común para las pruebas aplicadas por agencias privadas bajo el contrato de municipios o redes de colegios. Nótese que en este caso, estamos hablando de enlazar las pruebas específicas entre sí a través de una prueba nacional. Si el objetivo fuera enlazar una prueba específica con la prueba estandarizada nacional, podríamos recurrir a una *Tabla de Concordancia* (ver apartado 3.4), siempre y cuando se cumplieran los requerimientos para ese método.

3.3. Calibración: El mismo constructo evaluado en la misma población con pruebas de diferente fiabilidad

El término *calibración* ha sido utilizado con varias acepciones en la literatura psicométrica. Por ejemplo, Angoff (1971) lo emplea para referirse al tipo de enlace que actualmente se conoce como *Escalamiento Vertical* (ver apartado 3.5), mientras la mayoría de los especialistas en Teoría de Respuesta al Ítem (TRI) llaman "calibrar" al proceso de obtener los parámetros de los ítems analizados bajo ese enfoque psicométrico. En la taxonomía de Holland y Dorans (2006) *calibración* describe un tipo muy específico de alineamiento entre dos pruebas, que miden el mismo constructo y tienen niveles de dificultad similares, pero que difieren en fiabilidad³.

Habitualmente, en pruebas que miden el mismo constructo con niveles equivalentes de dificultad, la diferencia de fiabilidades es provocada por diferencias en longitud (número de ítems), por lo que no es sorprendente que el caso más típico de *calibración* sea el alineamiento de los resultados de una prueba con una versión abreviada de la misma.

Aunque es poco común en medición de rendimiento educativo, un ejemplo razonable sería el caso de una prueba de competencias o habilidades intelectuales de 100 ítems, para la cual se construye una forma abreviada de, por ejemplo, 20 ítems para servir como medida de tamizaje [screening]. En este caso, la *calibración* consistiría en construir una escala que permitiera poner el resultado basado en 20 ítems en la misma escala de la prueba de 100 ítems. Un problema práctico es que la *calibración*, por sí misma, no puede compensar la inevitable pérdida de fiabilidad y de validez causada por la reducción de la prueba original.

3.4. Concordancia: Constructos similares evaluados con pruebas de igual fiabilidad y dificultad en la misma población

Una consecuencia imprevista y cada vez más frecuente de la proliferación de mediciones estandarizadas es que muchas veces un mismo estudiante responde a varias pruebas destinadas a propósitos similares. Por ejemplo, un postulante podría examinarse en dos universidades distintas, cada una de las cuales

³ Este es otro caso en que las preferencias lingüísticas varían a ambos lados del Atlántico. Mientras el término anglosajón *reliability* ha sido traducido en España como *fiabilidad*, en Latinoamérica se utiliza preferentemente la expresión *confiabilidad*. En este artículo hemos optado por la versión ibérica dado que, en nuestro parecer, refleja con mayor precisión el significado del concepto y remite a un uso técnico del mismo, evitando las malas interpretaciones que surgen cuando se discuten cuestiones científicas con palabras de uso cotidiano.

aplica sus propias pruebas de selección. ¿Es posible establecer alguna equivalencia entre ambas pruebas? ¿Cómo comparar su rendimiento relativo y saber en cuál obtuvo un mejor desempeño?

En casos como estos, estamos frente a dos pruebas diseñadas para evaluar constructos similares, que muy probablemente fueron construidas siguiendo especificaciones técnicas distintas (a veces, sólo levemente distintas). Sin embargo, fuera de algunas diferencias en los subfactores evaluados o en los marcos conceptuales que las fundamentan, ambas pruebas pueden tener muchas similitudes técnicas, como por ejemplo, alta fiabilidad o número equivalente de ítems. Si además están destinadas al mismo propósito y son rendidas por miembros de la misma población, es posible *enlazar* sus puntuaciones a través de un *alineamiento por concordancia*. El resultado añade valor a ambas pruebas, al permitir que sus resultados se expresen en una métrica común (¡pero no equivalente!). Un ejemplo clásico de este tipo de problemas y su solución, son los estudios de *concordancia* entre la pruebas SAT y ACT en Estados Unidos (Dorans, 1999).

El producto final de este tipo de *alineamiento* suele ser una *Tabla de Concordancia*, como las que ofrece el College Board en su sitio web (www.collegeboard.com) para demostrar las correspondencias entre las puntuaciones de las pruebas SAT y ACT. Estas Tablas permiten a cualquier estudiante que se haya examinado en ambas pruebas comparar sus resultados en ellas. Por otro lado, aquellos estudiantes que sólo rindieron una de las pruebas pueden formarse una idea aproximada de cuál habría sido su rendimiento en la otra.

Pese a su evidente utilidad los estudios de *concordancia* no están exentos de crítica, pues resulta muy tentador usar sus resultados como si las pruebas fueran equivalentes y sus puntuaciones *intercambiables* (Pommerich, 2007). Como veremos en la sección 3 de este artículo, sólo los procedimientos de *equiparación* pueden garantizar (y sólo bajo el cumplimiento de supuestos bastante estrictos) que los resultados de una prueba sustituyan a los resultados de otra.

Por otro lado, debemos reconocer que la actual abundancia de tests y pruebas estandarizadas amerita ampliamente la realización de estudios serios de *concordancia* entre mediciones, para establecer hasta qué grado son comparables los resultados de pruebas distintas.

En ese sentido, además de la comparabilidad de resultados entre pruebas de selección universitaria administradas por distintas instituciones, otra área que se beneficiaría de estudios de *concordancia* es el alineamiento entre pruebas regionales y pruebas estandarizadas a nivel nacional. ¿Es posible establecer algún tipo de concordancia entre los resultados de la prueba nacional y las pruebas regionales? ¿Cómo evaluar el nivel de alineamiento entre ambas?

Un problema del mismo tipo, que suponemos se incrementará con los años, es la necesidad de establecer *alineamientos* entre las mediciones internacionales de rendimiento educativo (i.e. PISA, TIMSS, SERCE) y las respectivas mediciones nacionales que cada país haya desarrollado. En nuestra experiencia, una pregunta que se formula cada vez con mayor insistencia a los especialistas es ¿A qué puntuación en la prueba internacional X equivale el resultado que obtuvo este año la escuela "A" en nuestra prueba nacional? Cabe observar que en este último caso, la elección entre un estudio de *concordancia* o de *proyección* estaría determinada por el nivel de similitud conceptual entre los constructos evaluados por ambas pruebas y por la metodología y nivel del análisis de los datos (e.g. puntuaciones individuales versus puntuaciones agregadas, mediante valores plausibles, por ejemplo).

Los lectores interesados en profundizar en metodologías de *concordancia* se beneficiarán de la lectura del

capítulo de Pommerich (2007), que presenta una buena discusión sobre las ventajas y limitaciones de las técnicas disponibles. Otra fuente autorizada es el número especial de la revista *Applied Psychological Measurement*, editado por Pommerich y Dorans (2004). Para familiarizarse con los procedimientos estadísticos necesarios, lo mejor es revisar el texto de Kolen y Brennan (2004) y poner especial atención a las técnicas equipercenales, que se utilizan tanto en estudios de *equiparación* como de *concordancia*.

3.5. Escalamiento Vertical: Constructos similares evaluados con pruebas de igual fiabilidad pero distinta dificultad, en población diferentes

En parte debido a la legislación *No Child Left Behind* (NCLB) de Estados Unidos y su consecuente impulso a los modelos de crecimiento [*growth models*], y en parte gracias a la creciente demanda universal por incrementar el monitoreo del rendimiento de los estudiantes durante toda su escolaridad, en los últimos años ha cobrado fuerza la idea de alinear los resultados de las pruebas de rendimiento de distintos grados escolares, a través de estudios de *Escalamiento Vertical*.

Para contextualizar el problema, recordemos que muchos países aplican pruebas estandarizadas a nivel nacional para evaluar el rendimiento de sus estudiantes al término del año escolar, en distintos grados (por ejemplo, en 2°, 4°, 6° y 8° de primaria). Técnicamente, esas pruebas evalúan constructos similares (no son exactamente los mismos, dado que los cambios curriculares entre un nivel y otro introducen, necesariamente, variaciones en la definición del constructo), con la misma fiabilidad y requerimientos técnicos, pero en poblaciones distintas (estudiantes de distinto grado) y con diferentes rangos de dificultad (las pruebas de los grados superiores son, necesariamente, más difíciles que las de grados inferiores).

En un típico sistema nacional de evaluación, el resultado de los estudiantes se expresa en escalas de desviación, *relativas al desempeño promedio observado en cada grado*. La manera habitual de construir estas escalas es la siguiente. En la primera ocasión en que se administra la prueba en un grado determinado (por ejemplo 2° de primaria), se transforma la media y desviación estándar de las puntuaciones obtenidas a una media y desviación arbitraria, que definirá la métrica de la escala de evaluación para ese grado escolar. Por ejemplo, muchas pruebas estandarizadas utilizan escalas definidas por una media y desviación estándar de (250,50) puntos o de (500, 100) puntos, respectivamente. Supongamos que un sistema de evaluación elige una escala con media de 250 puntos para definir la métrica de sus pruebas en todos los grados evaluados.

En los años posteriores y bajo condiciones de *equiparación* apropiadas (ver sección 3), los resultados de cada grado representarán la distancia relativa respecto al promedio del año de origen de la escala. Por ejemplo, si el primer año una escuela obtiene un promedio de 270 puntos, puede concluir que se encuentra a 20 puntos por sobre el promedio de rendimiento a nivel nacional. Si al año siguiente obtiene 280 puntos, puede concluir con toda certeza que su promedio creció en 10 puntos y que ahora está a 30 puntos de distancia del promedio de origen (Naturalmente, si ese año el promedio nacional también mejoró y alcanzó 260 puntos, la escuela habrá progresado, pero en términos relativos su rendimiento no será mejor al del año anterior). De este modo, para cada grado evaluado, las escalas de desviación permiten medir el progreso anual de las escuelas respecto al rendimiento promedio de la población en el año en que comenzó el sistema de medición.

Ahora bien, un problema con este tipo de escalas es que los resultados de cada grado no son comparables entre sí, porque están anclados a su propio año de origen. Por lo mismo, su capacidad para monitorear el progreso individual de los estudiantes resulta muy limitada. Por ejemplo, si un estudiante

obtuvo 240 puntos en 2° de primaria y dos años más tarde logró 250 puntos en 4° grado, eso no significa que su rendimiento mejoró en 10 puntos, sino simplemente que en 2° se encontraba bajo el promedio de los otros estudiantes y en 4° tiene un rendimiento equivalente al promedio de los otros niños de su grado.

En otras palabras, la escala de resultados para cada grado no permite cuantificar el progreso relativo cada estudiante entre distintos años de escolaridad. El problema, por lo tanto, es: ¿Cómo *alinear* los resultados de las pruebas de cada grado, para que permitan mostrar el progreso o crecimiento relativo del rendimiento entre años de escolaridad? ¿Es posible *enlazar* las pruebas específicas para cada grado y transformarlas todas a una escala continua de crecimiento o progreso estudiantil?

Frente a este tipo de problema, la solución es llevar a cabo un tipo particular de *alineamiento*, denominado por *Escalamiento Vertical*, cuyo propósito es poner los resultados de cada uno de los grados en una escala común. Esta es llamada *Escala Vertical*⁴ y permite monitorear el rendimiento individual de cada alumno a lo largo del tiempo, idealmente a través de todo el ciclo escolar. De hecho, buena parte del impulso que han recibido los métodos de *Escalamiento Vertical* se justifican en la promesa de medir apropiadamente el progreso o crecimiento [*growth*] de los estudiantes, lo que a su vez permite contextualizar mejor los estudios de valor añadido [*value added*] tan en boga últimamente.

A diferencia de las escalas en que se reportan los resultados para grado (llamadas, en este contexto, *Escalas Intragrado* o *Escalas Horizontales*), las *Escalas Verticales* permiten alinear el rendimiento de todos los grados en una única escala global (por ejemplo, de 0 a 100 puntos) y así cuantificar el monto de aprendizaje que los estudiantes *ganan* entre un curso y el siguiente. De este modo, cumplen a lo menos con tres propósitos muy atractivos. Primero, permiten a los padres y profesores monitorear el progreso individual de sus estudiantes y emprender medidas remediales si es necesario. Segundo, ofrecen a los investigadores información muy valiosa para modelar los factores que determinan el crecimiento o progreso educativo de los alumnos [*growth models*]. Tercero, esos mismos datos, una vez procesados, pueden emplearse para cuantificar el valor añadido [*value added*] de los profesores o los colegios sobre el aprendizaje de los estudiantes y así retroalimentar la toma de decisiones y el diseño oportuno y eficiente de políticas educativas.

Naturalmente, el desarrollo de una *Escala Vertical* es bastante complejo técnicamente (mucho más que una *equiparación* tradicional) y requiere de un cuidadoso proceso de planificación. A partir del estudio de fuentes especializadas, de la revisión de *Escalas Verticales* construidas en Estados Unidos y de nuestra propia experiencia, creemos que hay seis puntos críticos a tener en cuenta en el desarrollo de una *Escala Vertical*.

El primero es decidir el nivel de análisis de los datos, esto es, reporte individual (para cada estudiante), reporte a nivel agregado (para investigación sobre crecimiento escolar, análisis del valor añadido de los profesores y escuelas, o para el diseño y evaluación de impacto de políticas educativas, por ejemplo) o ambos. La alternativa elegida condicionará directamente el diseño y los métodos de análisis psicométrico y puntuación e, indirectamente, el error en la escala de reporte de los resultados. Naturalmente, estos usos no son necesariamente contradictorios, pero su elección pasa por decisiones de política educativa y no por cuestiones técnicas, necesariamente.

⁴ A veces se utiliza la denominación *Escala Evolutiva* o *Escala de Desarrollo* [*Developmental Scale*]. Un problema con esos términos es que también se utilizan (en inglés y en castellano) para referir a pruebas de habilidades o de inteligencia para niños pequeños.

El segundo punto es que la validez, fiabilidad y precisión de una *Escala Vertical* es inversamente proporcional a la distancia entre los grados que se pretende alinear. La situación ideal es alinear grados adyacentes (5°, 6° y 7°, por ejemplo) o, en su defecto, diseñar un mecanismo para conseguir datos de calidad en grados consecutivos, idealmente con no más de 2 años de diferencia (por ejemplo, utilizar una medición en 6° grado como puente entre pruebas aplicadas en 4° y 8° grado). Si ello no es posible, resulta necesario aumentar el número y frecuencia de las mediciones, con los costos que ello supone (Haertel, 1991; Huynh & Schneider, 2004).

Un tercer elemento a considerar es que las *Escalas Verticales* son muy sensibles a la dimensionalidad de las pruebas, así que sólo son recomendables para aquellas áreas de aprendizaje como Lectura o Matemáticas (con alguna reserva) en que la habilidad subyacente no sufre modificaciones drásticas a lo largo de la escolaridad. El mismo motivo complica el desarrollo de Escalas Verticales en Ciencias (Lissitz & Huynh, 2003), donde el constructo subyacente aumenta progresivamente su especificidad a través del ciclo escolar, comenzando como "ciencias" en forma global, para llegar a convertirse en "física", "química" y "biología" como componentes diferenciados.

El cuarto punto es que el *Escalamiento Vertical* requiere adoptar previamente una *definición de crecimiento* que opere como marco conceptual para el diseño e interpretación de la escala. Siguiendo a Kolen y Brennan (2004) hay dos enfoques predominantes. En la definición de *dominio de crecimiento* se asume que la habilidad evaluada es un continuo de contenidos y habilidades, que se expresa en distintos niveles de complejidad en las respectivas pruebas para cada grado. Bajo esta definición, la construcción de la *Escala Vertical* requiere de una *prueba de escalamiento* que se aplica a todos los estudiantes de los cursos que se desean alinear (por ejemplo 5° a 8° grado) y sirve para enlazar posteriormente las respectivas *Escalas Horizontales* de cada grado. Por el contrario, en la *definición grado a grado*, se asume que hay contenidos y habilidades propias de cada grado evaluado y el progreso escolar se entiende como el logro sucesivo de las habilidades esperadas para cada grado. Si se adopta esta definición, el diseño de la *Escala Vertical* utilizará alguna variación de los diseños generales de enlace que se emplean para equiparar puntuaciones (Kolen & Brennan, 2004).

Por supuesto, además de su impacto en el diseño, el tipo de definición de crecimiento condiciona las interpretaciones posibles de los resultados y debe estar alineada con los marcos conceptuales que guían el desarrollo de las respectivas pruebas para cada grado. En este sentido, es necesario enfatizar que en toda *Escala Vertical* subyace una definición de crecimiento. A veces, esta definición se encuentra implícita en las decisiones que se toman para el diseño de la escala. Lo óptimo, sin embargo, es que la definición de crecimiento sea explícita al inicio del proceso, de manera tal que sirva de guía para alinear el diseño, análisis, reporte e interpretación de los resultados en forma coherente y válida.

Continuando con la idea, un quinto punto a considerar con detenimiento es el impacto de las decisiones psicométricas sobre los resultados de la Escala. Por ejemplo, cuestiones tales como el tipo de diseño utilizado, el número de ítems y la calidad de los mismos, el modelo de análisis psicométrico para el análisis y la puntuación o hasta el software utilizado pueden repercutir, con mayor o menor fuerza, sobre la validez de los resultados o la magnitud del avance educativo mostrado por los estudiantes evaluados por la *Escala Vertical* (Tong y Kolen, 2007).

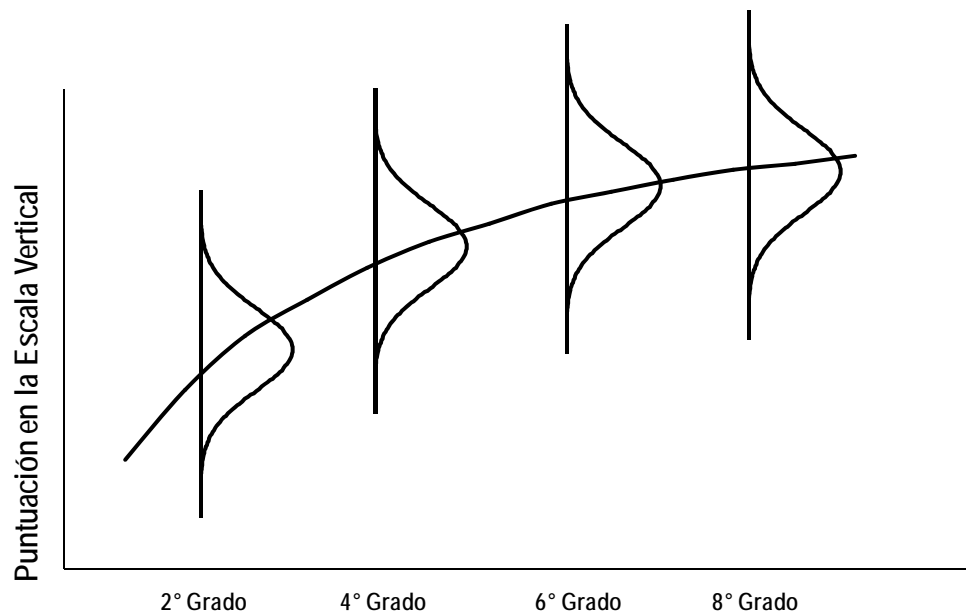
Finalmente, un último punto relevante es la articulación de la *Escala Vertical* con las respectivas *Escalas Horizontales* y con los estándares de aprendizaje, si éstos existen. La alternativa más recomendada es mantener ambas escalas separadas, esto es, reportar los resultados de las pruebas para cada grado en su

escala habitual (por ejemplo, con media 250 y desviación estándar 50) y definir una métrica independiente para la Escala Vertical (por ejemplo, una escala en el rango entre 0 a 100 puntos). La otra alternativa es situar todas las pruebas en una misma escala (por ejemplo, entre 200 y 800 puntos, donde la centena equivalga al grado evaluado y las decenas y unidades a la escala de rendimiento por grado, entre 0 y 100 puntos).

Cualquiera sea la opción que se escoja, debe cuidarse que la elección de la escala evite los resultados paradójicos, como que un desempeño muy deficiente en un grado superior se traduzca en la misma puntuación que un buen rendimiento en grados inferiores.

Este punto, que quizá es el más crítico para la finalidad última de una *Escala Vertical*, puede ilustrarse en la Figura 4. Allí se grafican los resultados hipotéticos obtenidos por estudiantes de 2°, 4°, 6° y 8° grado en una *Escala Vertical*. Asumiendo una distribución normal de las puntuaciones en la *Escala Horizontal* de cada grado, es obvio que los alumnos de más bajo rendimiento en 8° tendrán, en la *Escala Vertical*, en una puntuación similar a los alumnos de rendimiento medio en 2° grado. Por el contrario, los estudiantes de alto rendimiento en 2° grado podrían lograr puntuaciones similares a los examinados de rendimiento promedio en 8° grado. Un resultado de este tipo, que podría prestarse a muchos malentendidos, puede evitarse, o al menos minimizarse, si se pone el suficiente cuidado en la elección de la métrica para el reporte de los resultados.

FIGURA 4. EJEMPLO DE LA DISTRIBUCIÓN DE RENDIMIENTO POR GRADO EN UNA ESCALA VERTICAL



Hemos dedica algunas líneas adicionales al *Escalamiento Vertical* porque se trata de un tópico candente en la investigación y práctica psicométrica contemporáneas, que además creemos interesante de abordar en Latinoamérica. Para los lectores interesados en profundizar, una buena idea es partir documentándose sobre algunas de las *Escalas Verticales* desarrolladas en Estados Unidos, como por ejemplo, el *Stanford Achievement Test Series* (Harcourt Educational Measurement, 2003), los *Metropolitan Achievement Tests* (Harcourt Educational Measurement, 2002), el *TerraNova* (CTB/McGraw-Hill, 1996), o el *Florida*

Comprehensive Assessment Test (Hoffman, Wise, Thacker & Ford, 2003). Una vez comprendidos los alcances del problema, la mejor introducción técnica al *Escalamiento Vertical* es la sección 9.8 del texto de Kolen y Brennan (2004), que, por lo demás, es una referencia obligada sobre métodos de *enlace* de puntuaciones. Adicionalmente, en Young (2006) puede encontrarse una buena presentación de los diseños de recogida de información, y en Patz y Yao (2006) revisarse algunas de las técnicas de análisis disponibles para llegar a puerto. Para contar con una mirada crítica, el trabajo de Harris (2006) ofrece una buena revisión de algunos de los problemas prácticos que supone el construir una *Escala Vertical* (incluyendo la pregunta de si merece la pena el intento), mientras Briggs y Weeks (2009) demuestran empíricamente el impacto que pueden causar pequeñas decisiones sobre los resultados arrojados por la *Escala Vertical*.

4. EQUIPARACIÓN DE PUNTUACIONES

La *equiparación* de puntuaciones es el más importante y estadísticamente robusto de todos los tipos de *enlace*. Su propósito básico es poner en una misma escala los resultados de una o más pruebas que evalúan el mismo constructo, con la misma fiabilidad y bajo especificaciones técnicas similares, con el propósito de que éstos pueden ser utilizados en forma intercambiable.

Aunque hay muchas situaciones que pueden requerir de métodos de equiparación, las más habituales son cuatro: *equiparación* de cuadernillos de una misma prueba, *equiparación* de resultados entre años en una medición estandarizada, desarrollo de cuadernillos de prueba equivalentes para evaluar intervenciones educativas y *equiparación* de ítems de un Banco. A continuación describiremos brevemente cada uno de ellos.

4.1. Problemas típicos cuya solución requiere aplicar técnicas de equiparación

La finalidad más habitual de las técnicas de equiparación es poner en una misma métrica distintos *cuadernillos* de prueba. Este problema surge porque en muchos sistemas de medición se acostumbra subdividir las pruebas en varios cuadernillos diferentes (también llamados *formas* o *modelos* de pruebas) de modo que no todos los evaluados respondan exactamente los mismos ítems. Esta práctica puede cumplir varios propósitos. En las pruebas con resultados individuales de alto impacto para los respondientes (e.g. pruebas de selección universitaria, evaluación académica, certificación laboral o de acreditación de competencias) se construyen cuadernillos diferentes en un intento de controlar o al menos minimizar los efectos de la copia entre los estudiantes.

Por otro lado, en pruebas estandarizadas destinadas al monitoreo de resultados escolares a nivel agregado (e.g. NAEP) o en estudios internacionales de rendimiento escolar comparado (e.g. PISA, TIMSS, SERCE) un único cuadernillo de prueba es insuficiente para muestrear todos los aprendizajes esperados de los alumnos, lo que hace necesario distribuir el total de preguntas en varios cuadernillos diferentes, para optimizar los tiempos y costos de aplicación. De esta manera, cada alumno responde una parte del total de la prueba y la agregación de todos los alumnos permite reconstruir la prueba total. El resultado, en uno u otro caso, es que los alumnos respondan cuadernillos de prueba distintos, lo que obliga a desarrollar metodologías para asegurar que el resultado de la prueba no sea afectado por cuál haya contestado el estudiante.

Otro problema prototípico que requiere técnicas de *equiparación* se presenta en los sistemas nacionales de medición educativa que tienen entre sus propósitos el monitoreo del sistema escolar a lo largo de

varios años (e.g. NAEP de Estados Unidos). Diversas razones (cambios curriculares, seguridad de las pruebas, publicación de ítems, entre otros) aconsejan actualizar periódicamente el *Banco de Ítems* y obligan al sistema a desarrollar pruebas cuyo contenido (ítems) cambia con el paso de los años. En la práctica, esto exige desarrollar metodologías especiales para garantizar que los resultados sean perfectamente comparables entre un año y el que sigue, para así evaluar precisión cuánto ha mejorado el rendimiento (del país, región, estado, escuela) desde la última medición.

En otro contexto, la *equiparación* entre años es necesaria en caso que un mismo resultado requiera utilizarse en más de una ocasión. Un ejemplo típico de este problema afecta a las pruebas de selección universitaria, cuyos resultados habitualmente son válidos sólo para el año y el período de postulación en que fueron rendidas. Para extender la vigencia de sus resultados a varios períodos de postulación, es necesario garantizar, previamente, que los resultados de cada año son comparables (misma validez, misma fiabilidad y misma métrica) con los de otros años.

Otra variación de los problemas de *equiparación* entre ocasiones de medición se presenta en el contexto de estudios de impacto, experimentales o cuasi experimentales, con diseño pre-test y post-test, en que los que se mide el estado inicial de una variable antes de una intervención escolar y luego se mide el estado final de la misma variable, post intervención. A continuación, la magnitud del cambio relativo antes – después, contrastado con uno o más grupos de control, se utiliza para evaluar la efectividad de la intervención. Un problema bien documentado de estos estudios es que el uso de la misma medida en el pre test y el post test test, limita la validez de la investigación (Shadish, Cook & Campbell, 2002) debido al efecto de la *instrumentación* (sesgo de los resultados por memorización de preguntas, por ejemplo). Para evitar este problema se requiere contar con al menos dos versiones de la prueba que tengan un contenido diferente pero que al mismo tiempo sean completamente paralelas, es decir, midan con igual validez y fiabilidad. Para ello, la aplicación de procedimientos de equiparación es fundamental.

Finalmente, otra situación típica que demanda el uso de técnicas de equiparación es el desarrollo de *Bancos de Ítems*. La mayoría de los sistemas de evaluación estandarizada operan a través de un banco de preguntas que se actualiza periódicamente y a partir del cual se obtienen los ítems que integran las pruebas. Normalmente, estos *Bancos de Ítems* contienen ítems analizados con Teoría de Respuesta al Ítem y es necesario, antes de ingresarlos al *Banco*, que sus parámetros sean puestos en la misma métrica de los restantes ítems.

Es importante advertir que una misma prueba estandarizada puede enfrentar simultáneamente todas esas demandas y que la mayoría de los sistemas de medición requieren resolver problemas de equiparación bastante complejos, lo que exige aplicar varias técnicas de equiparación simultáneamente. En nuestra experiencia, no es exagerado afirmar que ningún sistema de evaluación estandarizado puede cumplir correctamente con su misión si no destina suficientes recursos a resolver los problemas de equiparación entre sus pruebas.

Nótese que los métodos de *equiparación* pueden clasificarse en varias subcategorías (e.g. Kolen & Brennan, 2004; Holland & Dorans, 2006), pero para comprenderlas es necesario manejar primero cuestiones técnicas referidas al diseño de recogida de datos y a los métodos estadísticos para *equiparar* las pruebas. Como ello escapa al alcance de este trabajo, obviaremos su presentación, aunque el lector interesado puede consultarlas en las fuentes antes señaladas.

4.2. Características definitorias de los enlaces por equiparación de puntuaciones

En el apartado anterior mencionamos algunas situaciones prototípicas que requieren aplicar técnicas de *equiparación*. Sin embargo aún si en esos casos utilizamos un diseño correcto y aplicamos las técnicas estadísticas apropiadas, no lograremos obtener una *equiparación* a no ser que se cumplan ciertas condiciones mínimas.

De hecho, en el desarrollo de la tipología de métodos de *enlace* que hemos venido presentando, la *equiparación* aparece como una forma especial de *alineamiento*, pero con supuestos mucho más estrictos. Estos supuestos se traducen en ocho requisitos fundamentales. Algunos son cualitativos, así que su evaluación es materia del juicio experto, lo que a veces complica decidir su cumplimiento. Otros son cuantitativos o pueden evaluarse parcialmente a través de procedimientos estadísticos. En uno u otro caso, es necesario insistir en que aún si aplicamos metodologías estadísticas altamente sofisticadas para equiparar dos pruebas, el resultados no será una equiparación válida si alguno de estos ocho supuestos no se cumple apropiadamente.

Los cuatro criterios básicos fueron sugeridos por Lord (1980):

4.2.1. El mismo constructo

Ambas pruebas deben medir el mismo constructo, definido de la misma manera. De lo contrario, el enlace no puede considerarse una *equiparación* legítima. La evaluación de este criterio es principalmente cualitativa, pero puede asistirse con técnicas como Modelos de Ecuaciones Estructurales para, por ejemplo, modelar la invarianza de una solución factorial. Dependiendo del diseño de equiparación, se podría evaluar invarianza total o parcial.

4.2.2. Equidad

Una vez que dos pruebas han sido equiparadas sus puntuaciones pueden utilizarse en forma intercambiable entre sí, de manera que sea indiferente para el resultado cuál de las dos responda el evaluado. Como puede apreciarse, este es requerimiento es teórico y, hasta cierto punto, una consecuencia de los demás. También es un prerequisite para la quinto condición.

4.2.3. Simetría de la transformación

La equiparación debe llevarse a cabo con una técnica que garantice la reversibilidad de los resultados. Esto es, que la función que permite transformar la puntuación de A en B, sea exactamente la inversa de la que permita transformar B en A. Esta condición es fácil de evaluar empíricamente. Por ejemplo, si de acuerdo a la función de equiparación 10 puntos de la prueba A equivalen a 8 puntos de la prueba B, entonces 8 puntos de la prueba B deben equivaler a 10 puntos de la prueba A, en caso que la función se aplique en sentido inverso. Volveremos a insistir en que el requerimiento de *simetría* es esencial en la definición de equiparación. Además, que impide que los enlaces por *predicción* puedan ser considerados una forma de equiparación y descarta el uso de técnicas de regresión tradicional del arsenal de procedimientos estadísticos para equiparar puntuaciones.

4.2.4. Invarianza de la transformación en subpoblaciones de interés

La transformación debe ser invariante para cualquier subpoblación de evaluados. Esto supone, en el fondo, que la equiparación no está sesgada (no produce resultados distintos) para diferentes grupos de interés (e.g. etnias, género, u otras). Otra forma de plantearlo, es que la *invarianza* exige que la función de transformación para equiparar ambas pruebas sea válida para cualquier examinado con independencia

de variables no relacionadas con el constructo que evalúa la prueba. En la práctica, este criterio también puede evaluarse empíricamente, simplemente subdividiendo la población en varios grupos, estimando las funciones de equiparación correspondientes, y finalmente verificando si éstas no difieren significativamente entre sí.

Además de estos cuatro criterios Holland y Dorans (2006) consideran un quinto que, desde su punto de vista, debería ocupar el segundo lugar en orden de importancia:

4.2.5. Igual fiabilidad

Las puntuaciones de ambas pruebas deben tener la misma fiabilidad. Aunque se trata de una condición implícita en el criterio 2 (*Equidad*) y por lo tanto resulta algo redundante, su importancia es que puede evaluarse fácilmente, simplemente comparando los respectivos coeficientes de fiabilidad. Holland y Dorans (2006) han demostrado que no basta con que ambas pruebas tengan igual fiabilidad, sino que además ésta debe elevada.

4.2.6. Las mismas inferencias

Las dos pruebas deben tener el mismo propósito y estar diseñadas para obtener el mismo tipo de conclusiones. Aunque este criterio está implícito en el primero (*mismo constructo*) es útil explicitar que el propósito debe ser el mismo si dos pruebas van a ser equiparadas (e.g. selección universitaria, monitoreo escolar).

4.2.7. La misma población

Ambas pruebas deben estar diseñadas para la misma población objetivo.

4.2.8. Las mismas propiedades psicométricas

Las dos pruebas deben estar construidas con la misma tabla de especificaciones, ser administradas en idénticas condiciones y ser equivalentes en todas sus propiedades métricas. Nuevamente, este criterio es relativamente redundante con los requisitos primero y quinto señalados anteriormente.

Si alguna de estas ocho condiciones no se cumple, el enlace entre las puntuaciones no podrá ser considerado una *equiparación* sino un *alineamiento*. Obviamente, en algunos casos se trata de condiciones dicotómicas, que además son fácilmente comprobables empíricamente (e.g. *misma fiabilidad* o *invarianza*). El problema es que en otros casos (e.g. *mismas propiedades psicométricas*) se trata más bien de un juicio cualitativo respecto a cuánto nos podemos acercar al cumplimiento del requisito.

Típicamente, la medición de temperatura ambiental, con termómetros graduados en escalas Celsius versus Fahrenheit, se utiliza como ejemplo idealizado de *equiparación*. Se trata del mismo constructo, la temperatura (condición 1), que es medido con dos instrumentos que sirven al mismo propósito (condición 6), que se aplican en problemas equivalentes (condición 7), y que son idénticos en todas sus características técnicas (condición 8), incluyendo su precisión (condición 5), y su validez en diferentes situaciones (condición 4). Se trata, en el fondo, de un mismo instrumento (termómetro) con versiones distintas, que difieren entre sí únicamente en la escala de medida. Por ello, ambas escalas de temperatura pueden *equipararse* mediante las transformaciones $F=(9/5)C+32$ o, en forma equivalente, $C=(5/9)(F-32)$. Como puede verse, la transformación es simétrica (condición 3) y, una vez conocida, permite que ambos instrumentos se utilicen en forma intercambiable, sin que ello afecte los resultados (condición 2).

A modo de síntesis, recordemos que todos los autores especializados en la materia insisten en que la

equiparación simplemente ajusta las diferencias en dificultad pero no en contenido de dos pruebas y, por lo tanto, equiparar no es sinónimo de igualar (Kolen y Brennan, 2004). Esta aseveración insiste, una vez más, en que los procedimientos estadísticos para equiparar las puntuaciones de las pruebas no garantizan por sí mismos que sea válido utilizar sus resultados en forma intercambiable, a no ser que primero se cumplan las condiciones antes señaladas.

5. ¿EQUIPARACIÓN, ALINEAMIENTO O PREDICCIÓN?

Aunque conceptualmente la distinción entre las tres metodologías de *enlace* es bastante clara, no siempre es fácil elegir cuál de ellas aplicar frente a un problema empírico. A continuación sugerimos cuatro criterios que pueden ayudar a tomar una decisión.

En primer lugar, la diferencia fundamental entre los tres tipos de *enlace* se desprende de la finalidad de cada uno de ellos. Esencialmente, el propósito de la *Equiparación* es *obtener puntuaciones intercambiables* entre medidas equivalentes de un mismo constructo. En cambio, el objetivo de los métodos de *Alineamiento* es transformar los resultados de medidas distintas a una *escala común*, para *obtener puntuaciones comparables*. Mucho menos ambiciosos, la meta que persiguen los estudios de *Predicción* es simplemente *predecir el resultado más probable en una medición*, para un individuo o grupo, a partir de otra información relevante (e.g. su puntuación en otra prueba). Por lo tanto, el punto de partida es preguntarse ¿Qué pretendemos lograr con nuestro *enlace*?

Una vez definido el objetivo y optado tentativamente por una metodología, la segunda consideración crítica es si se cumplen o no las condiciones técnicas que el método exige. Por ejemplo, en la sección 4.2 definimos la *equiparación* en función de ocho requerimientos básicos y comentamos que si alguno de ellos no se cumple, el *enlace* de las puntuaciones sólo alcanzaba el estatus de *alineamiento*. En el Cuadro 1 hemos organizado estos criterios en perspectiva comparada con las otras metodologías de *enlace*, en un intento de sintetizar las diferencias entre ellas.

CUADRO 1. DIFERENCIAS ENTRE EQUIPARACIÓN, ALINEAMIENTO Y PREDICCIÓN DE PUNTUACIONES

REQUERIMIENTOS TECNICOS	METODOLOGÍA DE ENLACE DE PUNTUACIONES		
	Equiparación	Alineamiento	Predicción
<i>Constructos</i>	Iguales	Similares, distintos o iguales	Similar o distintos
<i>Equidad</i>	Necesaria	No existe	No existe
<i>Simetría</i>	Necesaria	Necesaria	No existe
<i>Invarianza</i>	Necesaria	Necesaria	Deseable
<i>Fiabilidad</i>	Igual	Similar o igual	Similar (deseable)
<i>Inferencias</i>	Iguales	Similares o distintas	Distintas o similares
<i>Población objetivo</i>	Igual	Similar, igual o distinta	Similar o distinta
<i>Propiedades Psicométricas</i>	Iguales	Similares	Similares o distintas

Podemos ver que la *predicción* difiere de las otras dos metodologías en que no hay *simetría* en la función de enlace y, por lo tanto, tampoco es posible alcanzar el requerimiento de *equidad* entre las pruebas. Sin embargo, las otras condiciones pueden ser similares a las que exige un *alineamiento*. Este, por su parte, se diferencia fundamentalmente de la *equiparación* en que los *constructos* evaluados por ambas pruebas no requieren ser los mismos, lo cual supone que las características técnicas de las pruebas (*fiabilidad, inferencias, propiedades psicométricas*) tampoco sean iguales. De allí se desprende que el requerimiento de *equidad* es imposible de lograr.

Aplicando estos ocho criterios a la metodología o método específico que se ha escogido, es posible contar

con un segundo criterio para decidir si se trata de una opción recomendable.

En tercer lugar, es una buena idea evaluar empíricamente la magnitud de la asociación entre las puntuaciones de las pruebas que se desean *enlazar*. En este sentido, Dorans (2004) argumenta que se requieren correlaciones de al menos $r=.866$ para que la *concordancia* o la *equiparación* entre dos pruebas tenga un nivel de incertidumbre aceptable para fines prácticos. Con correlaciones de menor magnitud es cuestionable la exactitud de *enlace* obtenido, aunque ello dependerá de la fiabilidad de las pruebas y de la amplitud de la escala de puntuaciones. Sin embargo, es difícil defender algún tipo de *alineamiento* con correlaciones inferiores a $r=.70$ (es decir, un 49% de varianza común), así que en esos casos sería razonable optar por algún método de *predicción*. Una discusión interesante sobre el uso de coeficientes de correlación para orientar la decisión entre métodos de *enlace*, puede consultarse en un par de trabajos de Dorans (2000, 2004).

Por último, diferentes trabajos (Dorans, 2000; Dorans & Holland, 2000; Yin, Brennan & Kolen, 2004) han enfatizado la importancia de la *invarianza* del *enlace* como condición indispensable, tanto para la *equiparación* como para cualquier tipo de *alineamiento de puntuaciones*. En esos artículos pueden encontrarse algunos procedimientos estadísticos, de distinto nivel de complejidad, para evaluar el cumplimiento de la *invarianza*, así como algunas recomendaciones para su interpretación.

6. COMENTARIOS FINALES

El propósito de este trabajo ha sido presentar una síntesis de los distintos métodos disponibles para el *enlace* de puntuaciones, a partir de la taxonomía propuesta por Holland y Dorans (2006). Para terminar, creemos importante señalar los temas que no hemos abordado pero que resulta necesario conocer si se quiere profundizar en estas materias.

En concreto, omitimos el tratamiento de los aspectos metodológicos propiamente tales para cada tipo en *enlace*. Brevemente, cualquier metodología de *enlace de puntuaciones* tiene dos componentes básicos: un diseño de recogida de la información (de los examinados y de las pruebas, es decir, las características de las muestras utilizadas y las especificaciones técnicas y propiedades psicométricas de las pruebas aplicadas) y un conjunto de técnicas estadísticas para el análisis de los datos. Estas incluyen, a su vez, tres grupos de procedimientos que son aplicados en forma secuencial: técnicas para evaluar el cumplimiento de los supuestos de la metodología de *enlace*, técnicas para construir una función de transformación entre las pruebas (es decir, el *enlace* propiamente tal) y técnicas para evaluar el ajuste, calidad y precisión del *enlace* resultante.

En general, los diseños se dividen en dos grupos; aquellos que asumen que los examinados provienen de una *misma población* y aquellos que utilizan alguna *medida de anclaje* como en los ejemplos que ya se comentaron en la sección 3.2. A partir de esa distinción, la variedad de diseños utilizados en aplicaciones prácticas es inmensa. Kolen y Brennan (2004) ofrecen una descripción de los diseños más comunes para distintos métodos y Young (2006) sintetiza muy bien los diseños disponibles para el *Escalamiento Vertical*.

Por otro lado, las técnicas estadísticas están estrechamente ligadas al diseño y condicionadas por el modelo psicométrico con que se trabaje (Teoría Clásica de los Test o Teoría de Respuesta al Item). Generalizando, las técnicas clásicas construyen los *enlaces* a partir de las puntuaciones en el total de la prueba o en un *anclaje*, utilizando la distribución completa de puntuaciones (e.g. técnicas equipercenitiles)

o alguna medida sintética (e.g. *equiparación* lineal por la media). Las técnicas basadas en Teoría de Respuesta al Ítem construyen los enlaces a partir de todos o algunos de los ítems, mediante técnicas que transforman los parámetros de los ítems o que los intervienen directamente en la calibración. En cualquier caso, la presentación más completa y didáctica de las técnicas estadísticas también puede consultarse en el texto de Kolen y Brennan (2004), aunque Holland et al. (2007) ofrecen una síntesis recomendable para los lectores con mayor formación estadística.

En castellano, los textos introductorios a la psicometría de Martínez, Hernández y Hernández (2006) y de Muñiz (2001) contienen una buena presentación de los diseños y técnicas clásicas de *equiparación*, que son generalizables a los demás problemas de *enlace*. Finalmente, el capítulo de Navas (1996) ofrece un tratamiento completo de las metodologías de *equiparación*, incluyendo los diseños y las aproximaciones estadísticas clásicas y basadas en Teoría de Respuesta al Ítem.

REFERENCIAS BIBLIOGRÁFICAS

- American Psychological Association. (1985). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Angoff, W.H. (1971). Scales, norms and equivalent scores. En R.L. Thorndike (Ed.) *Educational Measurement (2dn Ed)*. Washington, DC: American Council on Education.
- Briggs, D.C. & Weeks, J. (2009). The impact of Vertical Scaling decisions on growth interpretations. *Educational Measurement: Issues and practice*, 28(4), 15-26.
- Camilli, G. (2006). Test Fairness. En R.L. Brennan (Ed.) *Educational Measurement (4th Ed)*. Wesport, CT: Praeger Publishers.
- CTB/McGraw-Hill. (1996). *TerraNova prepublication technical bulletin*. Monterray, CA: Author.
- Dorans, N.J. (1999). *Correspondences between ACT and SAT I scores*. College Board Report 99-1. New York. The College Board.
- Dorans, N.J. (2000). *Distinctions among classes of linkages*. College Board Research Note RN-11. New York. The College Board.
- Dorans, N.J. (2002). Recentering and realigning the SAT score distributions: How and why. *Journal of Educational Measurement*, 39, 59-84.
- Dorans, N.J. (2004). Equating, concordance and expectation. *Applied Psychological Measurement*, 28(4), 227-246.
- Dorans, N.J. & Holland, P.W. (2000). Population invariance and the equitability of tests: Basic theory and the linear case. *Journal of Educational Measurement*, 37, 281-306.
- Dorans, N.J., Pommerich, M. & Holland, P.W. (2007). *Linking and aligning scores and scales*. New York: Springer.
- Feuer, M.J., Holland, P.W., Green, B.F., Bertenthal, M.W. & Hemphill, F.C. (1999). *Uncommon measures: Equivalence and linkage among educational tests*. Report of the Committee on Equivalency and Linkage of Educational Tests, National Research Council. Washington, DC: National Academy Press.
- Flanagan, J.C. (1951). Units, scores and norms. En E.F. Linquist (Ed.) *Educational Measurement (1st Ed)*. Washington, DC: American Council on Education.

- Galton, F. (1888). Co-relations and their measurements, chiefly from anthropological data. *Proceedings of the Royal Society of London*, 45, 135-145.
- Haertel, E. (1991). *Report of TRP analyses of issues concerning within-age versus cross-age scales for the National Assessment of Educational Progress*. ERIC Clearinghouse Document Reproduction Service N° ED404367. Washington, DC: National Center for Education Statistics.
- Harcourt Educational Measurement. (2002). *Metropolitan Achievement Tests: Technical Manual (8th Ed)*. San Antonio, TX: Author.
- Harcourt Educational Measurement. (2003). *Stanford Achievement Tests Series: Spring technical data report (10th Ed)*. San Antonio, TX: Author.
- Harris, D.J. (2007). Practical issues in Vertical Scaling. En N.J. Dorans, M. Pommerich & P.W. Holland (Eds.) *Linking and aligning scores and scales*. New York: Springer.
- Hoffman, R.G., Wise, L.L., Thacker, A.A. & Ford, L.A. (2003). *Florida Comprehensive Assessment Tests: Technical Report on vertical scaling for reading and mathematics*. A HumRRO Report under subcontract to Harcourt Assessment, San Antonio, TX.
- Holland, P.W. (2007). A framework and history for score linking. En N.J. Dorans, M. Pommerich & P.W. Holland (Eds.) *Linking and aligning scores and scales*. New York: Springer.
- Holland, P. W. & Dorans, N.J. (2006). Linking and equating. En R.L. Brennan (Ed.) *Educational Measurement (4th Ed)*. Wesport, CT: Praeger Publishers.
- Holland, P. W., Dorans, N.J. & Petersen, N.S. (2007). Equating test scores. En C.R. Rao & S. Sinharay (Eds.) *Handbook of Statistics, Vol 26*. Amsterdam: Elsevier.
- Huynh, H. & Schneider, C. (2004). *Vertically moderated standards as an alternative to vertical scaling: Assumptions, practices and a odyssey through NAEP*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.
- Keeves, J. (1988). Scaling achievement test scores. En T. Husen & T.N. Postlethwaite (Eds.) *International Encyclopedia of Education*. Oxford: Pergamon.
- Kelley, T.L. (1914). Comparable measures. *Journal of Educational Psychology*, 5, 589-595.
- Kelley, T.L. (1923). *Statistical Methods*. New York: MacMillan.
- Kolen, M.J. (1988). Defining score scales in relation to measurement error. *Journal of Educational Measurement*, 25(2), 97-110.
- Kolen, M.J. (2004). Linking assessments: Concept and history. *Applied Psychological Measurement*, 28, 219-226.
- Kolen M.J. & Brennan, R.L. (2004). *Test equating, scaling and linking. Methods and practice (2th Ed)*. New York: Springer.
- Linn, R.L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83-102.
- Lissitz, R.W. & Huynh, H. (2003). Vertical equating for state assessment: Issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research and Evaluation*, 8(10).

- Lord, F.M. (1980). *Applications of Item Response Theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Martínez, M.R., Hernández, M.J. & Hernández, M.V. (2006). *Psicometría*. Madrid: Alianza Editorial.
- McGaw, B. (1977). The use of rescaled teacher assessments in the admission of student to tertiary study. *Australian Journal of Education*, 21, 209-225.
- Mislevy, R.J. (1992). *Linking Educational Assessments: Concepts, issues, methods and prospects* (Policy Information Report). Princeton, NJ: Educational Testing Service.
- Muñiz, J. (2001). *Teoría Clásica de los Test*. Madrid: Pirámide.
- Navas, M.J. (1996). Equiparación de Puntuaciones. En J. Muñiz (Ed.). *Psicometría*. Madrid: Universitas
- Pashley, P.J. & Phillips, G.W. (1993). *Toward world-class standards: A research study linking international and national assessments*. Princeton, NJ: Educational Testing Service.
- Patz, R.J. & Yao, L. (2007). Methods and models for Vertical Scaling. En N.J. Dorans, M. Pommerich & P.W. Holland (Eds.) *Linking and aligning scores and scales*. New York: Springer.
- Pommerich, M. (2007). Concordance: The Good, the Bad and the Ugly. En N.J. Dorans, M. Pommerich & P.W. Holland (Eds.) *Linking and aligning scores and scales*. New York: Springer.
- Pommerich, M. & Dorans, N.J. (Eds.), (2004). Concordance [Special Issue]. *Applied Psychological Measurement*, 28(4).
- Sackett, P.R. & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85, 112-118.
- Shadish, W., Cook, T. & Campbell, D. (2002). *Experimental and quasi-experimental design for generalized causal inference*. Boston: Houghton Mifflin.
- Sireci, S. G., & Zenisky, A. L. (2006). Innovative item formats in computer-based testing: In pursuit of improved construct representation. En S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Testing*. Mahwah, NJ: Lawrence Erlbaum.
- Tong, Y. & Kolen, M.J. (2007). Comparisons of methodologies and results in vertical scaling for educational achievement tests. *Applied Measurement in Education*, 20(2), 227-253.
- Yin, P., Brennan, R.L. & Kolen, M.J. (2004). Concordance between ACT and ITED scores from different populations. *Applied Psychological Measurement*, 28(4), 274-289.
- Young, M.J. (2006). Vertical Scales. En S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Testing*. Mahwah, NJ: Lawrence Erlbaum.

EVALUAR EL RENDIMIENTO INTERNO Y ACADÉMICO: UN DESAFÍO PARA LA MACRO Y LA MICRO POLÍTICA. LECCIONES A PARTIR DE UN ESTUDIO DE CASO

TO EVALUATE THE INTERNAL AND ACADEMIC PERFORMANCE: A CHALLENGE FOR THE MACRO AND MICRO POLICY . LESSONS DEPARTING FROM A CASE OF STUDY

Silvina Gvirtz y Ángela Inés Oría

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art7.pdf>

Fecha de recepción: 26 de marzo de 2010
Fecha de dictaminación: 26 de abril de 2010
Fecha de segundo envío: 27 de abril de 2010
Fecha de aceptación: 29 de abril de 2010

Este artículo explora las implicancias del rendimiento interno y académico sobre la política educativa. Se concentra principalmente en el problema de la repitencia concebida como la antesala del abandono y, por lo tanto, un asunto que reclama la atención de todos los actores del sistema comprometidos con la justicia educacional.

La repitencia es una estrategia pedagógica utilizada por maestros y escuelas, generalmente como acción correctiva del fracaso académico. Ocurre cuando un estudiante es retenido en el mismo año escolar en lugar de ser promovido al siguiente junto con sus compañeros de curso. En el plano macro-político, la repitencia es un indicador privilegiado del rendimiento interno del sistema educativo.

Un breve recorrido teórico vincula el rendimiento académico y la repitencia con el fracaso escolar y, en consecuencia, con el desafío acuciante de la justicia educacional. Sin embargo, los actores del sistema tienden a disociar estos términos. La repitencia es percibida como un problema de las escuelas. La justicia educacional como una función exclusiva del Estado. Las líneas intermedias de gobierno, instancias privilegiadas para incidir sobre la equidad (Gvirtz 2008: 15), mantienen muchas veces un papel pasivo frente a estos desafíos. Sin embargo, rendimiento académico, repitencia y justicia educacional se entrelazan. Atañen tanto al nivel macro como al micro político, demandando soluciones paralelas y combinadas.

En este artículo se observa el caso argentino y, en este contexto, 'Escuelas del Bicentenario', un Proyecto de articulación público-privado cuyo objetivo es colaborar con el Estado y los actores de distintos niveles del sistema en su función principalísima de brindar educación de calidad. Se comparten las estrategias de trabajo sobre la repitencia, los modos de evaluación académica y los logros obtenidos. Por último, se ofrecen algunas conclusiones.

1. EL SISTEMA EDUCATIVO ARGENTINO: DIAGNÓSTICO Y DESAFÍOS

Un diagnóstico del sistema educativo argentino señala la distribución desigual del bien educativo como uno de los principales desafíos de la política pública (Gvirtz 2005). En la Argentina, las escuelas son tanto más ricas en sus capitales físico, humano y social cuanto mayor es el nivel socioeconómico de sus alumnos. En otras palabras, las escuelas que atienden a la población más pobre del país son también las más vulnerables. Esto tiene efectos negativos acumulativos en la escolarización y calidad de los aprendizajes. En cambio, proveen un mejor servicio las escuelas que atienden a la población de quintiles de ingreso más altos. Esta 'discriminación educativa' (Braslavsky 1985) adquiere implicancias graves cuando se inserta en un contexto socio económico de creciente desigualdad y exclusión social.

Tal como refleja el cuadro 2, la asistencia a la escuela de niños entre 3 y 5 años varía según el nivel socio-económico.

CUADRO 1. TASA DE ASISTENCIA DE LA POBLACIÓN DE 3, 4 Y 5 AÑOS SEGÚN QUINTIL DE INGRESO PER CÁPITA FAMILIAR. TOTAL AGLOMERADOS URBANOS (%) 2005

	3 años	4 años	5 años
Quintil 1	13.6	40.8	84.4
Quintil 2	26	59.8	87.7
Quintil 3	33.1	69	93.4
Quintil 4	46.7	76.2	97
Quintil 5	63.9	92.2	96
TOTAL	28.5	59.9	89.1

Fuente: INDEC. Encuesta Permanente de Hogares, segundo semestre, 2005

Ya sea por falta de oferta o demanda de educación, asisten a la escuela solo un 13.6% de los niños de 3 años provenientes de familias del 1° Quintil de ingreso, contra un 63.9% de los niños del Quintil 5°. En sala de 4, más del doble de alumnos del Quintil 5° asiste al nivel inicial.

Los indicadores cobran relevancia a la luz de la investigación de Berlinski *et al* (2008), que analiza el efecto ulterior de la educación pre-primaria escolares de entre 7 y 15 años de edad. Los autores destacan las ventajas de la experiencia preescolar sobre el desempeño de los alumnos en edades tempranas y observan que dichos efectos aumentan a medida que los niños crecen. El efecto de la enseñanza preescolar sobre las perspectivas de culminación tanto de la primaria como la secundaria es significativamente positivo. Cae la tasa de repitencia en los primeros grados y se reduce la deserción escolar en los adolescentes.

El análisis de los indicadores de abandono y repitencia constatan este escenario de inequidad. Reflejan que el sistema educativo tiene escasa capacidad de retención y, en consecuencia, que no logra incluir a muchos niños y jóvenes en niveles educativos altos. El cuadro 2 sintetiza de manera elocuente la situación de desigualdad. El inicio tardío, el abandono y la repitencia en el nivel básico y medio aumentan en relación inversa a los quintiles de ingreso de la población.

CUADRO 2. INICIO TARDÍO, REPITENCIA Y ABANDONO (%) TOTAL DEL PAÍS

	Inicio Tardío	Repitencia primaria	Repitencia Secundaria	Abandono Secundaria
Quintil 1	5.2	24.8	32.2	30.3
Quintil 2	4.1	14.1	33.6	22.8
Quintil 3	3.2	6.5	27.1	18.3
Quintil 4	2.3	6.9	18.7	11.9
Quintil 5	3.0	4.7	12.4	6.7
TOTAL	4.0	14.4	24.8	18.4

Fuente: Siempre (2000) en (Llach 2006)¹

La expansión del sistema parece acarrear un proceso contradictorio de inclusión-expulsión de alumnos, afectando principalmente a aquellos de contextos sociales vulnerables.

Desde distintos ámbitos se intenta dar respuesta a este diagnóstico crítico. Programas compensatorios del Ministerio de Educación de la Nación asisten a números siempre crecientes de escuelas en condiciones de 'vulnerabilidad' y organizaciones de la sociedad civil despliegan redes de apoyo escolar, programas de becas y proyectos de intervención técnico-pedagógica. Las iniciativas particulares consiguen impactos positivos sobre algunas o varias escuelas, pero no alcanzan a expandir el impacto en gran escala.

Las políticas que emanan del Estado, en particular aquellas dirigidas a poblaciones en contextos de pobreza, son comúnmente llamadas políticas 'compensatorias'. Su acción principal consiste en dar apoyo a estudiantes a través de múltiples sistemas de becas: 'becas de retención', 'becas de inclusión', 'becas específicas para la inclusión y la retención'. Bajo el esquema de 'Inclusión' se despliegan el Programa Nacional de Becas Estudiantiles y el Programa Nacional de Inclusión Educativa (PNIE), que a su vez

¹ Siempre es el Sistema de Información, Evaluación y Monitoreo de Programas Sociales de Presidencia de la Nación. Lamentablemente, esta encuesta no se actualiza desde el año 2000.

contiene cuatro líneas adicionales de becas con características específicas: Todos a Estudiar, Volver a la Escuela, Inclusión de Chicos con Causa Judicial y PNIE Rural.²

Los Programas Compensatorios, de vasto alcance, se apoyan en un concepto 'tradicional' de justicia educativa. Actúan sobre la reparación del daño y realizan esfuerzos paralelos que muchas veces no consiguen (ni miden) resultados eficaces. En otras palabras, revierten el daño que ya aconteció, manteniendo la inequidad estructural. La repitencia en la Argentina no sólo es una constante, sino que ha ido en aumento en los últimos diez años.

CUADRO 3. EFICIENCIA INTERNA. TASA DE REPITENCIA (%) - TOTAL PAÍS.

	1996	1998	2001	2004	2006
EGB 1 y 2	5.8	6	6.1	6.4	6.1
EGB 3	8.9	8.3	7.6	10.3	12.8
Polimodal	5.6	5.3	4.8	7.5	8.1

Fuente: elaboración propia en base a DiNIECE – Ministerio de Educación Nacional

En 1996, el porcentaje de repitentes en los dos primeros ciclos de enseñanza general básica (EGB) fue del 5.8%. En 2006, ese valor asciende al 6.1%, alcanzando su pico en 2004, con un 6.4% de alumnos repitentes. En el nivel medio la situación es aún más grave. De un piso de 5.6% en 1996, la tasa de repitencia alcanza el 8.1% en 2006, verificando un salto pronunciado a partir del 2001.

A pesar de los distintos intentos de mejora, los diagnósticos señalan tendencias de mediano plazo poco alentadoras. Aún cuando algunos programas actuales del Ministerio dan pasos hacia adelante en la complejización de las políticas compensatorias³, continúan centrando sus esfuerzos en dar, en brindar compensaciones, recursos y atención prioritaria a sus destinatarios. Es loable y rinde frutos. Sin embargo, las intervenciones difícilmente incidan sobre el rendimiento interno y académico cuando los diseños no contemplen el monitoreo de resultados parciales y efectos finales. Desde el enfoque que concibe el acceso y permanencia en la escuela como un 'derecho' de la infancia, esto tiene implicancias graves. Cuando el derecho a la educación es vulnerado, se comete un acto de injusticia.

Se requiere redefinir conceptualmente a la 'política educativa'. Los desafíos educacionales no se resuelven solo en el nivel macro, sino a través de acciones articuladas y en simultaneo con el nivel micro-político.

A los actores que actúan en cada uno de estos niveles, les caben roles concretos en el cuidado del rendimiento interno, que es condición para la justicia educacional. Este es el quid de las políticas preventivas: aumentar la retención del sistema y reducir el volumen de políticas compensatorias.

El Proyecto 'Escuelas del Bicentenario' busca llevar a la práctica definiciones renovadas de la política y la justicia educacional. Se ofrece primero una breve introducción a sus objetivos y diseño institucional.

² La Dirección de Políticas Socioeducativas también ofrece apoyo a escuelas a través del 'Fondo Escolar', que financia proyectos institucionales elaborados por las escuelas y 'Textos de Estudio', que lleva a las escuelas libros y material didáctico en línea con los contenidos curriculares (Políticas Socioeducativas - URL). También se propone brindar apoyo y acompañamiento a las Jurisdicciones a través de 'Aportes Financieros a la Gestión Provincial' y 'Capacitación, asistencia técnica y material bibliográfico'. Sobre estas 2 líneas la información disponible en Internet y/u otros medios públicos es escasa o nula (Políticas Socioeducativas - URL).

³ Se destaca la experiencia del Programa Integral para la Igualdad Educativa (PIIE). 2003 – continúa.

Luego se describen las estrategias de trabajo sobre la repitencia, los modos de evaluación académica y los logros obtenidos. Por último, se ofrecen algunas conclusiones.

2. PROYECTO ESCUELAS DEL BICENTENARIO

'Escuelas del Bicentenario' es un proyecto que articula el sector público-estatal y privado con un doble objetivo:

- Colaborar con la mejora de escuelas primarias públicas que atienden a población vulnerable en todo el país.
- Contribuir con el diseño de políticas públicas, desarrollando dispositivos de mejora escolar sustentables y escalables.

El Proyecto no propone a las escuelas acciones *ad hoc*, sino que refuerza su funcionamiento esencial: Que los niños ingresen y se gradúen, aprendan, disfruten del conocimiento, puedan aplicarlo y continúen con éxito el nivel siguiente. Para eso acompaña a las escuelas integralmente, trabajando en el plano de la gestión institucional, la mejora académica, la salud escolar y la articulación con Institutos de Formación Docente de la zona.

Cada una de estas dimensiones se plasma en propuestas de intervención evaluables y hay distintas instancias donde se analizan los resultados del trabajo. Sin información sobre logros y desafíos, pierde buena parte del sentido el esfuerzo de implementación y construcción de alianzas entre sectores y niveles del sistema educativo.

El Proyecto reúne a representantes de distintas audiencias en sus órganos de dirección y desarrolla articulaciones con actores de distintos niveles en sus fases de implementación.⁴ Su diseño institucional es

⁴ Son parte de esta 'alianza': **Organismos Oficiales:** el Ministerio de Educación Nacional, que declara al Proyecto de Interés Nacional (Resolución N° 1011, 12 de Julio de 2007) y la Dirección Nacional de Políticas Socioeducativas; Ministerios de Educación de las Provincias de Chaco, Córdoba, Corrientes y Tucumán; el Consejo Provincial de Educación de Santa Cruz y la Dirección General de Cultura y Educación del Gobierno de la Provincia de Buenos Aires, que declara el Proyecto de Interés Educativo (Resolución N° 3243, 12 de Octubre de 2007); las Municipalidades de Campana, Ensenada, San Nicolás, Carlos Casares, Gobernador Virasoro, Barranqueras, Vilelas, Las Heras y Caleta Olivia. Participan asimismo Ministerios de Salud: Ministerio de Salud Pública de la Provincia del Chaco, Ministerio de Salud Pública de la Provincia de Corrientes, Ministerio de Salud de la Provincia de Tucumán - SIPROSA y Subsecretaría de Salud de Santa Cruz. Otras instituciones forman parte de la RED de Salud: Instituto 15 de Campana, Instituto Profesional de Educación Superior de Caleta Olivia, Universidad Nacional de Luján, Universidad Nacional de la Patagonia Austral, Escuela Municipal de Enfermería de Pilar, Instituto Campero, Instituto JIM y Universidad Nacional de Tucumán y PUEDES (Programa Universitario de Extensión y Desarrollo social). Por último, el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), que actualmente analiza la posibilidad de iniciar algunas líneas de investigación sobre el Proyecto. Dichos organismos cooperan con la viabilidad política, técnica y financiera del Proyecto, avalando la intervención, monitoreando su marcha y prestando recursos materiales y humanos para concretar eficazmente la intervención técnica. Algunos funcionarios integran el directorio del Proyecto, sirviendo como interlocutores claves para el análisis de la marcha del Proyecto y la discusión sobre sus posibilidades de escala a nuevas escuelas y territorios. **Empresas y Fundaciones:** Asociación Empresaria Argentina (AEA), Accenture, Aique, Allianz Group, Ángel Estrada, Avina, Banco Galicia, Banco Hipotecario, Bosques del Plata, Cámara de Comercio, Industria y Servicios de Carlos Casares, CIPPEC, Coca-Cola, Ejes de Comunicación, Deutsche Bank, Fundación Arcor, Fundación Bunge y Born, Fundación Hermanos Agustín y Enrique Rocca, Fundación Irsa, Fundación La Nación, Fundación Lúminis, Fundación MAPFRE, Fundación Mundo Sano, Fundación por Pilar, Fundación Vitoria Jean/Las Marías, Fundación YPF, Grupo Clarín, Grupo Los Grobo, Liberty Seguros, MBA-Lazard, Mercedes Benz, Muchnik, Azurralde, Jasper & Asoc., Pampa Energía, Price Water House, Pump, Roggio, San Miguel, Santillana, Telecom, Tenaris, Ternium Siderar, Tinta Fresca, TN & Platex, Tomás Hermanos y Zurich Las empresas colaboran financieramente, brindan asesoramiento académico-técnico y, en algunos casos, integran el directorio del Proyecto, comprometiendo tiempo, capacidad y esfuerzo en la definición de nuevas estrategias y el monitoreo de su marcha y resultados. **Organismos Intergubernamentales:** La Comisión Económica para América Latina (CEPAL) y el Fondo de Las Naciones Unidas Para la Infancia (UNICEF). Estos organismos apoyan la elaboración de informes y documentación del Proyecto, ofreciendo asimismo espacios para la discusión y el intercambio con especialistas de otros campos ligados a la política pública. El Proyecto tiene su sede en el Instituto Internacional de Planeamiento de la Educación (IPE/UNESCO) Sede Regional Buenos Aires y trabaja en asociación con la Escuela de Educación de la Universidad de San Andrés, que a través del Proyecto 'Escuelas del Futuro' (2000 - continúa), brinda un antecedente clave en esta materia.

en sí mismo el objeto de estudio de otro trabajo (Gvirtz y Oría, en prensa), donde cabe el análisis en detalle del modo específico en que este diseño aporta elementos que trascienden y a la vez potencian el contenido técnico de la mejora educativa. La alianza entre sectores aumenta el hábito mutuo de 'dar cuenta', 'tracciona' la producción de información, la identificación de logros y desafíos. Es además una iniciativa federal, en tanto constituye y capacita equipos técnicos *locales* para trabajar en las escuelas y distritos ubicados en distintas regiones del país.

Escuelas del Bicentenario lleva tres años y medio de funcionamiento. En este periodo se sumaron 129 escuelas, abarcando una población destinataria de 50.000 alumnos y 2000 actores del sistema, incluyendo supervisores, directores y maestros. Se elaboraron diagnósticos, se diseñaron y desarrollaron las propuestas de mejora escolar y se consolidaron equipos técnicos locales en cada Jurisdicción. Los logros finales esperados se organizan en tres dimensiones:

- Mejoras sustantivas en el rendimiento interno de las escuelas, medidas de acuerdo a variables tales como la tasa de graduación y promoción, y la disminución de tasas de repitencia, abandono y sobre-edad.
- Mejoras sustantivas en el rendimiento académico de los alumnos, medidas en el desempeño de los alumnos en las Áreas de Lengua, Matemática y Ciencias Naturales
- Mejoras sustantivas en el Área de Salud, medidas en el logro de articulaciones efectivas entre las escuelas y los centros de salud cercanos, mejoras sanitarias en las escuelas, registros del estado de salud de cada niño y realización de acciones de promoción y prevención en la escuela y la comunidad.

A medida que avanza la implementación del Proyecto, se advierten nuevas dimensiones de logros que no se subsumen a las categorías mencionadas. Dada su importancia, actualmente se elabora un sistema de evaluación amplio que busca ponderar el impacto del Proyecto en ordenes adicionales de la vida escolar.

A continuación se describen los resultados obtenidos a partir de estrategias planteadas por el Proyecto para disminuir la repitencia y mejorar la calidad de la enseñanza. Los resultados de tres años de trabajo sobre escuelas que ingresaron al Proyecto en el año 2006 sugieren tendencias alentadoras y asimismo dimensiones a fortalecer.

2.1. Repitencia

Según datos oficiales del Ministerio de Educación Nacional (DINIECE MECyT Relevamientos Anuales 2004 y 2005), la tasa de repitencia⁵ para todo el país considerando los primeros seis años de escolaridad primaria, es del 6,47%. En primer grado, el valor asciende al 9,68%. Este dato es inquietante, considerando que primer grado es el punto de entrada al nivel primario y, por lo tanto, el espacio donde se construyen los fundamentos y aprendizajes que influirán, positiva o negativamente, sobre el futuro aprendizaje de los niños. El campo académico presenta análisis interesantes del fenómeno de la repitencia (Brophy, 2006; Roderick, Nagaoka y Allensworth, 2005; Xia y Glennie, 2005, McCay, 2001; Thompson y Cunningham, 2000). Sus manifestaciones son complejas y abarcan dimensiones de orden técnico, cultural y político.

⁴ Tasa de repitencia: Porcentaje de alumnos que se matriculan como alumnos repitientes en el año lectivo siguiente. DINIECE-Ministerio de Educación.

La repitencia es comúnmente utilizada como 'solución' para niños y jóvenes que no alcanzan estándares académicos fijados por distintas audiencias: el Estado, la escuela, los docentes y, en algunos, los padres. Su uso es cultural y políticamente aceptado en un número considerable de países del mundo. Resulta, en parte, como consecuencia del modo en que se estructura la escuela moderna: graduada por edad. Por lo mismo, la repitencia no es un problema significativo en escuelas pluri-grado ubicadas generalmente en zonas rurales. Brophy (2006) distingue entre repitencia 'voluntaria' e 'involuntaria'. La primera resulta de una decisión de los padres y/o responsables del alumno. Se trata de una opción de los padres como paliativo a limitantes psicológicas, académicas o de provisión. La repitencia voluntaria no es estadísticamente significativa, ni suele traer consecuencias graves sobre la escolaridad de los estudiantes. En estos casos, la escuela y la familia actúan asociadamente.

La repitencia involuntaria, en cambio, es problemática tanto por su representación estadística como por sus efectos sobre la experiencia escolar del niño. En esta clase de repitencia se concentra este trabajo.

Un rasgo distintivo de la repitencia involuntaria es que no se manifiesta de manera homogénea en el total de la población escolar. Thompson y Cunningham (2000) observan que los alumnos varones repiten más que las mujeres, que los niños en contextos de pobreza repiten significativamente más que sus pares de segmentos socioeconómicos más altos. Esto ocurre, entre otros motivos, porque los criterios de promoción y de repitencia no son siempre claros ni objetivos. Siguiendo a Brophy (2006), 'los maestros en países en desarrollo, en general, no tienen acceso a estándares de logro detallados o a sus instrumentos de evaluación asociados (...) muchas decisiones tienden a basarse en observaciones arbitrarias o creencias, en lugar de criterios justificados' (p.9).

En la actualidad, los estudios sobre repitencia cuestionan la capacidad de este mecanismo para resolver problemas de rendimiento académico. Las investigaciones muestran que la repitencia empeora el rendimiento del alumno con el correr de los años, genera consecuencias afectivas indeseables para el niño y complica la gestión general de la escuela.

El estudio de Brophy (2006) señala que si bien estudiantes repetidores pueden mejorar su rendimiento en el corto plazo, desmejoran en el mediano y el largo. Desde el punto de vista pedagógico-didáctico, la repitencia no parece ser una herramienta eficiente para producir mejoras durables en el desarrollo académico del alumno. El segundo considerando refiere a los efectos de la repitencia sobre el autoestima del alumno. Un estudio de Frigerio (1992) realizado en la Argentina, señala que no sólo los niños empiezan a creer que repiten por falta de inteligencia, sino que sus familias refuerzan esa percepción. Este fenómeno es especialmente frecuente en comunidades en contextos de pobreza. Los menores no tienen posibilidades de pensar que la repitencia puede darse por motivos ajenos a su desempeño y su persona. El impacto de la repitencia sobre la institución escolar tampoco es desestimable. De acuerdo con Brophy (2006), la repitencia produce efectos económicos negativos sobre el sistema. Aumenta el número de vacantes esperado en los grados donde quedan retenidos los alumnos, demandando recursos adicionales a la institución, desde sillas y escritorios hasta libros y material didáctico.

Frente a este panorama, la política a veces pregona a favor de la 'promoción automática', estrategia que obliga la promoción de estudiantes al grado siguiente independientemente de sus esfuerzos y logros académicos. En la Argentina, hasta muy recientemente, ha habido decretos gubernamentales en algunas provincias prohibiendo la repitencia en los primeros grados. Sin embargo, los estudios desaconsejan esta medida: el problema es postergado, se mantiene o agudiza (Roderick, Nagaoka y Allensworth, 2005, Green y Winters, 2006). Maestros y directivos en escuelas con altas tasas de repitencia observan la

promoción automática como una amenaza. Sostienen, con razón, que ésta puede generar la baja en la calidad educativa y eliminar el incentivo al estudio.

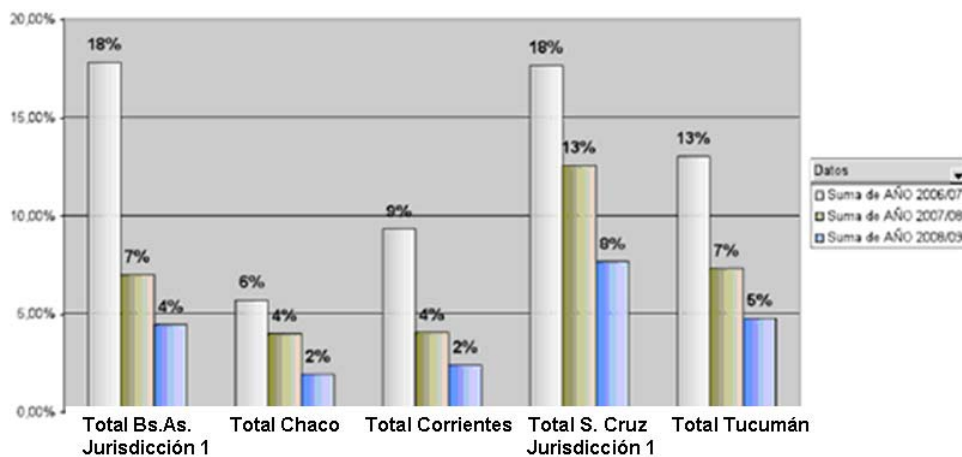
Frente a este panorama, urge explorar nuevas estrategias que garanticen aprendizajes de calidad para todos. Las soluciones pueden ser muy diversas. Se puede apuntar a la detección temprana de los niños con dificultades. Se puede promover el trabajo a contra-turno con docentes que brinden apoyo escolar, extensiones del horario de clase y trabajo en articulación con la familia.

Cabe reiterar, el desafío es fortalecer el accionar del sistema, evitar la exclusión y reducir el volumen de las políticas compensatorias. Las políticas de igualación requieren 'brindar mejores oportunidades a quienes tienen peores puntos de partida' (Braslavsky 1985).

Escuelas del Bicentenario acompaña a los directivos y maestros en su planificación y práctica pedagógica y les brinda herramientas para el seguimiento de indicadores, a fin de que tomen conciencia de la dimensión de estos problemas y analicen estratégicamente el modo de resolverlos.

El resultado de los primeros tres años de trabajo sobre esta variable clave del rendimiento interno, se percibe al establecer la comparación entre las tasas de repitencia del año 2006, 2007 y 2008⁶. Para cada escuela, se consideran todos los grados de la primaria. Los gráficos siguientes muestran patrones similares entre las escuelas ingresadas al Proyecto en el año 2006 y 2007.

GRÁFICO 1. REPITENCIA POR JURISDICCIÓN EN ESCUELAS INCORPORADAS AL PROYECTO EN 2006



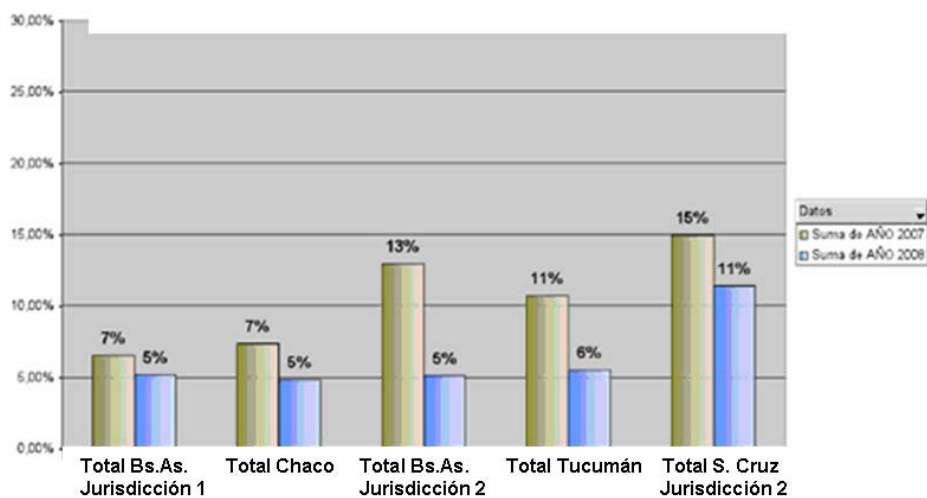
Fuente: elaboración propia

Este Gráfico presenta una evolución de los porcentajes de repitencia de las 20 escuelas incorporadas al Proyecto en 2006, comparando ese año (línea de base), el año 2007 (situación posterior al primer año de trabajo) y 2008 (situación posterior al segundo año de trabajo). Los resultados reflejan una disminución del 40% en 2007 respecto de 2006; y otro 39% en 2008 respecto del año anterior.

⁶ El Proyecto calcula la tasa de repitencia a partir de la composición de la matrícula total de la escuela a Diciembre y la cantidad de alumnos que luego del período compensatorio de Febrero del año próximo, han repetido el año.

La disminución en la tasa de repitencia es significativa, teniendo en cuenta que alcanzados ciertos niveles (menos del 10%), la mejora tiende a ser más lenta. Los análisis en profundidad de cada caso y jurisdicción están disponibles en los archivos del Proyecto, que servirán como base para la construcción de la memoria empírica y extracción de 'lecciones aprendidas' del Proyecto.

GRÁFICO 2. REPITENCIA POR JURISDICCIÓN EN ESCUELAS INCORPORADAS AL PROYECTO EN 2007



Fuente: elaboración propia

Este gráfico presenta una evolución de los porcentajes de repitencia de las 19 escuelas integradas al Proyecto en 2007, comparando ese año (línea de base) y el 2008 (situación posterior al primer año de trabajo). En términos generales, la mejora sigue el patrón de resultados del primer grupo de escuelas (2006).

En resumen, el punto de partida de la repitencia para las escuelas del Proyecto guarda relación con las tendencias generales observadas en poblaciones en condiciones de vulnerabilidad social. Este grupo particular de escuelas ha podido correrse de la media para mejorar sus prácticas de repitencia y, en consecuencia, sus indicadores de rendimiento interno.

TABLA 1. SÍNTESIS DE LA DISMINUCIÓN DE LA TASA DE REPITENCIA

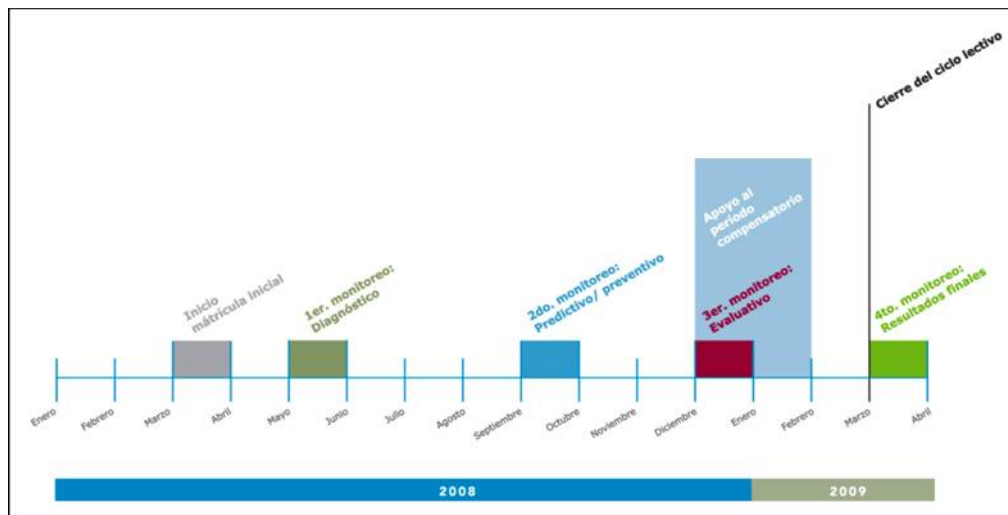
Primeras 20 Escuelas	Disminución tasa de Repitencia
1° Año de Intervención	40%
2° Año de Intervención	39%
En 2 años de Intervención	64%
Segundas 20 Escuelas	Disminución tasa de Repitencia
1° Año de intervención	39%
Total Proyecto	Disminución tasa de Repitencia
Ultimo Año de Intervención	39%

El trabajo sobre la repitencia y sus frutos dejaron 'lecciones aprendidas' que el Proyecto se propuso capitalizar. Para el segundo año de trabajo, se desplegó una estrategia preventiva sobre la base de cortes parciales de monitoreo del rendimiento interno. Esto permitió acompañar a los alumnos y las maestras a lo largo del año, identificando aquellos en riesgo de compensar o repetir con la necesaria antelación.

Esta estrategia consiste en recolectar información cuantitativa a través de una planilla informatizada o 'sistema de gestión' organizado en tres momentos durante el año. Los datos ofrecen información valiosa que permite tomar medidas preventivas, según se describe a continuación:

2.2. Cronograma Anual de Monitoreos

GRÁFICO 3. CRONOGRAMA ANUAL DE MONITOREOS



2.2.1. Primer Monitoreo. Diagnóstico

El primer monitoreo abarca los meses de marzo a mayo. Los datos corresponden al período comprendido entre el primer día de clases y el 31 de Mayo. Además de la matrícula inicial de la escuela, grados y secciones, en esta instancia los datos generales abarcan:

- 1) Datos del Registro: Son datos que se obtienen a partir del registro de clase e incluyen información sobre el movimiento de matrícula y la asistencia de los alumnos a la escuela. Los datos se calculan para el período de marzo a mayo en tablas independientes.
 - Movimiento: al completar el movimiento de alumnos 'entrados y salidos' (con pase y sin pase), la planilla de rendimiento interno calcula automáticamente la matrícula final mensual.
 - Asistencia: se completan los días hábiles del mes y los días efectivos de clase. Los días efectivos de clase se calculan considerando los días hábiles del mes menos los días en que las clases estuvieron suspendidas, cualquiera haya sido el motivo. Además de estos datos se completa la 'asistencia total' e 'inasistencia total'. La planilla calcula automáticamente el 'porcentaje de asistencia y asistencia media', en función de los días efectivos de clase.
- 2) Datos del Boletín: Datos que se obtienen a partir del último cierre del Boletín de Calificaciones. Aquí se completa la cantidad de alumnos aprobados y desaprobados en cada Área (Lengua, Matemática, Ciencias Naturales y Ciencias Sociales).
- 3) Resultados del Período: Arroja resultados en función de la información cargada en las planillas de movimiento, asistencia y alumnos desaprobados. Para los datos de entrados, salidos y matrícula, la planilla calcula el Total del Período (los tres meses). Se incluye además el dato de 'variación de matrícula', que indica el porcentaje de aumento (números positivos) o disminución (números negativos) de la matrícula final del período, en función de la matrícula inicial. En la sección de asistencia se calculan el total de días hábiles, días efectivos de clase, asistencia total e inasistencia

total. A partir de estos datos se calculan el 'porcentaje de asistencia y asistencia media' del período. En la última sección de la tabla se muestran los porcentajes de 'alumnos desaprobados en cada área'.

2.2.2. Segundo Monitoreo. Predictivo/Preventivo

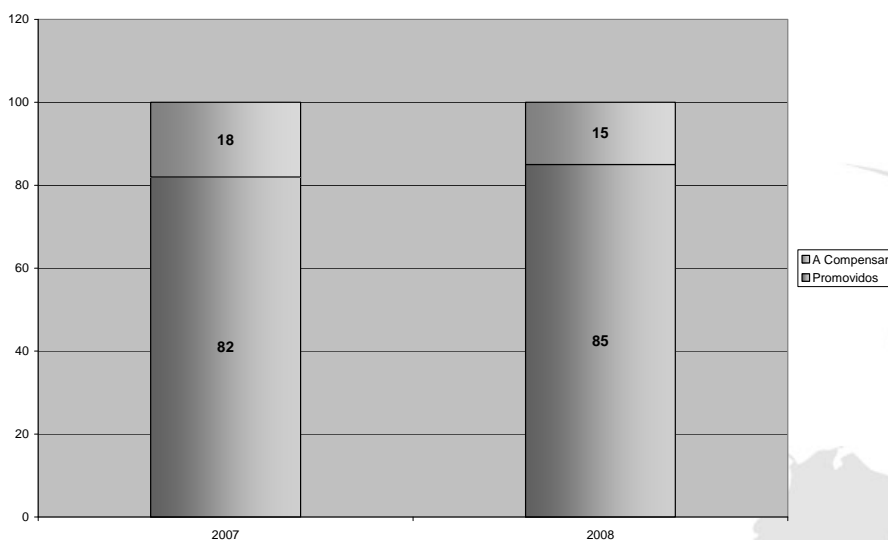
La estructura del segundo monitoreo es idéntica a la del primero. Abarca los meses de Junio a Agosto. Los datos a completar en esta sección corresponden al período comprendido entre el 1° de Junio y el 31 de Agosto. No se consignan datos acumulados. En los datos generales se incluye la matrícula al 31 de Mayo tomada del monitoreo anterior.

2.2.3. Compensación⁷

A partir de la estrategia 'preventiva', apoyada en los cortes parciales de monitoreo del rendimiento, se obtuvieron bajas en el porcentaje de alumnos 'a compensar' en Diciembre y/o Marzo y aumentos en el porcentaje de alumnos 'promovidos' directamente de año. Los gráficos a continuación muestran los porcentajes de alumnos promovidos sin necesidad de rendir exámenes en Diciembre/Febrero y los alumnos derivados al período compensatorio luego del primer y segundo año de trabajo en escuelas incorporadas al Proyecto en los años 2006 y 2007.

Se comparan datos del primer período compensatorio en 2007 y el primer período compensatorio en 2008, respectivamente.

GRÁFICO 4. ALUMNOS PROMOVIDOS Y A COMPENSAR A DICIEMBRE, EN ESCUELAS INCORPORADAS AL PROYECTO EN 2006

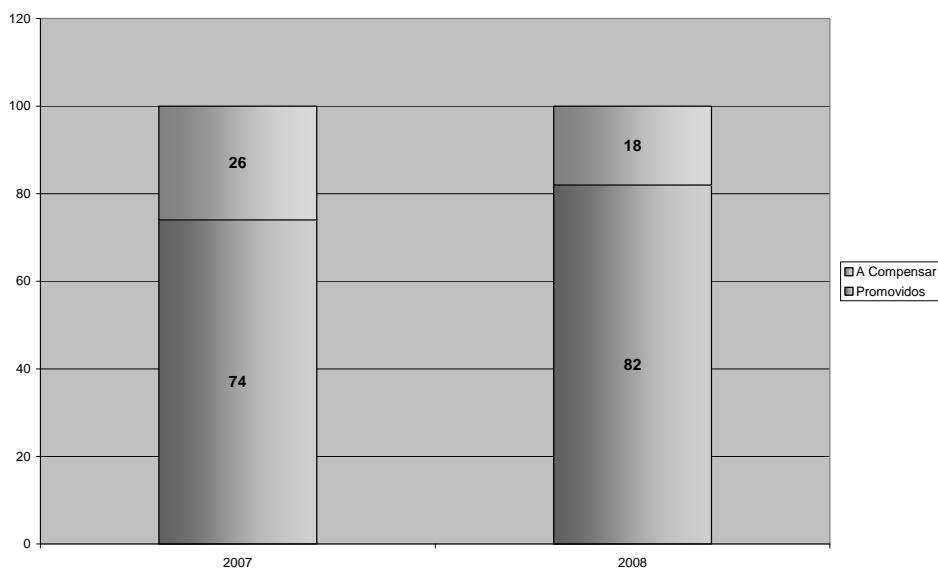


Fuente: elaboración propia

Para las escuelas incorporadas al Proyecto en 2006, el porcentaje de alumnos a compensar en Diciembre 2008, **disminuyó un 17 % respecto del año anterior.**

⁷ Alumnos que deben rendir examen en los meses de diciembre y/o febrero para promover el año.

GRÁFICO 5. ALUMNOS PROMOVIDOS Y A COMPENSAR A DICIEMBRE, EN ESCUELAS INCORPORADAS AL PROYECTO EN 2007



Fuente: elaboración propia

Para las escuelas incorporadas al Proyecto en 2007 al Proyecto, el porcentaje de alumnos a compensar en diciembre de 2008 **disminuyó un 29% respecto del año anterior.**

3. SEÑALES DE PROGRESO Y PERCEPCIÓN DE DIRECTIVOS SOBRE EL IMPACTO DEL PROYECTO

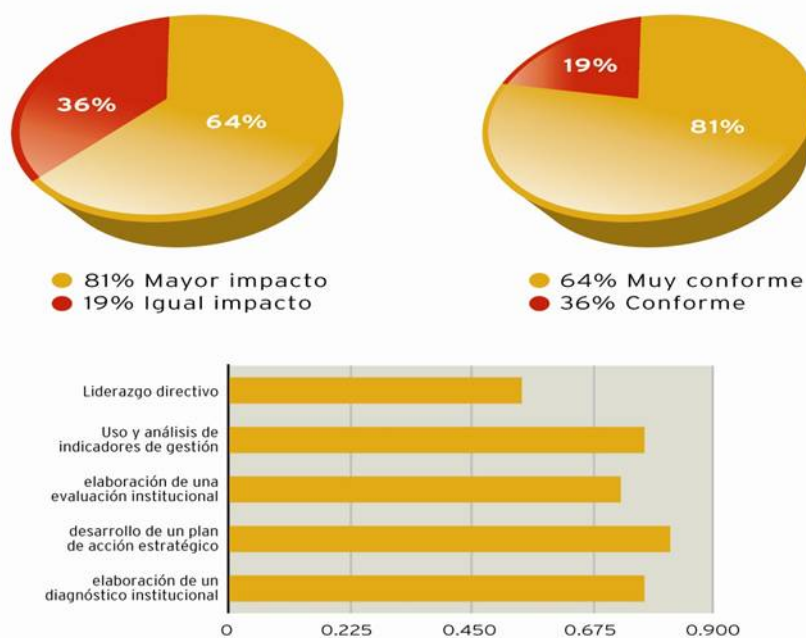
Las señales de progreso en el área de Gestión son complementarias a las variables cuantitativas. Reflejan capacidades instaladas en las jurisdicciones y escuelas y dan indicios de mejoras parciales.

- Mayor integración de los equipos Directivos e identificación de docentes líderes que acompañan la marcha del proyecto en la escuela;
- Replanteo del paradigma de la repitencia como estrategia pedagógica y esfuerzo conjunto con docentes para la detección y apoyo a los alumnos en riesgo pedagógico;
- Valoración de las herramientas de gestión y monitoreo: para la mejora escolar, para la identificación de alumnos en riesgo de abandono/repitencia, para la puesta en acto de estrategias preventivas del fracaso escolar;
- Puestas en común mensuales del avance de los proyectos institucionales y del proyecto escuelas del bicentenario en la escuela;
- Adopción del mecanismo 'acuerdos pedagógicos para la promoción';
- Adquisición de capacidades para realizar análisis situacionales y estratégicos, por parte de los directivos;
- Fortalecimiento del liderazgo directivo;

- Compromiso creciente de autoridades y supervisores con las distintas acciones promovidas por el Proyecto.

A partir de una encuesta aplicada a los equipos directivos de las escuelas incorporadas al Proyecto en el año 2006 luego del primer año de trabajo, se verificaron actitudes positivas hacia las propuestas del Proyecto en el área de Gestión Escolar. La encuesta indagó la percepción del impacto sobre la mejora escolar y el grado de satisfacción de los directivos. Se buscó asimismo diferenciar las habilidades desarrolladas y/o mejoradas a partir de la intervención del Proyecto.

GRÁFICA 6. PERCEPCIÓN DE DIRECTIVOS SOBRE EL IMPACTO DEL PROYECTO



Los resultados indicaron que un alto porcentaje de directivos consideraba 'muy bueno' el impacto del Proyecto sobre la mejora escolar, que la gran mayoría se encontraban 'muy conformes' con el Proyecto y que un alto porcentaje reconocía haber desarrollado y/o mejorado ciertas capacidades, entre las cuales se destacó la 'habilidad estratégica' y sugirió la necesidad de reforzar el trabajo sobre el 'liderazgo directivo', ya que el Proyecto la considera de suma importancia. A finales del 2010 se aplicará una nueva encuesta que medirá las percepciones de los directivos de todas las escuelas luego de cuatro años de trabajo.

La mejora en el rendimiento interno adquiere sentido cuando se acompaña de mejoras en el rendimiento académico, es decir, cuando los alumnos aprenden más y mejor. Por lo tanto, es clave acompañar el monitoreo del rendimiento interno con mediciones del rendimiento académico.⁸ El trabajo en simultáneo

⁸ El Proyecto Escuelas del Bicentenario refleja resultados interesantes en las Áreas Académicas. Se encuentran disponibles los Informes sobre resultados de evaluaciones de progreso en el Área de Lengua, Área de Matemática y Ciencias Naturales.

en estas dos dimensiones da cuenta de que las fórmulas 'alta repitencia = alta exigencia = alta calidad', y su opuesta, 'baja repitencia = baja exigencia = baja calidad', son falaces.

El siguiente apartado presenta el escenario de resultados en las áreas de mejora académica del Proyecto.

4. RENDIMIENTO ACADÉMICO

Antes de presentar esta dimensión del Proyecto, cabe una nota preliminar. La Argentina participa desde el año 2000 en estudios internacionales de evaluación de la calidad educativa con el fin de obtener información relevante sobre los aprendizajes de los alumnos.

Los datos del Segundo Estudio Regional Comparativo y Explicativo (SERCE 2006), en el que participaron 16 países de la región, 100.572 estudiantes de 3er. grado y 95.288 de sexto grado, muestran que la Argentina se ubicó en el séptimo lugar en comprensión lectora para los alumnos de 3er. grado, y en el octavo lugar para los alumnos de 6to. grado. En las Áreas de Matemática y Ciencias Naturales, los resultados tampoco fueron alentadores, ya que la Argentina pasó a ocupar la sexta y cuarta posición (en el área de Ciencias intervinieron solamente nueve países, y sus resultados fueron más bajos que el promedio general).

Estas evaluaciones también reflejaron que 7 de cada 10 alumnos no logran comprender o interpretar un texto o una consigna y 8 de cada 10 alumnos no son capaces de resolver un problema o una fracción. Los resultados del Operativo Nacional de Evaluación ONE 2005 también ponen de manifiesto debilidades en la calidad de la educación, en consonancia con los resultados de las pruebas Internacionales.

TABLA 2. PORCENTAJE DE ALUMNOS POR NIVELES DE DESEMPEÑO. TOTAL PAÍS (*)

	LENGUA			MATEMÁTICAS			CS. NATURALES		
	Alto	Medio	Bajo	Alto	Medio	Bajo	Alto	Medio	Bajo
3° Primaria	38.7	30.2	31.1	18.9	42.4	38.7	34.1	40.3	25.6
6° Primaria	32.5	37.8	29.7	13.5	46.1	40.4	23.5	39.5	37.0

Fuente: DINIECE en base a datos del ONE 2005

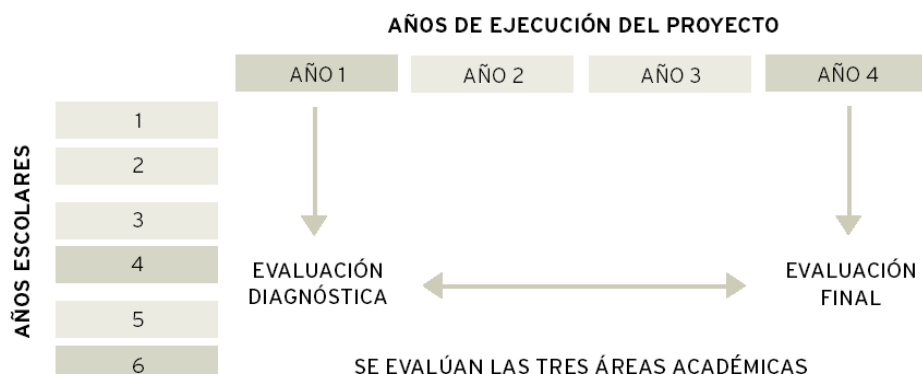
*Las referencias al significado de 'Alto', 'Medio' y 'Bajo' se encuentran en el anexo Nro. 1

La situación en la que se encuentra la Argentina alerta sobre la necesidad de trabajar para la mejora académica en las escuelas. A fin de complementar la información de estos estudios nacionales e internacionales y obtener datos más precisos sobre la población que atiende el Proyecto, el área de mejora académica diseña dispositivos de evaluación con instrumentos y metodologías acordes a los enfoques teóricos de sus disciplinas. Cabe una descripción de cada uno.

4.1. La evaluación de resultados: línea de base y línea de egreso

La evaluación de resultados se utiliza para medir el efecto de las acciones del Proyecto sobre el rendimiento académico de los alumnos cumplidos los 4 años de trabajo. Se aplica el mismo instrumento sobre una muestra representativa de alumnos de 4to y 6to año al comienzo y al final de la intervención. El producto de la comparación entre el diagnóstico inicial o línea de base y la situación final o línea de egreso constituye un indicador clave de las mejoras logradas en el rendimiento académico de los alumnos. El siguiente esquema (Gráfica 7) señala los momentos y grados en los que aplica el dispositivo.

GRÁFICA 7. EVALUACIÓN DE RESULTADOS: LÍNEA DE BASE Y LÍNEA DE EGRESO



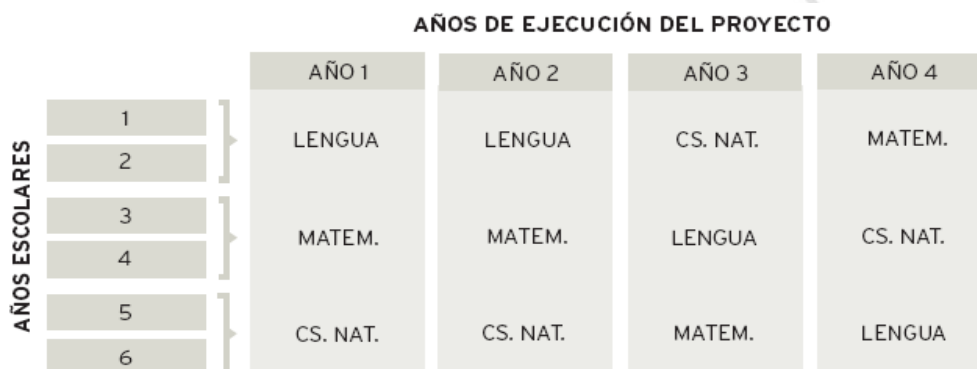
A diferencia del área de gestión, donde indicadores centrales como la repitencia, el abandono y la sobreedad pueden relevarse y compararse anualmente, la evaluación de resultados solo permite establecer comparaciones entre estados iniciales y finales al cierre de la intervención. Como se dijo al inicio, las primeras 20 escuelas incorporadas al Proyecto en el año 2006 completan su período de trabajo a finales del corriente 2010. Por lo tanto, los efectos de la intervención sobre esta variable del Proyecto se darán a conocer a finales de este año. No obstante, las áreas académicas utilizan otras estrategias para identificar mejora parciales a medida que avanza el Proyecto.

4.2. La evaluación de progreso

Este dispositivo se utiliza para obtener indicios y muestras parciales del progreso académico. Busca valorar avances en la forma en que los alumnos responden a situaciones didácticas propuestas, teniendo en cuenta las condiciones de enseñanza. Incluye una variedad de instrumentos de evaluación, como pruebas, entrevistas, cuestionarios, observaciones de clase, portfolios y análisis de cuadernos, que sirven como insumo para la capacitación y elaboración de la memoria empírica del Proyecto.

La evaluación de progreso aplica en los grados o años donde la capacitación es intensiva, siguiendo la secuencia de capacitación plasmada en el siguiente gráfico.

GRÁFICA 8. EVALUACIÓN DEL PROGRESO



El proyecto se implementa en las escuelas de manera paulatina pero completa. Según las características del establecimiento – urbanos o rurales, número de docentes, alumnos, grados y secciones - se diseña un esquema ajustado de trabajo. En las primeras 20 escuelas la capacitación a docentes se organizó en bloques de dos grados. El año 2007, Lengua trabajó en 1° y 2° y Matemática en 3° y 4°. El año 2008 incorporó Ciencias Naturales en 5° y 6°. En los años sucesivos, las áreas fueron rotando sobre los otros bloques hasta abarcar la totalidad de grados de la primaria.

Las pruebas de progreso se aplican a final de año en el grado donde funciona la capacitación de cada área. La información que se recaba permite realizar ajustes en las prácticas pedagógicas y modalidades de evaluación propuestas a los docentes. Estas pruebas señalan tendencias de mejora en los aprendizajes de los alumnos, ya que el modo en que responden y resuelven contrasta con las respuestas de los alumnos a ítems semejantes en las pruebas diagnósticas o línea de base. Dichos instrumentos no son comparables *stricto sensu*. Por lo tanto, constituyen buenos indicios de que las mejoras acontecen y sirven como insumos para orientar el trabajo de las áreas académicas.

En resumen, El Proyecto Escuelas del Bicentenario distingue dos grandes categorías de resultados: ‘capacidades instaladas’ y ‘efectos’. La primera abarca todas las mejoras cualitativas en las prácticas escolares y condiciones pedagógicas establecidas dentro del aula, el desarrollo personal de los alumnos y maestros, las actualizaciones y ampliaciones en los planes y contenidos curriculares, las mejoras en las relaciones entre los docentes, entre los docentes y la dirección, el fortalecimiento del rol directivo, la construcción de espacios de participación, la ampliación en los márgenes de autonomía e iniciativa escolar, entre otros. La segunda refiere a los resultados objetivos del Proyecto, organizados en dos dimensiones: rendimiento interno de las escuelas y rendimiento académico de los alumnos. El artículo trató la primera en más detalle y ofreció una descripción general del modo de evaluar la segunda. Se espera sistematizar y socializar los resultados finales al cierre del ciclo lectivo 2010.

5. CONCLUSIONES

A lo largo de este trabajo se expusieron hipótesis relacionadas con la repitencia y sus efectos. Asimismo se analizaron estrategias alternativas para su tratamiento. El análisis teórico y los resultados de tres años de trabajo con escuelas públicas en el Proyecto Escuelas del Bicentenario, permiten formular algunas conclusiones.

En primer lugar, la política educativa no debe entenderse como sinónimo de ‘macro política’. Cualquier estrategia de política debe operar en simultáneo en los tres niveles: macro, meso y micro. Si bien la ‘provincialización’ del sistema permite que las escuelas pasen a gobernarse desde las distintas jurisdicciones, ocurre al mismo tiempo una re-centralización del gobierno educativo en el nivel provincial. Las líneas de gobierno intermedias cuentan con espacios de acción y grados de autonomía limitados. Tampoco se fomentan espacios institucionalizados para atender las necesidades de equipos directivos en cada escuela y otros actores de la comunidad local. Si, como expresa la Ley Nacional de Educación (2006), se desea que padres y comunidades participen de la vida de la escuela, hacen falta diseños institucionales adecuados para que dicha participación sea posible. Si se desea que las escuelas puedan llevar adelante proyectos institucionales propios, hace falta hacer viables nuevas formas de organización escolar. Es preciso alentar y fortalecer la habilidad programática en el nivel local, en las escuelas y distritos. Se requiere involucrar a todos los niveles del sistema educativo en la construcción de justicia

educacional: aquellos que diseñan las políticas educativas, supervisores del nivel intermedio y miembros de la comunidad educativa (Joyce, Calhoun et al 1999).

Escuelas del Bicentenario plantea un dispositivo de trabajo encuadrado en la formación y el acompañamiento a los actores en la base del sistema. Los logros en la baja de la repitencia y la mejora en los aprendizajes se atribuyen centralmente al maestro, al director y al supervisor que conviven cotidianamente con la institución escolar y sus desafíos. Ellos son los protagonistas de los procesos de mejora y considerarlos en el diseño de la transformación es indispensable para conseguir la mejora.

Cuando la situación de repitencia es constante a lo largo de los años, el problema no es coyuntural sino estructural. La solución demanda una revisión profunda del modelo de provisión del sistema, que redunde en menos políticas compensatorias en el mediano o largo plazo.

En este marco, es necesario trabajar con sistemas de información que permitan supervisar los procesos de mejora escolar en forma paralela y articulada con la eficiencia interna del sistema. Resulta clave el diseño de sistemas y políticas de formación que hagan accesible tanto a los actores del nivel micro como a los del nivel meso y macro la lectura y comprensión de la información que les compete. La información produce efectos en las propias instituciones escolares.

En el diseño de acciones y políticas orientadas a la solución de la repitencia se juegan los derechos de la infancia a una buena educación y el desafío de la justicia educacional.

REFERENCIAS BIBLIOGRÁFICAS

- Braslavsky, C. (1985). *La discriminación educativa*. Buenos Aires, Grupo Editor Latinoamericano.
- Berlinski, S., Galiani, S., Manacorda, M. (2008) Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics* 92, 1416–1440.
- Brophy, J. (2006). *Grade repetition*. Educational Policy Series N° 6, 33 p. Paris, UNESCO-IIEP; Brussels, International Academy of Education.
- DINIECE and MECyT (Relevamientos Anuales 2004 y 2005).
- Joyce, B, Calhoun E. and Hopkins, D. (1999) *The new structure of school improvement. Inquiring schools and achieving students*. Open University Press, Buckingham. Philadelphia.
- Green, J. y Winters, M. (2006) *Getting Farther Ahead by Staying Behind: A Second-Year Evaluation of Florida's Policy to end Social Promotion*. Manhattan Institute for Policy Research, Civic Report No. 49.
- Frigerio, G. *Obstinaciones y estrategias: fracaso escolar y sectores populares en Argentina*. Propuesta Educativa, Año 4, n° 6, mar.1992.
- Gvirtz, S. (2005). *De la tragedia a la esperanza. Hacia un sistema educativo justo, democrático y de calidad*. Buenos Aires, Academia Nacional de Educación.
- Gvirtz, S. C. (2008). *Equidad y niveles intermedios de gobierno en los sistemas educativos: un estudio de casos en la Argentina, Chile, Colombia y Perú*. Buenos Aires, Aique Grupo Editor.
- Gvirtz, S. and A. Oría (en prensa). *La alianza público-privado para la mejora educacional. Proyecto Escuelas del Bicentenario*. Buenos Aires, Aique - IIPE/Unesco Sede Regional Buenos Aires.

- Llach, J. J. (2006). *El desafío de la equidad educativa. Diagnóstico y propuestas*. Buenos Aires, Editorial Granica.
- Ministerio Nacional de Educación (2006) Ley Nacional de Educación. http://www.me.gov.ar/doc_pdf/ley_de_educ_nac.pdf.
- McCay, E. (Ed) (2001) *Moving beyond retention and social promotion*. Bloomington, IN: Phi Delta Kappa International.
- Pérez Murcia, L. E., R. Uprimmy Yepes, et al. (2007). *Los derechos sociales en serio: Hacia un dialogo entre derechos y políticas públicas*. Bogotá, Centro de Estudios DeJuSticia.
- Roderick, M., Nagaoka, J. y Allensworth, E. (2005) *Is the glass half full or mostly empty? Endign social promotion in Chicago*, Yearbook of the National Society for the Study of Education, 104 (2), 223-259.
- Thompson, C. y Cunningham, E. (2000) *Retention and social promotion: Research and implications for policy*. New York: ERIC Clearinghouse on Urban Education, Teachers College, Columbia University.
- Xia, C. y Glennie, E. (2005) *Grade Retention: A Flawed Education Strategy*. Part One of Three Part Series on Grade Retention. Center for Child and Family Policy, Duke University.

MODELO EXPLICATIVO DEL BAJO RENDIMIENTO ESCOLAR: UN ESTUDIO CON ADOLESCENTES MEXICANOS

EXPLANATORY MODEL ON THE LOW ACADEMIC PERFORMANCE. A STUDY WITH MEXICAN ADOLESCENTS

Joaquín Caso Niebla y Laura Hernández Guzmán

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art8.pdf>

Fecha de recepción: 15 de agosto de 2009
Fecha de dictaminación: 24 de abril de 2010
Fecha de aceptación: 24 de abril de 2010

1. ANTECEDENTES

El creciente interés observado en las últimas décadas por la evaluación del aprendizaje a gran escala ha contribuido al desarrollo de acciones que permiten la descripción, caracterización y explicación de los niveles de logro educativo de los estudiantes. Producto de dichas acciones, hoy sabemos con niveles aceptables de certeza cual es el nivel de dominio que poseen los estudiantes en diversas áreas del conocimiento y en determinado nivel y grado escolar, los avances o retrocesos de los niveles de aprovechamiento escolar en determinado grupo poblacional, y el efecto diferenciado de las variables contextuales en el rendimiento escolar.

Con todo y lo alentador de este panorama, en México se ha subestimado el papel de las variables personales y escolares en el desempeño escolar de los estudiantes. Si bien existe un avance considerable al conocer los efectos que sobre el rendimiento académico ejercen el ingreso familiar, los niveles de escolaridad de los padres, las expectativas familiares, el capital cultural, las características de la escuela y la influencia del profesor, se sigue soslayando el papel de variables asociadas al funcionamiento psicológico de los individuos, determinadas como resultado de su interacción con el entorno escolar, familiar y social.

La relación del rendimiento académico con variables de esta naturaleza, entre las que se encuentran la autoestima, la asertividad, el establecimiento de metas, el consumo de sustancias, las actividades de estudio y la adaptación escolar, no se ha estudiado a cabalidad en este país. A pesar de que es común observar acciones de naturaleza preventiva hacia el interior de las instituciones educativas, éstas difícilmente se respaldan y fundamentan en los hallazgos de la investigación realizada en contextos educativos y en marcos conceptuales que emergen de ésta. A este respecto, existen diversos estudios a nivel internacional que vinculan al rendimiento académico en adolescentes con la autoestima, las habilidades asertivas, el consumo de sustancias, las habilidades de estudio, el establecimiento de metas y la adaptación escolar.

Una de las asociaciones que más se ha investigado es la existente entre la autoestima y el desempeño académico (DuBois, Bull, Sherman & Roberts, 1998; González-Pienda *et al.*, 2002; Hoge, Smit & Crist, 1995; Kaplan, Peck & Kaplan, 1994; Owens, 1994). Diversas investigaciones han informado de la comparación de estudiantes con alto y bajo rendimiento escolar donde éstos últimos presentan baja autoestima, sentimientos de ineficacia personal y ausencia de expectativas profesionales (DuBois *et al.*, 1998; Felner *et al.*, 1995; Harter, 1993; Hernández-Guzmán & Sánchez-Sosa, 1996; Hoge *et al.*, 1995). Estudios realizados por DuBois *et al.* (1998), Owens (1994) y Hoge *et al.* (1995), con muestras considerables de adolescentes, refuerzan parcialmente estos hallazgos al destacar correlaciones positivas significativas entre la autoestima y el rendimiento académico en sus respectivas investigaciones.

Otros estudios se han enfocado en estudiar la relación entre la autoestima y el fracaso escolar y el efecto de variables mediadoras. Tal es el caso de una investigación realizada con estudiantes de bachillerato ($n=3,148$) en el que Kaplan y colaboradores demostraron, mediante el modelamiento de los datos, que los mecanismos autoprotectores desarrollados por los estudiantes orientados a evitar el fracaso, pueden convertirse en el principal obstáculo del rendimiento académico (Kaplan *et al.*, 1994). Al manifestar conductas tales como desinterés en las exigencias escolares y la reducción del esfuerzo, les es posible atribuir su fracaso a la propia ausencia de esfuerzo y no a la falta de habilidad, por lo que protegen con ello sus sentimientos de valía personal, pero a la vez incrementan la probabilidad de que fracasen en sus estudios.

Por haber merecido la autoestima tanta atención en la literatura de investigación, se ha confirmado su interconexión teórica con variables como la asertividad. Los datos muestran asociaciones positivas y altas entre ambos constructos (Furr, 2005; Stake & Pearlman, 1980). Como la asertividad se refiere a la manifestación adecuada de opiniones y sentimientos en situaciones sociales, entraña la demostración de las habilidades sociales que posee la persona. No es de extrañar entonces que también se haya documentado, mediante correlaciones, el vínculo entre el rendimiento académico y la asertividad (Poyrazli, Arbona, Nora, McPherson & Pisecco, 2002; Spaight, 1987).

Estos hallazgos se han verificado en otros estudios que muestran cómo los programas que promueven el desarrollo de habilidades sociales en adolescentes tienden a favorecer sus relaciones interpersonales, la autoestima, la solución de problemas y la manifestación de conductas asertivas que les permiten rechazar el consumo de sustancias, además de que afectan positivamente su rendimiento académico (Fajardo-Vargas, Hernández-Guzmán, & Caso-Niebla, 2001; Shell, 1999; Stake & Pearlman, 1980; Thompson & Bundy, 1995; Wentzel, 1991).

A pesar de que los estudios revisados coinciden al destacar un efecto positivo entre el desarrollo de conductas asertivas y reducción en los índices de consumo de marihuana, alcohol, cigarro y otras sustancias, no se expresa de manera explícita las repercusiones que dicho entrenamiento pudiera tener sobre variables escolares tales como el rendimiento, la reprobación o el abandono escolar (Kamalanabhan, 1988; Rajendran & Kaliappan, 1991). Sin embargo, ya que el consumo de sustancias sí parece tener un efecto, podría suponerse una relación. Al igual que la autoestima, la asertividad se ha asociado con procesos autorregulatorios que involucran a las creencias de una persona en la elección de las conductas asertivas, en los que la expectativa de eficacia y de resultado están presentes (Thompson & Bundy, 1995). Sin embargo esta relación se encuentra débilmente documentada, situación que se suma a las limitaciones observadas en el estudio de este constructo.

El consumo de sustancias adictivas es otra de las variables que la literatura asocia, no sólo con la autoestima y la asertividad, sino también con el rendimiento académico y con el abandono escolar (Dishion, Kavanagh, Schneiger, Neilson & Kaufman, 2002; Wynn, Schulenberg, Maggs & Zucker, 2000). Se ha sugerido que la autoestima disminuida, la falta de asertividad, y el consumo de sustancias afectan negativamente el funcionamiento del adolescente en los contextos en los que se desenvuelve, incluyendo el escolar (Poyrazli *et al.*, 2002; Quatman & Watson, 2001; Wentzel, 1991) y, especialmente, la relación con su desempeño académico. El consumo de sustancias se ha estudiado principalmente con propósitos clínicos por lo que, las dificultades que pudieran presentarse al fundamentar su estudio en ambientes educativos, son elementos de los resultados de investigación ampliamente documentados en el ámbito de la psicología de la salud. Relaciones entre el consumo de sustancias y la dinámica familiar disfuncional, el estrés, conducta agresiva, depresión y otros trastornos afectivos se han documentado ampliamente, lo que sugiere un efecto importante en la disfunción psicológica de los individuos (Toray, Coughlin, Vuchinich & Patricelli, 1991).

Lo anterior se ha confirmado a partir de investigaciones que han evaluado programas preventivos de las adicciones en los que al reducir el consumo de sustancias en adolescentes, tienden a reducirse concomitantemente sus índices de ausentismo y mejoran sus calificaciones escolares en comparación con los grupos de control correspondientes (Botvin, 1986; Botvin, Griffin, Díaz, Scheier, Williams & Epstein, 2000). La aplicación de estrategias preventivas en contra de adicciones sugiere, además, cambios positivos en variables tales como la depresión, autoestima, estrés percibido, enojo y bienestar psicológico, variables que pudieran encontrarse, según sea el caso, como causas o consecuencias del bajo

rendimiento académico (Bennett, 1995; D'Amico, 2000; Dukes *et al.*, 1997; Ellickson, 1998; Snow *et al.*, 1992; Sussman *et al.*, 2002; Valentine *et al.*, 1998; Wynn *et al.*, 2000).

Otra de las variables comúnmente asociadas con el desempeño académico se encuentra en las habilidades de estudio. Esta relación goza de un apoyo considerable en investigación que sugiere el valor explicativo que éstas tienen sobre el éxito o fracaso escolar y en la caracterización de estudiantes con alto y bajo rendimiento (Lammers, Onweugbuzie & Slate, 2001; Proctor, Hurst, Prevatt, Petscher & Adams, 2006). La organización y concentración en el estudio, la capacidad para relacionar nuevos conocimientos con los existentes, la comprensión lectora y la capacidad para autorregular el aprendizaje son habilidades que la literatura ha asociado con el rendimiento académico de estudiantes (Lammers *et al.*, 2001; Ruban, 2000; Tuckman, 2003; Zhang & RiCharde, 1999). Entre las líneas de investigación desarrolladas en torno a las habilidades de estudio se encuentran aquéllas que comparan a estudiantes con bajo rendimiento escolar, con problemas de aprendizaje, y con dificultades en variables psicosociales, de aquéllos con desempeños aceptables. Los resultados indican que estudiantes con dificultades académicas cuentan con un pobre repertorio de habilidades para el estudio (Proctor *et al.*, 2006) y que las calificaciones pueden explicarse en buena medida por el peso explicativo que aporta el uso de estrategias de aprendizaje (Ruban, 2000).

La autorregulación del aprendizaje, proceso que comúnmente hace referencia a procesos autorregulatorios basados en las estrategias que un estudiante utiliza en la adquisición de nuevos conocimientos, en la forma en que un estudiante organiza, planea y administra de manera intencionada las actividades escolares y el tiempo asignado al estudio, se vincula estrechamente con el rendimiento académico (Lammers *et al.*, 2001; Tuckman, 2003).

Al respecto, algunos datos señalan que cuando se ha entrenado específicamente a estudiantes del bachillerato a desplegar habilidades de naturaleza autorregulatoria, sus calificaciones escolares tienden a mejorar (Reid, 1997; Tuckman, 2003), aumentan sus habilidades para la comprensión de lectura (Fuchs *et al.*, 2001), de las matemáticas (Ashman & Conway, 1993) y de las ciencias (Nelson, Smith & Dodd, 1992), a la vez que mejora su motivación hacia el estudio (Catello, 2000) y se reducen sus niveles de ansiedad (Sanghvi, 1996) y depresión (Litten, 1999).

El efecto de variables de naturaleza afectiva en el rendimiento académico y los posibles efectos mediadores con respecto a las habilidades de estudio deben estudiarse con mayor profundidad. Si bien no se ha documentado claramente si la motivación hacia el estudio promueve que los estudiantes le dediquen más tiempo y, posiblemente, echen mano de mejores estrategias de aprendizaje y habilidades de estudio, se reconoce el efecto positivo que ejerce en el desempeño escolar y las calificaciones escolares (Powell & Arriola, 2003; Shim & Ryan, 1995; Tavani & Losh, 2003). La motivación hacia el estudio se ha medido mediante el establecimiento de metas definido como los procesos asociados a estados finales que el individuo busca alcanzar mediante la regulación cognitiva y afectiva de su conducta (Eccles & Wigfield, 2002). El establecimiento de metas, en el contexto del desarrollo adolescente, se asocia con el rendimiento académico, la elección vocacional y la educación para el trabajo (Casullo & Cayssials, 1995; Giota, 2002; Heikkinen, 1997; Lupart, Cannon & Telfer, 2004).

Estudios realizados con adolescentes atribuyen a la formulación de metas efectos positivos en la motivación hacia las tareas académicas con resultados positivos en sus calificaciones escolares (Shim & Ryan, 2005; Wentzel, 1991), en sus niveles de dominio en áreas específicas del currículo escolar (Giota,

2002), en sus niveles de participación en clase (Sanders, 1995) y en la habilidad para escribir (Bogolin, Harris & Norris, 2003).

La adaptación escolar es una variable referida comúnmente por la comunidad educativa como un componente clave para la retención de estudiantes y de formar parte de los programas y acciones institucionales orientados a este propósito (Tinto, 1987; De la Orden *et al.*, 2001; ANUIES, 2001). A pesar de ello prácticamente no existen registros de investigaciones que documenten la existencia de dicha relación. Algunos la ubican como parte del sistema social de una institución, otros como indicador que describe los efectos de la naturaleza de las interacciones entre los diversos actores del sistema educativo, y otros como el sentido de pertenencia hacia la propia institución.

En la literatura psicológica se ha estudiado esta variable principalmente mediante el constructo "sentido de pertenencia escolar" (Certo, Cauley & Chafin, 2003; Osterman, 1998). De acuerdo con Osterman (1998), el sentido de pertenencia que un estudiante desarrolla hacia su comunidad escolar fomenta actitudes positivas hacia la escuela e impacta favorablemente en su rendimiento académico. Por el contrario, los estudiantes que no lo desarrollan presentan problemas de conducta, poco interés en los asuntos escolares, bajo rendimiento e índices de abandono escolar. Osterman sostiene que el sentido de pertenencia se ha asociado principalmente con el rendimiento académico y con los procesos psicológicos que inciden en el éxito escolar, a motivos y actitudes académicas, a actitudes personales y sociales y a niveles de compromiso e involucración personal. El sentido de pertenencia también se ha explicado a través de las necesidades de afiliación, por lo que pertenecer a redes de apoyo que permitan atenderlas pudiera contribuir al fortalecimiento de la autoestima, de los niveles de competencia social y en la percepción de eficacia personal.

Diversos estudios que han asociado el sentido de pertenencia con variables tales como la competencia social (Carlson *et al.*, 1999), la autoestima (Bishop & Inderbitzen, 1995), la autoeficacia (Osterman, 1998), el ambiente escolar (Certo *et al.*, 2003) y la motivación escolar (Lane, Marquardt, Meyer & Murray, 1997), lo que viene a confirmar lo propuesto por Osterman (1998). En tanto, aunque existe un acuerdo generalizado en la literatura en asociar al sentido de pertenencia escolar con el rendimiento escolar, los hallazgos de estudios que han explorado esta relación mediante índices de logro académico en dominios curriculares son contradictorios ya que, mientras unos señalan no haber identificado niveles de correlación entre dichas variables (Battistich, Solomon, Kim, Watson & Schaps, 1995), otros sugieren correlaciones moderadas entre constructos (Goodenow, 1993).

Si bien los presentes hallazgos sugieren la forma en como algunas de estas variables se vincula con el rendimiento académico, explicándolo y prediciéndolo, la relación que guarda con otras variables no es igual de clara. Es por ello, que resulta relevante someter a prueba un modelo que intenta explicar las relaciones entre la autoestima, la asertividad, el establecimiento de metas, las habilidades de estudio, la adaptación escolar y el consumo de sustancias, con respecto al rendimiento académico, a partir de los hallazgos de la investigación en esta materia.

2. MÉTODO

2.1. Participantes

Participaron 1581 estudiantes de una institución de educación media superior pública de la Ciudad de México, 850 mujeres (53%) y 731 hombres (47%), con edades fluctuaban entre los 15 y los 23 años. De éstos, 634 estudiaban el primer año escolar, 334 el segundo y 613 el tercero, 1018 en el turno matutino y 563 en el turno vespertino. Se seleccionó a los participantes mediante el método de muestreo aleatorio simple y se consideró como unidad de muestreo al grupo escolar de adscripción, lo que propició la participación de estudiantes pertenecientes a 51 grupos escolares de los 931 grupos que conforman este subsistema del bachillerato mexicano.

2.1.1. Instrumentos

- Escala de Integración y Adaptación Escolar (Méndez, 2003). Se conforma por trece reactivos con cinco opciones de respuesta tipo Likert. El instrumento pretende evaluar aspectos inherentes al proceso de integración y adaptación de un estudiante a la comunidad escolar. Ejemplos: *"me siento orgulloso de pertenecer a esta escuela"*, *"Conozco los servicios y apoyos que me ofrece mi escuela"*. Las cinco opciones de respuesta tipo Likert con que cuenta cada reactivo se califican de la siguiente forma: totalmente = 5, casi totalmente = 4, más o menos = 3, un poco = 2, nada = 1. Una calificación alta en la escala sugiere un nivel alto de sentido de pertenencia a la escuela. La escala presenta un índice de consistencia interna = .85
- Inventario de Autoconocimiento y Establecimiento de Metas (Osorno, Crespo, Arjona & Romero, 2003). Se conforma por 16 reactivos que evalúan dimensiones del proceso mediante el cual un individuo aprende a planear y proyectar sus metas mediante el esclarecimiento de sus capacidades y limitaciones cognitivas, afectivas y conductuales. Ejemplos: *"Antes de plantear una meta analizo los obstáculos que se me puedan presentar y la forma de enfrentarlos"*, *"Me quedan claros los pasos que tengo que seguir para alcanzar las metas que me he propuesto"*. Se responde a los reactivos con base a una escala de repuesta tipo Likert de cinco opciones (totalmente de acuerdo, parcialmente de acuerdo, indeciso, parcialmente en desacuerdo y totalmente en desacuerdo) calificando con 5 a siempre y 1 a nunca cuando el reactivo implique una expresión positiva del constructo y con 1 a siempre y 5 a nunca cuando la expresión sea negativa. Por lo tanto, a mayor puntaje registrado en el inventario, mayor capacidad en un individuo para planear y proyectar metas. El inventario presenta un índice de confiabilidad de =.81
- Prueba de Autoestima para Adolescentes (Caso, 2000). Escala dirigida a adolescentes mexicanos conformada por 20 reactivos con cinco opciones de respuesta tipo Likert con un índice de confiabilidad $\alpha = .88$. Permite evaluar al conjunto de juicios valorativos que un individuo hace de sí mismo con respecto a su desempeño, capacidades y atributos. Ejemplos: *"Me gusta la forma como me veo"*, *"Soy muy lento para realizar mi trabajo escolar"*. Los reactivos cuentan con cinco opciones de respuesta (siempre, usualmente, algunas veces, rara vez, nunca) y se califican con 5 las respuestas que impliquen expresiones relacionadas con una alta autoestima y con 1 las que impliquen baja autoestima. Una calificación alta en la escala, representa una mayor autoestima.
- Adaptación del Inventario de Asertividad de Gambrell y Richey (Guerra-Ramos, 1996). Con base al comportamiento del instrumento en estudios realizados con adolescentes mexicanos (Caso, 2000;

Fajardo, 2002), se utilizó una adaptación del inventario conformada por 23 reactivos mismos que evalúan la conducta de aquellos individuos capaces de expresar directa y adecuadamente sus opiniones y sentimientos (tanto positivos como negativos) en situaciones sociales. Ejemplos: Evalúa la probabilidad de que adoptes esta conducta.... "Disculparte cuando tienes la culpa", "Cortar la plástica con alguien que habla mucho". Los reactivos se responden a través de una escala tipo Likert con cinco opciones de respuesta por lo que se califican de la siguiente forma: siempre = 5, usualmente = 4, la mitad de las veces = 3, rara vez = 2 y nunca = 1. Un puntaje alto en el inventario supone la utilización de conductas asertivas. La adaptación del inventario presenta un índice de consistencia interna = .85.

- Cuestionario de Actividades de Estudio (Sánchez Sosa & Martínez Guerrero, 1993). Inventario dirigido a adolescentes mexicanos que permite evaluar las habilidades intelectuales empleadas en la adquisición de información para lograr un aprendizaje efectivo. Ejemplos: "Cuando estudio, señalo en el libro las ideas o conceptos más importantes (subrayo, anoto al margen, encierro párrafos, etc.)", "Siento que hay muchas cosas que me distraen en el lugar donde estudio". Los reactivos se apoyan en una escala tipo Likert con seis niveles de respuesta (siempre o casi siempre, muchas veces, poco más de la mitad de las veces, poco menos de la mitad de las veces, pocas veces, nunca o casi nunca) calificando con 6 a siempre o casi siempre y con 1 a nunca o casi nunca cuando el reactivo implique una expresión positiva del constructo y con 1 a siempre o casi siempre y 6 a nunca o casi nunca cuando la expresión sea negativa. Un puntaje alto en el cuestionario representa en el individuo la existencia de habilidades, actividades y situaciones personales que favorecen el estudio. Se utilizó una adaptación conformada por 47 reactivos (Caso, 2007). El inventario presenta un índice de consistencia interna de = .91
- Escala Frecuencia de Consumo del Cuestionario de Abuso de Sustancias (Gil, Wagner & Tubman, 2004). Para este estudio solo se aplicaron los siete reactivos correspondientes a la escala de Frecuencia de Consumo, mismos que permiten identificar el nivel de consumo de alcohol, marihuana, cocaína y sustancias que afectan la salud física y mental de un individuo. No se tienen identificados reportes que documenten su aplicación en nuestro país. Ejemplos: "¿Con qué frecuencia fumas?", "¿Con qué frecuencia tomas alcohol?". Los reactivos cuentan con siete niveles de respuesta calificados de la siguiente forma: Todos los días = 7, varias veces a la semana = 6, varias veces al mes = 5, varias veces en la vida = 4, lo hacía y lo dejé = 3, una sola vez en la vida = 2, nunca = 1. Un puntaje elevado en la escala implica mayor frecuencia de consumo de sustancias.
- Historial académico. Se consultaron los cardex de cada estudiante a fin de determinar su rendimiento académico mediante el promedio de calificaciones escolares acumulado al concluir el año escolar. Las calificaciones escolares representan la manera más comúnmente utilizada para evaluar el grado de cumplimiento de los objetivos de aprendizaje dentro del salón de clases (Brookhart, 1994; Lambating & Allen, 2002).

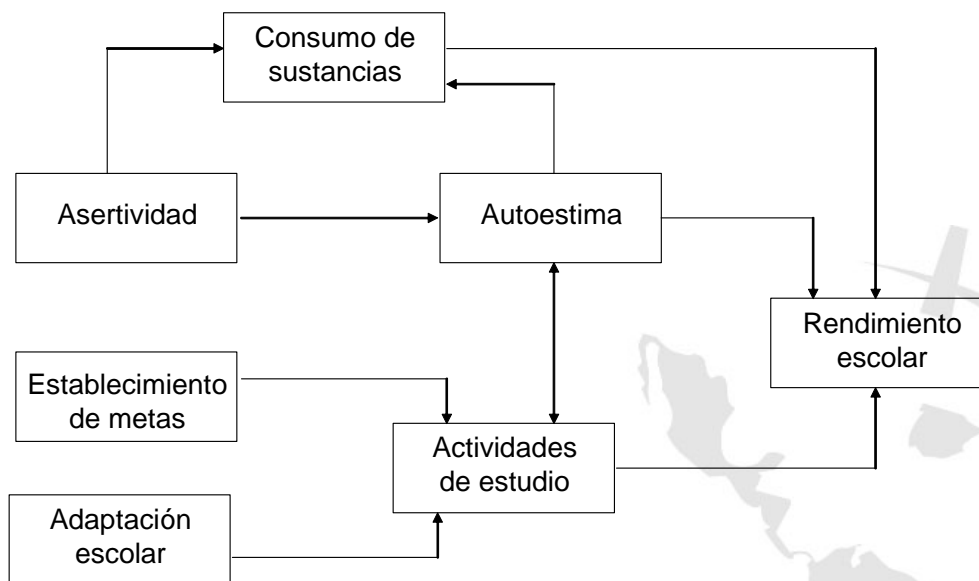
2.2. Procedimiento

Se integró la batería de instrumentos, se diseñó un cuadernillo, la hoja de lectura óptica y el instructivo para la administración de la batería, se capacitó al grupo de aplicadores que apoyarían esta actividad, se programó la aplicación en la totalidad de planteles y grupos involucrados, se aplicaron instrumentos, se procesaron las respuestas de los estudiantes con apoyo del lector óptico, se integró la base de datos y se realizaron los análisis estadísticos correspondientes. La aplicación de la batería tuvo una duración promedio de 90 minutos.

3. RESULTADOS

Con el propósito de calcular la contribución relativa de la autoestima, el establecimiento de metas, las habilidades de estudio, el consumo de sustancias, la asertividad y la adaptación escolar al rendimiento académico, propósito principal de este estudio, se sometió a prueba un modelo hipotético (Diagrama 1) que pretende explicar cómo contribuyen las diversas variables al rendimiento académico, con base en las siguientes consideraciones: a) la asertividad favorece los sentimientos de valía personal y el rechazo al consumo de sustancias; b) el consumo de sustancias afecta negativamente el funcionamiento del adolescente en el contexto escolar, particularmente su desempeño académico; c) la autoestima es un constructo vinculado principalmente a variables de naturaleza personal y escolar, y se ha estudiado como variable moderadora o mediadora del rendimiento académico; d) la organización y concentración en el estudio, el tiempo asignado al estudio y la capacidad para autorregular el aprendizaje, son habilidades asociadas al rendimiento académico de los estudiantes; y e) los aspectos emocionales que subyacen a procesos cognitivos son determinantes para el aprendizaje por lo que variables de naturaleza afectiva tales como la autoestima, el establecimiento de metas y las habilidades de estudio pudieran incidir en el rendimiento académico.

DIAGRAMA 1. MODELO HIPOTÉTICO INICIAL QUE INTENTA EXPLICAR LA RELACIÓN ENTRE VARIABLES INDIVIDUALES Y SU RELACIÓN CON EL RENDIMIENTO ACADÉMICO Y QUE INCORPORA A LAS VARIABLES ASERTIVIDAD Y ADAPTACIÓN ESCOLAR



El modelo se probó mediante el modelamiento de ecuaciones estructurales y con apoyo del programa AMOS 6.0. El modelamiento de ecuaciones estructurales lineales es una técnica multivariada que se utiliza para desarrollar y poner a prueba modelos teóricos acerca de un fenómeno o problema de interés. Combina aspectos del análisis de regresión múltiple y del análisis factorial para analizar relaciones de dependencia entre variables latentes y manifiestas así como valorar sus efectos causales (Cuevas & Aguilar, 1999).

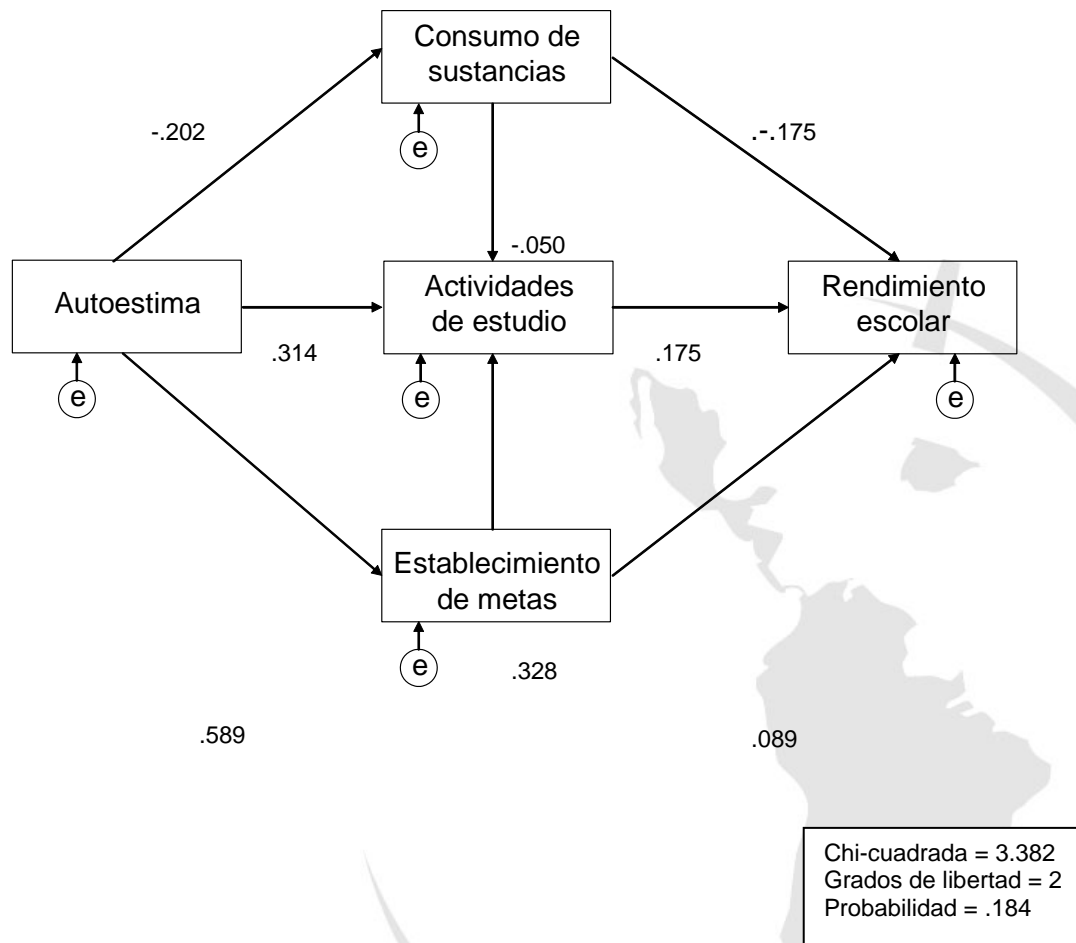
El análisis del modelo propuesto, representado por el diagrama no. 1, no registró valores que confirmen su ajuste ($X^2=437.718$, $gl=11$, $p=.000$) para dicha versión. Por tal motivo, se propuso un segundo modelo en el que se incluyen la autoestima, el establecimiento de metas, las habilidades de estudio y el consumo de sustancias, y en el que se excluyeron las variables adaptación escolar y asertividad debido a las

limitaciones registradas en torno a su operacionalización y a la falta de datos empíricos que respalden su relación con el rendimiento académico.

Además se atendieron las siguientes consideraciones: a) la autoestima como variable moderadora con efectos directos sobre el consumo de sustancias, las actividades de estudio y el establecimiento de metas; b) las actividades de estudio como variable mediadora entre el rendimiento académico y las variables autoestima, consumo de sustancias y establecimiento de metas; y c) el consumo de sustancias, la autoestima y el establecimiento de metas con efectos directos sobre el rendimiento académico (Diagrama 2).

Los resultados del modelo alternativo sugieren que la configuración propuesta presenta un nivel aceptable de ajuste ($X^2=3.382$, $gl=2$, $p=0.184$), un nivel de varianza explicada para el promedio de calificaciones del orden del 9.8%. Se registran medidas de ajuste aceptables para el modelo ($GFI = 0.999$, $AGFI = 0.994$ y $RMSEA = 0.021$).

DIAGRAMA 2. MODELO HIPOTÉTICO ALTERNO QUE INTENTA EXPLICAR LA RELACIÓN ENTRE VARIABLES INDIVIDUALES Y SU RELACIÓN CON EL RENDIMIENTO ACADÉMICO



Los efectos estandarizados directos, indirectos y totales entre estas variables y el rendimiento académico se incluyen en la tabla 1, destacando los efectos directos del orden de .175, .089 y -.175 de las actividades de estudio, el establecimiento de metas y el consumo de sustancias, respectivamente.

TABLA 1. EFECTOS DIRECTOS, INDIRECTOS Y TOTALES DE LAS VARIABLES CONSIDERADAS EN EL MODELO EN EL RENDIMIENTO ACADÉMICO

Variables	Efecto		
	Directo	Indirecto	Total
Actividades de estudio	.175	.000	.175
Establecimiento de metas	.089	.058	.147
Consumo de sustancias	-.175	-.009	-.184
Autoestima	.000	.178	.178

4. DISCUSIÓN

El modelo hipotético propuesto inicialmente, al someterse a prueba no mostró un ajuste aceptable. Era posible anticipar de los análisis previos la responsabilidad escasa de algunas variables en la explicación del desempeño académico. Por lo tanto el que el modelo no ajustara podría explicarse en términos de: a) la exclusión de la asertividad y adaptación escolar como variables que contribuyen al rendimiento académico; b) la inclusión de la autoestima como variable mediadora en el modelo, con efectos directos sobre el rendimiento académico, cuando los análisis previos sugerían su efecto moderador al registrar correlaciones de moderadas a altas con las variables consumo de sustancias, establecimiento de metas y actividades de estudio, y un bajo nivel de asociación con las calificaciones escolares; c) la propuesta de las actividades de estudio como variable mediadora entre establecimiento de metas, adaptación escolar y el rendimiento académico.

La participación de algunas variables en el modelo, de acuerdo con la literatura de investigación, no se sostuvo por lo que el presente trabajo derivó en un modelo hipotético alterno. El modelo resultante, con un nivel de ajuste aceptable, se presenta en el diagrama no. 2. En éste, la autoestima juega un papel moderador en el rendimiento académico, con un efecto directo sobre las actividades de estudio, el consumo de sustancias y el establecimiento de metas.

Los resultados sugieren que los sentimientos de valía personal y la percepción de eficacia personal, componentes centrales de la autoestima, contribuyen al desarrollo de habilidades autorreguladoras asociadas con la conducta académica, tales como la organización de las actividades de estudio y el empleo de estrategias de aprendizaje (Lammers *et al.*, 2001; Ruban, 2000; Tuckman, 2003; Zhang & RiCharde, 1999). Además, fungen como el soporte para la formulación de metas dirigidas a la realización y dominio de una determinada tarea, con efecto indirecto sobre el rendimiento escolar, relación ampliamente documentada (Ames, 1992; Shim & Ryan, 2005).

Además, se observa que la autoestima presenta un efecto directo sobre el consumo de sustancias y éste a su vez con el bajo rendimiento escolar. A este respecto se ha expuesto que los sentimientos de minusvalía personal afectan negativamente el funcionamiento del adolescente en los contextos en los que se desenvuelve, lo que pudiera desencadenar conductas de riesgo que afecten su rendimiento académico, entre ellas el consumo de sustancias (Poyrazli *et al.*, 2002; Quatman & Watson, 2001; Wentzel, 1991).

El modelamiento de ecuaciones estructurales, si bien contribuyó a la evaluación de los modelos propuestos y con ello a la confirmación de una serie de relaciones y efectos indicados entre variables en

torno al rendimiento académico, con implicaciones conceptuales considerables, sólo contribuyó a explicar un porcentaje reducido de la varianza asociada con las calificaciones escolares. Lo anterior, además de requerir la inclusión de variables contextuales en modelos explicativos del rendimiento académico en estudios futuros, también fundamenta la inclusión de variables personales en este tipo de estudios, hasta ahora subestimadas.

El modelo resultante, no sólo resulta congruente con los estudios que evalúan la relación entre las variables personales y escolares, sino que está de acuerdo con las perspectivas contemporáneas del desarrollo (Lerner, 2001). Tanto las explicaciones contemporáneas del desarrollo, como la investigación que las sustenta, indican que ciertas variables como la autoestima presentan un curso a lo largo del desarrollo de la persona que comienza con el nacimiento y que otras, como el consumo de sustancias, el establecimiento de metas y las actividades de estudio, dependen más de circunstancias asociadas con momentos específicos del ciclo vital, como lo es la adolescencia. Por lo tanto, el modelo es útil para explicar el rendimiento académico en la adolescencia tardía, periodo en el que entran en juego variables dadas por las circunstancias y oportunidades propias de esa etapa de la vida. Así, el modelo no podría explicar el rendimiento en otras edades, como la niñez, por ejemplo.

No debe olvidarse que el modelo resultante es limitado en cuanto su alcance explicativo, ya que deja una proporción muy grande de la varianza sin explicar. Los estudios subsiguientes tendrán forzosamente que integrar otras variables distales, como el nivel socioeconómico, y proximales, como la participación de la familia y las relaciones familiares, para explorar con mayor precisión la naturaleza multideterminada del rendimiento escolar. Independientemente de esta consideración, el modelo que mejor ajustó contribuye al estado del arte sobre las variables que predicen el rendimiento académico, con un grupo de variables que habían recibido menos atención.

Lo anterior, además de contribuir de manera importante a la construcción de cuerpos teóricos y modelos explicativos, permite aportar elementos a considerar en la fundamentación de programas preventivos en contextos educativos y en la instrumentación de acciones específicas que pudieran requerirse para la atención de la problemática asociada al bajo rendimiento escolar.

REFERENCIAS BIBLIOGRÁFICAS

- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- Ashman, A. & Conway, R. (1993). Teaching students to use process-based learning strategies and problem solving strategies in mainstream classes. *Learning & Instruction*, 3(2), 73-92.
- Asociación Nacional de Universidades e Instituciones de Educación Superior / ANUIES (2001). *La Educación Superior en el Siglo XXI*. México: ANUIES.
- Battistich, V. Solomon, D., Kim, D., Watson, M. & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, 32(3), 27-58.
- Bennett, G. T. (1995). An exploration of patterns of drug use and of the effectiveness of a substance abuse prevention program according to adolescents level of academic achievement. *Dissertation Abstracts International*, 55(11-A), 3091.

- Bogolin, L., Harris, L. & Norris, L. (2003). Improving student writing through the use of goal setting. Master of Arts Action Research Project. U.S. Illinois: Saint Xavier University.
- Botvin, G. (1986). Substance abuse prevention research: Recent developments and future directions. *Journal of School Health*, 56 (9) 369-74.
- Botvin, G., Griffin, K. W., Díaz, T., Scheier, L. M., Williams, C. & Epstein, J. (2000). Preventing illicit drug use in adolescents: Long term follow-up data from a randomized control trial of a school population. *Addictive Behaviors*, 25(5), 769-774.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279-301.
- Carlson, E., Sroufe, A., Collins, W., Jimerson, S., Weinfield, N., Henninghausen, K., Egeland, B., Anderson, F. & Meyer, S. (1999). Early environmental support and elementary school adjustment as predictors of school adjustment in middle adolescence. *Journal of Adolescent Research*, 14 (1), 72-94.
- Caso, J. (2000). Validación de un instrumento de autoestima para niños y adolescentes. Tesis de Maestría. México: Universidad Nacional Autónoma de México.
- Casullo, M. & Caussials, A. (1995). *Decisión vocacional y proyecto de vida*. Buenos Aires: Ed. Paidós
- Catello, J. (2000). The effect of the R.E.A.L. (Reinvention of Education for Active Learning) classroom experience on motivational orientation among middle school students. *Dissertation Abstracts International*, 60(7-A), 2348.
- Certo, J. L., Cauley, K. M. & Chafin, C. (2003). Students' perspectives on their high school experience. *Adolescence*, 38(152).
- Cuevas, C. & Aguilar, J. (1999). Metodología para la construcción de modelos estructurales lineales con variables ordinales. *Revista Mexicana de Psicología*, 16(2), 237-243.
- D'Amico, E. J. (2000). Adolescent risk-taking and prevention: Development of a new risk skills training program. *Dissertation Abstracts International*, 60(9-B), 4882.
- De la Orden, A., Oliveros, L., Mafokozi, J. & González, C. (2001). Modelos de investigación del bajo rendimiento. *Revista Complutense de Educación*, 12(1), 159-178.
- Dishion, T. J., Kavanagh, K., Schneiger, A., Neilson, S. & Kaufman, N. K. (2002). Preventing early adolescent substance use: A family-centered strategy for the public middle school. *Prevention Science*, 3 (3), 191-201.
- DuBois, D. L., Bull, C. A., Sherman, M. D. & Roberts, M. (1998). Self-esteem and adjustment in early adolescence: A social-contextual perspective. *Journal of Youth and Adolescence*, 27(5), 557-583.
- Dukes, R. L. Stein, J. A. & Ullman, J. B. (1997). Long-term impact of Drug Abuse Resistance Education (DARE): Results of a 6-year follow-up. *Evaluation Review*, 21(4), 483-500.
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values and goals. *Annual Reviews Psychological*, 53, 109-132.
- Ellickson, P. L. (1998). Preventing adolescent substance abuse: Lessons from the Project ALERT Program. En: *Social programs that work*. Crane, Jonathan (Ed.). Pp. 201-257.

- Fajardo-Vargas, V., Hernández-Guzmán, L & Caso-Niebla, J. (2001). La autoestima y su relación con la depresión, ansiedad y asertividad. *Revista Mexicana de Psicología*, 18(1), 150. Memorias del IX Congreso Mexicano de Psicología.
- Felner, D., Brand, S., DuBois, D., Adan, A. M., Mulhall, P. F & Evans, E. G. (1995). Socioeconomic disadvantage, proximal environment experiences, and socioemotional and academic adjustment in early adolescence: Investigation of a mediated effects model. *Child Development*, 66, 774-792.
- Fuchs, D., Fuchs, L. S., Thompson, A., Svenson, E., Yen, L., Otaiba, S., Yang, N., McMaster, K., Prentice, K., Kazdan, S., & Saenz, L. (2001). Peer-assisted learning strategies in reading: Extensions for kindergarten, first grade, and high school. *Remedial & Special Education*, 22(1), 15-21.
- Furr, M. (2005). Differentiating happiness and self-esteem. *Individual Differences Research*, 3(2), 105-127.
- Gil, A., Wagner, E. & Tubman, J. (2004). Culturally sensitive substance abuse intervention for Hispanic and African American adolescents: Empirical examples from the Alcohol Treatment Targeting Adolescents in Need (ATTAIN) Project. *Addiction*, 99, 140-150.
- Giota, J. (2002). Adolescents' goal orientations and academic achievement: Long-term relations and gender differences. *Scandinavian Journal of Educational Research*, 46(4), 349-371.
- González-Pienda, J. A., Nuñez, J. C., González-Pumariega, S., Alvarez, L., Rocés, C. & García, M. (2002). A structural equation model of parental involvement, motivational and aptitudinal characteristics, and academic achievement. *The Journal of Experimental Education*, 70(3), 257-287.
- Goodenow, C. (1992). School motivation, engagement, and sense of belonging among urban adolescent students. *Paper presented at the Annual Meeting of the American Educational Research Association*. San Francisco, CA., April 20-24.
- Guerra-Ramos, M. T. (1996) Estandarización del inventario de asertividad de Gambrill y Richey para población de la Ciudad de México. Tesis de licenciatura. México: Universidad Nacional Autónoma de México.
- Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. En R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard*. (Pp. 87-116). New York: Plenum Press.
- Heikkinen, A. (1997). Vocational education as a "Life Project". *Journal of European Industrial Training*, 21 (6/7).
- Hernández-Guzmán, L. & Sánchez-Sosa, J.J. (1996). Factores de riesgo y protectores que afectan el aprovechamiento escolar. *Psicología y problemática social*. Facultad de Psicología, UNAM.
- Hoge, D.R., Smit, E. K. & Crist, J. T. (1995). Reciprocal effects of self-concept and academic achievement. *Journal of Youth and Adolescence*, 24(3), 295-314.
- Kamalanabhan, T. J. (1998) Efficacy of a behavioural program for personality change and improvement in academic performance of school students. *Journal of Psychological Researches*, 32(1-2), 93.
- Kaplan, D. S., Peck, M. & Kaplan, H. B. (1994). Structural relations model of self-rejection, disposition to deviance, and academic failure. *Journal of Educational Research*, 87(3), 166-173.
- Lambating, J. & Allen, J. D. (2002). How the multiple functions of grades influence their validity and value as measures of academic achievement. Ponencia presentada en: *Annual Meeting of the American Educational Research Association*. New Orleans, April 1-5.

- Lammers, W., Onweugbuzie, A. & Slate, J. R. (2001). Academic success as a function of gender, class, age, study habits, and employment of college students. *Research in the schools*, 8(2), 71-81.
- Lane, C., Marquardt, J., Meyer, M. A. & Murray, W. (1997). Addressing the lack of motivation in the middle school setting. Master's Action Research Project. U.S. Illinois: Saint Xavier University.
- Lerner, R. M. (2001). Concepts and theories of Human Development. Lawrence Erlbaum Associates.
- Litten, K. (1999). A prevention program design for inner-city, high school adolescents to combat learned helplessness and depression. *Dissertation Abstracts International*, 60(3-B), 1306.
- Lupart, J. L., Cannon, E. & Telfer, J. A. (2004). Gender differences in adolescent academic achievement, interests, values and life-role expectations. *High Ability Studies*, 15(1), 26-42.
- Méndez, R. (2003). Curso de Integración y adaptación escolar. Manual del Alumno. México: Dirección General de Orientación y Servicios Educativos, UNAM.
- Nelson, J., Smith, D. & Dodd, J. (1992). The effects of teaching a summary skills strategy to students identified as learning disabled on their comprehension of science test. *Education & Treatment of Children*, 15(3), 228-243.
- Osorno, G.P., Crespo, C., Arjona, S. & Romero, P. (2003). Taller Proyecto de Vida. Manual del Alumno. México: Dirección General de Orientación y Servicios Educativos, UNAM.
- Osterman, K. (1998). Student community within the school context: A research synthesis. Ponencia presentada en: Annual Meeting of the American Educational Research Association. San Diego, April 13-17.
- Owens, T.J. (1994). Two dimensions of self-esteem: Reciprocal effects of positive self-worth and self-depreciation on adolescent problems. *American Sociological Review*, 59(3), 391-407.
- Powell, C. L. & Arriola, K. R. (2003). Relationship between psychosocial factors and academic achievement among African American students. *Journal of Educational Research*, 96(3), 175-181.
- Poyrazli, S., Arbona, C., Nora, A., McPherson, R. & Pisecco, S. (2002). Relation between assertiveness, academic self-efficacy, and psychosocial adjustment among international graduate students. *Journal of College Student Development*, 43(5), 32-42.
- Proctor, B., Hurst, A., Prevant, F., Petscher, Y. & Adams, K. (2006). Study skills profiles of normal-achieving and academically struggling college students. *Journal of College Student Development*, 47(1), 37-51.
- Quatman, T. & Watson, C. M. (2001). Gender differences in adolescent self-esteem: An exploration of domains. *The Journal of Genetic Psychology*, 62(1), 93-117.
- Rajendran, R. & Kaliappan, K. (1991). Efficacy of a behavioural programme in managing the academic stress and improving academic performance. *Journal of Personality & Clinical Studies*, 6(2), 193.
- Reid, E. (1997). Exemplary Center for Reading Instruction (ECRI). *Behavior & Social Issues*, 7(1), 19-24.
- Ruban, L. M. (2000). Patterns of self-regulated learning and academic achievement among university students with and with out learning disabilities. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 61(4-A), 1296.
- Sánchez-Sosa, J. J. & Martínez-Guerrero, J. I. (1993). Cuestionario de Actividades de Estudio (C.A.E.). México: Coordinación de Programas Académicos de Enseñanza Media Superior.

- Sanders, J. (1995). *Improving classroom behavior through a motivational intervention*. U.S. Illinois: Saint Xavier University.
- Sanghvi, C. (1996). Efficacy of study skills training in managing study habits and test anxiety of high test anxious students. *Journal of the Indian Academy of Applied Psychology*, 21 (1), 71-75.
- Shell, A. M. (1999). *Metacognitive intervention effects on attendance, achievement, and prosocial behavior among ninth and tenth grade students at-risk for school dropout*. Dissertation Abstracts International, 59 (11-A).
- Shim, S. & Ryan, A. (2005). Changes in self-efficacy, challenge avoidance, and intrinsic value in response to grades: The role of achievement goals. *Journal of Experimental Education*, 73(4), 333-349.
- Snow, D., Tebes, J., Arthur, M. & Tapasak, R. (1992). Two-year follow-up of a social-cognitive intervention to prevent substance use. *Journal of Drug Education*, Vol. 22(2), 101-114.
- Spaights, E. (1987). The relationship of assertiveness and the academic success of black students in predominantly white institutions of higher education. *Psychology: A Quarterly Journal of Human Behavior*, 24(3), 9-16
- Stake, J. & Pearlman, J. (1980). Assertiveness training as an intervention technique for low performance self-esteem women. *Journal of Counseling Psychology*, 27(3), 276-281.
- Sussman, S., Dent, C. W., Clyde, W., Craig, S. Ritt-Olsen, A. & McCuller, W. J. (2002). Development and immediate impact of a self-instructed curriculum for an adolescent indicated drug abuse prevention trial. *Journal of Drug Education*, 32(2), 121-137.
- Tavani, C. M. & Losh, S. C. (2003). Motivation, self-confidence, and expectations as predictors of the academic performances among our high school students. *Child Study Journal*, 33(3), 141-151.
- Thompson, K. & Bundy, K. (1995). Social skills training for young adolescents: Symbolic and behavioral components. *Adolescence*, 30 (119), 723-735.
- Tinto, V. (1987). *El abandono de los estudios superiores: Una nueva perspectiva de las causas del abandono y su tratamiento*. México: UNAM / ANUIES.
- Toray, T., Coughlin, C., Vuchinich, S. & Patricelli, P. (1991). Gender differences associated with adolescent substance abuse: Comparisons and implications for treatment. *Family Relations*, 40, 338-344.
- Tuckman, B. (2003). The effect of learning and motivation strategies training on college students' achievement. *Journal of College Student Development*, 44(3), 430-437.
- Valentine, J., Griffith, J., Ruthazer, R., Gottlieb, B. & Keel, S. (1998). Strengthening causal inference in adolescent drug prevention studies: Methods and findings from a controlled study of the Urban Youth Connection Program. *Drugs & Society*, Vol. 12(1-2), 127-145.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development*, 62(5), 1066-1079.
- Wynn, S. R., Schulenberg, J., Maggs, J. L. & Zucker, R. A. (2000). Preventing alcohol misuse. The impact of refusal skills and norms. *Psychology of Addictive Behaviors*, 14(1), 36-47.
- Zhang, Z. & RiCharde, R. S. (1999). Freshman academic achievement: A structural equation model. Ponencia presentada en: Annual Meeting of the American Educational Research Association. Montreal, April 19-23.

EL USO DE INFORMACIÓN EVALUATIVA EXTERNA CON FINES FORMATIVOS: EL CASO DE ESTABLECIMIENTOS EDUCACIONALES CHILENOS PARTICIPANTES DE SEPA

THE USE OF EXTERNAL EVALUATIVE INFORMATION WITH FORMATIVE
MEANS: THE CASE OF CHILEAN EDUCATIVE ESTABLISHMENTS
PARTICIPANTS OF THE SEPA

Andreas Hein y Sandy Taut

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art9.pdf>

Fecha de recepción: 21 de enero de 2010
Fecha de dictaminación: 02 de mayo de 2010
Fecha de segundo envío: 10 de mayo de 2010
Fecha de aceptación: 11 de mayo de 2010

Se vive en la actualidad en una época centrada en la necesidad que las organizaciones rindan cuentas de su quehacer (*accountability*). Las organizaciones educacionales no se encuentran exentas de esta exigencia y son objeto de numerosas evaluaciones externas a través de test estandarizados que buscan medir el proceso educacional. Los indicadores desarrollados a partir de las evaluaciones estandarizadas externas se encuentran en vías de ser cada vez más complejas y se han convertido en el reflejo del éxito de los establecimientos educacionales (Gareis, 2007; Koretz, 2008). Pese a ello la introducción de indicadores, no siempre contribuye directamente al mejoramiento de las prácticas educativas. A diferencia de lo que tradicionalmente se cree, la traducción de los resultados de evaluaciones externas en acciones concretas de mejoramiento es un proceso complejo, no lineal y cuyas implicancias son diversas.

En el presente trabajo, se desarrolla una investigación que indaga sobre las diversas formas en que se puede utilizar información evaluativa y los factores que facilitan y dificultan su uso, en un grupo de establecimientos educacionales que participan del Sistema de Evaluación de Progresos del Aprendizaje (SEPA), desarrollado e implementado por el Centro de Medición de la Pontificia Universidad Católica de Chile (www.sepauc.cl).

SEPA es un sistema de evaluación externo y consiste en un conjunto de pruebas que evalúan los aprendizajes logrados por alumnos de 1° a 8° año básico en Lenguaje y Matemática. Los clientes de SEPA son mayormente colegios individuales, o grupos de colegios asesorados y/o administrados por fundaciones. Las pruebas SEPA incluyen preguntas de alternativas múltiples que son corregidas por el Centro para mantener su confidencialidad, y de desarrollo que son corregidas por los profesores apoyados por una rúbrica.

Los resultados son reportados a través de una página web (<http://reportes.sepauc.cl>) a la cual tienen acceso los clientes y actores de los establecimientos en los que se aplica SEPA. La presentación de resultados permite a cada colegio compararse con otros, resguardando su identidad. Los resultados incluyen:

- (a) Los logros de aprendizaje (o el "estado" del aprendizaje) por subsector evaluado (Lenguaje y/o Matemática), por ejes temáticos (áreas de contenido mínimo obligatorio definidas por Marco Curricular) para el establecimiento, nivel, curso y alumno.
- (b) Si el establecimiento permanece más de un año en SEPA, se le informa respecto del "progreso" de aprendizaje entre dos años evaluados, a nivel de establecimiento, nivel, curso y alumno, brindando una visión longitudinal acerca del desarrollo del alumno.
- (c) Finalmente, a establecimientos con datos longitudinales, se les proporciona también índices de "valor agregado" para cada nivel evaluado. El valor agregado se obtiene a partir de una sofisticada técnica estadística que permite trazar la trayectoria de los aprendizajes en el tiempo y establecer el aporte efectivo que el colegio hace a ellos.

A través de SEPA, se espera que los establecimientos puedan disponer de información técnicamente validada que les permita monitorear sus resultados, y tomar decisiones de focalización de acciones pedagógicas a partir de las fortalezas y debilidades reportadas por la evaluación. Se anticipa que los datos proporcionados pueden ser utilizados para una diversidad de tareas, como por ejemplo, para el diseño o selección de actividades de capacitación para los profesores, modificación de planificaciones y estrategias pedagógicas empleadas en cada curso de acuerdo al perfil de rendimiento de sus alumnos, o desarrollo de actividades que atienden a las necesidades de alumnos particulares.

El sistema busca combinar las bondades de la rigurosidad que brinda la implementación de un sistema de evaluación externo a través de pruebas estandarizadas con la posibilidad que sean los actores escolares los que analicen e interpreten la información, y tomen las decisiones que estimen pertinentes según los procesos que estén dentro sus posibilidades. A través de este diseño se pretende que SEPA no sea percibido como amenazante y gatille respuestas defensivas en los destinatarios. Por otra parte, la regularidad con que se levanta la información (una vez al año) y el nivel de desagregación con que se entregan los resultados (por nivel, curso, alumno, subsector y eje temático) busca facilitar el hecho que la información pueda ser utilizada con fines formativos.

El artículo está estructurado de la siguiente manera: Primero revisamos la literatura relevante a nivel internacional respecto a la evaluación educacional, sus formas de uso, y factores que dificultan el uso de evaluaciones educacionales externas con fines formativos. Seguimos con la descripción del problema de investigación y de la metodología del estudio. Finalmente, se presentan los resultados y conclusiones del trabajo.

1. REVISIÓN DE LA LITERATURA

1.1. Distinciones importantes en la evaluación educacional

El proceso de enseñanza-aprendizaje es un proceso interactivo que requiere que los profesores tengan una idea de cómo progresan sus estudiantes y cuáles son las dificultades que enfrentan de modo de adaptar su trabajo a estas necesidades. En este proceso, la evaluación es esencial. La evaluación se refiere a todo proceso que genera información que puede ser utilizada para retroalimentar la práctica (Wiliam y Black, 1998). En este sentido Scriven (1967), diferencia entre *las evaluaciones sumativas* y *las evaluaciones formativas*.

El tipo de evaluación más conocido, es la *evaluación sumativa*. Las evaluaciones sumativas se desarrollan con el objetivo de medir que es lo que los estudiantes han aprendido al final de una unidad, para promover a los estudiantes de curso y para asegurarse que estos hayan cumplido con los requisitos mínimos para merecer una determinada certificación. Las evaluaciones sumativas se utilizan para producir información comparativa entre establecimientos en diversos contextos (Wiliam y Black, 1996). En un contexto de presión por que establecimientos educacionales demuestren resultados, las evaluaciones sumativas son crecientemente utilizadas para promover su responsabilización ("accountability") por estos (OECD, 2005). Sin embargo, el hecho se reciban resultados provenientes de evaluaciones sumativas, no quiere decir que estos sean utilizados para tomar decisiones orientadas al mejoramiento del desempeño de la organización. Por otra parte las evaluaciones se puede usar con fines formativos (OCDE, 2005), esto es, para redireccionar la gestión del establecimiento escolar.

La *evaluación formativa*, se refiere a un proceso interactivo y usualmente más frecuente, orientado a caracterizar el proceso y nivel de comprensión de los estudiantes de una materia determinada (OCDE, 2005; Ruiz-Primo y Furtak, 2006; Looney, 2007). A través de las evaluaciones formativas, se busca por ejemplo, que los profesores puedan ajustar sus estrategias de enseñanza en función de las necesidades emergentes de los estudiantes. A diferencia de la evaluación sumativa, más que para servir como indicador del desempeño de los estudiantes, la información producida busca ser utilizada para acomodar el proceso de enseñanza-aprendizaje (Wiliam y Black, 1998; Wiliam, Lee, Harrison y Black 2004; Wiliam, 2006). La evaluación formativa es considerada como una herramienta que permite mejorar efectivamente

el desempeño en los estudiantes (Wiliam y Black, 1998; Wiliam, Lee, Harrison y Black 2004) y desarrollar sus habilidades de “aprender a aprender” (OCDE, 2005).

Desde el sentido común se esperaría que la información que produce una evaluación formativa fuese analizada por el equipo responsable del establecimiento y utilizada directamente para tomar decisiones orientadas a mejorar los aspectos críticos detectados. Este razonamiento supone que la entrega de información es un ejercicio neutro, que su interpretación será uniforme y que la entrega de esta información no producirá consecuencias negativas para ninguno de los actores involucrados. Sin embargo, normalmente estos supuestos no se cumplen.

Otra distinción relevante en torno a los enfoques de evaluación educativa tiene que ver con quien está a cargo de esta o, en otras palabras, si el evaluador se encuentra dentro o fuera del programa o de la institución que será evaluado. Por un lado, los evaluadores externos no tienen ninguna relación permanente con el programa, por lo tanto no están subordinados a personas dentro de este. Su valor radica en que pueden ser independientes, por ende son más creíbles. En cambio en el caso de los evaluadores internos, se presume que están expuestos a ser presionados a hallar resultados positivos. Sin embargo las evaluaciones internas tienen una serie de ventajas, entre ellas, el que son menos amenazantes, permiten desarrollar la capacidad de auto-monitoreo de la organización, empoderando a los actores escolares. También promueven la reflexión interna y permiten incorporar información relevante del contexto, siendo así de mayor utilidad para el programa o la institución evaluada (Nevo, 2001; McDowell, Smailes, Sambell, K., Sambell, A. & Wakelin, 2008). En cambio, si bien las evaluaciones externas suelen ser más creíbles, también tienen mayores dificultades para incorporar información contextual relevante (Nevo, 2001). Usualmente las evaluaciones sumativas están a cargo de personas externas al programa, mientras que las evaluaciones formativas suelen desarrollarse en forma interna. Sin embargo, en el caso de SEPA tenemos un sistema de evaluación externa, pero con un enfoque en un uso formativo de sus procesos y resultados.

Como se puede apreciar, la evaluación no es un proceso unívoco, pueden existir diversos enfoques evaluativos según la finalidad y la ubicación del evaluador. Estas distinciones son relevantes porque son opciones para la implementación de evaluaciones en el ámbito educativo e influyen en la relación que establecen las organizaciones con la información evaluativa.

1.2. Formas de uso de la evaluación

Una larga tradición de investigación en el uso de información evaluativa sugiere que la relación entre el input de información sobre el desempeño y su uso para la toma de decisiones no es lineal, y que los posibles usos de la información son múltiples (Shula y Cousins 1997; Weiss, 1998; Alkin & Taut, 2003).

En este sentido Shula y Cousins (1997) estiman que es un error ignorar que el conocimiento es socialmente construido, y por ende inextricablemente ligado al contexto. Los autores observan que ante la entrega de información evaluativa, un destinatario tiene múltiples posibilidades de respuesta. Este puede optar por utilizar la información evaluativa entregada, abstenerse de utilizarla justificadamente, utilizarla en forma malintencionada (“misuse”), utilizarla en forma equivocada (por ejemplo usando datos de mala calidad) y utilizarla en forma abusiva (por ejemplo omitiendo la calidad de los datos para fines políticos). Los autores indican, que si bien desde una perspectiva externa el mal uso de la información parece ser algo reprochable, en la medida que se toma en cuenta el contexto, resulta cada vez más complejo hacer juicios taxativos sobre lo que es un buen o un mal uso.

Por su parte Weiss (1998), señala que no solamente es posible utilizar múltiples aspectos de una evaluación particular, sino que también se pueden utilizar las señales producidas por el mismo hecho de evaluar. El "usar" una evaluación implica considerar múltiples efectos posibles para múltiples usuarios de esta información. El fenómeno del uso puede ser muy complejo y cambiar para cada contexto organizacional. En este sentido Rich (1977) distingue el uso instrumental del uso conceptual. El uso instrumental refiere al uso tradicional, donde la información proporcionada es utilizada directamente para tomar decisiones. Según Weiss (1998) el uso instrumental se ve facilitado cuando los resultados son poco controvertidos, cuando los cambios sugeridos se encuentran dentro del repertorio de las posibilidades de acción de la organización y el ambiente del programa es relativamente estable o cuando la organización está en crisis. Por otra parte, el uso conceptual se refiere al uso de la información evaluativa con la finalidad de ampliar la comprensión de un problema determinado, pero no necesariamente para producir consecuencias concretas. Weiss (1998) extiende esta diferenciación, distinguiendo además el uso como herramienta de persuasión, es decir se utilizan los resultados para movilizar apoyo a una posición que ya se tiene sobre los cambios que se requiere implementar.

Finalmente Owen y Rogers (1999), distinguen entre el uso para legitimación ("legitimative use") y el uso simbólico ("symbolic use"). En el primer caso, los resultados de la evaluación se utilizan para respaldar o legitimar decisiones previamente tomadas, usualmente para demostrar que estas eran las correctas. Por otra parte, el uso simbólico busca legitimar acciones previas sin necesariamente referirse a los resultados de la evaluación. En este caso, la evaluación puede conducirse, por ejemplo como un acto simbólico que busca mejorar la imagen de la institución o del tomador de decisiones.

Los tipos de uso están anclados en el contexto de la organización destinataria de la información, por lo que sus características son fundamentales para comprender sus formas de uso. Por ello un evaluador debe ser sensible a las preocupaciones de las personas que están en los programas que evalúan (Weiss, 1998) y preocuparse por ayudarlos a interpretar los resultados. Pese a que hay factores fuera del control del evaluador, este tiene la responsabilidad de hacer lo posible por asegurar la utilidad y facilitar el uso de las evaluaciones que conduce (Joint Committee on Standards for Educational Evaluation, 1994).

Así como se ha podido observar, la información evaluativa puede ser utilizada de múltiples formas y finalidades. El uso instrumental para tomar decisiones con el fin de mejorar la calidad de los procesos de enseñanza aprendizaje es solamente uno de los usos posibles. El evaluador debe estar consciente de la multitud de usos, siempre tomando en cuenta el contexto en que opera.

1.3. Factores que dificultan el uso de la evaluación

Entendiendo que existen numerosas formas en que se puede utilizar la información evaluativa, se han descrito determinados factores que puede facilitar u obstaculizar el su uso. Si bien la evaluación (sumativa o formativa), se ha reconocido como fundamental para la promoción del progreso del aprendizaje, el convertir los resultados evaluativos en prácticas efectivas es un proceso que enfrenta numerosas dificultades (McDowell, Smailes, Sambell, Sambell, Wakelin, 2008; Wildawsky, 1972). La OCDE (2005) también reconoce que la evaluación formativa, pese a sus beneficios identificados, no es sistemáticamente practicada.

La introducción de un sistema de evaluación externa puede enfrentar resistencias que pueden frenar su uso, particularmente en el caso de los docentes. Se han descrito diversos factores asociados a la resistencia a la evaluación, tales como características del evaluado, por ejemplo: un menor número de años de formación, una mayor edad (Navarro, 2003) y el no haber tenido experiencias previas con

evaluaciones externas (Donaldson, Gooler & Scriven, 2002; Taut & Brauns, 2003). Asimismo Funderburg y Levy (1997) plantean que aquellos individuos que tienen un locus de control interno, reciben de mejor manera la retroalimentación dado que la sienten más aplicable a su comportamiento. Asimismo la evaluación puede percibirse como una amenaza para la autoimagen del evaluado (Donaldson, Gooler & Scriven, 2002). Se puede percibir que las evaluaciones proveen retroalimentación respecto al mismo y no solamente sobre el desempeño (Taut & Brauns, 2003). Por otro lado, muchos sujetos se resisten porque experimentan temor debido a que se anticipan a posibles consecuencias negativas de los resultados evaluativos (Donaldson, Gooler & Scriven, 2002). La resistencia también se asocia a comportamientos del evaluador o características del entorno en el cual la evaluación tiene lugar, conocidos como factores situacionales. Entre estos se cuentan la preexistencia de una cultura contraria a la evaluación, la existencia de conflictos de poder e interés, y la desconfianza del instrumento, de las características del evaluador, de las estimaciones que este hará, o de la interpretación que se hará de los resultados (Taut y Brauns, 2003). Un tercer grupo de factores asociados a la resistencia, tienen que ver con la evaluación propiamente tal. Por un lado se encuentra la ambigüedad de rol entre los evaluadores y quien financia la evaluación (Donaldson, Gooler & Scriven, 2002), o la ambigüedad asociada al desconocimiento o ausencia de estándares de los niveles de desempeño (Milanowski & Heneman, 2001). La desconfianza puede manifestarse también como desacuerdo con el enfoque, método y prácticas evaluativas utilizadas (Navarro, 2003; Taut & Brauns, 2003).

La revisión de la literatura sugiere que el uso de la información evaluativa depende de una serie de factores entre los cuales se pueden mencionar los siguientes:

- **Problemas de comprensión relativa al uso de datos cuantitativos.** El tipo y la calidad de información proporcionada para establecer los logros académicos de los alumnos ha resultado ser un factor relevante para su uso. Antiguamente las decisiones sobre la orientación de los procesos de enseñanza y aprendizaje se solían tomar fundamentalmente basadas en el juicio profesional los actores escolares. Se asumía entonces que estos tenían la responsabilidad y el derecho a tomar decisiones sobre los estudiantes, escuelas y la educación en general. Para ello se basaban en una combinación de conocimiento tácito y privilegiado del contexto, la sabiduría política, la experiencia y el análisis lógico. En este contexto los datos no jugaban un rol relevante en la toma de decisiones. En la actualidad esto ha cambiado y la presión por rendir cuentas por las acciones de los establecimientos hace necesaria la producción y el uso de datos cuantitativos. Esto ha enfrentado a los actores escolares a la necesidad de poder comprender estos datos y tratar de darles sentido (Earl y Fullan, 2003).

Se ha observado que los actores escolares, aunque altamente motivados por el cambio, suelen mostrar bajas capacidades para interpretar y manejar datos cuantitativos (Earl y Fullan 2003; Taut, Cortés, Sebastian & Preiss, 2009). Esto sería aun más complejo en el caso del uso de índices de mayor complejidad como el valor agregado (Santelices & Taut, 2009). Aunque en estos estudios se valora y se percibe la información como claramente presentada, es al momento de la interpretación concreta donde se observan claras limitaciones de comprensión.

- **Foco exclusivo en una evaluación solamente externa o interna.** Fullan (2006) estima que la sola introducción de indicadores de logro del aprendizaje (externamente, por parte de la política educacional) no es suficiente para promover el cambio en establecimientos educacionales, porque suelen no tomar en cuenta las condiciones locales. Asimismo, este tipo de indicadores suelen estar diseñados para satisfacer las necesidades de información de tomadores de decisiones, no estando conectados con lo que sucede al nivel del aula de clases, siendo poco informativos para los docentes.

Nevo (2001) postula que es necesario desarrollar un equilibrio entre evaluaciones internas y externas. Si bien la introducción de sistemas de evaluación externa estimula el desarrollo de evaluaciones internas, el impacto de un sistema de evaluación externo depende de la existencia previa de una cultura de evaluación. Equipos acostumbrados a evaluarse, tendrán menos dificultades para comprender el significado y relevancia de una evaluación externa, y estar menos a la defensiva.

- **Baja capacidad organizacional.** Las organizaciones suelen responder a las evaluaciones con acciones que se encuentran ya dentro de su repertorio de acciones posibles y hasta entonces conocidas (Shula y Cousins, 1997), existiendo usualmente una brecha relevante entre lo que se sabe hacer y lo que debería hacerse para mejorar (Elmore, 2008). La capacidad de la organización (Fullan, 2002) está dada por elementos como el liderazgo del director, la orientación al desarrollo de habilidades de los docentes, la calidad de la comunidad profesional, la coherencia del programa académico y la disponibilidad de recursos técnicos. Fullan (2002) agrega que la mayor parte de las estrategias de mejoramiento escolar se orientan pobremente a la construcción de capacidad escolar. En este sentido Elmore (2008) agrega que el desarrollo de conocimientos y habilidades tanto individuales como colectivas es un proceso que ocurre en forma gradual y sin seguir una trayectoria lineal.

- **Cultura de la evaluación.** Un elemento relacionado con la capacidad organizacional, tiene que ver con lo que se ha denominado la "cultura de la evaluación" (OCDE, 2005). Una cultura de evaluación facilita el uso de las evaluaciones y la introducción de cambios consecuentes de mejoramiento. En una organización en la cual se puede observar una cultura de la evaluación, los actores escolares utilizan la información para generar conocimiento sobre que funciona y porque, comparten dicho conocimiento con sus colegas y se preocupan por desarrollar sus habilidades para abordar un rango cada vez más amplio de necesidades de aprendizaje de sus estudiantes. En estas la evaluación formativa es considerada como una herramienta para alcanzar metas de alto desempeño, equidad y enseñanza de habilidades de "aprender a aprender" (OCDE, 2005).

Scheerens (2004) propone que existen obstáculos cognitivos, emocionales, interpersonales, estructurales y procedurales que pueden dificultar el uso de resultados de evaluación interna para el aprendizaje organizacional. Por ejemplo, si la información ofrecida es vaga y carece de una retroalimentación instrumental a nivel de aula, y si además el propósito de la evaluación es de llegar a juicios en vez de una orientación a la tarea, se tiende a observar ausencia de aprendizaje a través de la evaluación. A nivel psicológico, el autor menciona la ansiedad frente la evaluación en combinación con un sentido de baja auto-eficacia como un escenario negativo (ver también Taut & Brauns, 2003). Finalmente, a nivel estructural-procedural, la baja transparencia en la entrega de información, el autoritarismo en la toma de decisiones, la insuficiencia de espacios para discutir en equipo los resultados, son obstáculos relevantes para el uso.

De lo anterior es posible destacar que el tipo de uso de información evaluativa se ve afectado por numerosas variables que no tienen que ver solo con la calidad técnica de la información o del diseño de evaluación mismo. Esas variables "extra técnicas" – sobre todo organizacionales – tienen un rol determinante en el modo en que se utilizará la información.

3. EL PROBLEMA DE INVESTIGACIÓN

A partir de la literatura revisada, se espera que la introducción de SEPA enfrente dificultades respecto a su comprensión y uso por los establecimientos o fundaciones contratantes. La presente investigación tenía como objetivos (1) determinar el grado de comprensión de los conceptos de estatus, progreso, y valor agregado por parte de usuarios intencionados de SEPA; (2) describir el uso de los distintos tipos de información SEPA en los establecimientos participantes, en relación a su uso de prácticas evaluativas en general; (3) identificar las consecuencias de corto plazo de la entrega de la información SEPA; (4) identificar factores que facilitan o dificultan el uso de la información SEPA; y (5) formular recomendaciones que permitan mejorar el servicio de entrega de información SEPA.

Derivados de estos objetivos se buscó responder las siguientes preguntas de investigación:

- ¿Cuál es el grado de comprensión que tienen directivos y profesores respecto de los conceptos de estado, progreso y valor agregado?
- ¿Cuál es la utilidad percibida que tiene esta información para directivos y docentes? Se han cumplido las expectativas (por qué, por qué no)?
- ¿Para qué (propósitos) y cómo (procesos) se ha utilizado la información?
- ¿De qué depende el hecho que la información sea utilizada?
- ¿Qué consecuencias de corto plazo ha tenido la entrega de esta información en el establecimiento (positivas, negativas)?
- ¿Cómo se relaciona la información SEPA con otros tipos de información sobre el aprendizaje de los alumnos (p.ej., SIMCE; evaluaciones del profesor en la sala de clase)?
- ¿Cómo mejorar la claridad de la entrega de la información? ¿Qué tipo de información todavía les falta?

De modo de focalizar mejor el presente artículo, solo se presentarán en mayor detalle los resultados de las preguntas relacionadas con el uso de la información. Los resultados relacionados a la comprensión se encuentran en forma resumida en este paper, y en mayor detalle en Hein y Taut (2009).

4. METODOLOGÍA

4.1. Muestra

En el año 2008 se encontraban afiliados N=101 establecimientos a SEPA, con una total de cerca de N=25.000 alumnos tomando pruebas SEPA. N=34 de los establecimientos participantes (cubriendo algo más de N=5.000 alumnos) han participado por segundo año consecutivo, por lo que recibieron la información completa sobre estado, progreso y valor agregado. De estos 34 establecimientos, 9 son de dependencia municipal (Mun.), 12 de dependencia particular subvencionada (PS) y 13 de dependencia

particular pagada (PP)¹. Dentro del universo así definido, se desarrolló un screening contactando directamente a 21 establecimientos de la Región Metropolitana, para identificar establecimientos con altos niveles de uso instrumental (medidas o decisiones concretas) y establecimientos de bajo uso (ninguna medida o medidas poco concretas) de la información SEPA. Se buscó seleccionar en cada dependencia (municipal, particular subvencionado, particular pagado), dos casos de alto uso (A) y un caso de bajo uso (B).

La siguiente tabla muestra los detalles respecto a los N=7 establecimientos que finalmente se seleccionaron para el estudio.

FIGURA 1. DISTRIBUCIÓN DE ENTREVISTAS SEGÚN CARGO, ESTABLECIMIENTO, NIVEL DE USO Y SOSTENEDOR/ASESOR QUE CONTRATA EL SERVICIO

Cargo	Colegio	Dependencia	Uso	Sostenedor/asesor
Director	EC	Mun	A	Em.
Profesor matemáticas	EC	Mun	A	Em
Director	G	PP	A	Particular
Profesor matemáticas	G	PP	A	Particular
Director	RM	PP	A	Particular
Profesor matemáticas	RM	PP	A	Particular
Profesor lenguaje	RM	PP	A	Particular
Coordinador académico	SM	PP	A	EUC.
Profesor lenguaje	SM	PP	A	EUC.
Coordinador académico	SE	PS	B	EUC.
Profesor lenguaje	SE	PS	B	EUC.
Profesor matemáticas	EB	PS	B	EUC.
Coordinador académico	EB	PS	B	EUC.
Coordinador académico	IOM	PS	B	EUC.

En cada uno de los establecimientos seleccionados se entrevistaron al menos a dos personas (un directivo y un profesor de matemática o lenguaje con un curso evaluado) por cada establecimiento, 14 entrevistas en total. La saturación conceptual se alcanzó para cada establecimiento con dos entrevistas, y para los temas del estudio general se alcanzó con las 14 entrevistas.

Las limitaciones de la muestra alcanzada son los siguientes: (a) Por dificultades de acceso a establecimientos municipales (acceso restringido por la fundación asesora, la cual contrata el servicio de SEPA) solo se pudo incluir a N=1 establecimiento municipal en la muestra final; (b) La distribución entre dependencia y el nivel de uso en la muestra final no se encuentra adecuadamente equilibrada. Esto es, N=3 establecimientos particulares pagados muestran altos niveles de uso versus N=3 establecimientos particulares subvencionados con bajos niveles de uso.

Es importante hacer notar que en N=5 de estos establecimientos el servicio SEPA era contratado por un sostenedor o asesor externo (fundaciones) al establecimiento, quien daba uso adicional a los datos para perseguir sus objetivos respecto de los establecimientos evaluados.

¹ Los establecimientos de dependencia municipal son establecimientos públicos administrados por cada gobierno local. Los establecimientos de dependencia particular subvencionada son establecimientos pagados con subvención estatal y los establecimientos de dependencia particular pagada son establecimientos privados y pagados.

4.2. Recolección y análisis de datos

Se levantó información a través de la técnica de una entrevista semi-estructurada. Esta tuvo un carácter “interventivo”, ya que exploró primero reacciones y conocimientos iniciales ante un tema, luego entregó información, y finalmente exploró las reacciones ante la información entregada. Además de las entrevistas, se tomaron notas reflexivas y se elaboraron “mini-ethnografías” sobre las visitas en cada uno de los siete establecimientos en la muestra.

Las entrevistas se orientaron a través de una pauta que buscaba explorar los siguientes temas principales: 1) la caracterización del establecimiento (estrategias previas de promoción de progreso del aprendizaje), 2) los modos de uso de la información SEPA (incluyendo obstaculizadores y facilitadores), y 3) la comprensión de indicadores específicos entregados por SEPA (en base de materiales con resultados seleccionados y reales de cada establecimiento visitado).

Los datos son analizados basados en la metodología de la “Grounded Theory” (Glaser y Strauss, 1999). La utilización de este tipo de técnicas estuvo orientada a recabar información sobre las percepciones que los entrevistados tienen sobre el uso de SEPA en sus respectivos establecimientos. También se elaboró un modelo teórico que explica el fenómeno del alto o bajo uso, desde la propia perspectiva de los entrevistados, que es justamente el propósito de las técnicas planteadas por la Grounded Theory. Se aplicaron los tres tipos de codificación propuestas por ella (abierta, axial, selectiva). Para tales efectos se codificaron las respuestas de los entrevistados y se fue construyendo un sistema de categorías a través de un software especializado en análisis cualitativo (ATLAS.ti, versión 5.2). Una vez que la información fue codificada, se pudo graficar las relaciones encontradas entre las categorías.

4.3. Limitación de generalización

Es necesario explicitar que los hallazgos y conclusiones de este estudio se limitan a los establecimientos SEPA y no son extrapolables a todos los establecimientos educativos de Chile.

5. RESULTADOS

5.1. Resumen de resultados sobre comprensión de la información SEPA

En términos generales, es posible observar que la comprensión de los conceptos y de la información proporcionada suele ser parcial, particularmente en cuanto a los conceptos más novedosos y complejos. Los datos que con mayor facilidad se comprenden son los de estado o estatus, particularmente en el caso de las estadísticas expresadas en porcentajes de logro. Los usuarios están familiarizados con la expresión de rendimiento en porcentajes, por lo que los consideran fáciles de leer y concretos en cuanto a su interpretación. En cambio, la “escala vertical” suele no ser comprendida y, consecuentemente, tampoco los son los puntajes estandarizados para los cuales no se cuenta con herramientas de interpretación. Pese a ello, los entrevistados no consideran los porcentajes de logro son suficientes para tomar decisiones. Los indicadores de progreso muestran mayor dificultad inicial para ser comprendidos. Bastó una breve explicación orientada a proporcionar una norma de interpretación para interpretar los datos. Una vez comprendidos, se los percibe como útiles para lograr monitorear el progreso del establecimiento, los cursos y los alumnos.

Finalmente, en cuanto al índice de valor agregado, la mayor parte de los entrevistados entiende de su existencia, les parece inicialmente interesante, pero la comprensión del concepto y su interpretación es parcial. Una vez que se ha explicado, los datos de valor agregado se vuelven particularmente atractivos para directivos. En caso que los resultados fuesen buenos, se estima que estos pudiesen ser mostrados a los padres, como parte de una estrategia para promover la confianza de estos en la gestión del establecimiento.

5.2. Tipos de usos observados

Un primer uso observado, y el más común de todos, tiene que ver con el *uso conceptual*, a saber el *análisis cruzado de la información*. Gran parte del contacto con la información se da en el contexto del análisis que busca relacionar los resultados observados en SEPA con los resultados observados en las evaluaciones sumativas internas, otras evaluaciones externas (SIMCE, otras) y la experiencia de docentes y directivos. Los equipos testean hipótesis, ensayando explicaciones para comprender las variaciones observadas. En esta etapa, se pone a prueba el sistema, en términos que si los resultados son coherentes con lo esperado, se convierte en una importante fuente de legitimidad del instrumento. En general, todos entrevistados coinciden en que los resultados observados han sido coherentes con la experiencia, y cuando no ha sido así, estas discrepancias pudieron ser explicadas satisfactoriamente.

La segunda medida más utilizada, tiene que ver con un tipo de *uso instrumental*, a saber la *revisión de planificaciones* del establecimiento, nivel, curso o algún alumno particular. En la medida en que el sistema ha ganado credibilidad a través del análisis cruzado de la información, los actores toman decisiones al respecto. Lo más común es que se repitan o enfatizen determinadas materias dedicándoles un mayor número de horas, o se convoque a talleres de reforzamiento a grupos de alumnos particulares.

Un tercer uso observado (instrumental), donde la revisión cruzada arroja discrepancias con la impresión de los actores y de las evaluaciones sumativas internas, es la *revisión y mejoramiento de los mecanismos de evaluación interna*. En dos establecimientos fue necesario desarrollar capacitaciones en construcción de instrumentos de evaluación a los docentes para instalar la capacidad de construir instrumentos de evaluación educacional y utilizarlos con fines formativos. Anecdóticamente, en un establecimiento esto trae como consecuencia el descenso general del promedio de notas de los alumnos. En ninguno de los casos mencionados estos cambios fueron desarrollados espontáneamente. En ambos casos se requirió de apoyo o presión externa (según el caso) que obligara a los equipos a aceptar las problemáticas detectadas y comprometerse con el cambio requerido.

Una cuarta forma de uso (instrumental), tiene que ver con la *detección de alumnos con dificultades o con talentos*. En realidad más que detectar alumnos con dificultades particulares, en opinión de los entrevistados la información SEPA permite confirmar que alumnos que suelen tener mal rendimiento académico también obtienen pobres resultados en las pruebas SEPA. Dado que existe una relación relevante entre los resultados y lo que los profesores suelen esperar de sus alumnos, el aporte más relevante que hace el SEPA en este aspecto, tiene que ver con la detección de alumnos que tienen rendimientos discrepantes con lo esperado. Particularmente en un establecimiento también se logró detectar un talento insospechado a través de las pruebas SEPA. En este último caso se opta por potenciar las habilidades del alumno con asignaciones especiales en matemáticas. Como resultado final el alumno, previamente de rendimiento mediocre mejora su desempeño en todas las asignaturas.

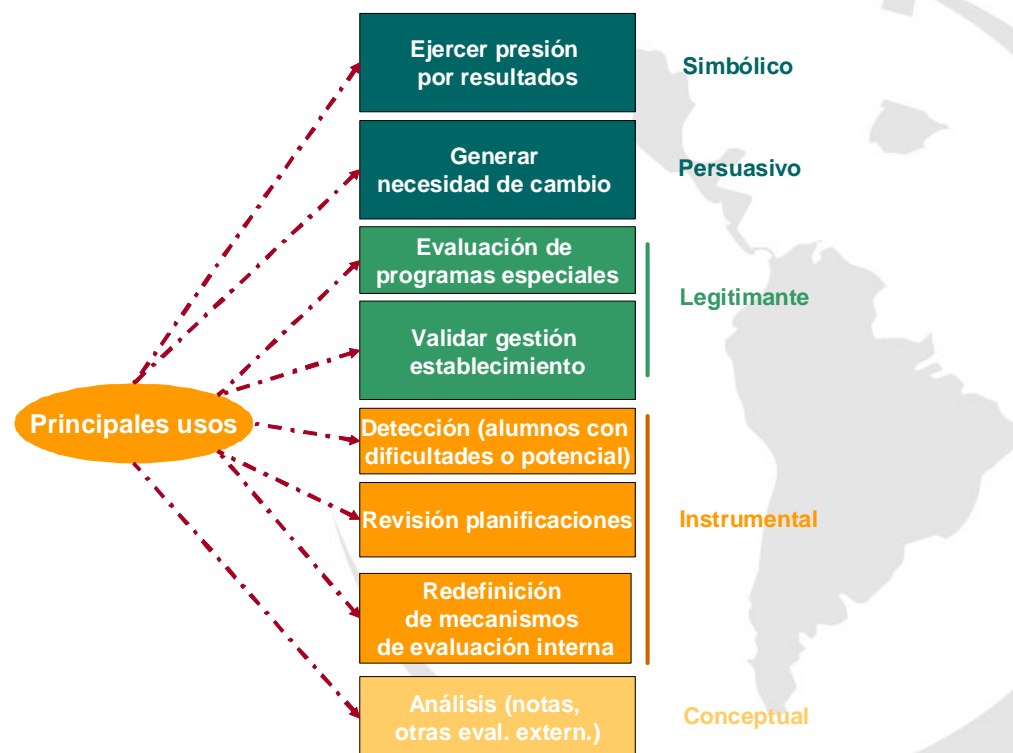
Una quinta forma de uso, de tipo legitimante, se observa en equipos directivos y tiene que ver con la *validación de la gestión de la organización*. Específicamente, en el caso que los resultados obtenidos son

valorados como positivos, estos resultados son utilizados para demostrar, que la gestión del establecimiento ha sido la adecuada. En específico, puede ser utilizada, ya sea para presentar los resultados ante los padres, o simplemente como forma de reafirmar internamente que el desempeño ha producido el efecto deseado. Esto cobra relevancia en establecimientos que son total o parcialmente financiados por apoderados, donde y mantener la confianza de los apoderados es clave garantizar la estabilidad de la matrícula.

Una sexta forma de uso tiene que ver con la *evaluación de los resultados de programas pedagógicos especiales* implementados con anterioridad a la introducción de SEPA. En principio si bien esto parece un uso instrumental, el desarrollo de esta evaluación tuvo un carácter legitimante, debido a que la instalación del nuevo programa pedagógico estuvo lleno de controversias.

Un séptimo uso detectado, esta vez desde los asesores externos-clientes de SEPA, tiene que ver con la utilización de los resultados generados para motivar a los equipos directivos y docentes para *generar la necesidad de un cambio (uso persuasivo)*. Las organizaciones externas que asesoran o de las cuales dependen los establecimientos, estaban particularmente interesadas en convencer a los directivos y docentes de los establecimientos, de que era necesario hacer un cambio. Más allá de estos casos, se observa una alta preocupación por el modo que se transmitirán resultados de la evaluación al cuerpo docente. Esto es así tanto en establecimientos con restricciones a la movilidad docente (establecimientos municipales) como en establecimientos que no tienen estas restricciones (particulares y particulares pagados). Se estima que si el cuerpo docente no legitima la evaluación, no se podrán implementar acciones a partir de esta. Es interesante destacar que la legitimidad que goza la institución a cargo de SEPA resulta ser un elemento que respalda a los equipos internos y externos encargados de "bajar" la información, particularmente en el caso que los resultados observados no son los esperados, momento en que se refieren mayores resistencias en los equipos docentes.

FIGURA 2. TIPOS DE USOS DE LA INFORMACIÓN SEPA DETECTADOS EN ESTABLECIMIENTOS DEL ESTUDIO



Finalmente un último uso observado tiene que ver con *la presión por resultados* (accountability) aunque en un formato inesperado. En un caso particular, se pudo observar como el director de un establecimiento era asesorado por un organismo externo cuyos honorarios eran financiados por un donante conseguido por este mismo director. Al tomar conocimiento de los índices no favorables de valor agregado, el director utilizó esta información para presionar al asesor externo y responsabilizarlo por la falta de resultados, a tres años del inicio de la consultoría. Podría decirse que es un uso de tipo *simbólico*, ya que si bien se refiere a los datos, tiene que ver más bien con el envío de una "señal" más que en acciones concretas (como por ejemplo lo hubiese sido despedir a la consultora).

La Figura No. 2 ofrece un resumen gráfico de los usos detectados.

5.3. Condiciones que favorecen o dificultan el uso de información evaluativa

Se pudieron hallar una serie de condiciones que pueden facilitar o dificultar el uso de la información evaluativa. En primera instancia vale la pena destacar cual es la relación que se observa entre el uso de la información SEPA y la estrategia previa que tenía la escuela para promover el progreso del aprendizaje de los alumnos. En segunda instancia se identifican factores particulares que facilitan o dificultan el uso de la información SEPA.

5.3.1. Relación entre el tipo de uso y la estrategia previa de progreso del aprendizaje del establecimiento

- **Estrategia previa de promoción del progreso del aprendizaje.** Frente a los desafíos que se les plantea a los establecimientos para promover el progreso del aprendizaje de sus alumnos, estos desarrollan diversas estrategias. Una "estrategia" consta de al menos tres elementos, a saber de herramientas de detección de alumnos o cursos que muestran dificultades en determinados contenidos, las estrategias de análisis/toma de decisiones de intervención y finalmente cuales son los tipos de respuestas disponibles existentes. Los establecimientos suelen tener una estrategia que varía en cuanto a su formalidad y capacidad de respuesta personalizada a las necesidades de alumnos particulares.

En cuanto a los mecanismos de detección, los establecimientos recurren principalmente al uso de evaluaciones estructuradas (sumativas o formativas internas o externas) o la detección a través de profesionales de apoyo (por ejemplo el departamento de orientación). Los establecimientos que utilizan pruebas sumativas internas o de evaluaciones externas con fines formativos, suelen tener mayores niveles de uso de la información SEPA y muestran un discurso fuertemente orientado a la tarea. Estas herramientas permiten monitorear el progreso de los estudiantes, tomar decisiones para la implementación de medidas remediales y comunicar los resultados a los apoderados y a los alumnos haciéndolos partícipes de la metas de aprendizaje del establecimiento. Los establecimientos con resultados positivos de VA, tienden a monitorear el progreso con herramientas estructuradas de evaluación.

Los resultados de estas evaluaciones son analizados por diversos actores de la organización escolar (profesores, coordinadores, consejos escolares). Se observa multiplicidad de estrategias que involucran diferencialmente a distintos actores según las características y aprehensiones propias de los equipos directivos.

Una tercera característica relevante de la estrategia tiene que ver con el repertorio de medidas remediales de las que dispone el establecimiento. Las medidas más comunes desplegadas, una vez que se ha detectado alguna discrepancia respecto del progreso esperado, son las siguientes:

- *Revisión de planificaciones:* Se refiere a la modificación de la planificación para alumnos que manifiestan alguna dificultad. En los establecimientos que tienen procesos más formales de evaluación formativa, se hace una revisión grupal regular de las planificaciones de cursos y niveles involucrando al profesor, al coordinador, y a otros profesores pares del nivel.
- *Innovación en métodos pedagógicos:* Hay ocasiones en que se siente la necesidad de desarrollar innovaciones en los métodos pedagógicos para abordar determinado tema o determinado subsector. En estos casos, los establecimientos suelen apoyar la capacitación de los docentes en técnicas determinadas para los ámbitos que se van a reforzar.
- *Desarrollo de reforzamientos grupales:* Otra medida común observada es el desarrollo de reforzamiento dirigido a temas y grupos particulares. El reforzamiento se implementa fuera de las jornadas habituales de clases. Esta suele ser una medida excepcional que regular y suele estar orientada a grupos más que a alumnos específicos.
- *Contrato con padres de alumnos de bajo rendimiento:* Cuando los problemas de rendimiento de un alumno se vuelven relevantes se involucra a los padres. Se asegura su participación a través de "contratos" donde padres, hijos y el establecimiento se comprometen a determinadas acciones orientadas a mejorar el rendimiento del alumno.
- *Inter-consulta con profesionales de apoyo:* En casos de mayor complejidad, el problema suele ser "externalizado" del aula, es decir se busca apoyo externo para abordar el caso consultando profesionales internos o externos del establecimiento (psicólogo, psicopedagogo, orientador, neurólogo) si es que estos están disponibles.

Los establecimientos tienden a preferir el uso de medidas remediales con un foco grupal, opción que facilita la gestión de respuestas y responden a problemas generalizados en el establecimiento, nivel o curso. La capacidad de los establecimientos de responder a necesidades particulares de los alumnos es usualmente limitada.

Establecimientos con problemas de rendimiento, suelen no tener una estrategia articulada de progreso (particularmente de evaluación y monitoreo). Esto se observa con mayor claridad en dos establecimientos, que comparten además una orientación que enfatiza las necesidades afectivas de los alumnos por sobre sus resultados de aprendizaje.

- **Uso de la información SEPA y estrategias previas de promoción de progreso del aprendizaje.** La estrategia de progreso de aprendizaje con que cuenta el establecimiento previo a la introducción de SEPA parece ser determinante del nivel de uso que se le dará a la información. El tipo de respuestas que implementarán estarán dadas por el repertorio previo de respuestas posibles que tiene el establecimiento. Esto es consistente con la descripción de Nevo (2001).

Dadas las estrategias previas del establecimiento es posible observar tres escenarios, esquematizados en la Figura N°3.

- (1) En un primer escenario, la información proporcionada por el sistema SEPA encaja en la estrategia previa, lo que facilita el uso completo de la información para la toma de decisiones. Esto es que el establecimiento tiene un repertorio de respuestas posibles que permite diseñar respuestas a nivel de establecimiento, curso y alumnos particulares. No solamente se utiliza la

información a nivel de establecimiento o de curso, si no que se diseñan respuestas para alumnos particulares.

(2) En un segundo escenario, la información proporcionada por el sistema SEPA encaja parcialmente en la estrategia previa. Se cuenta con un repertorio de respuestas de orientación grupal (niveles, cursos o grupos), sin contemplar respuestas para alumnos particulares. Cuando se observan problemas de progreso en alumnos particulares, la tendencia es a externalizar la respuesta.

(3) En un tercer escenario, la información proporcionada por SEPA revela una discrepancia relevante con la percepción previa del desempeño del establecimiento, lo que obliga a que el staff revise sus procedimientos. Queda la impresión que cambios relevantes en la estrategia existente solamente son posibles con apoyo externo (mediador), especialmente si el establecimiento se ha visto cuestionado por su bajo rendimiento.

FIGURA 3. RELACIÓN ENTRE EL USO DE SEPA Y LA ESTRATEGIA PREVIA DEL ESTABLECIMIENTO



5.3.2. Factores específicos asociados a un alto uso de la información SEPA

En esta sección se hace una síntesis de los factores que parecen afectar el nivel de uso de la información. Estos pueden ser clasificados según su relación con condiciones previas del establecimiento, o con características propias de SEPA. La síntesis de estos factores se desarrolla en la figura N°4.

Entre las condiciones previas del establecimiento asociadas a un alto uso de la información SEPA, se cuentan:

- *Un alto nivel de motivación de los docentes:* Factor asociado al compromiso y voluntad de los docentes de implementar cambios, particularmente en el nivel de la gestión del trabajo de aula.
- *Un alto nivel de confianza entre directivos y docentes:* Este factor facilita la bajada de la información, la comunicación y discusión de medidas remediales. Un alto grado de confianza no implica que la información se maneje abiertamente, y el manejo abierto de la información

tampoco está necesariamente asociada a un alto uso. Por ejemplo en un establecimiento se observa alto uso (en docentes y directivos) pese a que la información se da a conocer solo parcialmente por temor a un posible mal uso.

- *El interés del director o sostenedor:* Este resulta fundamental para introducir la necesidad de cambio en la organización. El director/sostenedor define la agenda, determina las prioridades y establece los énfasis con que se abordarán determinados temas, incluyendo la evaluación formativa. Si, el interesado (director/sostenedor) debe tener el poder para tomar decisiones y definir consecuencias por incumplimiento.
- *Existen consecuencias asociadas:* Este se refiere a que la información se utiliza para tomar decisiones concretas. Si esto no se da, el análisis de la información se convierte en un ejercicio irrelevante para la práctica docente dando lugar a una situación de bajo uso de la información.
- *El uso de la evaluación formativa ya se encuentra instalado:* Si ya existe el hábito de utilizar herramientas de evaluación formativa, la introducción de SEPA resulta mucho más fácil. Si SEPA compite con otro sistema de evaluación interna, se usa el que tenga consecuencias asociadas.
- *Se dispone de equipos competentes y con tiempo para analizar la información:* La disponibilidad de equipos profesionales con capacidad de interpretación de datos numéricos, competencias informáticas adecuadas (manejo web y planilla de cálculo), y tiempo suficiente para desarrollarlo, resulta ser un elemento que facilita la decodificación de la información entregada a través de la web.
- Como ya se mencionaba anteriormente, la disponibilidad de *mediación interna o externa* que acompañe a los equipos en el procesamiento de la información, parece ser un facilitador asociado a un posterior uso de la información. El que actúa como mediador no solamente debe poder comprender los datos, sino que dominar un rango de habilidades blandas que le permitan manejar y confrontar las resistencias emergentes en el proceso.
- Como también se describió anteriormente en detalle, la forma en que se organiza previamente la estrategia de progreso del aprendizaje del establecimiento, puede facilitar o restringir la aplicación de la información proporcionada. Establecimientos que solamente contemplan un repertorio de respuestas centradas en los grupos, aprovechan menos la información proporcionada.
- Finalmente, en cuanto a la *orientación filosófica de la escuela*, esta facilita el uso, en la medida que centre su atención en el progreso del aprendizaje. Establecimientos centrados solamente en acoger a los alumnos suelen presentar mayores dificultades de rendimiento. Esto puede estar asociado también con el hecho que estos establecimientos reciben a alumnos con historias de vida más desfavorecidas. Otros establecimientos equilibran ambas orientaciones.

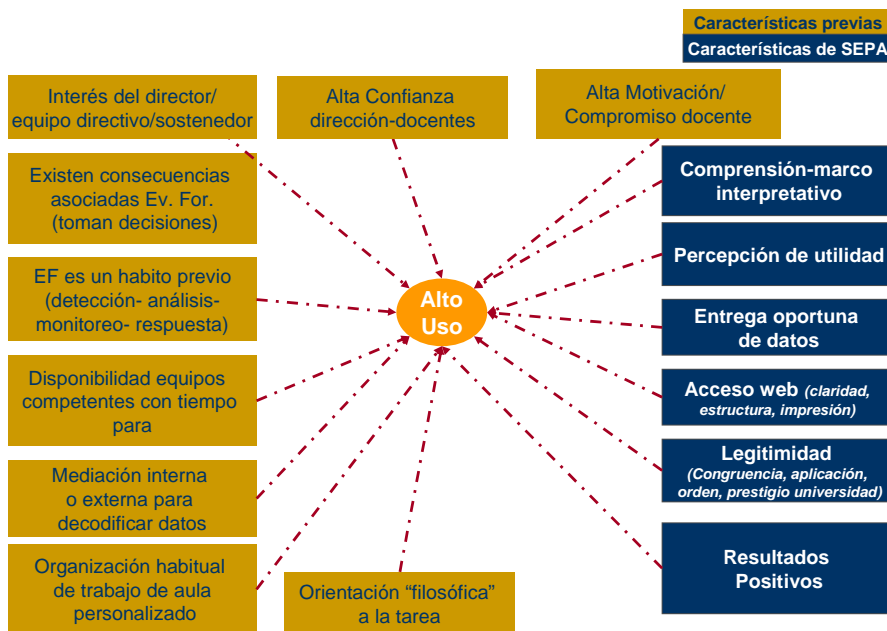
En cuanto a las características asociadas a SEPA se pueden identificar las siguientes:

- *Comprensión de la información:* Se pudo observar que el uso de la información se facilita en la medida que al menos un representante del establecimiento comprende adecuadamente la información proporcionada. Esta variable parece ser crítica. Se constata que la mayor parte de los actores (incluso los alto uso), comprenden solo parcialmente la información proporcionada los conceptos introducidos. Los actores enfrentan diversas dificultades para dar sentido a la

información (procesarla y relacionarla con hechos, procedimientos, conceptos, interpretaciones, ideas, observaciones y juicios propios). Solamente una vez que la información ha sido procesada, es que es posible que sea utilizada (Alavi & Leidner, 2001). La novedad de la herramienta, la falta de tiempo, la forma de organización de los contenidos de la página web (supone que las personas exploran la web en forma secuencial: leer primero la sección definiciones, luego los datos, luego los anexos) y la falta de competencias de análisis dificultan la comprensión. En cambio, el contar con asesoría externa (o equipo interno competente) facilita la comprensión y de este modo el uso.

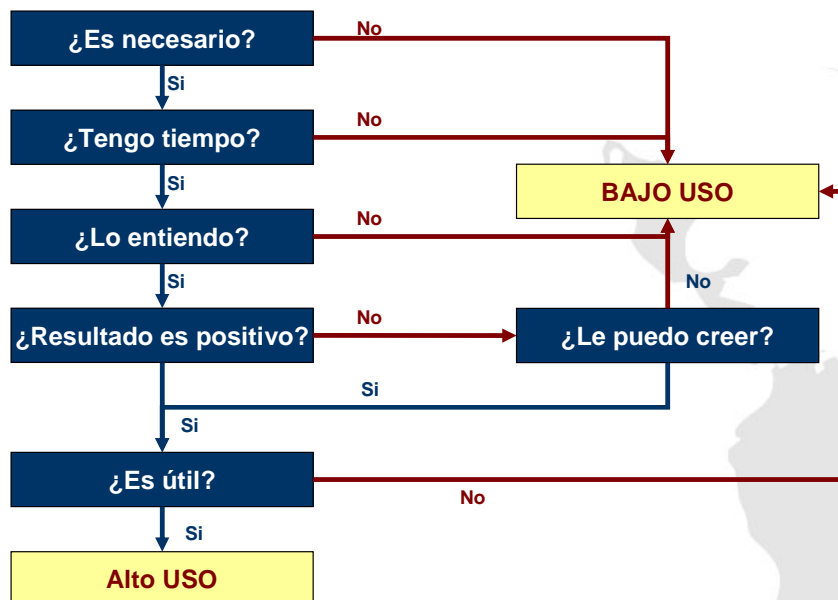
- *Percepción de utilidad de la información:* Se observan usos diferenciados en los actores según las necesidades. Profesores estiman que la información más útil es aquella que desglosa el rendimiento por alumno y /o por subsector. Pese a ello, la especificidad con que SEPA describe los sub-ejes no les parece suficiente para tomar decisiones pedagógicas. En cambio la corrección de preguntas de respuesta abierta se percibe como muy útil, porque permite derivar fácilmente acciones remediales. Por otra parte la información sobre el establecimiento o sobre el valor agregado, es más valorado por directivos que por docentes.
- *Entrega oportuna de la información:* Como distintos tipos de resultados son entregados en dos fechas distintas (diciembre y marzo), aquellos entregados en el segundo momento ("progreso") son menos utilizados. En parte esto se debe a que ya se han analizado los resultados de "estado" y se han tomado decisiones, pero por otra parte se debe a que la planificación del año suele cerrarse en febrero. Si los resultados llegan en marzo es difícil que sean considerados.
- *Acceso web y diseño de la plataforma de entrega de información:* Algunos elementos asociados al acceso de la web como la estructura de la organización de la información pueden llegar a dificultar el acceso y comprensión de los datos. Se observan inconvenientes específicos con la calidad con que se puede imprimir la información (por ejemplo tablas cortadas), y con la exportación de archivos, funcionalidad que no todos los actores conocen. Aun en establecimientos de alto uso, se observan dificultades asociadas al manejo de herramientas informáticas. Si bien el acceso web es valorado, algunos actores preferirían tener la opción de acceder a un informe escrito.
- *Legitimidad:* La legitimidad que tiene SEPA para los destinatarios de la información, es un elemento muy relevante para su uso. Si la herramienta no es creíble, esta no es utilizada. Las fuentes de legitimidad, son la congruencia de los resultados de SEPA con otros instrumentos o experiencia, el orden de la presentación de los datos, el orden con que se desarrolla el proceso de levantamiento de los datos y el prestigio de la institución que desarrolla el sistema. Si hay dudas en alguno de los niveles anteriores, el prestigio de la universidad contribuye a dotar de legitimidad de la herramienta. Esto es un elemento particularmente valorado en el caso que se requiera convencer a los equipos sobre la necesidad de un cambio.
- Un último elemento tiene que ver con el hecho que los resultados *sean interpretados como positivos*. Si esto es así, su uso se facilita. Resultados percibidos como negativos son sometidos a un mayor escrutinio orientado a desacreditar la herramienta de modo de no tener que enfrentar a necesidad de revisar sus procesos.

FIGURA 4. SÍNTESIS DE FACTORES ASOCIADOS A UN ALTO USO DE LA INFORMACIÓN SEPA



5.3.3. Resumen: Como los usuarios determinan si utilizan o no la información SEPA

FIGURA 5. PREGUNTAS CRÍTICAS QUE SE FORMULAN DOCENTES AL ENFRENTARSE A LA INFORMACIÓN SEPA



Pareciera ser que existe una secuencia de preguntas internas que los usuarios finales se hacen al momento de confrontarse con información evaluativa de SEPA. Uno de los elementos más relevantes tiene que ver la necesidad de uso, usualmente mandatada por el director o el sostenedor. Como se mencionó anteriormente, el mandante debe ser una figura con poder para definir consecuencias. Esto se

pudo observar en establecimientos donde convivían dos sistemas de evaluación. Aquel sistema de evaluación que tenía consecuencias para el sostenedor, era el sistema más utilizado. En una segunda instancia entran en juego las condiciones efectivas de tiempo y acceso a la información para poder procesar la información. Luego entran en juego factores de comprensión de los datos. En este momento la mediación se vuelve crucial. Finalmente, se evalúa si los resultados son amenazantes, y si lo son, si estos son creíbles. Solamente en una última instancia se evalúa la aplicabilidad y utilidad de los datos (desagregación, calidad de la descripción de las categorías, etc.). En la figura N°5 se resumen las preguntas internas en formato de flujograma.

6. CONCLUSIONES

Tal como se anticipó, si bien se han tomado diversos resguardos para que el sistema sea lo menos amenazante posible y efectivamente brinde información útil para los actores escolares, la relación que los actores establecen con la información evaluativa es altamente compleja y depende de las características de la organización y del sistema educativo.

Tal como se puede apreciar en la literatura, los usos que se le pueden dar la información evaluativa son múltiples, y solamente una fracción de estos tiene que ver con un uso instrumental directo. Pese a ello, la mayor parte de los usos observados tienen una relación directa o indirecta con la meta de mejorar las prácticas pedagógicas del establecimiento. La multiplicidad de usos pone de manifiesto que la introducción de información es solamente una parte del proceso necesario para influir el cambio necesario en las organizaciones escolares.

En este contexto resalta la relevancia de las características de la organización escolar y sus particularidades como centrales para comprender cuáles serán las respuestas probables a la introducción de un sistema como SEPA. Tal como fue posible anticipar a través de la revisión bibliográfica, existen características fundamentales de la organización que tienen un alto peso al momento de definir como se responderá al input de información evolutiva. Dentro de estas, se encuentran el repertorio previo de respuestas anclado en la estrategia pedagógica previa de la organización (Shula y Cousins, 1997), la capacidad de los equipos internos en general y específicamente de comprender e interpretar la información presentada (Earl y Fullan, 2003; Elmore 2008), la instalación previa de sistemas internos de evaluación (Nevo, 2001) asociados a la presencia de una cultura de la evaluación (OCDE, 2005; Scheerens, 2004).

Por estas razones, y dado que la introducción de información evaluativa a una organización nunca es un ejercicio neutro (aunque una organización particular haya accedido voluntariamente a someterse a esta evaluación), es que este proceso debiese considerar cuáles son las características de la organización escolar que recibirá la información. Organizaciones que adhieren voluntariamente al programa (o cuentan con un mandante con poder efectivo), que cuentan con equipos competentes en análisis de datos y habilidades blandas para comunicar y manejar resistencias, que cuentan con sistemas o al menos prácticas evaluativas internas, que cuentan con una cultura de la evaluación que se desarrolla en un marco de confianza, estarán en una mejor posición para aprovechar la información evaluativa desarrollada por SEPA.

En cambio, organizaciones que no cuenten con alguna de estas características, posiblemente pueden requerir apoyo externo para desarrollar cambios tendientes a mejorar el desempeño organizacional. Se

observa que una estrategia que parece tener resultados tiene que ver con el apoyo de una organización externa que pueda suplir capacidades organizacionales importantes tales como la capacidad de análisis o de capacidad de autocrítica observada en organizaciones con culturas de evaluación instaladas. En otras palabras, en este tipo de organizaciones se requiere que la información pueda ser mediada por actores competentes externos que permitan decodificar, interpretar, contener las ansiedades y apoyar el desarrollo de capacidades internas de trabajo.

En este sentido, es posible coincidir con Elmore (2008), en que la construcción de capacidad de responder ante sistemas de evaluación externa como SEPA es un problema fundamental asociado a su introducción y aprovechamiento. En este sentido, el sistema SEPA funciona sobre la base del supuesto, que los establecimientos educacionales harán un uso instrumental de los datos, sin considerar sus capacidades instaladas, suposición que no se confirma. Si el objetivo del sistema es contribuir al mejoramiento de la calidad de los procesos de enseñanza aprendizaje, entonces el sistema debiese considerar la incorporación de medidas que permitan facilitar este proceso.

Este proceso podría apoyarse de diversas maneras, tales como ofrecer asesoría individualizada, brindar cursos temáticamente focalizados o generar una plataforma de intercambio de experiencias entre actores y organizaciones miembros de SEPA. De este modo se podría desarrollar una plataforma de servicios agregados que, no solamente confronten al establecimiento con un estándar de desempeño, si no que además permita acceder a herramientas para cerrar las brechas detectadas. Esta plataforma de servicios se puede desarrollar además en función del tipo de capacidad que un establecimiento debe desarrollar en orden de aprovechar la entrega de información evaluativa externa. En el futuro, sistemas similares debiesen considerar, desde el inicio, el problema de la facilitación del proceso de uso de la información evaluativa como parte del paquete de servicios ofrecidos.

REFERENCIAS BIBLIOGRÁFICAS

- Alavi, M. & Leidner, D.E., (2001). Review: knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly* 25 (1), 105–136.
- Alkin, M., & Taut, S. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*. (29), 1-12.
- Black, P. & William, D. (1998): *Inside the black box: Raising standards through classroom assessment*. Phi delta Kappan, October, 139-148.
- Earl, L. & Fullan, M (2003). Using data in leadership for learning. *Cambridge Journal of Education* 33 (3), November, 383 – 394.
- Elmore, R. (2008). *School reform from the inside out: Policy, practice and performance*. Harvard Education Press, Cambridge M.A. Canada.
- Fullan, M, (2002). Principals as leaders in a culture of change. *Educational Leadership, Special Issue*. May
- Fullan, M, (2006). Change theory: A force for school improvement. *Seminar Series Paper N°. 157*, November. Centre for Strategic Education. Jolimont, Australia.
- Funderburg, S. A. & Levy, P. E. (1997). The influence of individual and contextual variables on 360-degree feedback system attitudes. *Group & Organization Management*, 22 (2), pp.210-35.
- Gareis, C. (2007). Reclaiming an important teacher competency: The lost art of formative assessment. *Journal of Personnel Evaluation in Education*, 20, 17–20.

- Glaser, B. & Strauss, A. (1999). *The discovery of grounded theory: Strategies for qualitative research*. Sociology Press.
- Hein, A. & Taut, S. (2009). *Informe de un estudio de comprensión y uso de la información SEPA*. Documento interno MIDE UC, Pontificia Universidad Católica de Chile, Santiago, Chile. Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage.
- Koretz, D. (2008). *Measuring Up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.
- Looney, J. (2007). Formative assessment in adult language, literacy and numeracy. *Assessment in Education: Principles, Policy & Practice*, 14 (3), 373 — 386.
- McDowell, L. Smalies, J., Sambell, K., Sambell, A. & Wakelin, D. (2008). Evaluating assessment strategies through collaborative evidence-based practice: Can one tool fit all? *Innovations in Education and Teaching International* 45 (2), May, 143–153.
- Milanowski, A.T. & Heneman III, H.G. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: A Pilot Study. *Journal of Personnel Evaluation in Education*, 15(3), 193-212.
- Navarro, J. C. (2003). La evaluación y las actitudes de los docentes frente a ella: Dificultades y alternativas de política. En *Evaluar las evaluaciones: Una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: IIPE-UNESCO, pp. 147-164
- Nevo, D. (2001). School evaluation: Internal or external?. *Studies in Educational Evaluation*, 27 (2), 95-106.
- OCDE (2005). *Formative assessment: improving learning in secondary classrooms*. Paris, Francia. Disponible en línea http://www.oecd.org/document/39/0,3343,en_2649_35845581_34365415_1_1_1_1,00.html
- OCDE (2008). *Measuring Improvements in learning outcomes. Best practices to assess the value added of schools*. Paris, France. Disponible en línea http://www.oecd.org/document/54/0,3343,en_2649_39263231_41701046_1_1_1_1,00.html
- Owen, J.M. & Rogers, P. (1999). *Program evaluation: forms and approaches* (2 Ed.). London: Sage.
- Patton, M.Q. (2008): *Utilization-focused evaluation*. Fourth Edition SAGE Publications, Inc., California.
- Rich, R.F. (1977). Uses of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C.H. Weiss (Ed.), *Uses of social research in public policy* pp. 199–211. Lexington, MA: D.C.
- Ruiz-Primo, M. & Furtak, E. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44, (1), 57–84
- Santelices, V. & Taut, S. (2009). Comprehension and use of value-added school indicators by parents and teachers. *Paper presented at the Annual Conference of the European Educational Research Association*, September 28-30, 2009, Vienna, Austria.
- Scheerens, J. (2004). The evaluation culture. *Studies in Educational Evaluation*, 30, 105-124.
- Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, and M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.

- Shula, L.M., & Cousins, J.B. (1997). Evaluation use: theory, research, and practice since 1986. *Evaluation Practice*, 18 (3), 185-208.
- Taut, S. & Alkin, M. (2003). Program staff perceptions of barriers to evaluation implementation. *American Journal of Evaluation*, 24, (2), 213-226.
- Taut, S. & Brauns, D. (2003). Resistance to evaluation: A psychological perspective. *Evaluation*, 9, (3), 247-264.
- Taut, S., Cortés, F., Sebastian, C. & Preiss, D. (2009). Evaluating school and parent reports of the national student achievement testing system (SIMCE) in Chile: Access, comprehension, and use *Evaluation and Program Planning*. 32, 129–137
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation* 19(1), 21-33.
- William, D. & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22, (5), 537-548.
- William, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education*, 11, 49–65.
- Wildawsky, A. (1972). The self-evaluating organization. *Public Administration Review*, 32(5), 509-520.

EL TRANSITO A LA VIDA POST SECUNDARIA. LA EXPERIENCIA DE LA PRIMERA GENERACIÓN DE EGRESADOS DE UN LICEO DE UNA ZONA POPULAR EN SANTIAGO

THE JOURNEY TO THE POST HIGH SCHOOL LIFE. THE EXPERIENCE OF THE FIRST GENERATION OF ALUMNI GRADUATED FROM A LYCEUM OF AN UNPRIVILEGED AREA IN SANTIAGO

Nelson Paulus, Esteban Geoffroy y Manuela Mendoza

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art10.pdf>

Fecha de recepción: 20 de marzo de 2010
Fecha de dictaminación: 15 de abril de 2010
Fecha de segundo envío: 29 de abril de 2010
Fecha de aceptación: 06 de mayo de 2010

A través del presente artículo se pretende compartir algunos de los hallazgos derivados del estudio exploratorio que en materia de tránsito a la vida post-secundaria realizó el departamento de Investigación e Información Pública del Consejo Nacional de Educación durante 2009.

El estudio que da origen a este artículo indagó en las percepciones, interpretaciones y en el sentido que atribuyen a los eventos ocurridos tras su egreso de la educación secundaria, jóvenes egresados de cuarto medio de un liceo municipal de la comuna de Peñalolén durante 2008.

La forma en la cual tomamos contacto con estos jóvenes puede ser entendida, temporalmente, en la siguiente secuencia

- 1) Abril de 2008: A los cuartos medios del Liceo se les aplicó una encuesta. Esta buscaba indagar en las expectativas que los jóvenes poseían de su egreso mientras cursaban el cuarto año de educación media. Adicionalmente, los datos derivados de esta encuesta permitieron caracterizar la cohorte de la cual fueron extraídos los entrevistados.
- 2) Diciembre de 2008: Los jóvenes estudiados egresan de cuarto medio.
- 3) Abril a Julio de 2009: Los jóvenes estudiados son entrevistados para conocer sus percepciones, interpretaciones y sentido atribuido los eventos ocurridos en su tránsito entre cuarto medio y sus inicios en la vida post secundaria.

La necesidad de estudiar este grupo de jóvenes surgió en tanto el Consejo Nacional de Educación instaló un proyecto de apoyo a la orientación vocacional en su liceo, el cual operó entre 2008 y 2009. Sin embargo, a poco andar de este proyecto se constató que en el marco de esta intervención, la problemática del tránsito a la vida post-secundaria no había sido abordada de manera directa; ya que no se estaba estudiando el fenómeno desde la perspectiva de quienes están viviendo el proceso. Es por esta razón que en el curso posterior de la actividad surgió con gran interés la necesidad de explorar el significado que los estudiantes egresados de la educación media dan a este proceso, y la forma en la cual lo dotan de sentido, identificando factores incidentes y situaciones significativas en ella.

Lo anterior cobra importancia en tanto el proceso de tránsito post educación secundaria, constituye, cada día más, un objeto de estudio relevante en materia de políticas públicas, ya que en tanto este proceso sea adecuadamente comprendido, será posible entregar recomendaciones que permitan reducir el impacto negativo de fenómenos tales como la deserción temprana, la cual, entre otros aspectos, representa un alto costo económico, tanto para el sistema en general, como para las familias.

No obstante lo anterior, y más allá de la reflexión en materia de impacto económico, creemos que explorar la temática del tránsito y abordar las características de este proceso resulta de primera importancia, dado que, cómo fenómeno, presenta una rica multidimensionalidad, no siempre comprendida desde las perspectivas de las cifras.

Por esta razón, el estudio del cual se desprende el presente artículo no busca cuantificar el fenómeno, sino que aproximarse a él desde una perspectiva que permita comprender la experiencia de los sujetos. Así, a través de este trabajo, se intenta ilustrar algunos factores y situaciones que los sujetos entrevistados identifican como relevantes en su proceso de tránsito a la vida post secundaria, intentando luego relacionar dichos factores con trayectorias proyectadas por los mismos sujetos.

A continuación, se presenta el marco conceptual que guió la elaboración de este estudio, en segundo lugar una caracterización general de los jóvenes estudiados, en tercer lugar se presenta la estrategia metodológica, luego los principales hallazgos del estudio, para finalizar con un apartado de conclusiones.

1. MARCO CONCEPTUAL: TRÁNSITO Y TRAYECTORIAS

La pertinencia del proceso de tránsito desde la educación media a la vida post secundaria como objeto de estudio tiene sus fundamentos teóricos en la constatación de que este proceso se hace más problemático entre jóvenes que corresponden a la primera generación de sus familias en acceder a estudios superiores.

Esta situación ha sido documentada por la investigación internacional, a través de investigadores como Patricia M. McDonough, en su libro *"Choosing Colleges: How Social Class and Schools Structure Opportunity"* (1997). La investigadora norteamericana señala que: "los estudiantes que corresponden a la primera generación de su familia en acceder al *college*, comienzan a pensar acerca de asistir al *college* mucho más tarde que aquellos estudiantes cuyos padres han accedido al *college*, y sus ideas al respecto son gatilladas por el personal de sus escuelas, específicamente profesores y orientadores. Los estudiantes cuyos padres han asistido al *college* frecuentemente comienzan su preparación para el *college* en la escuela primaria, tomando los cursos correctos y manteniendo buenas calificaciones. Además, sus familias les han transmitido información acerca de los diferentes tipos de *colleges* y universidades" (McDonough, 1997:6).

De lo anterior se desprende que la carencia de referentes familiares por parte de los jóvenes de primera generación hace pertinente la intervención pública que apoye su tránsito hacia un nivel educativo superior, lo cual vuelve relevante la investigación que entregue insumos para el desarrollo de políticas al respecto.

Por lo demás, posee relevancia estudiar el fenómeno del tránsito a la vida post secundaria y las transiciones entre distintos momentos de la secuencia educativa, en tanto dichas secuencias han ido mutando en el tiempo, de manera tal que aquellas trayectorias que hasta hace poco eran consideradas transiciones típicas, han variado o bien ya no son las únicas: "la tradicional estructura *lineal* de transición, definida por una secuencia culturalmente y socialmente reproducida, en que de estudiar se pasa a trabajar, ahí al matrimonio y la crianza de hijos, todo con plazos estrictos, con edades prescritas, ha ido cediendo terreno a nuevas formas de hacerse adulto, nuevas *formas de transición*, con otras estructura, con otro orden en la secuencia y otros tiempos para cada paso" (Dávila y otros, 2008:70).

Lo anterior ha constituido a estos procesos en un objeto de estudio relevante en el campo de la investigación educativa, distinguiendo el concepto de transiciones del de trayectorias. Dávila y otros (2008) nos indican que mientras la transición es un proceso inevitable, común a todo individuo y presente en todo momento histórico como lo es el paso de la niñez a la adultez, la trayectoria se encuentra fuertemente ligada a los procesos y contextos sociales con los que se articula el proceso educativo: "La *trayectoria* está puesta en otro plano, en el plano social, de las posiciones que van ocupando los sujetos en la estructura social, o lo que es igual, en el campo de las relaciones de poder entre los grupos sociales. No es la secuencia que producen las distintas fases de generación de nuevos individuos adultos lo que importa al análisis de trayectorias, sino las posiciones estructurales y las disposiciones subjetivas que producen –en el doble sentido de *ser producto* y de producir – esos cambios de condición" (Dávila y otros, 2008:73).

Oyarzún e Irrazabal (2003) han descrito algunos tipos de trayectorias educativas que son de gran utilidad para el análisis del tránsito a la vida post secundaria en que se encuentran inmersos los jóvenes a los que hace referencia el presente artículo. Fueron dos tipos de trayectorias definidas por estos autores, los que iluminaron la investigación que da origen a este artículo. En primer término, las trayectorias lineales entendidas por los autores como aquellas caracterizadas por relaciones causa-efecto y que cuentan con etapas claramente diferenciadas, las cuales nunca se sobreponen porque los ámbitos y espacios de la vida están socialmente organizados por distintas instituciones sociales. Así, un ejemplo de trayectoria lineal es la de aquel joven que al salir de cuarto medio se dedica exclusivamente a estudiar, o bien, exclusivamente a trabajar, dejando atrás definitivamente el mundo de los estudios, sin combinar nunca ambas actividades.

En contrapartida, y cada vez más frecuentes, se encuentran los que estos mismos autores definen como trayectorias sincronizadas, entendidas como el desarrollo simultáneo de actividades de estudio y trabajo (Oyarzún e Irrazabal 2003), seguidas principalmente por jóvenes que deben trabajar para financiar sus estudios. Estos mismos autores - siguiendo una definición de Machado Pais (2002) - dan cuenta de lo que se ha conocido como trayectorias yo-yo, aquellas en las cuales los jóvenes van y vuelven de una posición a otra, como es el caso del joven que pasa de estudiante a trabajador y luego vuelve a ser estudiante. Esta conceptualización será utilizada de manera instrumental para comprender los casos de jóvenes incluidos en este artículo.

2. CARACTERIZACIÓN GENERAL DE LOS JÓVENES ESTUDIADOS

En términos generales, los jóvenes estudiados corresponden a un perfil particular que podemos definir como: urbanos, de sectores socioeconómicos medios bajos y cuyos padres, por lo general, no accedieron a la educación superior.

2.1. Urbanos y de sectores socioeconómicos medios bajos

Los jóvenes entrevistados en esta investigación, son estudiantes de un liceo municipal de la comuna de Peñalolén ubicado en una de los sectores más populosos de la región metropolitana. Este sector fue clasificado en 2008 en el nivel económico medio-bajo por el Ministerio de Educación, a partir de la puesta en juego de tres variables: años de escolaridad de los padres, ingresos del hogar y niveles de vulnerabilidad social de los alumnos.

En este contexto, el rango medio-bajo corresponde a aquellos liceos en los cuales la mayoría de los apoderados declara tener diez años de escolaridad y un ingreso del hogar que varía entre \$190.001 y \$300.000 y que cuentan con un rango de entre 45,1% a 62,5% de estudiantes en condición de vulnerabilidad social según JUNAEB¹.

2.2. Primera Generación que accedería a estudios superiores

Como se mencionó en el apartado anterior, la clasificación socioeconómica del MINEDUC nos indica que estos jóvenes eran alumnos de un liceo donde el grueso de los padres alcanza una escolaridad de diez años, lo que indica que en promedio sus padres no poseen los doce años de escolaridad mínima para

¹ Junta Nacional de Auxilio Escolar y Becas.

obtener la licencia media imprescindible para cursar estudios superiores. Estos datos permiten suponer que los estudiantes del grupo abordado podrían convertirse, en lo eventual, en la primera generación en acceder a estudios superiores dentro de sus familias.

Lo anterior se constata empíricamente con una encuesta aplicada al grupo específico (cuartos medios 2008) al que pertenecen los jóvenes entrevistados, a quienes se consultó por el nivel educacional alcanzado por sus padres. En efecto, se constató que el 43,3 % de los encuestados señala que sus madres no alcanzaron a completar la enseñanza secundaria, es decir, no vivieron el hito que sus hijos se preparaban a enfrentar a fines de ese año. Si sumamos el 43,3% que señala que sus madres completaron la enseñanza media pero no siguieron estudios superiores, nos encontramos con un 86,6% de alumnos cuyas madres no transitaron hacia los estudios superiores.

Respecto a los padres, un 46,7% indica que sus padres no completaron los estudios secundarios, siendo un dato relevante que un 20% indica que su padre no terminó siquiera la enseñanza básica. Sólo un 10 % indica que su padre accedió a alguna modalidad de estudios superiores.

Los datos anteriores corroboran que estamos ante estudiantes que, en lo eventual, podrían ser la primera generación de sus familias nucleares en acceder a la educación superior. Es más, muchos de estos jóvenes provienen de familias en que sus padres no tan solo no alcanzaron la educación superior, sino que en muchos casos no completaron los estudios secundarios e incluso en algunos casos tampoco la educación primaria. Vale decir, alumnos provenientes de familias deprivadas no sólo económicamente sino que también en términos de capital cultural, en tanto poseen menos referentes y apoyos para elaborar su estrategia de tránsito a la vida post-secundaria.

A partir de una recodificación realizada, se constató que sólo un 33,3% de los jóvenes encuestados afirma que sus dos padres concluyeron los estudios secundarios, un 16,7% que sólo uno de ellos lo hizo, mientras que un 33,3% afirma que ninguno de los padres concluyó sus estudios secundarios. De esta manera, se puede afirmar que al menos un tercio² de los encuestados corresponde a la primera generación de sus familias que enfrentó el hito de egresar de cuarto medio.

3. ESTRATEGIA METODOLÓGICA

Para la realización del estudio se tomó, en primer término, información recogida a través de una encuesta a los jóvenes estudiados mientras cursaban el cuarto año de educación media. Esta se aplicó en abril de 2008, destacándose entre los asuntos consultados: las expectativas y aspiraciones sobre continuidad de estudios; importancia asignada a la formación entregada en el liceo para la vida post-secundaria; nivel educacional de sus padres, etc. Esta encuesta fue aplicada a la totalidad de los jóvenes que cursaban cuarto medio en el liceo durante abril de 2008 (treinta alumnos). Estos datos permitieron caracterizar a los jóvenes estudiados, a la vez que aportaron antecedentes que permiten contextualizar los hallazgos presentados a través de este artículo.

² El restante 16,7% no entregó esta información

Durante el primer semestre del año 2009 (con posterioridad a su egreso de la educación secundaria) se tomó una muestra de jóvenes a los cuales se aplicó una entrevista³, para la realización de esta se diseñó una pauta semi-estructurada, cuyo diseño incorporó las dimensiones y los aspectos generales que se debía tratar en la entrevista, con el objeto de profundizar en la experiencia de tránsito de los entrevistados.

Para efectos de la realización del estudio cualitativo, la muestra teórica⁴ seleccionada, se componía de dieciséis casos, de las cuales se logró realizar doce entrevistas, de estos doce jóvenes egresados entre cinco y siete meses antes de las entrevistas:

- Cuatro era mujeres y ocho hombres.
- Siete habían accedido a alguna modalidad de estudios superiores, tres cursaban preuniversitarios y dos no seguían estudios formales ni informales.
- Cinco trabajaban al momento de la entrevista
- Tres habían accedido a becas de estudios.

4. ALGUNOS HALLAZGOS DE LA INVESTIGACIÓN

En esta sección se informan los principales hallazgos respecto a los jóvenes estudiados. En primer término se entrega la caracterización general que de los cursos abordados, se puede realizar a partir de la encuesta aplicada en abril de 2008 en materia de expectativas. Luego se trata con mayor detalle las entrevistas realizadas a doce de estos estudiantes unos meses después de haber egresado de la enseñanza secundaria (entre mayo y julio de 2009).

4.1. Encuesta expectativas (abril de 2008): Altas expectativas de continuar estudios

A pesar de que la mayoría de los padres de estos jóvenes no llegaron a la educación superior, se observan altas expectativas de continuidad de estudios post-secundarios:

En primer término ante la pregunta ¿qué te gustaría hacer el año siguiente a salir de cuarto medio? el 47% señala que desea estudiar; el 46 % trabajar y estudiar; mientras que un 7% desea hacer el servicio militar. Sumando las dos alternativas que contemplan estudiar se puede constatar un altísimo interés de

³ En este sentido, cabe destacar que aquellos jóvenes que no egresaron o que eventualmente se retiraron de la enseñanza media antes de egresar de ella, no fueron considerados en el universo a partir del cual, finalmente, se seleccionaron las personas a entrevistar.

⁴En tanto se trata de un estudio cualitativo, la construcción de la muestra no tiene por objeto constituir un reflejo de la realidad general de todos los estudiantes de 4º medio egresados durante el año 2008 a nivel nacional. De manera distinta, lo que se pretende es abordar la diversidad de situaciones contenidas en el caso específico de la 1ª Cohorte de egresados de 4º medio de liceo estudiado. Así visto, la construcción y posterior constitución de una muestra busca dar cuenta de la variedad presente al interior de la cohorte, y de rescatar la diversidad de subjetividades presentes en el grupo de jóvenes que componen nuestro objeto de estudio, previniendo de esta manera posibles sesgos. Para la realización de entrevistas se definió una muestra teórica de 16 casos, en la que se buscaba representar la variedad del grupo y reducir sesgos en función de los siguientes criterios:

Sexo: Criterio que aparece como relevante especialmente en contextos socioeconómicos bajos, en los cuales regularmente se observan diferencias en las oportunidades y expectativas cifradas en hombres y mujeres.

Escolaridad de los padres: Para observar su influencia en la conformación de las percepciones y las atribuciones de sentido hechas por los estudiantes. Además la investigación internacional demuestra que los hijos de padres con mayores niveles de escolaridad comienzan a pensar antes y por ende se preparan mejor para las elecciones académicas que tendrán que tomar al salir de la secundaria.

Ranking al interior del curso; En base al promedio NEM: Aspecto relevante para eventualmente evaluar el peso de los logros académicos en la conformación de atribuciones de sentido.

parte de estos jóvenes por la continuidad de estudios (93%). Al preguntarles respecto a lo que creían que harían efectivamente, nos encontramos con que la opción “estudiar” (de manera exclusiva) baja en comparación a lo que querían hacer, desde un 46% a un 37%, mientras que la opción “trabajar y estudiar” aumenta desde un 47% a un 60%.

Vale decir, nos encontramos con que la gran mayoría de los jóvenes cree que podrá seguir estudiando (97%) pero estas expectativas se ajustan a algún tipo de juicio que les indica que, para efectos de la factibilidad de dicho proyecto, no podrán hacerlo de manera exclusiva, principalmente porque deberán financiarlo, ya sea parcial o íntegramente.

Así vemos que en abril de 2008, se vislumbraba claramente en el horizonte de los jóvenes encuestados la posibilidad de seguir trayectorias sincronizadas que combinaran el trabajo con los estudios.

4.2. Entrevistas semiestructuradas (abril a junio de 2009)

El ejercicio de volver a abordar a algunos de estos jóvenes después del egreso, a través de una entrevista, pretendía que los propios involucrados en el proceso describieran su situación actual, connotando aquellos aspectos que ellos mismos identifican como más relevantes en el marco del tránsito desde la educación media, a la vida post-secundaria, tránsito muchas veces asociado a un paso de acercamiento a una vida más adulta. De la misma forma, nos interesaba presentar el modo en el cual los sujetos reprocessan esta situación y proyectan a partir de ello, nuevos escenarios.

En la base de la ejecución de este ejercicio está el entendido que la forma en la cual los distintos sujetos son capaces de dotar de sentido el proceso contingente que les ha tocado vivir, y consecuentemente la manera en la cual son capaces de procesar escenarios futuros a partir de esto, constituye, en lo eventual, un elemento altamente configurador de las trayectorias futuras.

En concreto, a continuación se exponen las percepciones, interpretaciones y el sentido que atribuyen a los eventos ocurridos en su tránsito a la vida post secundaria clasificados según las siguientes temas: condiciones que influyeron en la decisión respecto a la continuidad de estudios, relato de su experiencia de tránsito, valoración y sobrevaloración de los preuniversitarios, criterios y mecanismos de información utilizados para decidir, nuevos mecanismos para entender la deserción, valoración de la modularidad académica, imaginario juvenil sobre la educación superior, etc.

4.2.1. Condiciones que influyeron respecto a la continuidad de estudios

Una dimensión claramente distinguible al interior de la exploración del proceso de tránsito, es aquella que nos remite a la incorporación de los ex estudiantes de enseñanza media, a sus nuevas actividades. Al interior de esta dimensión distinguimos lo que se creía que iba a ser de lo que efectivamente pasó.

“...yo no pensaba estudiar después de salir de cuarto medio, pensaba trabajar y de ahí al otro año empezar a estudiar [...] yo quería hacer cursos cortos como de lo mismo, pero no estudiar.”

*“... mi hermana iba a entrar a estudiar, la que estudia contabilidad, la que estudió, iba a entrar a estudiar recursos humanos, y yo vi, trajo un folleto, y en ella salía ingeniería en informática. **Así que fui y me matriculé.**” - Consultada por cómo llegó al IP -*

En este sentido, cabe destacar, que las decisiones en temas de incorporación reconocen elementos diversos en su constitución, aunque por lo general no se observan razones puras. Vale decir, en la construcción de la argumentación que explica las decisiones para incorporarse eventualmente en algún programa de estudios, no apunta a la optimización de un criterio único, sino a una solución de equilibrio

entre dos o más de ellos, por ejemplo financiamiento (beca) y opción efectiva por una carrera que si bien no es la deseada, pertenece a la misma área.

"...yo quería irme a medicina. Entonces enfermería me dijeron y con beca y toda la cuestión. Y fui a ver y pregunté si de enfermería me puedo pasar a medicina, y me dijeron no. Pero me dijeron que de bachillerato en ciencias puedes pasarte si eres de los diez mejores promedios. Y yo conocía el bachillerato de la Católica, que ahora ya no está, entonces dije "debe ser divertido", pero no."

"...siempre me gustó la música, el arte, [...]... como lo que más se acercaba a lo que a mí gustaba era la informática, así que por eso llegué ahí."

"En realidad yo no quería estudiar preparación física, quería estudiar... no sé, algo que tuviera que ver con enfermería y cosas así."

O bien, considerando adicionalmente la variable tiempo...

"... por plata y también por lo que era prolongación de tiempo no lo hallé tan conveniente. Y preparación física igual se asemejaba un poco a ayudar a las personas y cosas con respecto a lo físico. Aparte que era corta y más barata, me pareció buena opción. En realidad es bastante buena la carrera. A mí me gusta."

*"...yo quería estudiar dibujante proyectista. La única que había era la UTEM. Al final quedé aprobado, quedé en la UTEM, pero decidí no optar ahí porque con **los cambios que ha habido de construcción** y toda esa cosa **no hay mucho campo**⁵, y además iba a terminar como en cinco años [...] la carrera era de dos, no, es que es la UTEM, y siempre están como... por los paros, muchos, muy problemático."*

No obstante, las combinatorias antes aludidas muchas veces ceden frente a la contundencia del argumento financiero:

"...pero me salió la beca y como que cambió todo y ahí tomé la decisión de estudiar..."

"...qué más... me iba a meter directamente a ingeniería comercial pero me metí al técnico por la cuestión de la beca." - Informado de la posibilidad de construir progresivamente la carrera de Ingeniería Comercial

Conforme a lo anterior, la forma en la cual la variable **financiamiento** es capaz de estructurar el proceso de tránsito, no debe ser subestimada, ya sea para dar continuidad directa a un proyecto educativo en educación superior, o bien proyectar una instancia de tránsito previa a este (preuniversitaria):

"Nosotros igual habíamos averiguado otros preuniversitarios,⁶ pero... el C... por ejemplo, son 60 mil que hay que pagar, entonces ya ahí uno los va dejando de lado..."

Para algunos de ellos, instancias de financiamiento como las becas siempre estuvieron en la retina:

*"Es que siempre me esforcé por tenerla... **sabía** que si me iba bien el colegio iba a tener buenas notas. Entonces trataba de no conformarme con un seis o un cinco, sino que llegar a siete." - Consultada por la obtención de una beca -*

"Sé de las becas por mi apellido, porque yo tengo apellido indígena, entonces a esas becas puedo postular para que me den comida o lo que me da el gobierno, de que me da una plata por estar estudiando, pero si yo tengo arriba de 5,5."

⁵ Entrevista realizada el día 24 de junio de 2009; en medio de la última crisis económica que en la ciudad de Santiago, por ejemplo, reconoció como hito emblemático, la paralización de las obras del nuevo centro comercial Costanera Center.

⁶ Instituciones educativas no pertenecientes al sistema formal cuyo objeto es preparar a los postulantes a la educación superior para rendir la Prueba de Selección Universitaria (PSU)

Mientras que para otros resulta una auténtica sorpresa...

"...las universidades privadas empiezan a llamar, entonces no sé cómo tienen mi número pero... y me llegaron correos que decían... yo no sé, en verdad no recuerdo haber hecho algo como inscribirme en una universidad ni nada, o a lo mejor en esas ferias donde uno dejaba los datos... y me llamaban y me llamaban y me decían usted quedó en esto, usted quedó en esto otro, y tiene esta beca, y tiene esto otro."

"Y justo al E... el papá le dio este dato de la Universidad Católica, de pagar cuatro mil pesos y ahí decidimos al tiro no perder la oportunidad." - Para un preuniversitario –

En este sentido la disponibilidad o no disponibilidad de una beca como recurso de financiamiento pareciera ser altamente estructurador de la forma en la cual se proyecta el futuro.

"Yo ahí empecé a pensar qué podía estudiar, igual me puse súper feliz porque ni tenía una solamente, tenía como cuatro o cinco, y aparte tenía aprobado el crédito, el fondo solidario entonces fue como "oh tengo que aprovecharlo no puedo dejar pasar esta oportunidad".

"...a ver. Lo primero fue cuando hicimos lo de las becas, aquí mismo. Ya, hicimos todo, vimos cuánto necesitamos de becas, hicimos todo ese trámite. Después, empecé a elegir las carreras. [...] Ahí, ya, estaba viendo, y empecé a ver por lo que me gustaba y las universidades que me había dicho mi hermano que eran buenas, y ahí empecé a averiguar."

*"...entonces me fui pa'allá. Igual yo tenía planeado ya estudiar eso, igual me ayuda una **beca** que tengo igual. [...] Con eso no se hace tan pesado."* - Consultado por cómo fue que terminó estudiando en INACAP -

El reverso de la medalla queda reflejado cuando el financiamiento determina un límite de imposibilidad. Por lo menos para ese año en específico.

*"...porque es muy caro y no me alcanzaba. Mi hermano me iba a ayudar pero mi mamá tiene problemas entonces... ella tiene un nódulo en el pecho, es un cáncer dormido, eso es. Por eso nosotros teníamos que hacerle biopsia, mamografía y todo es mucho gasto. Más que mi mamá tiene problemas con mi papá entonces no daba plata, y mi hermano era el único que nos ayudaba con eso. Entonces yo tenía que trabajar sí o sí **y no pude juntar plata.**"*

*"...en el último paso yo ya tenía todos los datos listos, y yo lo hice eso en mi casa, como que lo terminé en mi casa, pero en verdad no lo terminé porque al último yo debía poner imprimir, que era el certificado de que yo había postulado a las becas, entonces ese certificado yo no lo saqué, y ahí fue el problema [...] entonces **ahí jodí todo** porque no me salió ninguna beca, y yo igual tenía buenas notas así que me podía haber salido la beca de excelencia académica y todas esas cosas, que me hubieran ayudado más a pagar la universidad o cualquier cosa que hubiera entrado."*

No obstante lo anterior, las becas no aparecen en el imaginario como el único dispositivo de financiamiento, pues la conceptualización financiera del apoyo familiar es algo muy importante.

*"Y al E... los papás ni lo pescan. O sea el papá está con otra mujer ahora y creo que le da como 20 lucas mensuales. Y de repente él pide algo pa'ir, no sé poh, pa'la micro, y no están ni ahí con él. Entonces igual ahí sí que es difícil. Yo le doy gracias a Dios porque por lo menos tengo a mis papás que me apoyan. Que cualquier cosa que sea estudio van a estar así conmigo. **Al E... lo están empujando casi pa' que busque pega no más.**"*

De ahí la concepción de la educación como única dote y herencia de los padres a sus hijos, *leit motiv* ya tradicional, aparece casi como natural.

"... mi mamá y mi papá, más mi papá me dice siempre que yo tengo que estudiar y que lo único que me puede dejar él como herencia es el estudio. Bueno, yo creo que cada papá le dice eso a su hijo."

"...siempre me gustó la música, el arte, pero no me dejaron estudiar eso [...] me dijeron "yo no te voy a andar una cuestión que después no te deje como mucho beneficio, o sea que no te sirva para nada para tu futuro", eso me dijo."

"...siempre me dijeron que lo que ellos me podían dar era el estudio, nada más, que el esfuerzo tenía que ponerlo yo y las ganas de aprender. Y después, cuando creciera les iba a decir gracias mamá y papá, aunque no lo entendiera cuando chico, lo que dicen siempre."

4.2.2. Relato de su experiencia de tránsito

Cabe destacar, que incluso en el escenario positivo, de una eventual incorporación, las cosas no siempre son como se les proyecta, o mejor dicho, la inserción en nuevos contextos educativos les permite a los recién egresados disponer de cierta **perspectiva** para aproximarse a algunas situaciones...

"Igual ahora me arrepiento, uno se arrepiente por querer buscar un futuro mejor. Sí, porque ahora no estoy en una universidad tradicional, igual sale caro estar en un instituto.[...] bueno, el año pasado como que uno no se veía mucho la presión que ejerce después estar estudiando en el nivel superior. Uno como que se preocupa de leerse con los compañeros y como que rara vez alguno se preocupa de las notas, como yo. No me preocupaba mucho.."

*"...el colegio no es tan malo. Mi curso en realidad no era tan malo, no era malo. Entonces, como que ahora veía que tenía compañeras de la universidad que eran de colegios buenos, de colegios privados, unos que hicieron **preu** también aparte, en la Católica, y después en la prueba sacaron un poquito menos puntaje que yo. Y yo decía, Guau! Igual el colegio no era tan malo, o no sé si eso fue el Penta. O de repente veíamos temas y yo decía "pero si esto lo vimos en el colegio", y preguntaba alrededor mío y la mayoría no lo había visto en el colegio, entonces decía que igual no era tan malo. Y después cuando me encontraba con la M... por ejemplo, o con la S... y yo decía "oye, no sé si te ha pasado como que ven temas que nosotros igual hemos mirado en el colegio y el resto de los compañeros no en la U", y me decían "no, si nuestro colegio no era tan malo, nuestro curso no era tan malo. Depende del curso."*

La incorporación de nuevas perspectivas, para otros determina una suerte de disonancia entre aquello que eran sus expectativas previas al egreso, razonablemente coherentes en el marco de evaluación que imponía su realidad intra-liceo, y el desempeño entendido fuera de estos acotados marcos.

"Entonces por eso yo al dar la PSU... todos mis compañeros sacaron 400 y tantos puntos, sí poh, y yo saqué 512 parece, algo así, y era el más alto poh. Entonces uno se da cuenta que a veces los mismos preuniversitarios que hacían en el colegio no lo ha pescado, y ahora uno se da cuenta que lo necesitaba poh, que no los tomó en serio, y ahora uno se da cuenta que había que tomarlos en serio. Y por eso mismo estoy haciendo esto también."

Derivado de este último enunciado, se desprende que uno de los principales elementos a partir de los cuales los entrevistados explican lo que pasó, y que tras el egreso constituye uno de los hitos focales hacia los cuales los recién egresados reorientan su atención, queda constituido por los preuniversitarios. Surge así la necesidad de profundizar respecto al imaginario depositado en el accionar de estas instituciones, y principalmente las expectativas depositadas en ellos.

4.2.3. Valoración de los preuniversitarios

La influencia de los preuniversitarios puede ser abordada de distintas maneras. Por ejemplo, cabe destacar las atribuciones de eficacia depositadas en este tipo de instituciones, en tanto algunas de las

situaciones observadas como malogradas -a lo menos momentáneamente- son explicadas en términos de no haber asistido a uno de ellos. El reverso está constituido por aquél tipo de comunicaciones que se atribuyen éxito en la instancia de tránsito (a lo menos en términos relativos) y sin haber hecho preuniversitario

"...me fue bien igual, porque igual yo no hice preu, nada."

Asimismo, los preuniversitarios constituyen la instancia transitoria por excelencia, que permite compaginar el ingreso a una vida post secundaria, con el proceso que permite, en muchos casos, reordenar los recursos educativos en proceso de un nuevo intento a incorporarse a la educación superior.

"...estoy en el preuniversitario debido a que tuve un problema con las becas..."

"...cuando fui al preuniversitario como que me sentía mal por no haber entrado a la universidad. Pero después, cuando me di cuenta que había harta gente que no había entrado a la universidad, igual como que no me sentía tan mal porque decía "Ah, no fui la única". Entonces como que igual decía "ah, a lo mejor ellos tuvieron otros problemas". Y me pongo a hablar a veces con algunas amigas y me dicen "sí, si tení que estar tranquila, a lo mejor este no era el año pa'entrar y el próximo año será."

"...estoy en un preuniversitario, como decía el E..., y a ver si nos va bien en la PSU. Y... nada... cuando salimos de cuarto medio o sea igual costó aceptarlo un poco, después cuesta adaptarse a no hay más colegio, a levantarse temprano."

"...o trabajaba o iba a hacer un preuniversitario pa'seguir estudiando. Y como en la PSU me dan dos años gratis, tenía que darla poh. La fui a dar como por tener la experiencia, todo, ese año..."

De manera similar, es posible destacar la incidencia de los preuniversitarios de una manera un poco más sofisticada, y que se asocia a la incorporación de nuevas perspectivas, lo cual requiere explicaciones adicionales.

En términos generales, uno podría aludir a una suerte de asimetría de información existente en materia de proyecciones de rendimiento PSU, de la cual estudiantes de realidades como las del liceo estudiado serían afectados. En este sentido, podríamos referir al hecho que la puesta en contacto con instancias preuniversitarias o universitarias determina la incorporación de un tipo de información, o perspectiva de análisis, que es capaz de reestructurar fuertemente las expectativas de quienes, hasta ese momento, no habían tenido acceso a ella.

Así visto, podríamos observar que la información que nutre los sistemas de expectativas provienen de contextos escindidos en donde uno de ellos (el ampliado) configura una clara situación de asimetría informativa a los estudiantes de liceo.

En prácticamente todas las entrevistas es posible rastrear conductas racionales, entendiendo racionalidad en un sentido amplio de acciones dotadas de sentido; o mejor dicho, no se observan francos despropósitos en la atribución de la racionalidad de la acción, aunque sí racionalidades dignas de un escrutinio más detallado. De lo antes señalado, no es ajeno el tópico de los preuniversitarios.

"Sí poh, si quiero estudiar algo, y para eso quiero dar la PSU, y pa'dar la PSU bien tengo que estudiar en el preu, porque con lo que yo tenía no me alcanzaba. Así que por eso estamos haciendo eso."

"...igual estoy estudiando en un preuniversitario porque pasan física, cosas así, materia de masa, volumen, entonces igual en mecánica yo he sabido que... yo tengo igual gente que sabe mecánica y me dice que esa materia es súper buena y que igual me interesa la mecánica. Entonces yo por eso igual estoy estudiando, más por física, en ese sentido."

4.2.4. Criterios y mecanismos de información utilizados para decidir

A partir de lo antes visto es posible afirmar que aquello que pueda resultar problemático en el marco del tránsito a la vida postsecundaria no se debe a falta de criterios para procesar la información y tomar decisiones -pues se observa un conjunto amplio de ellos, en su gran mayoría bastante razonables- sino en la circulación de información precisa, pero no respecto al escenario que deben enfrentar, sino respecto a ellos mismos.

En este sentido, podríamos pensar que quienes acceden al circuito preuniversitario en forma previa a la rendición de la PSU, y dicho de mejor manera, en forma paralela al tercero y/o cuarto medio, cuentan con una ventaja comparativa en tanto rápidamente pueden ajustar sus expectativas en términos de la unidad de medida estándar en la materia: puntajes obtenidos en los ensayos. En quienes han tenido la oportunidad de optar tempranamente a estas instancias preuniversitarias, hay continuidad entre el marco informativo de su contexto local y de su contexto ampliado. No así en quienes no han tenido esta oportunidad.

"Yo no hubiera podido hacer un preuniversitario al mismo tiempo que el colegio porque salían del colegio ellos y al tiro se iban al preuniversitario. Y yo pensaba "No. Salir del colegio y al tiro ir al preuniversitario, y después llegar a la casa, y si al otro día tenía prueba... ". Entonces igual yo veía que era mucho stress, y decidí que no era mejor tomar un preuniversitario, y dar así, con lo que sabía no más la prueba."

"...fuimos a ver a C..., a todos los preu, al P..., y todos son muy caros pa'nosotros. Y ya po, cómo lo hacemos, y al final terminamos hablando allá y había un preuniversitario allá que hacían y cobran más barato. Allá en la misma Universidad Católica, son estudiantes. Son estudiantes de Derecho, de Medicina, y saben harto. Y... más barato poh, más barato y es bueno."

Así, si bien podemos decir que circula información de distinta naturaleza, (carreras, aranceles, etc.), desde fuentes y por canales distintos (familiares amigos, conocidos, medios de comunicación, internet, etc.), no circula toda la información necesaria para tomar de decisiones certeras. En efecto, no siempre cuentan con la adecuada **retroalimentación** que les permita saber qué tan bueno es su rendimiento desde una perspectiva de información estandarizada como la que provee el contexto ampliado que constituye los circuitos de información preuniversitaria.

La exploración desarrollada en el estudio refrenda la siguiente idea: parece que no circula toda la información necesaria para tomar decisiones, en tanto la sola concepción de informar la oferta de programas y sus características (matrícula, aranceles, requisitos, duración, etc.) resulta insuficiente. Por ejemplo, los eventuales postulantes no tienen información respecto a qué tipo de postulación efectiva pueden respaldar, no hablando en términos financieros -variable que los entrevistados parecen tener suficientemente clara -ni en términos de los puntajes requeridos (información que también circula) sino en términos de una aproximación realista a cuál es el puntaje que razonablemente pueden esperar tener.

Como un correlato explícito de lo antes expuesto, surge en el marco de la investigación la interrogante por la recurrente preferencia por la carrera de medicina en varias de las entrevistas realizadas

"Estaban todos casi decididos por medicina como que le gustaba. Por lo menos en mi curso la mayoría iba por medicina"

"...sí, pero directo a medicina porque no quiero otra cosa..."

"Bueno a ellos les gustaba que estudiara eso como que ooh! Siempre cuando le comentaban a los amigos y todo que iba a estudiar medicina y todo como que se maravillaban con eso..."

Evidentemente lo anterior es una variable mucho más crítica, cuando lo que se pretende es optar a instituciones tradicionales y selectivas (estimadas como aquellas que exigen un resultado PSU y unas notas de enseñanza media de excelencia) antes que a instituciones no selectivas, donde esta asimetría de información se vuelve menos importante. No obstante, la opción de adscribirse en instituciones no selectivas, por parte de este tipo de estudiantes, trae aparejada otro tipo de interrogantes para los postulantes, como la que se verá a continuación.

4.2.5. Nuevos mecanismos para entender la deserción

La falta de retroalimentación con respecto al contexto de rendimiento y la incorporación del perfil de estudiantes entrevistados en instituciones no selectivas, permite arrojar algunas luces respecto a una nueva dimensión de porqué desertan algunos buenos estudiantes de, por ejemplo, universidades menos selectivas y/o menos exigentes: se trata de expectativas defraudadas, generadas en un marco claro de asimetría de información.

Esta reflexión surge de manera colateral a partir del análisis del enunciado en el que una entrevistada expresa que no había quedado en medicina por no haber hecho preuniversitario.

"...mi mamá me dijo "estudia", y yo no quería entrar a nada, y tampoco me alcanzó mi puntaje para medicina, porque no hice un preu, no hice nada."

El caso resulto paradójico para el equipo investigador en tanto el puntaje obtenido por esta persona distaba mucho de estar cerca del suficiente para optar a la menos exigente de las carreras de medicina. No obstante lo anterior, dejando de lado el tema del puntaje, y a partir de los datos generales de la encuesta y de la cohorte, pudimos observar que en tanto desarrollásemos una aproximación acotada exclusivamente al universo de sentido educativo que constituye el liceo, la pretensión de quien emite esa comunicación es bastante razonable ya que se trata de una alumna cuyas notas de enseñanza media y su rendimiento PSU eran excelentes en el marco de su promoción, probablemente perteneciente al 5% superior de su generación: segunda en la PSU de Lenguaje en toda la cohorte; cuarta en la PSU de Matemáticas y cuarta en el promedio de notas de enseñanza media, de entre cuarenta y tres estudiantes que componen la cohorte de egresados.

En estas circunstancias las expectativas de la entrevistada pueden ser observadas desde dos perspectivas:

La primera, dada por lo que ha sido su desempeño en su contexto educativo local (liceo) determina que la pretensión de estudiar medicina, si bien es ambiciosa⁷, no es algo irreal.

La segunda perspectiva de análisis está dada por un contexto de sentido ampliado, como es el que determina las distintas instancias preuniversitarias. Estos pueden ser preuniversitarios formales, cursos preuniversitarios internos a los propios establecimientos educacionales o bien instancias de ensayos al interior de los cuales los puntajes PSU circulan como una información que en términos de retroalimentación es capaz de informar lo realista de las aspiraciones educativas.

Para el caso de nuestra enunciante, en el marco de este contexto ampliado, sus pretensiones son claramente irreales, lo cual es refrendado por aquello que señalan otros entrevistados

"...uno de repente llega al preu y se da cuenta de que hay cosas en el colegio que no han pasado, en materia, en la asignatura."

⁷ En Chile la carrera de Medicina es altamente selectiva, por lo cual en cualquier contexto socioeconómico es ambicioso pretender acceder a ella.

Así, en el marco de este caso preciso surge la interrogante por un mecanismo de deserción, hasta el momento inexplorado en el marco de los antecedentes que informan el presente estudio. El mecanismo de deserción aludido se formaliza en la siguiente secuencia:

- Se trata de un alumno excelente en su contexto educativo local, pero con expectativas razonablemente limitadas respecto a su desempeño en el marco de la educación superior.
- Que la situación antes referida determina que un estudiante de estas características posea razonables aspiraciones (en el marco de su contexto educativo local) de pretender alcanzar carreras altamente exigentes; no obstante el hecho que existe la limitación efectiva que la incógnita respecto a su eventual desempeño PSU representa para incorporarse efectivamente en estos programas más demandados.
- Si dejamos en suspenso la variable financiamiento, podemos ver que se abre para estos alumnos un sector de la oferta que si bien no se caracteriza por ser selectivo ni de alta exigencia académica, sí es capaz de ofrecer carreras altamente demandadas.
- Incorporado este alumno en ese contexto, ve que comparte aulas con muchos alumnos que además de compartir un desempeño PSU insuficiente para optar a las carreras selectivas más altamente demandadas, muchas veces no poseen tampoco el perfil de alumno de excelencia, que nuestro alumno poseía en un contexto local.
- Así visto, podríamos suponer en nuestro alumno un serio compromiso con su proceso educativo; y podemos observar, con mediana razonabilidad, que probablemente el programa educativo en el cual se inserta no responde a sus expectativas de desempeño, y que deba hacerse cargo de deficiencias de entrada de naturaleza distinta a las insuficiencias que ella presenta.

"Entonces yo decía "si esto se ve en el colegio", y a veces la profesora preguntaba y nadie sabía de nada, y de hecho los compañeros como que me tenían mala algunos porque yo siempre respondía todo. Entonces la profe preguntaba algo, yo esperaba un rato, y no, nadie respondía, entonces iba yo. Es que de verdad era como hacer un preu casi. Yo no hice preu pero igual supongo que es así porque vi cómo se estudia."

De esta forma, la deserción observada en el presente caso, no se explica por temas económicos, ni de rendimiento, ni siquiera por temas vocacionales (pues en definitiva en el caso se identificaba una estrategia de cambio interno de Bachillerato a Medicina), sino de desmotivación derivada de expectativas de vida académica defraudadas.

4.2.6. Valoración de la modularidad académica

Originalmente, en los resultados de la encuesta, observamos que un porcentaje mayoritario de integrantes de la cohorte tenía entre sus aspiraciones y expectativas trabajar y estudiar, lo que es consistente con uno de los hallazgos de las entrevistas cualitativas: estos jóvenes tienen incorporado en su imaginario la construcción curricular por módulos. Así, más de un entrevistado señaló que su estrategia sería obtener un título técnico en el marco de un programa que le permitiera continuar luego estudios de ingeniería de manera vespertina.

"...qué más... me iba a meter directamente a ingeniería comercial pero me metí al técnico por la cuestión de la beca.[...] Igual se pueden convalidar los ramos y seguir hasta ingeniería comercial en INACAP. Te dan esa opción. [...] después administración, después ingeniería en administración de empresas y después ingeniería comercial un año más. Son 5 años."

Conceptualmente, podemos verificar que los jóvenes estudiados tienen asumido que empezarán a vivir lo que algunos autores⁸ denominan trayectorias sincronizadas, es decir el desarrollo simultáneo y/o alterno de actividades de estudio y trabajo. Podríamos levantar la hipótesis que en los sectores populares y de reciente incorporación a la educación terciaria, los estudios superiores no son una etapa de preparación para la vida adulta, sino que parte de ella dada la superposición de los roles de estudiante y trabajador.

"...sí, buenos ramos. Ahora hace poco nos dijeron que si terminábamos en una materia completa este semestre podíamos salir con IT1 que es como una certificación. Eso es lo que tenemos que ir acumulando nosotros, certificaciones, pa'que las empresas vean 'aah ya, usted está certificado'. Y ahí nosotros si nos certificamos podemos ya trabajar en algo, podemos ya estar fijo trabajar en algo, sin haber terminado la carrera, con esa certificación."

Ahora bien, a partir de las entrevistas la idea de trayectorias sincronizadas puede ser complementada con una serie de figuras tales como trayectorias alternas o trayectorias yo-yo como se identifican en la bibliografía⁹ (se trabaja para juntar dinero para después estudiar durante un periodo) trayectorias modulares, o bien la combinación de estas.

"...seguir trabajando y juntar un ahorro, a ver si pago la matrícula y después... ahí voy a ver cómo la hago"

Vale decir, si bien la idea de trayectorias sincronizadas es ilustrativa principalmente de la realidad los estudiantes de programas vespertinos, es posible observar algún tipo de trayectorias de mayor complejidad.

4.2.7. Imaginario juvenil sobre la educación superior

También se pudo observar que la educación superior se ubica en el imaginario de los jóvenes como un dispositivo eficaz que permite un acceso casi directo a otros bienes o aspectos valorados de la vida. Así, para nuestros entrevistados, la educación superior aparece como garante de calidad de vida, como proveedora de status, además la referida imagen de herencia y única dote de parte de sus familias.

"...algo importante el estudio obviamente, porque uno sin estudios no es nadie. Y si quedo sin estudios pasaría a ser alguien más del montón, y nunca me ha gustado eso. Preferible siempre ser sobresaliente por lo bueno y no por lo malo."

Porque si yo saliera de cuarto ¿qué soy? Nada. O sea, salí de cuarto, terminé la enseñanza media, la básica, todo, pero como que siento que uno está vacío. Entonces si sigue estudiando y tienes un título ya eres alguien..."

Los estudios superiores aparecen como garantes de calidad de vida en tanto las credenciales académicas permiten una inserción menos precaria al mercado laboral, en el sentido que un profesional o técnico tendría mejores ingresos y mayor estabilidad laboral.

"...trabajar en el día, estudiar en la noche, cosas así. Es como súper importante los estudios, porque si uno tiene sólo cuarto medio a lo más que puedo optar es a trabajos menores con remuneración de a lo más doscientos mil, trescientos mil pesos mensuales."

⁸ Irrazabal y Oyarzún (2003)

⁹ Machado Pais (2002)

"...porque ganan cualquier plata. Si yo me medía por lo que ganaban plata. Siempre me ha gustado la plata, así que... ganar harta plata." - Consultado por que le gustaba la carrera –

Lo anterior se traduce en calidad de vida en tanto contribuiría a evitar situaciones angustiosas asociadas a la pobreza, por lo cual uno de los principales retornos de la educación sería vivir con tranquilidad.

"...y tener mis hijos y que puedan... que pueda darle lo que ellos quieran. Debe ser fome pa'un papá que su hijo le pida algo y no tenerlo, no poder dárselo. Entonces por eso es algo que... el estudio es algo que ayuda porque sino los trabajos de ahora, pa'la gente que no ha estudiado, más de doscientas lucas no... y es como mucho doscientas lucas. Entonces con doscientas lucas uno, anda apretado."

"En cambio, cuando uno ya tiene el título y una carrera rentable ya puede optar a mayor... pa'vivir tranquilo. Es como mi meta. No quiero ganar mucho sino que vivir tranquilo, relajado, y no estar preocupado de que 'ah, tengo que pagar la cuenta de la luz."

En el imaginario de estos jóvenes, las credenciales académicas que entrega la educación superior también aparecen como proveedoras de estatus. La posesión de un título es valorada no tanto en función de los conocimientos y competencias laborales asociadas a ella, sino que por su función ordenadora en la escala social. Hay una valoración esencialista del título técnico y/o profesional ya que esta certificación por sí misma permitiría distinguir entre quienes "son alguien en la vida" y quiénes no.

"...yo quiero estudiar porque sé que es la única forma de surgir no más, pero no porque me guste, o me guste leer, no."

"...porque es más conocimiento, más experiencia. Con el título igual tú vai a llegar al trabajo y te van a ver... ah, ingeniería estudió este cabro. No te van a ver como una persona normal y corriente."

"...sí, es que era... igual tenía como prestigio, y además puedo seguir estudiando en la A...B..., eso es como lo que me motiva, porque igual tener en estos tiempos un título de instituto no es lo mismo que tener un título universitario."

Finalmente - y posiblemente influenciados por el discurso de sus padres- observamos que la educación es percibida por los jóvenes como la única herencia que recibirán de sus padres, lo que le da un valor especial. En efecto, es una herencia no solo porque refleja un gasto monetario efectuado por sus padres, sino que por el protagonismo de estos últimos en tanto facilitadores y motivadores en su proceso educativo.

*"...mi familia me tenía una **cuenta** hace mucho tiempo porque ya sabían que yo iba pa' ese camino, mejor que mis hermanos, tenía mejores notas. Entonces empezaron como... no sé, de que yo era pequeño. Primero me la tenían para la casa, para que yo tuviera algo de capital pa'empezar, pero como vieron que me daba pa'la educación superior, la tiraron para eso..."*

5. CONCLUSIONES

A lo largo del presente artículo, hemos observado varios de los principales hallazgos del estudio que le dio origen, los que entregan luces para enriquecer la discusión y reflexión respecto a los elementos relevantes –en perspectiva de estos jóvenes- para comprender el tránsito entre la educación media y su vida post secundaria.

En primer lugar se observó el financiamiento como un estructurador fundamental en la opción de continuar o no estudiando. Es cierto que el creciente número de becas y ayudas estatales surgidas en

Chile durante los últimos años es percibido como relevante y un facilitador por parte de estos jóvenes, no obstante lo cual los jóvenes siguen percibiendo el apoyo de sus padres como prioritarios, esto último es probablemente indicio de la relevancia que tienen las familias chilenas en el financiamiento del sistema de educación superior chileno.

A su vez los padres no tan solo soportan materialmente a los jóvenes, sino que siguen siendo un sostén anímico y moral fundamental en su vida post secundaria.

No obstante lo anterior, el hecho que varios de los jóvenes estudiados estén empezando a desarrollar trayectorias sincronizadas, vale decir la combinación de estudios con trabajo, lo que permitiría hipotetizar que para los jóvenes de este perfil socioeconómico y cultural los estudios superiores no son sólo una etapa de preparación para la vida adulta, sino que parte de ella está dada por la superposición de roles de estudiante y trabajador y por ende es ya parte de su vida de adultos. La evidencia obtenida corrobora la pertinencia de los marcos conceptuales utilizados y abre la puerta para el desarrollo de otros estudios de caso que permitan robustecer la conceptualización sobre trayectorias juveniles.

Se observa que estos jóvenes, provenientes de sectores socioeconómicamente vulnerables y sin referentes familiares respecto al mundo de la educación superior, se ven afectados por asimetrías de información. Principalmente respecto a su *"peso relativo"* en la competencia por un cupo en carreras e instituciones altamente selectiva. Así a través de la información recogida es posible enterarse de la existencia de un grupo importante de jóvenes que se planteaba la posibilidad de estudiar medicina sin elaborar un plan alternativo, siendo que sus antecedentes y preparación indicaban que objetivamente era improbable que lo logaran.

Al parecer los preuniversitarios cumplen un rol remedial respecto a esta asimetría de información en tanto al someter a los jóvenes a constantes ensayos les va otorgando una adecuada información respecto a su *"peso específico"* dentro del proceso que enfrentaran y los pone en contacto con fuentes de información más diversas que las que encuentran solo en el liceo (compañeros de otros liceos o incluso jóvenes que ya han pasado por la universidad). En términos normativos aparece como relevante pensar en la formulación de políticas públicas que reduzcan esta asimetría de información, calibrando a los jóvenes en términos de su potencial desempeño en el proceso de selección universitaria, independientemente de que puedan acceder a preuniversitarios.

Otra constatación interesante es el caso de que una joven de alto rendimiento en su liceo, que desertó de una carrera universitaria no por mal rendimiento, sino que por la frustración de enfrentarse a bajos niveles de exigencia tanto de sus profesores como de sus compañeros, este caso abre una interrogantes ¿es este un fenómeno extendido? O ¿solamente corresponde a acasos marginales?, lo cual debería indagarse con nuevas investigaciones que vayan más allá de las hipótesis tradicionales manejadas en los estudios de deserción y retención en la educación superior

La alta valoración que se observa tienen estos jóvenes respecto a la educación superior, da cuenta de la importancia de seguir profundizando respecto a estas interrogantes.

REFERENCIAS BIBLIOGRÁFICAS

Dávila, O.; Ghiardo F.; Medrano, C. (2008). Los Desheredados. Trayectorias de vida y nuevas condiciones juveniles. Viña del Mar: CIDPA Ediciones.

- Machado Pais, J. (2002). Laberintos de vida: paro juvenil y rutas de salida (jóvenes portugueses). *Revista de Estudios de Juventud*, 56.
- McDonough, P.M. (1997). *Choosing Colleges: How social class and schools structure opportunity*. Albany: State University of New York Press.
- Oyarzún, A. e Irrazabal, R. (2003). Comportamiento de las trayectorias educacionales y laborales en jóvenes estudiantes. *Revista Última Década*, 18.



**PASADO, PRESENTE Y FUTURO. PROCESOS Y
PROCEDIMIENTOS EN LA EVALUACIÓN DE LOS
DIRECTORES Y DIRECTORAS DE LOS CENTROS ESCOLARES
DE ANDALUCÍA**

**PAST, PRESENT AND FUTURE. PROCESSES AND PROCEDURES IN THE
EVALUATION OF THE DIRECTORS OF THE SCHOOLS OF ANDALUSIA**

Esther Lorenzo Guijarro y Antonia Ramírez García

Revista Iberoamericana de Evaluación Educativa 2010 - Volumen 3, Número 2

<http://www.rinace.net/riee/numeros/vol3-num2/art11.pdf>

Fecha de recepción: 14 de marzo de 2010
Fecha de dictaminación: 06 de mayo de 2010
Fecha de segundo envío: 12 de mayo de 2010
Fecha de aceptación: 13 de mayo de 2010

Las sociedades evolucionan exigiendo mejoras en todos los aspectos y sectores que las componen, en ellas los centros escolares ocupan un lugar de privilegio y a los que se vuelven todas las miradas debido a su condición de garantes de la cultura y transformadores de la sociedad. A ellos, a los centros escolares, se les pide, quizás, por este motivo, mejores resultados, mayor eficacia, en definitiva, más calidad en la educación que reciben sus alumnos y alumnas. La responsabilidad de dar respuesta a estas demandas no queda diluida en la colectividad del profesorado que ejerce su docencia en el centro escolar, sino que recae en una figura determinada, el director o directora del centro, que cada vez más goza de mayor autonomía para la gestión del mismo.

Los tiempos cambian y las funciones del director se hacen cada vez más complejas en esa búsqueda del ideal educativo, de la eficacia, de los resultados, ya no se trata de administrar los recursos y ejercer la jefatura del personal, o de representar a la Administración educativa, ahora se le reclama ser un líder pedagógico, mantener el equilibrio entre la estabilidad y el cambio, conservar las fortalezas del centro escolar y de sus profesores y eliminar las debilidades que pudieran tener, rentabilizar las oportunidades que le ofrece el entorno y desintegrar las amenazas del mismo, todo ello en aras de la calidad de la educación.

Pero, al mismo tiempo que las exigencias y la autonomía se incrementan, también la necesidad de rendir cuentas; en este sentido, la Administración pública educativa muestra un interés creciente en la evaluación como medio para la mejora cualitativa de los servicios que gestiona. La evaluación es considerada en la actualidad como un instrumento privilegiado de la administración y de la política educativa, como una herramienta eficaz para la gestión y la mejora de la calidad de la enseñanza. En los procesos evaluativos el director o directora también ocupa una posición central, pues no podemos olvidar, sus funciones. A lo largo de estas páginas realizaremos un recorrido histórico respecto a la figura y funciones del director de un centro escolar, analizaremos los procesos y procedimientos de evaluación en que se ven envueltos los directores y directoras de la Comunidad Autónoma de Andalucía y plantearemos una nueva evaluación de los mismos a través del criterio *liderazgo*, propuesto por el modelo de excelencia EFQM que comenzará a aplicar en un futuro inmediato en nuestros centros educativos.

1. LA DIRECCIÓN DE LOS CENTROS EDUCATIVOS DE EDUCACIÓN PRIMARIA. HACIENDO UN POCO DE HISTORIA

Murillo y Gómez (2006) realizan un repaso sobre la evolución de la dirección escolar en España centrándose en los cambios que ha experimentado la normativa que la regula; en este sentido, el tránsito del siglo XIX al XX marca el inicio de la dirección escolar en educación primaria. En 1896 comienza a extenderse en España el modelo de escuela graduada y con ella surge la necesidad de crear una figura institucional que se encargue de coordinar a los maestros y maestras y distribuir responsabilidades; una figura, por tanto, con competencias diferentes a la de los docentes.

La primera norma referente a la dirección escolar en España es la Real Orden de 9 de marzo de 1918, por la que se establece el Reglamento de Escuelas Graduadas. De ahí surgen los llamados *maestros-directores* que, sin necesidad de una formación específica, alternaban su labor docente con los trabajos propios de la dirección. El Reglamento de Escuelas, de 23 de septiembre de 1898, habló por primera vez de la escuela graduada y de los directores, así como algunas de las tareas que se les encomendaba, entre

ellas la de *cuidar que existiese armonía entre los maestros* y de *la disciplina general de la escuela*. Posteriormente, la Ley de 1945 y el Estatuto de Magisterio de 1947 establecían que el director era el representante legal de la escuela y superior inmediato de los profesores; su selección y quedaron definidas en el Estatuto, en él se señalaba que la selección del director se realizaba mediante oposición restringida, en el caso de grupos escolares de más de seis secciones, y por nombramiento del Ministerio en el caso de grupo de menos de seis unidades. En la selección mediante oposición, los candidatos deberían ser maestros con más de cinco años de servicio, o bien Licenciados en Pedagogía o en Filosofía y Letras con un año de servicio; todos ellos con Informe favorable de la Inspección sobre su competencia, moralidad y aptitud.

Los directores al no tener docencia constituían un *cuerpo*, de este modo, el Decreto 985 de 20 de abril de 1967 estableció el *Reglamento de Cuerpo de Directores Escolares*, en el que el director se configura como el representante legal del centro y el superior inmediato de los profesores. El acceso al cuerpo es libre y pasa por un curso selectivo de formación al que se accede tras un examen de ingreso; al mismo tiempo, las tareas pedagógicas del director o directora se amplían (programa, libro de textos, temporalización de los contenidos o información a las familias); en este sentido, podemos apreciar que la figura del director o directora cobra una dimensión diferente, pues se le adjudica un perfil profesional determinado y constituye un cuerpo diferenciado.

Con la Ley General de Educación de 1970, los requisitos para ser director o directora se centran en la antigüedad en el cuerpo (tres años). Esta Ley elimina el *cuerpo de directores* como grupo diferenciado de los docentes, incorpora al cargo de director o directora la función docente y elimina su carácter vitalicio.

2. LAS FUNCIONES DEL DIRECTOR O DIRECTORA DE UN CENTRO DE ENSEÑANZA OBLIGATORIA

Las funciones directivas y sus respectivas dimensiones se encuentran encaminadas a dar cumplimiento al objetivo de promover, impulsar, coordinar, supervisar y controlar todas las actividades de la institución escolar.

Nuestro actual sistema educativo plantea la necesidad de proporcionar una educación de calidad a todo el alumnado partiendo de la diversidad de sus intereses, características y situaciones personales. Ello obliga a reconocer a los centros una capacidad de decisión que afecta tanto a su organización como a su modo de funcionamiento. Los centros, tal y como se reconoce en la normativa actual vigente, han de poseer un margen propio de autonomía que les permita adecuar su actuación a sus circunstancias concretas. Esta progresiva descentralización conlleva cambios en la dirección del centro escolar, concediendo a esta figura mayor protagonismo y capacidad de decisión traduciéndose ésta en un aumento de sus funciones. Esta nueva fisonomía requiere de mayor especialización del director del centro educativo, ya que realizará su intervención en cuestiones que van más allá de lo escolar. A lo largo de las distintas normas legales que han regulado nuestro sistema educativo podemos apreciar esta transformación de las funciones (vid. figura 1).

Si analizamos este desarrollo legislativo, apreciamos cómo las funciones asignadas al director han ido ampliándose modificando el perfil de éste. De la LODE (1985) a la LOPEGCE (1995) las competencias directivas se insertan en el ámbito burocrático (controla obras, servicios y suministros, autoriza gastos de acuerdo con presupuesto, ordena pagos y visa las certificaciones y documentos oficiales) y ejecutivo (ejerce jefatura personal y adopta resoluciones disciplinarias, garantiza el cumplimiento de las leyes o

convoca y preside reuniones y actos académicos). Con la posterior ley, la LOCE (2002), el ámbito competencial se amplía a la integración (fomenta el clima escolar que favorezca el estudio, favorece la convivencia en el centro e impulsa la colaboración familia) y a la innovación (promueve planes de mejora de la calidad del centro y proyectos de innovación e investigación educativa, impulsa procesos de evaluación interna del centro y colabora en las evaluaciones externas y con la evaluación del profesorado), otorgando a la dirección un carácter pedagógico y mediador del que antes carecía. Con la ley vigente, la LOE (2006), los campos de actuación descritos son similares, aunque se amplían las funciones dentro de ellos, como la responsabilidad que se le otorga en la dirección pedagógica del centro o la potestad disciplinaria del personal del centro. La LEA (2007), por su parte, sigue las mismas directrices planteadas a nivel estatal.

FIGURA 1. EVOLUCIÓN DE LAS FUNCIONES DE LA DIRECCIÓN ESCOLAR

Funciones	1985 LODE	1990 LOGSE	1995 LOPEGCE	2002 LOCE	2006 LOE	2007 LEA
Representar al centro ante la Administración.	X	X	X	X	X	X
Representar a la Administración en el centro.			X		X	X
Cumplimiento de la normativa.	X	X	X	X	X	X
Dirigir y coordinar las actividades del centro.	X	X	X	X	X	X
Ejercer la jefatura del personal del centro.	X	X	X	X	X	X
Convocar y presidir los actos académicos y las reuniones de todos los órganos colegiados del Centro.	X	X	X	X	X	X
Ejecutar los acuerdos adoptados por los órganos colegiados del centro.		X	X	X	X	X
Autorizar los gastos de acuerdo con el presupuesto del centro, ordenar los pagos y visar las certificaciones y documentos oficiales del centro.	X	X	X	X	X	X
Proponer el nombramiento de los cargos directivos.	X	X	X	X	X	X
Proponer el cese de los cargos directivos.			X	X	X	X
Realizar las contrataciones de obras, servicios y suministros.			X	X	X	X
Adquisición de bienes.						X
Promover planes de mejora de la calidad del centro.				X	X	X
Promover proyectos de innovación e investigación educativa.				X	X	X
Impulsar procesos de evaluación interna del centro y colaborar en las evaluaciones externas.				X	X	X
Colaborar en la evaluación del profesorado.				X	X	X
Favorecer la convivencia en el centro e imponer medidas disciplinarias.			X	X	X	X
Impulsar la colaboración con las familias, con instituciones y con organismos que faciliten la relación del centro con el entorno.				X	X	X
Fomentar un clima escolar que favorezca el estudio y el desarrollo de cuantas actuaciones propicien una formación integral en conocimientos y valores de los alumnos.				X	X	X
Responsable de la dirección pedagógica.					X	X
Responsable de que el equipo directivo, establezca el horario que corresponde a cada área, materia, módulo o ámbito.					X	X
Proponer requisitos de especialización y capacitación profesional respecto de determinados puestos de trabajo docentes del centro incluida la aprobación del gasto.					X	X
Potestad disciplinaria del personal.					X	X
Potestad en la toma de decisiones en lo que se refiere a las sustituciones de las ausencias del profesorado.					X	X

LODE (Ley Orgánica del Derecho a la Educación), LOGSE (Ley Orgánica de Ordenación General del Sistema Educativo), LOPEGCE (Ley Orgánica de Participación, Evaluación y Gobierno de los Centros Docentes), LOCE (Ley Orgánica de Calidad de la educación), LOE (Ley Orgánica de Educación), LEA (Ley de Educación de Andalucía).

Fuente: Elaboración propia

3. LA EVALUACIÓN DE LOS DIRECTORES Y DIRECTORAS EN LOS CENTROS ESCOLARES

Cantón (2001), considera que la mejora es el primer paso hacia la calidad y que por tanto, la evaluación se convierte en una actividad indispensable para identificar los diferentes elementos que contribuyen al logro de los objetivos educativos, orientando la mejora de la práctica.

La descentralización y la autonomía pedagógica deja a los centros educativos un margen de libertad más amplio para gestionar sus recursos; sin embargo, esta política va acompañada de una evaluación más rigurosa de éstos. En este sentido, la LOE (2006) considera la evaluación un elemento fundamental para la mejora de la educación, y de forma más específica, refiriéndose a la evaluación de la función directiva, para mejorar el funcionamiento de los centros educativos.

A lo largo de estos años se ha detectado cierta reticencia a ocupar los cargos directivos en los centros escolares, quizás por el incremento de sus funciones, tal y como hemos podido apreciar, pero, tal vez, también, por la evaluación a la que el director o directora de un centro se ve sometido, ya que son varios los aspectos y momentos en los que el cumplimiento de estas funciones van a ser evaluadas (vid. figura 2).

FIGURA 2. ASPECTOS Y MOMENTOS DE LA EVALUACIÓN DE LA DIRECCIÓN ESCOLAR

CUÁNDO SE EVALÚA	QUÉ SE EVALÚA	PARA QUÉ SE EVALÚA
Antes del ejercicio	Los requisitos exigidos para ser director/a	Selección del director/a
Durante el ejercicio	La fase de prácticas	Habilitación para ser director/a
Después del ejercicio	El desarrollo de la función directiva	Renovación del cargo Consolidación económica

Fuente: Elaboración propia

3.1. El proceso de selección de los directores y directoras

Cuando un centro educativo se crea requiere una persona que lo dirija, cuando un director o directora se jubila, acaba su mandato o es destituido se exige que otra persona lo sustituya; en las diferentes posibilidades que hemos nombrado hay dos opciones posibles, la elección o la selección.

En nuestro país, apunta Martínez Lobato (2006) que, a pesar de los cambios introducidos al modelo electivo de directores de la LODE (1985), el modelo selectivo de la LOCE (2002) y la actual LOE (2006) carece de los necesarios atractivos para que el profesorado se presente y comprometa en el desempeño de este cargo. Por su parte, Batanaz (2006:4-5) recoge de Murillo, Barrio y Pérez Albo (1999) la opinión de directores e directoras sobre las causas que con mayor frecuencia reducen la presencia de candidatos a la dirección de un centro escolar, éstas son:

- a) La dificultad de desempeñar un complejo trabajo para el que no se sienten preparados.
- b) La insuficiencia de las compensaciones que se reciben por ello.
- c) El temor a comprometerse en un rol poco aceptado por sus compañeros.
- d) La falta de poder real y de autonomía.
- e) El escaso apoyo recibido de la administración educativa.
- f) El escaso reconocimiento social de la función.

La primera fase en el camino a la dirección de un centro educativo es, tal y como expresa Martínez Lobato (2006), elegir o seleccionar al futuro director o directora; en este trabajo nos centraremos en el modelo de selección de directores en centros públicos; en este sentido, éste ha cambiado de forma sustancial, de ser una figura seleccionada exclusivamente por los miembros de la comunidad escolar (docentes, padres y madres y alumnado) a través de sus representantes en el consejo escolar (LODE, 1985), ha pasado a un modelo en el cual, en primer lugar, el aspirante a la dirección tiene que estar acreditado, es decir, previamente seleccionado por la administración en función de sus méritos, desarrollo de un proyecto de dirección, formación y experiencia; en segundo lugar, ser evaluados por una comisión y, finalmente, ser seleccionados para ejercer el cargo en cuestión. En la figura 3 podemos apreciar cómo, al igual que sucedía con las funciones del director, los requisitos necesarios para que el director sea seleccionado han ido creciendo.

FIGURA 3. EVOLUCIÓN DE LA LEGISLACIÓN DE APLICACIÓN A LA SELECCIÓN DE LOS CANDIDATOS A LA DIRECCIÓN ESCOLAR.

CONDICIONANTE						
	REQUISITOS	ANTIGÜEDAD	ACREDITACIÓN	PROYECTO DE DIRECCIÓN	PROCEDIMIENTO DE SELECCIÓN	DURACIÓN DEL CARGO
1985 LODE	- Profesor/a del centro. - 1 año de permanencia en el centro. - 3 años de experiencia docente.	- 1 año de permanencia. - 3 años de antigüedad.	No	No	Consejo Escolar	3 años
1990 LOGSE	- Funcionario de carrera. - Destino definitivo. - 1 año de permanencia en el centro. - 3 años de experiencia docente.	1 año de permanencia. - 3 años de antigüedad	No	Programa de Dirección	Consejo Escolar	3 años
1995 LOPEGCE	- Funcionario/a de carrera. - Acreditación. - 1 año de permanencia en el centro. - 5 años de antigüedad en el cuerpo.	- 1 año de permanencia. - 5 años de antigüedad.	Acreditación de méritos.	Programa de Dirección	Consejo Escolar	4 años
2002 LOCE	- Funcionario de carrera. - Destino definitivo. - 5 años de antigüedad en el cuerpo. - 1 año de permanencia en el centro.	- 1 año de permanencia. - 5 años de antigüedad.	Acreditación de méritos.	No se especifica	Comisión: Administración, padres/madres, docentes.	3 años
2006 LOE	- Funcionario de carrera. - Destino definitivo. - 5 años de antigüedad en el cuerpo. - 1 año de permanencia en el centro.	- 1 año de permanencia. - 5 años de antigüedad.	Acreditación de méritos.	Programa de Dirección	Comisión: Administración, padres/madres, docentes.	4 años

LODE (Ley Orgánica del Derecho a la Educación), LOGSE (Ley Orgánica de Ordenación General del Sistema Educativo), LOPEGCE (Ley Orgánica de Participación, Evaluación y Gobierno de los Centros Docentes), LOCE (Ley Orgánica de Calidad de la educación), LOE (Ley Orgánica de Educación).

Fuente: Elaboración propia

En la Comunidad Autónoma de Andalucía el proceso de selección y nombramiento de directores y directoras de centros docentes públicos ha quedado regulado en el Decreto 59/2007, de 6 de marzo. Este Decreto favorece el acceso a la dirección de las personas más cualificadas y mejor valoradas por la comunidad educativa para desarrollar un proyecto basado en la participación de todos los sectores, en el conocimiento del propio centro docente y su entorno, de su realidad social, económica, cultural y laboral y en la consecución de los objetivos educativos que se proponga, independientemente de que el futuro director o directora pertenezca o no a dicho centro educativo. En la figura 4 podemos apreciar los cuatro aspectos fundamentales del proceso de selección: requisitos, concurso de méritos, evaluación y nombramiento.

FIGURA 4. PROCESO DE SELECCIÓN DE DIRECTORES Y DIRECTORAS EN ANDALUCÍA

REQUISITOS	Tener una antigüedad de al menos cinco en el cuerpo.	
	Haber impartido docencia directa como funcionario/a de carrera, durante un período de cinco años.	
	Tener una antigüedad de un año en el centro.	
	Presentar un proyecto de dirección	Conocimiento del Plan de Centro, de su realidad social, económica, cultural y laboral.
SELECCIÓN (concurso de méritos)		Desarrollo de actuaciones que favorezcan la igualdad de oportunidades entre hombres y mujeres.
		Desarrollo de actuaciones para la mejora de la convivencia escolar.
		Estrategias de intervención para los objetivos educativos y para la mejora de los resultados escolares que se pretendan lograr.
	Méritos académicos	Titulaciones universitarias.
		Impartición o asistencia a cursos de formación o perfeccionamiento sobre la organización y el funcionamiento escolar o la dirección de centros docentes públicos.
		Publicaciones sobre la organización y el funcionamiento escolar o la dirección de centros docentes públicos.
	Méritos profesionales	Antigüedad como funcionario de carrera.
		Servicios efectivos prestados en el centro docente a cuya dirección se opta.
		Desempeño de cargos directivos o de coordinación docente.
		Participación como coordinador/a o como profesorado colaborador en actividades de formación del profesorado.
EVALUACIÓN (Comisión evaluadora)	Comunidad educativa	Dos representantes de los padres y madres del alumnado (miembros del Consejo Escolar).
		Dos representantes del alumnado (si procede).
		Representante del colectivo integrado por el personal de administración y servicios y el personal de atención educativa complementaria.
	Administración educativa	Representantes del profesorado en número igual a la suma de los representantes de los otros sectores de la Comunidad Educativa (miembros del consejo escolar).
		Inspector o inspectora de educación de referencia del centro docente (asumirá la presidencia).
NOMBRAMIENTO	Cuatro años, que podrán renovarse por un solo período de igual duración previa evaluación positiva.	

Fuente: Decreto 59/2007, de 6 de marzo por el que se regula el procedimiento para la selección y nombramiento de los directores y las directoras de los centros docentes públicos, a excepción de los universitarios. Elaboración propia.

El éxito o el fracaso de una u otra forma de acceso, elección o selección, depende, según Martínez Lobato (2006), no sólo de quienes eligen o de las comisiones que seleccionan, sino que también viene determinado del posterior seguimiento que se haga de los proyectos de dirección presentados y de la evaluación que de los directores (bien sean elegidos o seleccionados) se haga.

3.2. Procedimientos de recolección de información

El director o directora cuando ha sido seleccionado, comienza su mandato, pero a lo largo del mismo también va a sufrir un proceso de evaluación: durante el año de prácticas, al finalizar su mandato para renovar un periodo más, y al consolidar la retribución económica que conlleva el ejercicio de la dirección (de carácter voluntario). De este modo, la Orden de 20 de junio de 2007, por la que se regula el proceso de formación inicial de los directores y las directoras de los centros docentes públicos dependientes de la

Consejería de Educación de la Junta de Andalucía, establece todo el procedimiento que se va a seguir durante el primer año de ejercicio, éste ha quedado recogido en la figura 5.

FIGURA 5. PROCEDIMIENTO DE EVALUACIÓN DE LA FUNCIÓN DIRECTIVA EN ANDALUCÍA DURANTE EL AÑO DE PRÁCTICAS.

¿PARA QUÉ EVALUAR?	Para valorar el ejercicio de la función directiva realizado por el director o la directora a lo largo de todo el proceso de su programa de formación inicial.	
¿QUÉ EVALUAR?	<ol style="list-style-type: none"> 1. Coordina la elaboración, seguimiento y evaluación del Proyecto Curricular y el Plan Anual del Centro. 2. Intensifica el funcionamiento de los órganos de gobierno, coordinando las actuaciones del equipo directivo, del Claustro de Profesorado y del Consejo Escolar. 3. Ejerce la dirección pedagógica y la coordinación docente, impulsando el trabajo en equipo de profesores y profesoras, apoyando sus iniciativas y distribuyendo tareas y responsabilidades de manera equilibrada. 4. Promueve la implicación del profesorado en su formación y perfeccionamiento, fomenta su participación en actividades de innovación educativa y estimula la difusión de las experiencias y buenas prácticas que se realizan en el centro. 5. Organiza la oferta educativa del centro de manera que se tengan en cuenta las necesidades de todo el alumnado, especialmente de aquellos alumnos y alumnas que tienen mayores dificultades. 6. Fomenta la reflexión sobre las consecuencias metodológicas y sociales de la organización de los espacios y los tiempos y establece criterios pedagógicos para su aplicación en el centro. 7. Organiza y gestiona los recursos humanos, materiales y económicos de manera adecuada a las características del centro y procura una utilización intensiva de los mismos. 8. Promueve la existencia de normas de convivencia claras y consensuadas para abordar la resolución pacífica de los conflictos. 9. Promueve las prácticas coeducativas y genera iniciativas que favorezcan la igualdad entre los sexos en el desarrollo del proyecto educativo y en las relaciones dentro de la Comunidad escolar. 10. Impulsa las evaluaciones externas, las del propio centro y su profesorado, promoviendo la formulación de propuestas de mejora en relación con las mismas. 11. Favorece la apertura del centro a su entorno, así como la colaboración recíproca entre el centro y otros centros educativos, organismos, instituciones y entidades del entorno. 12. Toma iniciativas para aunar los intereses de los distintos sectores de la comunidad educativa y favorecer un buen clima de relaciones interpersonales. 13. Estimula la creación de asociaciones de alumnos y alumnas y de padres y madres del alumnado, y toma medidas para favorecer su participación en la vida del centro. 	
¿QUIÉN EVALÚA?	<p>La Comisión Provincial de Evaluación de formación inicial formada por:</p> <ul style="list-style-type: none"> - La persona titular de la Jefatura de Servicio de Ordenación Educativa de la provincia correspondiente. - Un inspector y una inspectora de educación que hayan intervenido en el desarrollo del programa de formación inicial. - Un director y una directora de centro docente público dependiente de la Consejería de Educación. - Un director o una directora de Centro del Profesorado de la provincia, designados por la persona titular de la Delegación Provincial de Educación. - La persona o personas que hayan realizado la coordinación pedagógica de formación para la dirección en la correspondiente provincia. - El Coordinador o la Coordinadora Provincial de Formación. 	
¿CÓMO SE EVALÚA?	Memoria de dirección sobre diversos aspectos	<ul style="list-style-type: none"> - Valoración de la formación recibida. - Reflexión sobre sus actuaciones en el ejercicio de la función directiva durante el periodo. - Relación de dicho ejercicio con los planteamientos recogidos
	Informes sobre el ejercicio de la función directiva de:	<ul style="list-style-type: none"> - La persona que ha ejercido la tutela. - El Consejo Escolar del centro. - El inspector o inspectora de referencia. - La dirección del centro del profesorado de referencia.
¿CÓMO SE EVALÚA?	Al finalizar el año de prácticas.	

Fuente: Orden de 20 de junio de 2007, por la que se regula el proceso de formación inicial de los directores y las directoras de los centros docentes públicos dependientes de la Consejería de Educación de la Junta de Andalucía. Elaboración propia.

FIGURA 6. PROCEDIMIENTO DE EVALUACIÓN DE LA FUNCIÓN DIRECTIVA EN ANDALUCÍA PARA LA CONSOLIDACIÓN DE LA RETRIBUCIÓN ECONÓMICA.

¿PARA QUÉ EVALUAR?	<ul style="list-style-type: none"> - Analizar el desarrollo de la función directiva. - Estimular y apoyar la mejora de su práctica. - Estudiar y enjuiciar todos los ámbitos de sus actuaciones. 	
¿QUÉ EVALUAR?	<p>a) Dirección y coordinación de la actividad del centro, en especial, en lo relativo a la planificación de la actividad del centro, elaboración de la programación general y de la memoria anual del centro, gestión de los recursos humanos, materiales y económicos; seguimiento, control y evaluación de la actividad del centro.</p> <p>b) Dinamización de la participación, en especial, en lo relativo al cumplimiento de las obligaciones derivadas de la presidencia de los órganos colegiados, información e impulso de los mecanismos de participación a los distintos sectores de la comunidad escolar.</p> <p>c) Impulso del clima de convivencia en el centro, en especial, en lo relativo a la atención a los miembros de la comunidad escolar, aplicación de la normativa reguladora de los derechos y deberes del alumnado, desarrollo de normas y procesos que enmarcan las relaciones entre los miembros de la comunidad escolar y el tratamiento de conflictos.</p> <p>d) Impulso de proyectos de mejora, en especial, en lo relativo al desarrollo de programas de innovación, fomento de la participación del profesorado en planes y actividades de formación, y</p> <p>e) Ejercicio de las competencias en materia administrativa y de personal, en especial, en lo relativo a la atención a los trámites y requerimientos administrativos, información puntual de las alteraciones en la prestación del servicio, control de asistencia del personal del centro.</p>	
¿QUIÉN EVALÚA?	<p>Comisión de Evaluación de la Delegación Provincial formada por:</p> <ul style="list-style-type: none"> -La persona titular de la Delegación Provincial o aquella en quien delegue, que ostentará la presidencia. -La persona titular de la Jefatura del Servicio de Inspección. -La persona titular de la Jefatura del Servicio de Ordenación. -La persona titular de la Jefatura del Servicio de Gestión de Recursos Humanos. -Un Inspector o Inspectora del Área Estructural de Evaluación. -Dos directores o directoras de centros docentes públicos, designados por la persona titular de la Delegación Provincial de Educación, que no participen en el proceso de evaluación. -Un funcionario o funcionaria de la Delegación Provincial con rango de jefe de sección, que actuará como Secretario o Secretaria, con voz y sin voto. 	
¿CÓMO SE EVALÚA?	Inspección educativa	- Seguimiento del ejercicio durante los cuatro años mediante un registro individual de actuaciones que posibilitará emitir un informe de evaluación.
	Director o directora	Memoria de autoevaluación.
	Consejo Escolar	Valoración del director o directora del centro mediante votación.
¿CÓMO SE EVALÚA?	Durante el proceso y al finalizar el mandato.	

Fuente: Decreto 59/2007, de 6 de marzo, por el que se regula el procedimiento para la selección y nombramiento de los Directores y Directoras de los Centros Docentes Públicos, a excepción de los universitarios, recoge los mismos principios y actuaciones. Elaboración propia.

Al terminar el año de prácticas y/o el mandato, el director o directora ha de ser evaluado para comprobar el cumplimiento de su proyecto de dirección (parcial o total dependiendo del caso), éste no constituye una declaración de intenciones, sino un proyecto de trabajo que ha debido llevarse a cabo y que al finalizar el periodo correspondiente exige ser evaluado.

Por su parte, el Decreto 141/2001, de 2 de junio, por el que se regula la consolidación parcial del complemento específico de los directores de los centros docentes públicos de Andalucía, establece en su artículo 3 los requisitos para la consolidación del complemento específico, entre los que se encuentra la obtención de una valoración positiva del desempeño del cargo mediante un procedimiento de evaluación de sus funciones, como veremos a continuación (vid. figura 6). Posteriormente, el Decreto 59/2007, de 6 de marzo, por el que se regula el procedimiento para la selección y nombramiento de los Directores y Directoras de los Centros Docentes Públicos, a excepción de los universitarios, recoge los mismos principios y actuaciones.

El porcentaje de consolidación, que dependerá del tiempo de ejercicio del cargo de director, será el siguiente:

- a) Por un período igual o superior a cuatro años, pero inferior a ocho, de desempeño del cargo: el 25% del complemento.
- b) Por un período igual o superior a ocho años, pero interior a doce: el 40% del complemento.
- c) Por un período igual o superior a doce años: el 60% del complemento.

La consolidación de las retribuciones del complemento económico recibido por parte del director o directora puede suponer un atractivo para la ocupación de este cargo; no obstante, son más los inconvenientes, como ya expresamos con anterioridad, los que merman la posibilidad de que candidatos más y mejor preparados ocupen la dirección escolar. Pero, a pesar de todas las dificultades, Tejero y González (2009) han detectado cuatro dimensiones que en mayor o menor medida miden la satisfacción laboral en la dirección escolar, por orden de importancia éstas son: satisfacción hacia las personas de apoyo a la gestión (equipo directivo y personal de secretaría) con un 4,2, satisfacción hacia el profesorado, satisfacción con el alumnado y las familias, de forma equitativa, con un 3,44 y 3,43 respectivamente y satisfacción con las condiciones laborales con un 3,11 en un rango de respuesta de 1 a 5.

4. LA INCORPORACIÓN DEL MODELO EFQM A LA FUTURA EVALUACIÓN DE LOS DIRECTORES Y DIRECTORAS

La función directiva es un factor clave en la mejora de los centros educativos, especialmente en la gestión del cambio. Hay un cierto consenso en que los centros que tienen capacidad para mejorar dependen, de manera significativa, de equipos directivos que contribuyen activamente a que su centro aprenda a desarrollarse, superando los retos y dificultades a los que tienen que enfrentarse. Los distintos modelos de calidad en la gestión que se están implantando en la administración pública, recogen como factor fundamental del éxito de la misma el ejercicio del liderazgo en la organización de que se trate. Del mismo modo, Serafín Antúnez (2000) señala que la dirección *es el elemento clave para promover o impedir cambios en los centros escolares*, tanto por el lugar que ocupan como por la información y relaciones que pueden manejar.

El liderazgo educativo se caracteriza por dotar de un sentido común a la organización escolar y por influir en el comportamiento de ésta, teniendo como "norte" la mejora de la calidad, que se expresa en un mayor y mejor incremento de los aprendizajes de los estudiantes. En este sentido, es posible distinguir dentro de las prácticas de los directivos, aquellas que son propias del liderazgo de aquellas que no lo son: mientras que estas últimas pueden considerar ciertas rutinas administrativas asociadas a la dirección de la organización, las primeras ponen el acento en aquella influencia sobre otros que permite avanzar hacia la consecución de metas grupales de carácter educativo (Robinson, 2009).

El liderazgo escolar se ha convertido en una prioridad de la política educativa a nivel internacional; desempeña una función decisiva en la mejora de los resultados escolares al influir en las motivaciones y en las capacidades de los maestros, así como en el entorno y el ambiente escolares. El liderazgo escolar eficaz es indispensable para aumentar la eficiencia y la equidad de la educación (Pont et al. 2008).

La importancia del liderazgo en el aprendizaje de los alumnos y alumnas ha sido destacado por Leithwood y Hopkins (2008) en los estudios cualitativos que han realizado sobre escuelas eficaces. Por su parte, la consultora McKinsey publicó en 2007 el informe internacional " *How the world's best-performing school systems come out on top*" ("Cómo los mejores sistemas escolares del mundo se ponen a la cabeza"). El estudio refleja que si no se cuenta con un director eficiente, las escuelas tienen pocas probabilidades de poseer una cultura con altas expectativas, o de esforzarse en pos de la mejora continua. Las pruebas sugieren que un fuerte liderazgo educativo es particularmente relevante a la hora de lograr mejoras. Las reformas llevadas a cabo en Boston, Inglaterra y Singapur demuestran que un buen liderazgo escolar es esencial para poner en práctica cambios sustanciales y rápidos que lleven a convertir a sus directores en impulsores de mejoras en la instrucción. En general, el desarrollo de líderes de instrucción eficientes en las escuelas requiere, a su juicio, de tres elementos:

4.1. Conseguir a los docentes apropiados para desempeñarse como directores

Los sistemas con más alto desempeño ofrecen los incentivos apropiados para que los mejores docentes se postulen para puestos de liderazgo e implementan procesos efectivos para seleccionar a los mejores entre todos los postulantes. La manera en que lo hacen depende principalmente de si la selección del director es centralizada (controlada por el distrito escolar o el ministerio) o descentralizada (controlada por las propias escuelas). En Singapur, por ejemplo, los salarios de los directores son altos, en parte como reconocimiento de las exigencias del puesto, y también para atraer a los mejores candidatos. Como parte del exigente proceso de selección de directores, los candidatos deben pasar por un Centro de Evaluación, que establece una serie de ejercicios cuidadosamente diseñados para provocar conductas observables relativas a las competencias que requiere el puesto. Los candidatos que demuestran potencial para el puesto asisten a un programa de seis meses desarrollado por el Instituto Nacional de Educación en el que son evaluados continuamente por el equipo de capacitación. Sólo los candidatos aptos para desempeñarse como directores y ser asignados a escuelas específicas son designados directores.

4.2. Desarrollar habilidades de liderazgo en instrucción

Aunque el líder educativo cuente con competencias necesarias para desempeñar el puesto, también es necesario suministrar un conjunto adecuado de habilidades para convertirse en líderes eficientes. Básicamente, todos los líderes escolares exitosos contemplan el mismo repertorio de prácticas de liderazgo básicas. Los mejores sistemas educativos implementan un modelo de desarrollo coherente y alineado (basado frecuentemente en un modelo de aprendizaje) que ayuda a los líderes educativos presentes y futuros a desarrollar estas prácticas.

Una vez que el sistema educativo ha identificado y formado a las personas apropiadas con las habilidades necesarias, debe entonces estructurar sus funciones, expectativas e incentivos para asegurarse de que sus directores se concentren en el liderazgo instructivo y no en la administración del establecimiento. Esto contrasta con los sistemas educativos en que muchos directores dedican la mayor parte de su tiempo a tareas que no se relacionan directamente con la mejora de instrucción en sus escuelas, limitando así la capacidad para hacer realidad mejoras concretas en los resultados de los alumnos. Los sistemas que pretenden emplear a sus directores como impulsores de reformas prevén que estos se conviertan en instructores excelentes y destinen la mayor parte de su tiempo a entrenar a los docentes.

4.3. Selección y desarrollo de líderes de instrucción efectivos

El entrenamiento es efectivo como intervención, pero puede adquirir mayor relevancia una vez que las escuelas hayan desarrollado la cultura de entrenamiento y desarrollo que lo sostendrá. Para lograrlo, algunos sistemas educativos se han asegurado de que sus líderes educativos sean también 'líderes de instrucción' y han puesto en práctica mecanismos para seleccionar los mejores docentes como directores. Luego los capacitan para convertirse en líderes de instrucción que dedicarán buena parte de su tiempo a entrenar y aconsejar a los docentes. Los directores de escuelas pequeñas de la mayoría de los sistemas con más alto desempeño dedican el 80 por ciento de su tiempo a mejorar la instrucción y exponer un conjunto de conductas que cimentan la capacidad y motivación de sus docentes, para mejorar en forma constante su propia instrucción.

Para Álvarez (2001), todos los modelos de gestión de calidad: modelo Deming (Japón), modelo Baldrige (EE.UU.), modelo EFQM (Europa), etc., consideran el ejercicio de un cierto tipo de liderazgo como un factor básico de la calidad. De acuerdo con Martínez y Ruipérez, (2005) la finalidad del modelo aplicado al ámbito educativo es conocer la realidad del centro mediante la autoevaluación, y reflexionar sobre ella, lo que posibilitará elaborar planes y estrategias de mejora. Estas dos características, junto con la implicación de todos en el proceso de autoevaluación y la toma de decisiones, se consideran factores que inciden en la mejora de la calidad de los servicios que el centro educativo presta a la sociedad.

La Ley 17/2007, de 10 de diciembre, de Educación de Andalucía (LEA), regula que la evaluación del Sistema Educativo Andaluz se orientará a la mejora permanente del mismo y al aprendizaje satisfactorio y relevante del alumnado que contribuya a su éxito escolar. Además, establece que dicha evaluación se extenderá a todos los ámbitos del Sistema Educativo regulados en la misma Ley y se aplicará sobre los procesos de aprendizaje y resultados del alumnado, la actividad del profesorado, los procesos educativos y la función directiva.

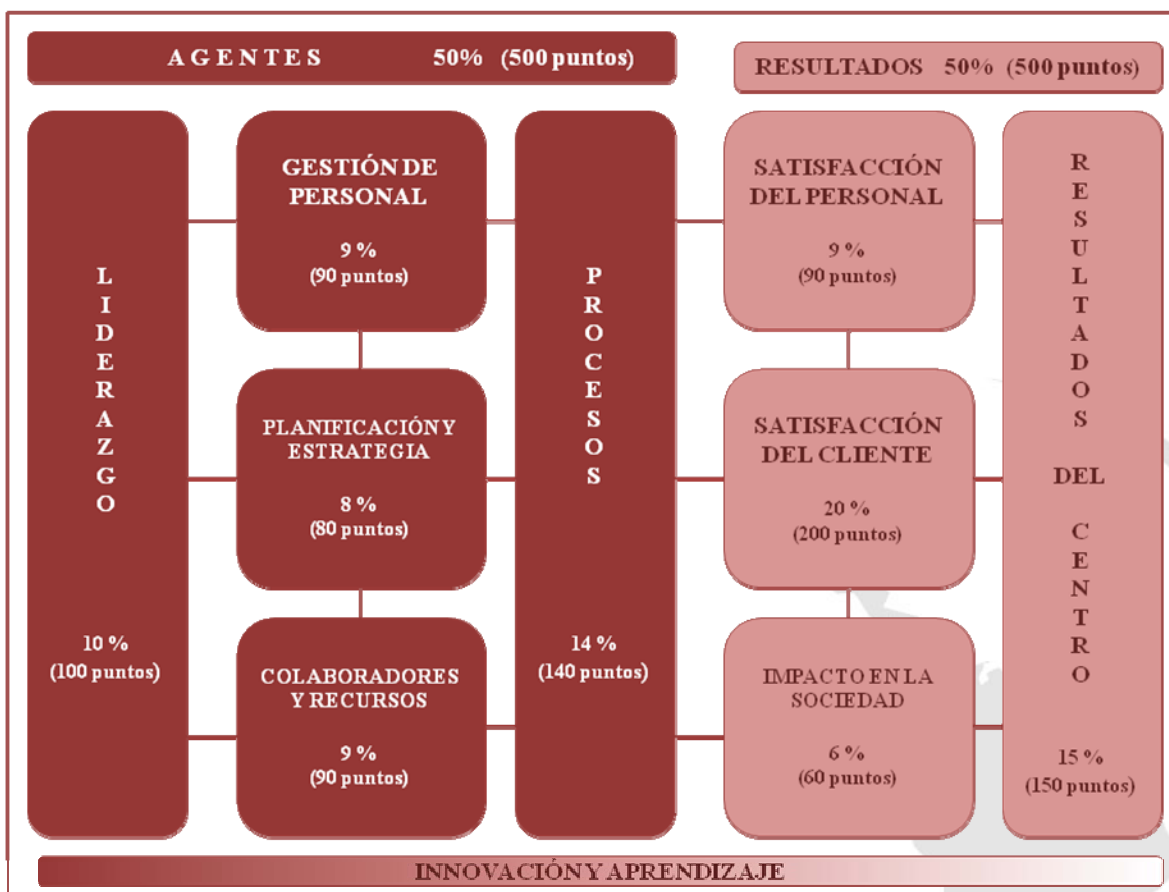
Esta evaluación será realizada por la recién creada Agencia Andaluza de Evaluación Educativa (AGAEVE), que establece como objetivo prioritario la implantación del modelo de gestión EFQM (AGAEVE, 2009). Los principios y criterios técnicos del modelo integrado de evaluación del Sistema Educativo Andaluz que sustenta este plan estratégico de actuación de la AGAEVE suponen una evaluación rigurosa y sistemática basada en la detección de necesidades, el establecimiento de objetivos e indicadores de resultados, la formulación y desarrollo de procesos, la elaboración de instrumentos, el registro de evidencias, la autoevaluación y evaluación externa.

La evaluación externa o institucional se realizará, entre otras, sobre el ejercicio de la dirección de los centros docentes, que se centrará en las buenas prácticas directivas sobre liderazgo, organización y el funcionamiento de los centros.

La adaptación de este modelo a los centros educativos integra nueve criterios, cinco de ellos, los del grupo "agentes facilitadores", se ocupan de la dirección del centro educativo, de lo que hace y de cómo lo hace; los cuatro restantes, integrados en el grupo de "resultados", versan sobre lo que consigue el centro a través de los criterios "agentes facilitadores" (vid. figura 7). De todos ellos nos interesa destacar un criterio concreto de los agentes facilitadores, nos estamos refiriendo al liderazgo, que puede definirse como el comportamiento y la actuación del equipo directivo y del resto de los responsables que guían al centro educativo hacia la mejora continua. Al mismo tiempo, este criterio refleja cómo todos los que tienen alguna responsabilidad en el centro educativo promueven y facilitan la consecución de los fines y objetivos, desarrollan los valores necesarios para alcanzar el éxito e implantan todo ello en el centro, mediante las acciones y los comportamientos adecuados, estando implicados personalmente en asegurar que el sistema de gestión hacia la mejora continua se implemente en el centro.

En la figura 7 se aprecia que el liderazgo es el primero de los criterios que aparece con una ponderación del 10%, destacando, por tanto, como un importante agente facilitador dentro del modelo.

FIGURA 7. MODELO EUROPEO DE GESTIÓN DE LA CALIDAD



Fuente: MEC (2001). Modelo Europeo de Excelencia. Madrid: MECD.

Los criterios se desglosan en subcriterios y, estos a su vez en áreas, lo que configura un sistema de indicadores cuya existencia o no en el centro educativo permite tener un conocimiento del mismo basado en hechos y no sólo en opiniones o percepciones. La figura 8 recoge los subcriterios y áreas del criterio Liderazgo. Nos hemos centrado en él por su vinculación directa con el ejercicio de la dirección escolar; sin

embargo, no podemos olvidar que el buen funcionamiento del centro educativo implica la participación y evaluación de todos y cada uno de los criterios del modelo EFQM, en los que se reflejaría de forma nítida no sólo el liderazgo del director o directora en aspectos relacionados con la gestión administrativa del centro, sino también los resultados de su gestión pedagógica, aspecto de suma importancia para Leithwood y Hopkins (2008) y Robinson (2009).

FIGURA 8. SUBCRITERIOS Y ÁREAS DEL CRITERIO LIDERAZGO.

SUBCRITERIOS	ÁREAS
1a. Desarrollo de los fines, objetivos y valores por parte del equipo directivo y de los otros responsables, y actuación de estos teniendo como modelo de referencia un planteamiento de mejora continua.	<ul style="list-style-type: none"> -Desarrollan fines y objetivos del centro. -Desarrollan principios éticos y valores de la mejora continua. -Mejoran y toman decisiones en asuntos de liderazgo. -Estimulan la creatividad, innovación y la colaboración.
1b. Implicación personal del equipo directivo y de los otros responsables para garantizar el desarrollo e implantación de los procesos de mejora continua en el centro.	<ul style="list-style-type: none"> -Adecúan la estructura del centro para apoyar la implantación de su planificación y estrategia. -Aseguran que se desarrolle e implante un sistema de gestión, evaluación y mejora de los procesos -Aseguran que se implanten procesos para revisar y mejorar las actividades mediante la creatividad, innovación y los resultados del aprendizaje.
1c. Implicación del equipo directivo y de los otros responsables con los beneficiarios del servicio educativo, con otros centros educativos e instituciones del entorno y con la Administración Educativa.	<ul style="list-style-type: none"> -Establecen relaciones de colaboración con agentes externos al centro. -Establecen y participan en actividades conjuntas de mejora. -Participan en actividades, conferencias y seminarios fomentando y apoyando, en particular, la mejora continua.
1d. Reconocimiento y valoración oportuna por parte del equipo directivo y de los otros responsables de los esfuerzos y los logros de las personas o instituciones interesadas en el centro educativo.	<ul style="list-style-type: none"> -Comunican personalmente los fines, objetivos, valores, planificación y estrategia y metas de la organización a las personas que la integran. -Son accesibles, escuchan activamente y responden a las personas que integran el centro. -Ayudan y apoyan a las personas a realizar sus planes, y a alcanzar sus objetivos y metas. -Reconocen, oportuna y adecuadamente, los esfuerzos de individuos y equipos, de todos los niveles de la organización.

Fuente: MEC (2001). Modelo Europeo de Excelencia. Madrid: MECD. Elaboración propia

De acuerdo con Batanaz (2009:2), *el director, como líder, tiene una alta capacidad para influir sobre la buena marcha de la escuela y sobre su mejora. Esa posibilidad de influencia debe ser ejercida como una práctica del liderazgo que comprometa a cuantas personas trabajan en la institución escolar.* De este modo, el rasgo determinante de la función directiva radica en su condición de líder, para la conducción al cambio creador. Si ese ejercicio de liderazgo que corresponde al director no se pone en práctica, la

institución perderá la mayor parte de las posibilidades que tiene de ser una organización coherente con los fines que tiene asignados y que ya apuntamos al inicio de estas páginas.

Desde un punto de vista optimista, el centro escolar se convertiría en una organización más o menos bien gestionada que ha renunciado a hacer frente a los desafíos que le afectan de forma creativa, quizás esto es lo que está sucediendo actualmente en muchos de los centros de la red pública española, ya que el convencimiento de creerse en el buen camino porque determinadas variables propias de la gestión se cumplen satisfactoriamente en ellos, tales como la asistencia puntual del profesorado a sus obligaciones docentes, el cumplimiento de los horarios o la puntual cumplimentación de las formalidades académicas y/o burocráticas, hacen perder de vista la necesidad de una evaluación, empezando por la del propio director o directora del centro escolar, lo que posibilitará un auténtico liderazgo y un compromiso continuo con la mejora, la eficacia y la calidad de la educación.

5. CONCLUSIONES

El avance hacia la descentralización está provocando que los centros escolares sean más autónomos en su toma de decisiones y se les otorgue mayor responsabilidad en los niveles de logro y resultados obtenidos por sus alumnos y alumnas en cuantas evaluaciones se lleven a cabo en el sistema educativo. Como consecuencia de esta tendencia, la función directiva escolar se define cada vez más por un exigente conjunto de labores que incluyen, entre otras, la administración financiera, la gestión de recursos humanos y el liderazgo para el aprendizaje.

El desarrollo legislativo ha dado sobradas muestras del cambio que se ha ido desarrollando en el perfil de la dirección escolar, exigiendo a éste nuevas funciones y aumento de tareas que han de desempeñar, lo que nos lleva a solicitar la necesidad de profesionalizar la dirección educativa, la creación de un cuerpo de directores y directoras, que, a la par que docentes, sean gestores formados específicamente que den respuesta de forma satisfactoria a las continuas demandas sociales y políticas. La propia Federación de Asociaciones de Directivos de Centros Educativos Públicos (FEDADI) lo solicita: diferenciación del puesto de director o directora en la plantilla de los centros educativos.

La formación es requisito imprescindible, amén de muchos otros que hemos analizado, en la selección de los candidatos a la dirección, opción que preferimos a la elección de los mismos. Para que la balanza se decante por la elección es necesario dignificar social, profesional y económicamente la dirección, al mismo tiempo que exigirse una evaluación y unos resultados que inviten a la responsabilidad que esta función lleva consigo. Por ello Martínez Lobato (2006) postula un modelo participativo y de liderazgo pedagógico compartido de la dirección escolar, en el que primen los proyectos y los resultados, la autonomía y la evaluación. Está suficientemente demostrada la conexión entre liderazgo pedagógico y escuelas eficaces. Consideramos que el factor crítico para la mejora continua y sostenida de los centros, enfocada a los aprendizajes de los alumnos y las alumnas es la capacidad interna de mejora, es decir, la competencia colectiva de la escuela para dar respuesta a los retos que se le presentan y gestionar los cambios de modo eficaz. La dirección es un elemento clave en el incremento de esa capacidad, para establecer una cultura de entendimiento y colaboración entre los distintos miembros de la comunidad educativa, para elaborar e impulsar proyectos que generen ilusión y compromiso y para crear equipos de trabajo coherentes con ellos.

La aplicación del Modelo de Excelencia EFQM en los centros escolares cuenta con una serie de ventajas, entre ellas:

- a) Permite incorporar un plan de calidad como base para el desarrollo de mejora continua de la institución educativa.
- b) Fomenta y facilita la reflexión de todos los miembros de la comunidad educativas: profesorado, familias, alumnado, etc.
- c) Posibilita descubrir las debilidades y fortalezas de la institución, fundamentadas en hechos y datos, lo que servirá de base para los planes de mejora.
- d) Genera actitudes positivas para responder a una sociedad cambiante y demandante de calidad.

La forma en que la Administración Educativa plantee su incorporación a la organización de los centros y los recursos con los que éstos cuenten para ponerla en marcha condicionará una mayor o menor aceptación de la comunidad educativa, especialmente del profesorado, que será el sector más afectado en el proceso evaluativo.

La evaluación de los centros educativos ha de entenderse como un medio para un fin y éste no es otro que lograr centros de calidad a través de un sistema que posibilite controlar, valorar y tomar decisiones en un proceso continuo y sistemático para alcanzar los objetivos que todo centro educativo determina respecto a los distintos ámbitos de gestión y sectores de la comunidad educativa que lo integran. En cuanto a la evaluación del criterio *liderazgo*, éste puede constituirse un referente fundamental, sin olvidar los demás criterios, un punto de partida para iniciar el cambio, a pesar de que, tal y como apunta Bolívar (1997), el tema del liderazgo pedagógico no parece preocupar excesivamente a la comunidad educativa española, frente a lo que ocurre en países de nuestro entorno.

Sin embargo, a nuestro entender, la función fundamental que debe asumir el director o directora de un centro educativo es el liderazgo en todos los ámbitos en los que mantiene atribuciones, lo que le posibilitará crear un compromiso continuo con la mejora, la eficacia y la calidad de la educación. La eficacia exige, la calidad presiona, los directores y directoras han de "estar a la altura" y la evaluación se convierte en el instrumento para comprobarlo.

REFERENCIAS BIBLIOGRÁFICAS

- AGAEVE (2009). *Plan estratégico de Evaluación General del Sistema Educativo Andaluz: 2009-2012*. Disponible en: http://www.juntadeandalucia.es/educacion/nav/contenido.jsp?pag=/Contenidos/OEE/evaluacion/PlanEstrategicodeEvaluacion2009_12&vismenu= [Consultado el 8 de marzo de 2010].
- Álvarez, M. (2001). El Liderazgo de los procesos de mejora. En I. Cantón. *et al.* (eds.). *La Implantación de la Calidad en los Centros Educativos*. Madrid: Editorial CCS.
- Antúnez, S. (2000). *La acción directiva en las instituciones escolares*. Barcelona: ICE/Horsori
- Barber, M. y Mourshed, M. (2007). *How the world's best performing school systems comes out on top*. USA: McKinsey & Company. Disponible en http://www.mckinsey.com/clientervice/Social_Sector/our_practices/Education/Knowledge_Highlights/~media/Images/Page_Images/Offices/SocialSector/PDF/Como_hicieron_los_sistemas_educativos.ashx . [Consultado el 11 de mayo de 2010].

- Batanaz Palomares, L. (2006). La función directiva, elemento clave para el desarrollo de las instituciones educativas. *Avances en supervisión educativa. Revista de la Asociación de Inspectores de Educación de España*, 2. Disponible en: http://www.adide.org/revista/images/stories/pdf_n4/ase04m03_batanaz.pdf?phpMyAdmin=BJkT-tBEqKxal12hom7ikt6vVu2. [Consultado el 7 de febrero de 2010].
- Cantón, I. (coord.) (2001). *La implantación de la calidad en los centros educativos: una perspectiva aplicada y reflexiva*. Madrid: Editorial CCS.
- Ley 17 julio de 1945 sobre educación primaria (Boletín Oficial del Estado del 18 de agosto de 1945).
- Decreto del 24 de octubre de 1947 sobre el Estatuto de Magisterio. (Boletín Oficial del Estado del 17 de enero de 1948).
- Decreto 193/1967 de 2 de febrero de 1967 por el que se aprueba el texto refundido de la Ley de Enseñanza Primaria (Boletín Oficial del Estado del 13 de febrero de 1967).
- Decreto 985/1967 de 20 de abril sobre reglamento del Cuerpo de Directores. (Boletín Oficial del Estado del 17 de mayo de 1967).
- Decreto 141/2001, de 2 de junio, sobre consolidación parcial del Complemento Específico de los Directores de los Centros Docentes Públicos de Andalucía (Boletín Oficial de la Junta de Andalucía, número 75, de 3 de julio).
- Decreto 59/2007, de 6 de marzo, por el que se regula el procedimiento para la selección y nombramiento de los directores y las directoras de los centros docentes públicos, a excepción de los universitarios (Boletín Oficial de la Junta de Andalucía, número 59, de 23 de marzo).
- Federación de Asociaciones de Directivos de Centros Educativos Públicos. Aportaciones desde FEDADI al pacto educativo. Disponible en: <http://www.fedadi.org/wp-content/uploads/2010/02/Aportaciones-desde-FEDADI-al-pacto-educativo.pdf>. [Consultado el 8 de marzo de 2010].
- LEA (2007). Ley 17/2007, de 10 de diciembre, de Educación de Andalucía (Boletín Oficial de la Junta de Andalucía número 252, de 26 de diciembre de 2007).
- Leithwood, H. y Hopkins, D. (2008). Seven strong claims about succesful school leadership. *School Leadership and Management*, 28 (1), pp. 27-42.
- Ley de 17 de julio de 1945, sobre Educación Primaria (Boletín Oficial del Estado del 18 de julio de 1945).
- Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa. (Boletín Oficial del Estado del 6 de agosto de 1970).
- Ley Orgánica 5/1980 de la Jefatura del estado, de 19 de junio, por la que se regula el Estatuto de Centros docentes, (Boletín Oficial del Estado del 27 de junio de 1980)
- LOCE (2002). Ley Orgánica 10/2002, de 23 de diciembre, de Calidad de la educación (Boletín Oficial del Estado, número 307, de 24 de diciembre de 2002).
- LOE (2006). Ley Orgánica 2/2006 de Educación (Boletín Oficial del Estado de 4 de mayo de 2006).
- LOGSE (1990).Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (Boletín Oficial del Estado de 4 de octubre de 1990).
- Martínez Lobato, E. (2006). La dirección escolar: ¿elegir o seleccionar? *Avances en supervisión educativa. Revista de la Asociación de Inspectores de Educación de España*, 4. Disponible en

http://adide.org/revista/index.php?option=com_content&task=view&id=93&Itemid=32. [Consultado el 25 de febrero de 2010].

Martínez Mediano, C. y Riopérez Losada, N. (2005). El modelo de excelencia en la EFQM y su aplicación para la mejora de la calidad de los centros educativos. *Educación XXI*, 8.

Ministerio de Educación y Ciencia. (2001). *Modelo Europeo de Excelencia*. Madrid: MECD.

Murillo Torrecilla, F., J. y Gómez Martín, J. C. (2006). Pasado, presente y futuro de la dirección escolar en España: entre la profesionalización y la democratización. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 4. Disponible en <http://www.rinace.net/arts/vol4num4e/art6.htm>. [Consultado el 12 de enero de 2010].

Orden de 20 de junio de 2007, por la que se regula el proceso de formación inicial de los directores y las directoras de los centros docentes públicos dependientes de la Consejería de Educación de la Junta de Andalucía, (Boletín Oficial de la Junta de Andalucía, número 230, de 25 de noviembre).

Orden Ministerial de 10 de febrero de 1967 sobre reglamentos de centros Estatales de enseñanzas Primaria. (Boletín Oficial del Estado del 20 de febrero de 1967).

Pont, B., et al. (2008). *Improving School Leadership. Policy and Practice*, 1. París: OECD.

Real Orden de 9 de marzo de 1918, por la que se establece el Reglamento de las Escuelas Graduadas.

Robinson, V. (2009). Fit for Purpose: An Educationally Relevant Account of Distributed Leadership. *Studies in Educational Leadership*, 7, pp. 219-240.

Tejero González, C. M^a y Fernández Díaz, M^a J. (2009). Medición de la satisfacción laboral en la dirección escolar. *RELIEVE*, 15, 2, pp. 1-16. Disponible en: http://www.uv.es/RELIEVE/v15n2/RELIEVEv15n2_1.htm. [Consultado el 28 de febrero de 2010].