

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 sub-dimensión 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 ítemes 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

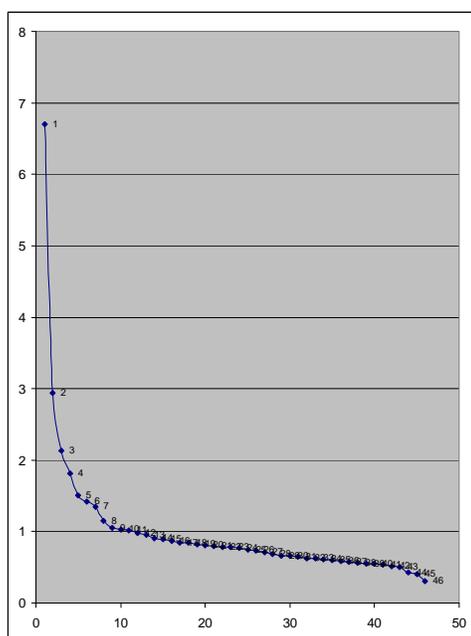


TABLA 6. ÍTEMS CON PESOS SOBRE 0.35 POR FACTOR

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

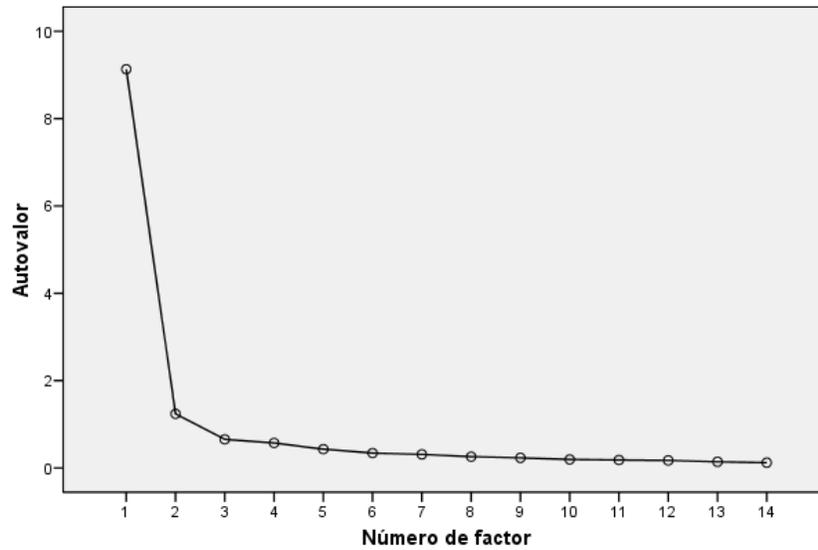
A5. Análisis factorial del Cuestionario del Profesor

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