

**PROPUESTA DEL DEPARTAMENTO DE EVALUACIÓN MEDICIÓN Y REGISTRO  
EDUCACIONAL (DEMRE)  
AL CONSEJO DE RECTORES DE LAS UNIVERSIDADES CHILENAS (CRUCH)  
DE MEJORAS A LA PSU  
Mayo 2017  
INTRODUCCIÓN**

El objetivo de este documento es proponer un conjunto de mejoras a la PSU, junto con los antecedentes y estudios en los que se fundamentan. Varias de estas propuestas requieren de ser analizadas y socializadas con expertos y *stakeholders* relevantes que deberían ser convocados para este fin, en conjunto por el CRUCH y el DEMRE. La anticipación con que se deben anunciar los cambios a un instrumento de tan altas consecuencias como es la PSU, hace aconsejable iniciar este proceso de discusión con prontitud.

La PSU se originó hace 15 años, en un momento histórico distinto del actual, en materia educativa. Chile había hecho una considerable inversión a lo largo de una década, en reformar el currículo escolar, y se optó por vincular la oportunidad de revisión de las pruebas de selección universitaria de entonces (PAA y PCE) con la necesidad de recabar información acerca de los resultados de la implementación del nuevo currículo.

Actualmente se discute una Reforma a la Educación Superior que incluye un nuevo sistema de admisión, más amplio e inclusivo, que involucra a los distintos tipos de instituciones del nivel terciario: Universidades, Centros de Formación Técnica (CFT) e Institutos Profesionales (IP). Este solo hecho, el que se define un sistema que acoga de mejor modo nuestra realidad actual, produce una pérdida de valor del sistema y los instrumentos vigentes. La actual batería PSU resulta estrecha e insuficiente para la amplitud y complejidad de las nuevas necesidades. Estas necesidades y la incapacidad de la PSU para satisfacerlas, ya se hicieron presente. En los dos últimos procesos de admisión se debió seleccionar a postulantes a cupos PAC, para lo cual no se utilizaron los puntajes PSU. La política de gratuidad de los estudios superiores que se está implementando, obliga a incorporar procesos objetivos y transparentes de admisión en un conjunto de instituciones que incluye a CFTs e IPs, para los cuales tampoco se ha utilizado la PSU, pues resulta inadecuada para la población que accede a ellos y para sus fines de selección. Incluso en las universidades que actualmente participan en el Sistema Único de Admisión (SUA), la población de postulantes “*rezagados*” (que egresaron de la enseñanza media en años anteriores al proceso de selección), corresponde a un tercio del total de postulantes y va en creciente aumento. Para todos ellos, una prueba fuertemente alineada al currículo vigente, introduce distorsiones que los perjudica y deteriora la calidad de la selección.

Adicionalmente, la falta de respuesta pronta a cuestionamientos serios y de fondo a la PSU que son públicos – como son los diagnósticos de las auditorías internacionales realizada por la consultora Pearson Education en 2013 y, anteriormente, por el Educational Testing Service (ETS) en 2005 – han

generado una reacción crítica, por parte de expertos, periodistas, actores sociales y, con particular fuerza, los movimientos estudiantiles, hacia la PSU y al sistema de admisión en su conjunto. En este escenario de cuestionamientos y confrontación, las máximas autoridades educacionales han anunciado el fin de la PSU.

Recientemente el DEMRE entregó al CRUCH una minuta detallada acerca del estado de avance de la implementación de las recomendaciones del Informe Pearson. De las 124 recomendaciones contenidas en este informe, DEMRE reconoce responsabilidad propia en 113, de las cuales 8 son de responsabilidad compartida con el CRUCH, quedando 11 recomendaciones fuera de su ámbito de acción. No todas estas recomendaciones tienen igual importancia ni están todas vigentes en la actualidad. De todas aquellas recomendaciones vigentes que incumben al DEMRE, ninguna ha sido desatendida y más del 91% se encuentra cumplida, entre ellas, todas las de mayor importancia de este conjunto de recomendaciones.

Entre las más importantes recomendaciones del Informe Pearson se encuentran aquellas que escapan al ámbito de acción del DEMRE: la necesidad de explicitar un Marco de Evaluación claro, que fije los propósitos de la PSU y considere la población a la que está dirigida, revisar la dependencia curricular estricta (centrar las pruebas en los contenidos y aptitudes o habilidades relevantes para el éxito universitario), llevar a cabo estudios para la reducción de temarios, resolver el problema de desajuste de la PSU de Matemáticas, establecer comparabilidad entre distintas aplicaciones y destinar esfuerzos a mejorar la limitada validez predictiva de todas las PSU. Otra recomendación importante, que como las anteriores, exceden las atribuciones del DEMRE, se refiere a reportar por separado las tres componentes de la prueba de Ciencias (Biología, Química y Física), recomendación basada en una contundente crítica técnica a la definición de esta prueba y al reporte de un puntaje único, de significado ambiguo.

Existe un conjunto pequeño pero importante de recomendaciones que involucran al DEMRE, que ameritan una discusión más amplia y una exploración de opciones alternativas que lleven al mismo fin. Por ejemplo, la solución al problema de comparación interanual.

Las propuestas de mejoras a la PSU que siguen, se insertan en el contexto expuesto.

1. **Síntesis y reducción de temas.** Las tres primeras recomendaciones del Informe Pearson y las recomendaciones 117, 118 y 119, se refieren a la necesidad de revisar las definiciones y especificaciones de una batería de pruebas que carece de un verdadero Marco de Evaluación. El informe recomienda revisar su alineación con los Contenidos Mínimos Obligatorios (CMO) del currículum de enseñanza media humanista-científica (EMHC) – que difiere del de la modalidad técnico-profesional (TP) en los dos últimos años de la EM – y la realización de estudios acerca de los efectos de limitar esta referencia a los dos primeros años de EM. La razón para ello es que los currículos prescritos para HC y TP son iguales hasta segundo medio. El informe también recomienda **redirigir el énfasis de la PSU al éxito universitario** e incluir una revisión internacional experta de las especificaciones, de manera que éstas y los nuevos marcos conceptuales sean sometidos a validación externa. La recomendación destacada en negro, tiene directa relación con una preocupación en la que el mencionado informe es insistente y cuya importancia es crucial: **la limitada validez predictiva de la PSU.** Este problema es claramente estructural, originado en la definición misma de la PSU, nacida con el doble propósito de seleccionar para el ingreso a la universidad y medir cobertura curricular. El segundo propósito necesariamente ha

limitado al primero.

El DEMRE ha realizado estudios y consultas, en este sentido, a un Comité Técnico Internacional constituido en Octubre 2016 y a expertos en cada una de las áreas de evaluación de las PSU (ver detalle en anexos). Las recomendaciones generales apuntan a reducir y sintetizar los extensos temas en categorías respecto de las cuales las pruebas puedan entregar información válida para la toma de decisiones. Las consideraciones de validez deben incluir la *oportunidad de aprender* que haya tenido la población a la que está dirigida la prueba y la relevancia de los aspectos evaluados para *predecir el éxito universitario*. Las recomendaciones particulares para cada prueba originan propuestas también específicas que se exponen más adelante. Por tratarse de una reducción y síntesis de los temas existentes, pensamos que su implementación no requiere de anuncios muy anticipados, si su discusión se realiza prontamente, incluye a los actores relevantes y resuelve efectos secundarios indeseados, como podría ser una desatención a los niveles más avanzados del currículum de EM implementado. Por ejemplo, una reducción de contenidos a los dos primeros años de EM podría ser acompañada de una mayor ponderación de las notas de 3° y 4° medio en el cálculo del NEM. Si se arriba a una solución como esta en el curso de este año, su implementación podría realizarse en el proceso 2020.

2. **Dos pruebas de Matemática.** La PSU de matemática tiene un grado de dificultad inadecuado

para la población que la rinde, resultando muy difícil para la mayor parte de la población y muy fácil para aquellos individuos con mayores conocimientos matemáticos, no logrando discriminar en la cola derecha de la curva de distribución. Ambos problemas se agravan al considerar la gran cantidad de contenidos a evaluar por esta prueba.

Durante el año 2016, DEMRE encargó un estudio externo (se anexa) para evaluar la confiabilidad y la validez predictiva de la prueba de Matemática en el escenario de eliminar preguntas referidas a los contenidos de 3° y 4° año de Enseñanza Media. Dicho estudio, utilizó datos del proceso de Admisión 2013 que habían sido utilizados por un estudio de validez predictiva realizado por el SVA (Grau, 2016). Asimismo, el equipo técnico del DEMRE replicó el análisis con datos de la admisión 2017 donde ya no se considera el descuento por respuestas erradas como en 2013. En ambos estudios se verifica que la confiabilidad de la prueba no se veía afectada al reducir los contenidos evaluados. Con 55 preguntas de primero y segundo medio, con similares características a las aplicadas en las pruebas del 2013 y 2017, se obtiene una confiabilidad tan alta como la que se obtiene con las 75 preguntas que se aplican en la actualidad.

El estudio externo encargado por el DEMRE concluye además que el uso del puntaje basado en contenidos básicos no comprometería la ya limitada validez predictiva de la prueba.

PSU Matemática 2017

FORMAS			
Confiabilidad (Alfa de Cronbach)	0.91	0.92	0.93
Total PSU Matemática (75 preguntas)	0.93	0.93	0.93
Eliminando ítems 3° y 4° medio (55 preguntas restantes)	0.91	0.92	0.91
			0.92

Por otra parte, al analizar las Curvas de Información del Test con modelos IRT, tanto de los procesos 2013 como 2017, se concluye que los ítems avanzados (de 3° y 4° Medio) aportan información en la cola derecha de la distribución, donde se realizan las decisiones relevantes para algunas carreras de base matemática, aunque sería necesario añadir ítems de mayor demanda cognitiva (recomendación 119 del Informe de Pearson).

Atendiendo a lo anterior, se recomienda desarrollar una **Prueba de Matemática (obligatoria)** que abarque únicamente contenidos hasta 2° año de Educación Media, para lo cual bastarían 60 ítems operativos. En segundo lugar, se propone **desarrollar una Prueba de Matemática avanzada (electiva)**, que debiese abarcar contenidos de 3° y 4° año de Enseñanza Media, y estaría compuesta por 40 ítems operativos. En ambos casos, la revisión de temarios propuesta en el punto 1 debe considerar los aspectos de validez recomendados. El DEMRE cuenta con ítems piloto que podrían servir a estos propósitos. De tomarse una decisión en este sentido y ser anunciada este año, el plazo para su implementación tampoco debería exceder el proceso 2020.

3. **Reducción de la PSU de Lengaje y Comunicación.** Por diseño, se determinó que la PSU de Lengaje y Comunicación mide los ejes curriculares de *Lectura y Escritura*. Sin embargo, la medición de la *Escritura* se realiza de manera indirecta a través de preguntas de selección múltiple. Esta modalidad no es una forma adecuada de realizar esta medición, a juicio de los expertos externos consultados. El DEMRE ha desarrollado pruebas de escritura de respuesta abierta y las ha aplicado a población recién ingresada a las universidades de Chile y de O'Higgins, con alta confiabilidad inter-jueces de correctores entrenados. Comparando esos resultados con aquellos obtenidos por los mismos postulantes en las secciones de *Conectores y Plan de Redacción* de la PSU de Lengaje, no se observan relaciones aceptables entre ambas mediciones, por lo que lo medido por la PSU de Lengaje en estas secciones no corresponde a la dimensión evaluada por la prueba de escritura. En consecuencia se recomienda **reducir la prueba de lenguaje** a contenidos referidos solo a comprensión lectora, es decir **eliminar las secciones de Conectores y Plan de redacción**. Para ello se propone que la PSU de Lengaje, midiendo solo *Comprensión lectora y Vocabulario*, sea de 60 preguntas operativas, al igual que la PSU propuesta en Matemática.

Por otra parte, si bien resulta relevante y factible evaluar directamente Escritura, mediante el desarrollo de un texto por parte del postulante, hay que considerar el alto costo monetario y en tiempo de corrección, además de aspectos de transparencia y objetividad requeridos en pruebas de altas consecuencias con fines de selección.

4. **Separar las pruebas de Ciencias.** Desde la introducción de la prueba única de Ciencias se sabe que su definición y cálculo de puntaje tienen graves problemas técnicos, como lo informó el estudio encargado por el DEMRE a Nancy Lacourly, la opinión de expertos del ETS y, más recientemente, el grupo de expertos de la consultora Pearson. Entre otros problemas, el constructo medido es distinto para cada disciplina y es inadecuado utilizar un puntaje común para reportar y tomar decisiones de selección. Esto produce que carreras asociadas a distintas disciplinas científicas seleccionen estudiantes sin saber qué conocimientos específicos tiene el estudiante en dicha disciplina. En la actual prueba de Ciencias no hay comparabilidad entre módulos electivos, tanto en lo relativo a su construcción como su dificultad. Al igual que los informes previos, la recomendación de Pearson es desarrollar pruebas separadas para Biología, Física y Química, considerando las diferentes poblaciones que rinden las PSU. El DEMRE comparte esta propuesta e insiste en la necesidad de su implementación. Como tránsito a esa estructura se podría considerar que las tres compartan un *módulo común de competencias científicas generales* y una parte específica de cada área científica (Biología, Física y Química). La prueba de Ciencias TP debería basarse en ese módulo común de competencias científicas generales, extendido. De hecho, con las actuales pruebas, es posible reportar un puntaje de la materia específica seleccionada. Por ejemplo, si un postulante rinde la prueba de Ciencias Física, es posible entregar información respecto de su desempeño en las 44 preguntas que corresponden a esta materia. La confiabilidad en cada materia científica en comparación al total de la prueba no disminuye sustancialmente, como lo muestra el análisis de la confiabilidad del subconjunto de preguntas específicas a cada disciplina de cada una de las opciones elegibles

5. **Pilotajes en aplicación oficial más programa de fomento de buenas prácticas y sanciones efectivas.** De acuerdo a la recomendación del informe Pearson de pilotear ítemes en aplicaciones oficiales, actualmente se incluyen 5 ítemes de este tipo en 3 de las 4 PSUs. Estos ítemes no se pueden utilizar posteriormente en pruebas oficiales, debido a que se vuelven públicos en reproducciones que realizan personas que rinden las pruebas, incluidos profesores de preuniversitarios. Estas instituciones, a su vez, difunden estas reproducciones como propaganda de la calidad de los servicios que ofrecen. En marzo del 2016, el DEMRE hizo una solicitud formal al SVA para eliminar estos 5 ítemes que alargan innecesariamente las pruebas, de por sí, demasiado extensas, la que fue rechazada. Desde el año 2016, el DEMRE realiza procesos de pilotaje de ítemes que incluyen un mecanismo de control de la diferencia fundamental entre rendiciones sin consecuencias y rendiciones de altas consecuencias (incluyendo en el pilotaje una PSU rendida y realizando *equating*). Sin embargo, el estándar internacional para pruebas de tan altas consecuencias como la PSU recomienda el pilotaje en aplicaciones oficiales. Esto podría realizarse si se pilotearan no 5 sino 150 a 200 ítemes por prueba en cada aplicación oficial, a través no de 4 formas distintas sino 10 a 20 formas distintas por prueba, cada una con 10 a 15 preguntas en pilotaje. Esto se basa en el gran número de estudiantes que rinden las pruebas y el menor tamaño de las muestras que se necesitan para pilotear válidamente un ítem. Al incluir un número tan alto de distintas formas e ítemes en cada

De este modo, sin hacer ningún cambio ni en la actual prueba de Ciencias ni en los procesos de selección, se podrían entregar puntajes separados para las tres pruebas de ciencia de la modalidad científico humanista en el proceso 2018, con el solo fin de entregar información relevante al postulante y a las instituciones que lo matriculen. Hacemos notar que el procedimiento mediante el cual se calcula actualmente el puntaje único de Ciencias (*linking con equiperceniti encadenado*), fue recomendado por el informe Pearson como medida transitoria. Este procedimiento supone condiciones que no se cumplen para dicha prueba y por lo tanto no resuelve el grave problema ocasionado por su diseño, y no se dispone de una mejor solución alternativa en este contexto. El diseño actual de la única prueba de Ciencias se ha conservado por 15 años y no se puede considerar superado, por la incorporación de la prueba de Ciencias para los egresados de la modalidad Técnico-Profesional. Resulta así de extrema urgencia definir una agenda que permita transitar a la separación de las pruebas de Ciencias.

Contrabilidad (Alfa de Cronbach)		Prueba Ciencias (80 Preguntas)		Sección Materia Específica (44 preguntas)	
PSU Ciencias-Biología	0.94	0.94	0.90	PSU Ciencias-Biología	0.90
PSU Ciencias-Física	0.95	0.95	0.92	PSU Ciencias-Física	0.92
PSU Ciencias-Química	0.95	0.95	0.91	PSU Ciencias-Química	0.91

actualmente por los postulantes de la modalidad científico-humanista, en la PSU de Ciencias 2017, a pesar de que el número de preguntas es casi la mitad (44 de un total de 80).

7. **Nombrar un Comité Técnico Internacional y un Comité Asesor Nacional.** El contexto de cambios y

efervescencia descrito en la introducción, además de las necesidades originadas en las propuestas previas, no solo exceden el ámbito de acción del DEMRE. La institucionalidad actual de la PSU resulta insuficiente –y se vería inútilmente tensionada– para darle un curso fluido y transparente a la implementación de mejoras de máxima calidad técnica, resguardando la fe pública en procesos de gran impacto institucional y en la vida de las personas.

La PSU es una prueba de importancia nacional que tiene consecuencias no solo en la selección de postulantes al SUA, en la actualidad y en el futuro inmediato. Las mejoras a la PSU que se discutan e implementen, influirán fuertemente en el tránsito a un sistema de admisión a la educación superior reformado, mucho más amplio y complejo. El momento actual aconseja discutir estas mejoras en un ámbito que no se limite a los representantes de instituciones involucradas en el actual Sistema Único de Admisión, como es el caso del Comité Consultivo aprobado por el CRUCH en agosto de 2016 en el marco del reglamento interno del SUA, que a la fecha aún no se ha implementado, y donde la participación de la *sociedad civil* se reduce a 3 estudiantes (uno universitario y dos secundarios). La discusión de las mejoras a la PSU aquí propuestas, necesita abrirse a audiencias más amplias, que incluyan a expertos e investigadores del área, y a actores educacionales, externos e independientes del actual sistema de admisión universitaria. El DEMRE propone para tal fin, la creación de un Comité Asesor Nacional, que incluya a expertos que provengan de los grupos de interés principales: Vicerrectores Académicos o Directores de Docencia de Universidades, Centros de Formación Técnica e Institutos Profesionales, Directores o profesores de establecimientos educacionales de distintas dependencias administrativas, incluyendo la modalidad TP, miembros del Consejo Nacional de Educación, investigadores, estudiantes y representantes del Ministerio de Educación.

En el último proceso de selección, el DEMRE convocó, con fondos propios, a un grupo de expertos internacionales y constituyó su Comité Técnico Internacional. Además contrató los servicios de la *International Association for the Evaluation of Educational Achievement (IEA)*, como contraparte técnica, para los análisis de las PSU rendidas, las decisiones de aceptación de ítemes para el cálculo del puntaje final y los procedimientos de equiparación entre formas de una misma prueba, proceso que se realiza en un tiempo muy acotado, y cuyas consecuencias son altísimas. Estimamos que este control de calidad de alto nivel técnico y de total independencia de los organismos involucrados en la prueba, es imprescindible y debe ser institucionalizado por el CRUCH, conocido y divulgado en las audiencias relacionadas y a nivel del público general.

A ambos comités técnicos, nacional e internacional, se les solicitará que evalúen las nuevas pruebas que se desarrollan en el DEMRE en el marco del proyecto Fondef, con el auspicio del Ministerio de Educación: *“Habilidades y Competencias para la Educación Superior. Desarrollo de instrumentos y propuestas que aporten a un nuevo Sistema de Admisión”*. Estas nuevas pruebas deberían permitir resolver el problema de adecuación antes mencionado, aminsonar las persistentes brechas que la PSU ha exhibido desde su origen y

mejorar la validez predictiva al centrarse en la medición de competencias relevantes para los estudios superiores con una referencia limitada a los aspectos medulares – y estables en el tiempo – del currículo escolar. Si bien el CRUCH no fue una de las instituciones asociadas en la postulación del proyecto, en el DEMRE entendemos que tiene un alto interés en su desarrollo y en los resultados que se obtengan, como quedó de manifiesto en la sesión plenaria de enero en Arica, donde se lo presentó.