



Fotografía tomada de <https://www.flickr.com/photos/nordlandcollege/> Licencia CC: Atribución - No comercial - Compartir igual

Los aprendizajes al ingreso en un Centro Universitario Regional

Resumen

Las políticas de regionalización y descentralización de la Universidad de la República (Udelar) han traído nuevos desafíos a los alcances de la enseñanza y el aprendizaje. El perfil de los estudiantes que ingresan a los Centros Universitarios Regionales es particular, son estudiantes pertenecientes a contextos culturales propios de la región en la que se encuentra inserto el centro educativo. En este artículo se estudia la generación de ingreso 2012 del Centro Universitario de la Región Este (CURE) en cuanto a sus características sociodemográficas y particularmente en relación con los aprendizajes en determinados dominios del conocimiento.

La información se obtuvo a partir de la aplicación de un cuestionario estructurado en dos partes: la primera orientada a obtener datos con el fin de elaborar un perfil social y educativo del estudiante y su entorno, y la segunda orientada a la evaluación de aprendizajes. Esta última se estructuró en tres dimensiones evaluativas: Comprensión y aplicación, Aplicación de conocimientos y Resolución de problemas.

Para establecer el punto de corte de la prueba se utilizó una aplicación del método de Angoff.

Menos de un tercio de los estudiantes de la generación de ingreso alcanzaron el nivel aceptable. Los estudiantes mayores de 35 años lograron un mejor desempeño con respecto a los más jóvenes. También aquellos cuya madre tiene un mejor nivel educativo obtuvieron mejores resultados.

Palabras claves:

evaluación diagnóstica, evaluación de aprendizajes, generación de ingreso.

Antecedentes

La regionalización y la descentralización son un eje de las reformas que está impulsando la Udelar. La regionalización vincula los conceptos de descentralización, territorio y redes. La descentralización se plasma en la creación de centros regionales que actúan en un territorio determinado, vinculándose con los agentes locales para construir y llevar a cabo propuestas comunes (Udelar, 2012).

En 2007 se crea el Centro Universitario de la Región Este (CURE), que tiene como zona de influencia los departamentos de Maldonado, Rocha y Treinta y Tres, a los que se suma en 2013 Lavalleja. Tuvo como carac-

terística diferencial que se conformó desde el inicio como un centro regional con tres sedes. Maldonado ha concentrado la mayor población estudiantil y opciones de formación. Dada su capacidad de absorber flujos migratorios de otros lugares, este departamento presenta una serie de tensiones y complejidades sociales (Veiga, Fernández, Lamschetein y Rivoir, 2012) que también se reflejan en la educación superior. Maldonado es el departamento con mayor proporción de población joven, aunque poco calificada. No obstante, esto no es un impedimento para acceder al mercado laboral debido a la fuerte demanda.

Pilar Rodríguez

Centro Universitario de la Región Este. Universidad de la República, Uruguay
prodriguez@cure.edu.uy

Martina Díaz

Centro Universitario de la Región Este. Universidad de la República, Uruguay
lmdiaz@gmail.com

Analía Correa

Centro Universitario de la Región Este. Universidad de la República, Uruguay
acorrea@oce.edu.uy

Citación recomendada:

RODRÍGUEZ, Pilar, DÍAZ, Martina, CORREA, Analía (2015). «Los aprendizajes al ingreso en un centro universitario regional». En: *InterCambios*, vol. 2, n.º 1, junio.

La Universidad en el interior tiene que desarrollar sus actividades y cumplir con sus funciones teniendo en cuenta el contexto cultural de la región.

El CURE está recibiendo estudiantes eminentemente locales. Más de la mitad de los estudiantes de ingreso residen en el departamento de Maldonado; los que residen en Rocha, Montevideo y Canelones representan alrededor del 30 % del total (UAE, 2012).

Es muy significativo, y el porcentaje ha ido en aumento desde 2011, el hecho de que muchos de los estudiantes del centro sean la primera generación de universitarios en el hogar. En 2011 eran el 62,6 %, en 2012, el 63,1 % y en 2013 representan el 67,3 % de los estudiantes que ingresan (UAE, 2013a).

El promedio de edad de los estudiantes de ingreso ha descendido, pero igualmente es más alto que en el resto de la Udelar. En 2011 se situaba en 26 años, y en 2012 y 2013 bajó a 24 años. Esto se asocia al hecho de que actualmente se captan estudiantes recién egresados de educación media, mientras que hace dos años casi la mitad de los estudiantes de ingreso habían dejado de estudiar por lo menos dos años.

Desde el establecimiento de la Unidad de Apoyo a la Enseñanza (UAE), los docentes y coordinadores de carrera han demandado tener un diagnóstico tanto sociodemográfico como de los aprendizajes obtenidos para orientar y adecuar el proceso de aprendizaje e implementar los nuevos planes de estudio.

Es común en Estados Unidos, Asia y Europa certificar los aprendizajes al egreso de la educación media mediante pruebas estandarizadas que sirven entre otros objetivos para pautar el acceso a la educación superior e incluso para certificar competencias para el mercado laboral. En nuestro país no contamos con este tipo de evaluaciones, por lo tanto es necesario realizarlas al inicio del siguiente ciclo educativo.

Introducción

Las dificultades en el ingreso son comunes a todos los estudiantes universitarios y han llevado a que se estudie el fenómeno en sus distintas facetas. Sin embargo, el contexto cultural y social de estos estudiantes de ingreso del CURE amerita estudios particularizados. Las ofertas de enseñanza de este centro regional son muy diversas y tienen como referencia académica a variados servicios. Se han puesto en marcha innovaciones educativas como los ciclos iniciales optativos.

Uno de los factores de riesgo asociados al fracaso en la universidad se identifica con los aprendizajes adquiridos. En este supuesto, las distintas facultades han diseñado y aplicado pruebas de carácter diagnóstico. El CURE comenzó a partir de 2011 a aplicar su prueba diagnóstica con el fin de identificar los aprendizajes de los estudiantes que ingresan.

A esos efectos, la UAE se propuso crear un instrumento fiable para medir los aprendizajes en determinados dominios del conocimiento; se valió de los últimos estudios relacionados con el desarrollo de pruebas de conocimientos para la creación de un instrumento propio. Para el equipo es el comienzo de una línea de trabajo que ha contado con amplio desarrollo en los últimos años en todo el mundo.

Este trabajo sistematiza la descripción de la prueba e incluye las variables relevadas, las especificaciones de los tests evaluatorios, el establecimiento de los estándares de desempeño, así como el análisis de los resultados obtenidos.

Las pruebas diagnósticas

Las pruebas diagnósticas o criterios consisten en la colección de información sobre el aprendizaje de los estudiantes con alguno de los siguientes propósitos primarios: identificar las fortalezas y debilida-

des de los estudiantes; ayudar a los educadores en la planificación de los cursos; auxiliar a los estudiantes a guiar su propio aprendizaje, rever su desempeño y adquirir capacidad de autoevaluarse, y fomentar la autonomía y responsabilidad de aprendizaje (Andrade y Cizek, 2010).

Estas pruebas pueden tener formatos muy diversos. La característica distintiva de una prueba diagnóstica, más allá del formato, es que su diseño y objetivo principal apuntan a la adquisición de información con el fin de adaptar la enseñanza y el aprendizaje a las necesidades de los estudiantes (Black y Wiliam, 1998).

Metodología

Descripción de la prueba

La evaluación diagnóstica (ED) 2012 se estructuró en dos partes: la primera orientada a obtener datos sociodemográficos de los estudiantes con el fin de elaborar un perfil, y la segunda orientada a la evaluación de aprendizajes.

Entre las variables relevadas en la primera parte estuvieron la procedencia de los estudiantes en cuanto a departamento de origen y lugar donde cursaron educación media, edad, situación familiar, manejo de herramientas informáticas, medio de transporte, nivel educativo propio y de sus padres, último año en el que cursaron estudios, repetición, situación laboral, etc.

La evaluación de aprendizajes se estructuró en tres partes: Comprensión y aplicación, Aplicación de conocimientos y Resolución de problemas.

Especificaciones del test y los ítems: formato, extensión

Se conformó la ED con un total de 26 ítems de evaluación con preguntas cerradas: múltiple opción y verdadero o falso. Para todos los ítems se establecieron estándares de contenido.

Dentro de la ED se agruparon tres grandes consignas de evaluación:

- **Comprensión y aplicación:** dentro de esta dimensión se evaluó comprensión lectora; se les presentaron a los estudiantes tres textos y se les formularon seis preguntas orientadas a valorar distintas habilidades adquiridas.
- **Aplicación de conocimientos:** esta segunda dimensión apuntó a evaluar sintaxis, léxico, ortografía, conocimientos generales sobre ciencias sociales.
- **Resolución de problemas:** se le solicitó al estudiantado resolver ejercicios que contenían cálculos numéricos, logaritmos, límites, derivadas, resolución de un sistema de ecuaciones lineales, cálculo de sumatoria, concepto de raíz, resolución de una ecuación de segundo grado, interpretación de gráficas, cálculo de probabilidades simples y compuestas.

Prácticamente todos los ítems creados fueron testeados previamente en estudiantes con el mismo nivel educativo y similar perfil sociodemográfico.

Establecimiento de estándares de desempeño: punto de corte

Los temas fundamentales de desarrollo de las pruebas criterios son dos: el análisis y especificación del dominio educativo como universo de medida a partir del cual se originan las pruebas, y el desarrollo de estándares o sistema de interpretación de puntuaciones dirigido a establecer un juicio de valor acerca de la calidad del aprendizaje (Jornet Meliá y González Such, 2009).

Un estándar es un punto en la escala de puntuaciones de una prueba que sirve para clasificar a quienes fueron examinados en categorías que reflejan diferentes niveles de ejecución en relación con los objetivos o competencias medidos por la prueba. Existe una gran diversidad de

métodos para establecer estándares y puntos de corte; para elegir el más adecuado se deben considerar algunos factores tales como la importancia de las decisiones que se tomarán; el tiempo, las fuentes y recursos disponibles; la capacidad de los jueces especialistas con los que contamos (algunos métodos requieren mayor conocimiento del dominio del contenido, mientras que otros requieren mayor conocimiento de los estudiantes que serán examinados), y la pertinencia del método para el tipo de prueba que estamos elaborando (Leyva Barajas, 2011).

Han sido múltiples los intentos realizados hasta la fecha por exponer y valorar los métodos para abordar este problema. Jornet y Backhoff (2008) los clasifican de la siguiente manera:

- métodos de juicio, basados en el juicio que realizan expertos acerca de los ítems, los sujetos o las tareas;
- métodos empíricos, basados prioritariamente en el comportamiento empírico de la prueba;
- métodos mixtos, que conjugan el juicio de expertos con otras informaciones de carácter empírico.

En este trabajo se utilizó un método de juicio basado en la idea de sujeto límite, definido como aquel que obtiene una puntuación media (o mediana) entre los dos grupos que pueden considerarse como criterio: aptos/no aptos. Para identificar esa puntuación límite (o punto de corte —*passing score*—) se parte del análisis lógico de los ítems que componen la prueba.

Se siguieron en este trabajo las etapas para el desarrollo de estándares de desempeño que describen Hambleton (1998) y Hambleton, Jaeger, Plake y Mills (2001). Se utilizó una aplicación del método de Angoff, método de juicio comúnmente usado en el contexto de licenciaturas y diplomaturas. Este se ha convertido en el método más

cuidadosamente estudiado y ha sido incluido en docenas de estudios comparativos con otros métodos. Es apropiado para ítems de formato de múltiple opción.

Un procedimiento sistemático para decidir cuál será el mínimo puntaje para aprobar consiste en tener en mente a una persona hipotética de aptitudes mínimamente aceptables y pasar por todos los ítems del test decidiendo en cada uno si tal persona podría contestar correctamente (1) o no (0). La suma de todos los puntajes de los ítems es el puntaje mínimo de aprobación o punto de corte de la prueba (Cizek y Bunch, 2007).

Para el establecimiento del punto de corte en esta prueba se escogieron dos grupos de jueces: uno para evaluar los ítems de Comprensión y Aplicación de conocimientos y otro para evaluar los ítems de Resolución de problemas. En su selección se tuvo en cuenta su formación y su experiencia docente en educación superior o educación media.

Resultados

Perfil de la generación de ingreso

Asistieron a realizar la ED 363 estudiantes de un total de 656 inscritos en el año 2012.¹ Si bien se contó con la declaración de obligatoriedad de la Comisión Directiva del CURE, solo el 55,3 % de los inscritos acudieron a la convocatoria.

El 42,8 % de los estudiantes que respondieron son de sexo masculino y el 57,2 % de sexo femenino. Las edades de los inscritos oscilan entre los 17 y los 66 años, con un promedio de 23 años. En el gráfico 1 se presentan los porcentajes de estudiantes según la franja etaria.

De variadas procedencias, prevalecen los estudiantes que residen en Maldonado (53,2 %), Rocha (11,8 %), Montevideo (9,2 %) y Canelones (8,7 %) (gráfico 2).

¹Agradecemos a los docentes que colaboraron en la aplicación de la ED: Clara Núñez, Franco Teixeira, Gastón de León, Guillermo Goyenola, Juan Pablo Pacheco, Laura Brum, Natalia Verrastro, Paula Laporta y Rossana Cantieri.

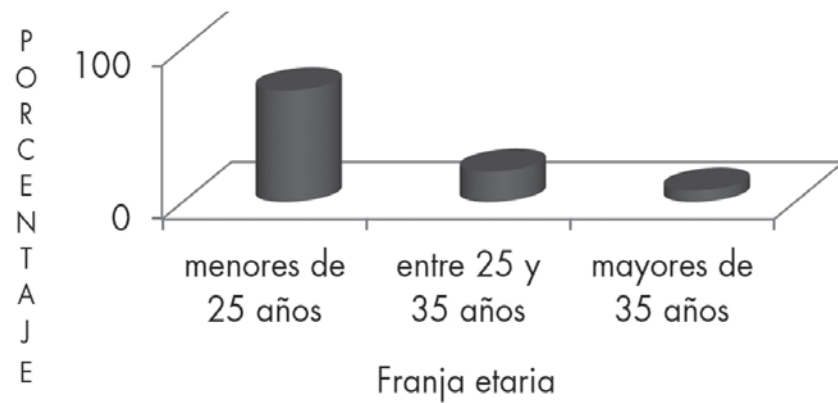


Gráfico 1. Porcentajes de estudiantes según franja etaria.

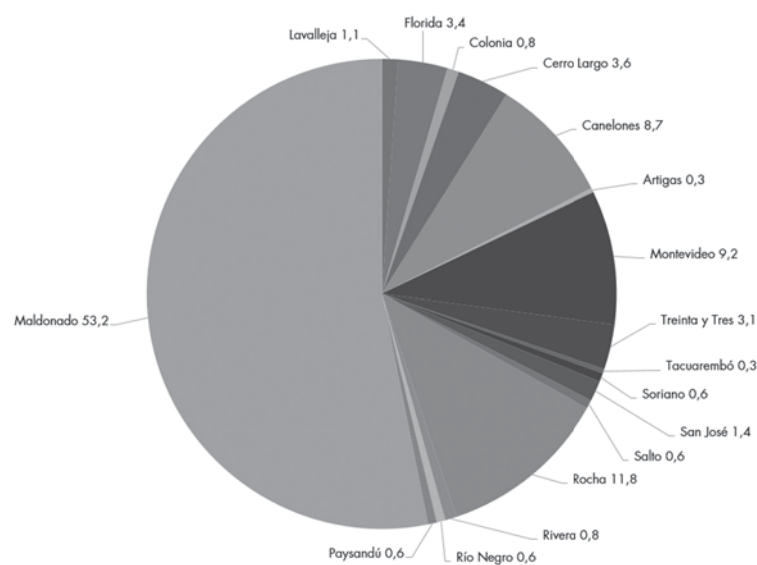


Gráfico 2. Departamentos de procedencia.

Tabla 1. Carrera o programa en el que se inscribieron los estudiantes.

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
CIO orientación Social	67	18,3	18,9	18,9
CIO orientación Ciencia y Tecnología	2	0,5	0,6	19,4
Licenciatura en Educación Física	69	18,9	19,4	38,9
Licenciatura en Diseño del Paisaje	44	12	12,4	51,3
Licenciatura en Gestión Ambiental	18	4,9	5,1	56,3
Licenciatura en Lenguajes y Medios Audiovisuales	4	1,1	1,1	57,5
Licenciatura en Turismo	50	13,7	14,1	71,5
Tecnólogo en Administración y Contabilidad	76	20,8	21,4	93
Tecnólogo en Informática	15	4,1	4,2	97,2
Tecnólogo en Telecomunicaciones	8	2,2	2,3	99,4
Tecnicatura en Vóleybol	2	0,5	0,6	100
Total	355	97	100	
No contesta	11	3		
Total	366	100		

En la tabla 1 se muestra a qué carreras pertenecen los estudiantes que participaron.

La mayoría de los estudiantes provienen de la enseñanza secundaria pública (64,6 %) y de la enseñanza secundaria privada (16,4 %); el restante 19,1 % proviene de UTU (7,5 %), universidad (7,2 %), magisterio, IPA, CERP (3,2 %) y extranjero (1,2 %).

Dada la consulta sobre el nivel educativo máximo alcanzado, es de hacer notar que el 69,3 % de los estudiantes que ingresan cuentan con el bachillerato diversificado de secundaria en su diversidad de orientaciones, el 14,8 % cursaron otras carreras universitarias, el 8,8 % realizaron estudios en educación terciaria no universitaria (IPA, CERP, magisterio) y un 7,1 % realizaron cursos en UTU (educación media superior de UTU 5,7 %, y cursos técnicos de UTU 1,4 %).

Más de la mitad de los estudiantes cursaron su último nivel de estudios en 2011 (50,6 %), un 16,6 % lo hizo en 2010 y el restante 32,9 % en los años anteriores.

En cuanto a los niveles educativos previos a la universidad, el 98 % no repitió ningún año en primaria, mientras que un 1,7 % repitió un año y un 0,3 % repitió dos años. El 95,8 % no repitió años en ciclo básico, el 3,1 % repitió un año y el 1,1 % repitió dos años. Por último, el 22,6 % no repitió ningún año en bachillerato y el 77,4 % sí repitió un año.

Análisis descriptivos de la ED

Frecuencias por ítem

Fueron calculadas las frecuencias de respuestas correctas para cada ítem por separado (gráfico 3) así como también para los tres grupos de ítems (Comprensión y aplicación, Aplicación de conocimientos y Resolución de problemas). Por lo tanto, fueron calculados los porcentajes de individuos que obtuvieron un nivel aceptable para la prueba total y para cada grupo de ítems (gráfico 4).

Se establecieron tres niveles de desempeño identificados como «no alcanza el nivel aceptable», «nivel aceptable» y «nivel destacado». Los puntos de corte establecidos para la prueba se describen en el apartado correspondiente.

Solo el 30 % de los estudiantes que realizaron la ED obtuvieron un nivel aceptable. Un 40 % obtuvo nivel aceptable para el grupo de ítems de Comprensión, un 53 % obtuvo nivel aceptable para el grupo de ítems de Aplicación de conocimientos, y solamente un 22 % de los estudiantes obtuvieron nivel aceptable para los ítems de Resolución de problemas (gráfico 4).

También se analizó el porcentaje de individuos de cada franja etaria (menores de 25, de entre 25 y 35, y mayores de 35 años) que obtuvieron nivel aceptable. Se observó que los mayores de 35 fueron el grupo que tuvo un porcentaje más alto de estudiantes con niveles aceptables (gráfico 5).

Asimismo, se comparó el porcentaje de estudiantes con nivel aceptable en la prueba según el nivel educativo máximo alcanzado por ellos previamente (bachillerato diversificado o UTU, estudios terciarios). Se observó que de los estudiantes cuyo nivel máximo alcanzado era bachillerato o UTU un 28 % obtuvo nivel aceptable, y que de los que habían realizado previamente algún curso universitario un 37 % lo alcanzó.

En cuanto al nivel educativo familiar, se analizó el desempeño según el nivel educativo máximo alcanzado por la madre. Para ello se tomaron dos grupos de referencia: 1) estudiantes cuya madre hubiera aprobado uno o dos años de secundaria o UTU (49 %), y 2) estudiantes cuya madre hubiera completado los estudios secundarios, o cursado estudios terciarios universitarios o no universitarios (51 %). Aquí se observó que de los estudiantes pertenecientes al grupo 1 un 28 % obtuvo nivel aceptable, y de los pertenecientes al grupo 2, un 33 % alcanzó ese nivel.

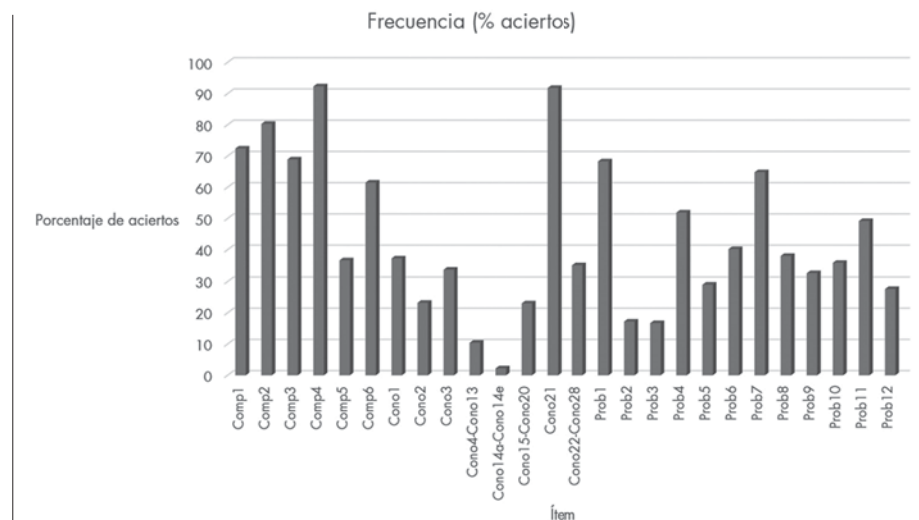


Gráfico 3. Porcentaje de aciertos para cada uno de los ítems descritos.

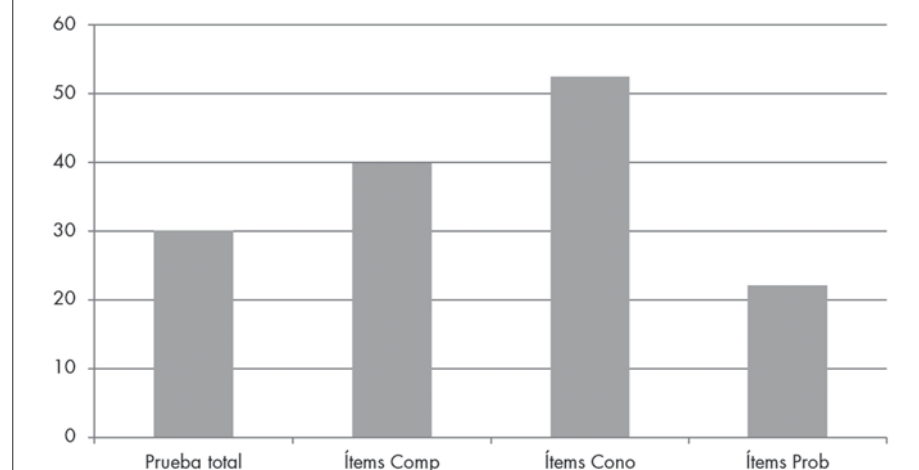


Gráfico 4. Porcentaje de individuos que obtuvieron nivel aceptable para la prueba total y para cada grupo de ítems.

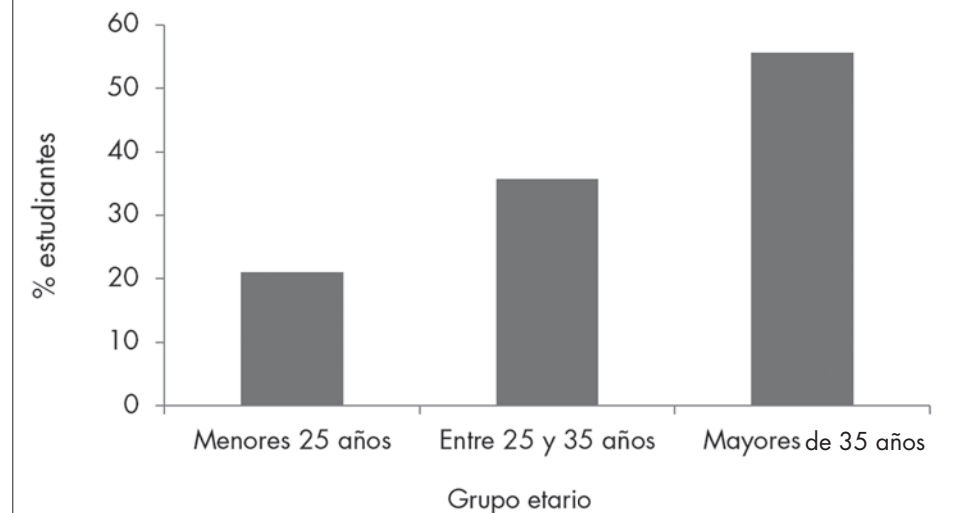


Gráfico 5. Porcentaje de estudiantes que obtuvieron nivel aceptable en la prueba según grupo de edad.

En este aspecto también se analizó el desempeño de los estudiantes que tenían algún integrante del hogar con estudios universitarios realizados (38 %) y de los que no lo tenían (62 %). En este caso, se observó que cuando existía al menos un integrante con nivel universitario en el hogar, un 26 % de los estudiantes obtuvo nivel aceptable, mientras que cuando no existía, el porcentaje de estudiantes con nivel aceptable fue de un 32 %.

Análisis del ítem

Se calcularon los *índices de dificultad* (ID) corregidos para cada ítem. Se entiende por índice de dificultad de un ítem la proporción de sujetos que lo aciertan de aquellos que han intentado resolverlo. Así, a medida que el índice aumenta (entre 0 y 1), significa que ha habido más aciertos de los individuos que intentaron resolverlo (el ítem es más fácil). En este caso se presentan los índices de dificultad corregidos, dado que, al ser los ítems de múltiple opción, es necesario corregir los efectos del mero azar, para lo cual se utilizó la siguiente fórmula:

$$ID = \frac{A - E}{N(K-1)}$$

donde:

- A: es el número de sujetos que aciertan el ítem,
- E: es el número de sujetos que fallan el ítem,
- K: es el número de alternativas del ítem,
- N: es el número de sujetos que intentan resolver el ítem.

En negrita se muestran los índices mayores de 0,5, en negro los índices de entre 0 y 0,5 y en gris los índices menores de 0.

En la figura 1 se observan los índices de dificultad corregidos para cada una de las preguntas de todos los ítems, separados por grupos [Comprensión (a), Aplicación de conocimientos (b) y Resolución de problemas (c)]. Allí se observa que en los ítems de Comprensión el 67 %

Ítem	ID corregido	Ítem	ID corregido	Ítem	ID corregido
COMP1	0,65	CONO1	0,11	PROB1	0,78
COMP2	0,72	CONO2	-0,11	PROB2	-0,06
COMP3	0,39	CONO3	0,22	PROB3	-0,05
COMP4	0,90	CONO4	0,41	PROB4	0,46
COMP5	0,08	CONO5	0,78	PROB5	0,21
COMP6	0,51	CONO6	0,84	PROB6	0,26
		CONO7	0,67	PROB7	0,64
		CONO8	0,01	PROB8	0,26
		CONO9	0,80	PROB9	0,15
		CONO10	0,74	PROB10	0,20
		CONO11	0,81	PROB11	0,37
		CONO12	0,52	PROB12	0,06
		CONO13	0,50		
		CONO14A	0,76		
		CONO14B	-0,17		
		CONO14C	-0,05		
		CONO14D	0,29		
		CONO14E	0,34		
		CONO15	0,43		
		CONO16	0,56		
		CONO17	0,72		
		CONO18	0,46		
		CONO19	0,80		
		CONO20	0,56		
		CONO21	0,94		
		CONO22	0,89		
		CONO23	0,73		
		CONO24	0,92		
		CONO25	0,91		
		CONO26	0,55		
		CONO27	0,41		
		CONO28	0,96		

Figura 1. Tablas de índices de dificultad corregidos para cada ítem de los tres grupos de la prueba.

presentó índices de dificultad mayores de 0,5, en los ítems de Resolución de problemas solamente un 17 % fue mayor de 0,5 y en Aplicación de conocimientos un 63 % fue mayor de ese valor.

Cabe aclarar que, dado que fue utilizada la fórmula para el cálculo del índice de dificultad corregido, los ítems que no llegan a tener siquiera la cantidad esperada de aciertos al azar tienen un índice de dificultad corregido negativo. Estos ítems deben ser analizados para establecer por qué no se llega a la cantidad de aciertos al azar esperada, si es un problema en la formulación de los distractores o del propio ítem. También puede ocurrir

que un distractor coincida con un conocimiento apropiado por los sujetos en forma incorrecta.

Asimismo, fueron calculados los *índices de discriminación*, en este caso las correlaciones biserials puntuales (bivariadas) de los ítems por separado y de los grupos. Las correlaciones miden cómo están relacionadas las variables o los órdenes de los rangos. En este caso fueron calculadas las correlaciones de Pearson. Las correlaciones de todos los ítems (menos siete) fueron significativas al nivel de 0,05 o 0,01 % contra el resultado total de la prueba. A su vez, cuando fueron correlacionados los aciertos por grupo de ítems (Compren-

sión, Aplicación de conocimientos y Resolución de problemas) con el resultado total del test, se observó que fueron significativas al nivel de 0,01 %.

Fiabilidad

El *coeficiente alfa de Cronbach* es un coeficiente de consistencia interna basado en el promedio de las correlaciones entre los ítems. Es una cota inferior de la fiabilidad. Entre las ventajas de esta medida se encuentra la posibilidad de evaluar cuánto mejoraría (o empeoraría) la fiabilidad de la prueba si se excluyera un determinado ítem. Cuanto más cercano a 1 es el coeficiente, mejor es la correlación entre los ítems analizados.

Se analizaron todos los ítems por separado y se obtuvo un coeficiente alfa de Cronbach de 0,652. Luego, al realizar el análisis de la fiabilidad del resultado total del test y el total de los aciertos de cada uno de los tres grupos globales de ítems (Comprensión, Aplicación de conocimientos y Resolución de problemas), se obtuvo un coeficiente de 0,731. Esto muestra que al analizar los ítems como grupos la correlación con los resultados del test mejora. Además se analizaron las correlaciones de los resultados del test con las combinaciones de dos de los grupos, y se observó que la correlación mayor (y el único caso que mejoró con respecto al análisis del resultado del test con los tres grupos de ítems) se obtuvo para el caso del test y los ítems de Aplicación de conocimientos y Resolución de problemas (0,804).

Punto de corte

El *punto de corte* obtenido para la prueba fue 39 (en un total de 50); para los tres grupos de preguntas, Comprensión, Aplicación de conocimientos y Resolución de problemas, fue 5 (en 6), 25 (en 32) y 9 (en 12), respectivamente, como se muestra en la tabla 2.

Se establecieron tres niveles de desempeño identificados como «no

	Punto de corte	No alcanza nivel aceptable	Nivel aceptable	Nivel destacado
Prueba total (50)	39	puntaje < 33	33 ≤ puntaje ≤ 47	puntaje > 47
Ítems Comp (6)	5	puntaje < 5	puntaje = 5	puntaje > 5
Ítems Cono (32)	25	puntaje < 21	21 ≤ puntaje ≤ 31	puntaje > 31
Ítems Prob (12)	9	puntaje < 10	7 ≤ puntaje ≤ 10	puntaje > 10

Tabla 2. Punto de corte y límites de los diferentes niveles de desempeño para la prueba total y para cada grupo de ítems.

alcanza el nivel aceptable», «nivel aceptable» y «nivel destacado», para los valores de la prueba encontrados en los rangos indicados en la tabla 2. Los valores límites de estos rangos fueron elegidos tomando en cuenta los valores obtenidos por los jueces que dieron la menor y mayor puntuación a los grupos de ítems (límites inferior y superior del nivel aceptable, respectivamente).

Conclusiones

Diferenciaremos entre las conclusiones con relación al instrumento y las conclusiones sobre los resultados obtenidos a partir de la aplicación del instrumento.

Con relación al instrumento

Si bien para la construcción del instrumento de evaluación se siguieron las recomendaciones de la literatura especializada de los últimos años, la herramienta lograda dista mucho de ser ideal. Consideramos importante continuar trabajando en el perfeccionamiento del instrumento y en la ampliación y profundización de estudios relacionados con este tema.

Se ha trabajado durante 2013 y 2014 en la creación de un instrumento de evaluación basado en competencias, en la ampliación del banco de ítems, en el establecimiento de estándares de contenido acordados

con docentes de educación media y en el establecimiento de niveles de desempeño. Además, se varió el procedimiento de establecimiento del punto de corte; primero se utilizó el método sí/no introducido por Impara y Plake (1997), para finalmente pasar a un método mixto, de forma de relativizar el peso del juicio de los expertos.

Del análisis de los ítems de la prueba, puede concluirse que las dimensiones Aplicación de conocimientos y Comprensión y aplicación presentaron un índice de dificultad mayor de 0,5 en más del 60 % de sus ítems, lo que significa que más del 50 % de los estudiantes respondieron correctamente en cada uno de ellos. En cuanto a los ítems correspondientes a Resolución de problemas, en los que solo el 22 % superó tal valor, se revisó si su formulación y contenido fueron adecuados y se decidió sobre su remoción o mantenimiento en el banco de ítems. También se analizaron y compararon los resultados en esta dimensión en las siguientes ediciones, en las que se observó la persistencia de la dificultad para acertar ítems relacionados con competencias lógico-matemáticas (UAE, 2013b). No obstante esto, los ítems correlacionan adecuadamente, y por dimensión la prueba tiene una buena consistencia interna, como se confirmó con el cálculo del coeficiente alfa de Cronbach. Por lo tanto, quedaría

para próximos estudios utilizar otros métodos para el cálculo del índice de dificultad que sean independientes de la muestra, como los que utiliza la teoría de respuesta al ítem.

Con relación a los resultados obtenidos

La ED fue respondida por el 55 % del total de inscritos. De este dato inmediatamente se desprende la necesidad de la obligatoriedad de la evaluación diagnóstica con el fin de obtener un muestreo en el que no existan dudas acerca de la representatividad de la población evaluada.

Acerca del perfil sociodemográfico encontrado para la generación de ingreso, cabe resaltar algunos datos. La población estudiantil que elige estudiar en el CURE es eminentemente local. Si bien lentamente se viene acrecentando la proporción de estudiantes de otros departamentos (UAE, 2012, 2013a), aún no ha logrado volverse un centro educativo de referencia para determinadas áreas del conocimiento. Con seguridad se ha cumplido el objetivo de acercar la educación universitaria a la población del interior, lo que también se aprecia en el promedio de edad de los estudiantes de ingreso, un poco más alto que en el resto de la Udelar.

Es preocupante la proporción de estudiantes que dicen haber repetido al menos un año en la educación media; particularmente la repetición en el ciclo inmediato anterior es un antecedente de fracaso que puede influir negativamente en la motivación para continuar los estudios. A su vez, es alentador que el 62,3 % de los estudiantes sean la primera generación en el núcleo familiar que llega a la universidad. Sin embargo, este factor puede influir en la finalización del ciclo educativo. Hay evidencias de las dificultades que presentan los estudiantes cuyos padres no han accedido a la universidad para completar los estudios universitarios (Choy, 2001).

Visto que menos de un tercio de los estudiantes que realizaron la

evaluación diagnóstica obtuvieron un nivel aceptable (30 %), puede concluirse que el 70 % de los estudiantes que ingresaron al CURE en 2012 evidenciaron carencias significativas en las áreas básicas de formación. Es de destacar que de los tres grupos de ítems (Comprensión, Aplicación de conocimientos y Resolución de problemas), fue en los de Resolución de problemas que se obtuvo el porcentaje menor de estudiantes con nivel aceptable, solamente un 22 %, mientras que en los ítems de Comprensión el 40 % alcanzó un nivel aceptable y en los ítems de Aplicación de conocimientos un 53 %. La mayor proporción de desempeños aceptables se dio en la dimensión Aplicación de conocimientos (alrededor del 53 %). El tipo de aprendizajes mejor adquiridos se relaciona con lo memorístico, no con el análisis y la interpretación. Las mayores dificultades se presentan en las dimensiones de Comprensión de textos y Resolución de problemas. Ambas dimensiones son básicas para asegurar un buen desempeño en el nivel universitario, pero lo relacionado con la comprensión de textos es esencial.

Claramente se aprecia la necesidad de reforzar los conocimientos previos de los estudiantes en las áreas de ciencias exactas, específicamente en matemáticas. En las sucesivas ediciones de aplicación de la prueba se acentúa la diferencia entre la proporción de aceptables en la dimensión Resolución de problemas y las demás dimensiones. Este dato es congruente con los resultados que arrojan otras pruebas diagnósticas aplicadas a estudiantes que ingresan a la universidad (UEFI, 2012; Mussio y Martinotti, 2013) y también con los resultados de PISA 2012, en la cual más de la mitad de los alumnos de 15 años de Uruguay no lograron los niveles básicos de competencias en matemáticas, y solo poco más de un 1 % se encuentran en los niveles más altos (ANEP-Programa PISA Uruguay, 2013). Según Montt (2013), esto debe ser materia de preocupa-

ción, porque una gran proporción de estudiantes no saben aprovechar las herramientas matemáticas para solucionar problemas muy simples, y muy pocos pueden usar las matemáticas para problemas complejos.

Es mayor el porcentaje de estudiantes con mejor nivel educativo propio y de su madre que obtienen un nivel aceptable. El nivel educativo de los restantes integrantes del núcleo familiar parece no influir positivamente en el desempeño en la prueba como sí lo hace el nivel educativo de la madre. Esto es coherente con la literatura internacional que relaciona el nivel educativo de la madre como el de mayor incidencia en el futuro desempeño de los hijos (Marchesi, 2000; Porcel, Dapozo y López, 2010).

Al comparar el rendimiento de los estudiantes divididos en grupos por edad, se observó que los que tuvieron mejor calificación fueron los mayores de 35 años. El porcentaje de estudiantes que obtienen un nivel de desempeño aceptable en la prueba aumenta con la edad. Los menores de 25 años fueron los que alcanzaron menor porcentaje de aprobación. Entre las personas de más edad seguramente se encuentran aquellas que tienen más formación; aunque la persona no haya continuado estudiando, el ingreso al mundo laboral parece aportar ciertas competencias que los más jóvenes no han alcanzado. Sin embargo, hay que apuntar los resultados de la edición siguiente, en la que la edad no parece hacer un aporte sustancial; específicamente en la dimensión Resolución de problemas, los mayores de 35 años obtienen la proporción menor de aceptables (UAE, 2013b).

Dado que el ingreso es el momento en el que se interceptan dos culturas institucionales y académicas de características extremadamente diferentes, evaluaciones de este tipo, consecutivas en los diferentes años lectivos, que puedan ser comparables, son sumamente necesarias para la identificación de carencias y la im-

plementación de planes de nivelación a fin de que los estudiantes puedan desarrollarse en iguales condiciones sea cual sea su situación inicial.

Para finalizar, cabe mencionar la necesidad de establecer políticas específicas para esta población estudiantil que ingresa y de reforzar las

que ya ha implementado la Udelar, como Progesa, para que tengan mayor alcance en los centros universitarios del interior.

Referencias bibliográficas

- ANDRADE, H. L. y CIZEK, G. J. (2010). *Handbook of formative assessment*. Nueva York: Routledge.
- ANEP-PROGRAMA PISA URUGUAY (2013). *Uruguay en PISA 2012. Informe preliminar*. ANEP.
- BLACK, P. y WILLIAM, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- CHOY, S. (2001). *Student whose parents did not go to college. Postsecondary Access, persistence and attainment*. Washington DC: National Center for Education Statistics. US Department of Education.
- CIZEK, G. J. y BUNCH, M. B. (2007). *Standard Setting. A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, California: Sage Publications.
- HAMBLETON, R. K. (1998). Setting performance standards on achievement tests: Meeting the requirements of Title I. En HANSCH, L. N. (ed.). *Handbook for the development of performance standards: Meeting the requirements of Title I*. Washington, DC: Council of Chief State School Officers, 97-114.
- HAMBLETON, R. K.; JAEGER, R. M.; PLAKE, B. S. y MILLS, C. N. (2001). *Handbook for setting standards on performance assessment*. Washington DC: Council of Chief State School Officers.
- IMPARA, J. C. y PLAKE, B. S. (1997). Standard setting: An alternative approach. *Journal of Educational Measurement*, 34, 353-366.
- INSTITUTO NACIONAL DE ESTADÍSTICA (2011). *Censo 2011*. Disponible en <www.ine.gub.uy>.
- JORNET, J. M. y BACKHOFF, E. (2008). *Modelo para la determinación de niveles de logro y puntos de corte de los exámenes de la calidad y el logro educativos (Excale)*. Colección Cuadernos de Investigación. México: INEE, 30.
- JORNET, J. M. y GONZÁLEZ SUCH, J. (2009). Evaluación criterial: determinación de estándares de interpretación (EE) para pruebas de rendimiento educativo. *Estudios sobre Educación*, 16, 103-123.
- LEYVA BARAJAS, Y. E. (2011). Un reseña sobre la validez del constructo de pruebas referidas al criterio. *Perfiles Educativos*, 23, 131-154.
- MARCHESI, A. (2000). Un sistema de indicadores de desigualdad educativa. *Revista Iberoamericana de Educación*, 23, 1-22.
- MONTT, G. (2013). No sorprende el resultado de Uruguay en las pruebas PISA. Entrevista de Gonzalo Terra. Archivo *El País*, 13/12/2013. Sección Información.
- MUSSIO, I. y MARTINOTTI, L. (2013). Informe sobre prueba diagnóstica aplicada a estudiantes que ingresan a la Facultad de Ciencias Sociales. Documento n.º 12-13. Departamento de Economía, Facultad de Ciencias Sociales, Udelar.
- PORCEL, E. A.; DAPOZO, G. N. y LÓPEZ, M. V. (2010). Predicción del rendimiento académico de alumnos de primer año de la Facena (UNNE) en función de su caracterización socioeducativa. *Revista Electrónica de Investigación Educativa*, 12, 2. Disponible en <http://redie.uabc.mx/index.php/redie>.
- UEFI (Unidad de Enseñanza de Facultad de Ingeniería) (2012). *Informe de la Herramienta Diagnóstica al Ingreso. Generación 2012*. Facultad de Ingeniería. Udelar.
- UNIDAD DE APOYO A LA ENSEÑANZA (UAE) (2012). *Informe perfil de ingreso de la generación 2012*. CURE.
- (2013a). *Informe perfil de ingreso de la generación 2013*. CURE.
- (2013b). *Resultados sobre la evaluación diagnóstica 2013*. CURE.
- UNIVERSIDAD DE LA REPÚBLICA (UDELAR) (2012). *Hacia la reforma universitaria. La universidad en el interior*, n.º 14. Montevideo: Rectorado Universidad de la República.
- VEIGA, D.; FERNÁNDEZ, E.; LAMSCHETEIN, S. y RIVOIR, A. (2012). *Crecimiento económico y desigualdad social en Maldonado*. Facultad de Ciencias Sociales, Centro Universitario de la Región Este, Udelar.