

Validez de predicción

DE LOS CRITERIOS DE ADMISIÓN, POR GÉNERO,
A LA UNIVERSIDAD DE PUERTO RICO

Rafael J. Colorado Laguna, Ph.D.

Facultad de Ciencias Naturales
Universidad de Puerto Rico
Recinto de Río Piedras
rafaeljcolorado@gmail.com

Lillian Corcino Marrero, Ed.D.

Facultad de Educación
Universidad de Puerto Rico
Recinto de Río Piedras
l_corcino19@hotmail.com

RESUMEN

Esta investigación consistió de un estudio de validez para determinar posibles diferencias de género en las puntuaciones en las pruebas de aptitud y de aprovechamiento, el conjunto de predictores de éxito universitario, y el poder de predicción relativa de cada predictor. Los resultados constataron tendencias registradas en los Estados Unidos respecto a diferencias de género en las puntuaciones en pruebas de aptitud y aprovechamiento. Por otro lado, parcialmente corroboraron tendencias de género respecto a la magnitud de los coeficientes de correlación entre predictores y criterios. Los tres mejores predictores de éxito universitario para mujeres y varones, a corto y a largo plazos, coinciden en dos de los predictores: el promedio de escuela superior y la puntuación en la prueba de aptitud verbal. El tercer mejor predictor para mujeres fue, a corto y largo plazos, la puntuación en la prueba de aprovechamiento en inglés.

Palabras clave: estudio de validez, criterios de admisión, predictores, coeficiente de validez, género

ABSTRACT

This research consisted of a validity study to determine gender differences with respect to aptitude and achievement tests scores, a set of college success predictors, and the relative predictive power of each predictor. Results confirmed nationwide findings with respect to tests scores. However,

tendencies with respect to correlation coefficients' magnitude were partially confirmed. The three best predictors of short time and longtime college success, for each gender, coincided in the first two: high school academic average and verbal aptitude test score. The third best predictor of college success for females, short term and long term, was the English achievement test score.

Keywords: admissions criteria, gender, predictors, validity coefficient, validity study

■ Introducción

La mayoría de las instituciones universitarias en los Estados Unidos y Puerto Rico utiliza criterios de admisión para determinar el ingreso de aquellos estudiantes egresados de escuela superior que aspiran a continuar estudios universitarios. Los criterios de admisión tradicionalmente han sido de carácter académico. Hawkins y Clinedinst (2006) reportan que los principales criterios de admisión en los Estados Unidos han sido: el promedio académico de escuela superior (GPA), el promedio académico en cursos preparatorios para estudios universitarios, las puntuaciones en pruebas estandarizadas de aptitud y de aprovechamiento, y el rango percentil del estudiante respecto a su clase al momento de graduación de escuela superior.

Al igual que en los Estados Unidos, los criterios de admisión a la Universidad de Puerto Rico (UPR) han sido de tipo académico. Menéndez (1995) informa que, a partir de los años 1970, la UPR ha confeccionado diversas fórmulas de admisión que han incluido los siguientes tres componentes: el promedio académico de escuela superior, las puntuaciones en las pruebas de aptitud verbal y aptitud matemática, y las puntuaciones en las pruebas de aprovechamiento en español, matemáticas e inglés como segundo idioma. Tanto las pruebas de aptitud como las de aprovechamiento han sido confeccionadas y administradas por el College Entrance Examination Board (CEEB).

A partir de 1994, la fórmula de admisión de la Universidad de Puerto Rico, conocida como el *Índice General de Solicitud* (IGS), ha consistido de un promedio ponderado de los siguientes tres componentes: el promedio de escuela superior (50%), la puntuación en la prueba de aptitud verbal (25%) y la puntuación en la prueba

de aptitud matemática (25%). A pesar de que en un momento lo hiciera, actualmente la fórmula no incluye ninguna de las puntuaciones en las tres pruebas de aprovechamiento.

El método comúnmente empleado para evaluar la efectividad de los criterios de admisión respecto a algún indicador de éxito universitario se conoce con el nombre de *estudio de validez*. En un estudio de validez los criterios de admisión juegan el papel de variables independientes y se les denomina *predictores*. Por otro lado, el indicador de éxito universitario constituye la variable dependiente y se conoce con el nombre de *criterio*. Un estudio de validez es un estudio de *correlación* e incluye los siguientes análisis estadísticos: el cómputo de *coeficientes de correlación simple* entre cada uno de los predictores y el criterio ($-1 \leq r_{xy} \leq 1$), la generación de diversas ecuaciones o *modelos de regresión lineal múltiple* y el cómputo de *coeficientes de correlación múltiple* entre los valores *predichos* y los valores *observados* del criterio, conocidos, muy apropiadamente, como *coeficientes de validez* ($0 \leq R \leq 1$).

El coeficiente de correlación lineal múltiple, o coeficiente de validez, permite estimar el porcentaje de variabilidad o dispersión del criterio, que puede ser explicado a partir del conjunto de predictores. De otra parte, según expresado por Punch (2004), los objetivos de los modelos de regresión lineal múltiple son: estimar qué fracción o porcentaje de la variabilidad del criterio puede ser explicada por un conjunto de dos o más predictores y comparar el peso o contribución relativa de cada uno de los predictores para explicar la variabilidad del criterio. Mientras mayor sea el porcentaje de variabilidad del criterio explicada por el conjunto de predictores, mayor será el valor de predicción conjunta de los predictores. Por otro lado, mientras mayor sea el peso o contribución relativa de cada predictor, mayor será el poder de predicción del predictor respecto a los predictores restantes.

Borg y Gall (1983) destacan la importancia que puede tener el examinar subgrupos de la muestra al llevar a cabo un estudio de validez. Con frecuencia el coeficiente de validez y la validez de predicción relativa de los predictores para subgrupos son diferentes a los de la totalidad de la muestra. De igual manera, la ecuación de regresión múltiple para un subgrupo resulta ser diferente a la ecuación de regresión para la totalidad de la muestra. Por ejemplo,

se ha encontrado que las puntuaciones en pruebas de aptitud son mejores predictores del promedio académico de las mujeres que de los hombres (Borg & Gall, 1983). Aquella variable que, al utilizarse para dividir una muestra en subgrupos, genera correlaciones y ecuaciones de regresión diferentes para los subgrupos se conoce con el nombre de *variable moderadora*.

Entre las variables moderadoras utilizadas con mayor frecuencia en los Estados Unidos están: raza o grupo étnico, género y nivel socioeconómico (Bridgeman, McCamley-Jenkins & Ervin, 2000; Burton & Ramist, 2001; Ramist, Lewis, & McCamley-Jenkins, 1993; Zwick, 2007). Recientemente, se han considerado otras variables moderadoras, tales como la presencia de algún tipo de discapacidad (Zwick, 2007). En un estudio de validez llevado a cabo en la UPR, Menéndez (1996) incorporó como variables moderadoras: el área o disciplina de estudios, el género, el tipo de escuela superior de procedencia (pública o privada), el ingreso familiar, el nivel de educación de los padres y la ocupación de los padres.

Estudios de validez llevados a cabo en los Estados Unidos consistentemente muestran diferencias por género respecto a: las puntuaciones en las pruebas de aptitud y de aprovechamiento, los coeficientes de correlación y la capacidad de predicción de los modelos de regresión múltiple. Los varones, apunta Zwick (2007), obtienen mejores puntuaciones en las pruebas de aptitud del SAT (Scholastic Aptitude Test, SAT I verbal y SAT I matemáticas) y en las pruebas del ACT (Academic College Test, matemáticas, ciencias y puntuación total). Por el contrario, las mujeres tienden a obtener mejores puntuaciones en las pruebas de lectura e inglés del ACT. Un resumen de investigaciones llevadas a cabo por el CEEB (1998) indica que las mujeres de duodécimo grado tienden a desempeñarse mejor en pruebas verbales y escritas, mientras que los varones exhiben mejor aprovechamiento en pruebas de ciencias, matemáticas y destrezas mecánicas.

Varias razones se han aducido para explicar estas diferencias. Para empezar, se argumenta que un porcentaje mayor de varones se ha inclinado por áreas de estudio tales como las ciencias naturales, matemáticas, ingeniería y administración de empresas, áreas que requieren tomar cursos preparatorios en ciencias y matemáticas en escuela superior. Recientemente, sin embargo, este argumento ha

perdido validez dado el aumento en la proporción de mujeres que muestran interés en estas áreas de estudio y también toman estos cursos preparatorios. Otras razones para explicar las diferencias entre géneros en las puntuaciones obtenidas en las pruebas estandarizadas incluyen: razones de tipo socioeconómico, sesgo en la confección de las pruebas, diversidad de intereses y aspiraciones e influencias sociales, y hasta diferencias biológicas.

Por otro lado, el College Board (1998) resalta que los resultados de cientos de estudios de validez llevados a cabo en los Estados Unidos confirman que la correlación del SAT con el promedio de primer año de estudios universitarios y con notas de cursos individuales es mayor para las mujeres que para los varones. La explicación tradicional ha sido que los varones son más propensos a faltar a clases y a no entregar asignaciones y trabajos; es decir, son menos responsables que las mujeres, lo que hace menos predecible su promedio académico de primer año y sus notas en cursos específicos. Para sustentar este argumento, Zwick (2007) indica que en instituciones universitarias altamente selectivas, en las que tanto los varones como las mujeres están comprometidos con el cumplimiento de sus labores académicas, los coeficientes de validez tienden a ser muy similares.

Utilizando como criterio de éxito universitario el promedio académico de primer año de estudios universitarios, Ramist, Lewis y McCamley-Jenkins (1993) llevaron a cabo un estudio de validez en los Estados Unidos con una muestra de 46,379 estudiantes procedentes de 38 instituciones universitarias. Los coeficientes de correlación registrados resultaron ser mayores para las mujeres que para los varones. De igual manera, utilizando el promedio de primer año como criterio, Bridgeman, McCamley-Jenkins y Ervin (2000) realizaron un estudio de validez con datos de una muestra de 48,039 estudiantes procedentes de 23 universidades de los Estados Unidos recopilados a mediados de los años 1990. Incluyeron como variables moderadoras género y grupos étnicos (blancos, afroamericanos, hispanos y asiáticos). Sus resultados pueden resumirse de la siguiente manera: el promedio de escuela superior resultó ser el mejor predictor para la totalidad de la muestra y para todos los grupos étnicos, excepto para los asiáticos, cuyo mejor predictor, tanto para varones como para mujeres,

fue la suma de las puntuaciones en las pruebas de aptitud; para todos los grupos étnicos, excepto para los asiáticos, los coeficientes de correlación para las mujeres fueron mayores, en menor o mayor grado, que los de los varones, y la magnitud de los coeficientes de correlación por grupo étnico, en orden descendente, fue la siguiente: asiáticos, blancos, hispanos y afroamericanos.

En lo que respecta a las diferencias en las correlaciones por género, todos los grupos, excepto el asiático, se conformaron a la tendencia de que las mujeres registran una correlación mayor que los varones. Evidentemente, algunas explicaciones para estas diferencias no aplican para el caso de los asiáticos. Podría ser que tanto los varones como las mujeres asiáticas seleccionan áreas difíciles de estudio y, además, son igual de responsables asistiendo a clases y entregando sus asignaciones y trabajos de clases.

La tercera diferencia entre varones y mujeres consistentemente registrada en la literatura de los estudios de validez está relacionada con la capacidad de predicción y el peso relativo de los predictores del modelo de regresión generado para toda la muestra, al compararse con los modelos de regresión correspondientes para cada uno de los posibles valores de la variable moderadora. Decimos que un modelo de regresión *sobrepredice* (*subpredice*) un criterio, si los valores predichos del criterio por el modelo de regresión son *mayores* (*menores*) que los valores observados del criterio. Los predictores tradicionales, promedios de escuela superior y puntuaciones en las pruebas estandarizadas tienden a subpredecir el desempeño académico de las mujeres y a sobrepredicir el desempeño de los varones.

Ramist, Lewis y McCamley-Jenkins (1993) examinaron diferentes variables moderadoras, incluyendo el género del estudiante. El criterio de éxito universitario utilizado fue el promedio académico de primer año. La Tabla 1 resume los resultados de este estudio. Se ha restado el *valor observado* del *valor predicho* (*valor predicho - valor observado*) de forma tal que una diferencia positiva significa sobrepredicción mientras que una diferencia negativa significa subpredicción.

De la Tabla 1 se pueden derivar las siguientes conclusiones:

1. Para todas las combinaciones de predictores, las mujeres evidenciaron subpredicción, y los varones, sobrepredicción.

Tabla 1

Sub y sobre predicciones, por género, para promedio de primer año

Predictor	Género	
	Femenino	Masculino
HSGAP	- 0.02	+ 0.02
SAT-V	- 0.06	+ 0.06
SAT-M	- 0.10	+ 0.11
SAT (Total)	- 0.09	+ 0.10
HSGPA + SAT	- 0.06	+ 0.06

Notas. Un signo negativo representa sub predicción; un signo positivo representa sobre predicción

HSGPA = Promedio de escuela superior

SAT = Prueba de Aptitud (V = Verbal, M = Matemáticas)

2. La diferencia entre las predicciones son menores cuando solo se incluye como predictor el promedio de escuela superior.
3. La diferencia entre predicciones es mayor cuando se incluye cualquiera de las dos partes de la prueba de aptitud (SAT-V o SAT-M) o el total de las dos partes de esta prueba (SAT-T).

En un estudio llevado a cabo con estudiantes de 23 universidades bastante selectivas, Bridgeman, McCamley-Jenkins y Ervin (2000) utilizaron como predictores de promedio de primer año el promedio de escuela superior y las puntuaciones en el SAT. Utilizaron como variable moderadora el género, por grupo étnico (blancos, afroamericanos, hispanos y asiáticos). Los resultados mostraron que el patrón de sobrepredicción para los varones, para todos los grupos étnicos, es más regular y marcado que el patrón de subpredicción de las mujeres. Por ejemplo, los resultados no mostraron subpredicción para el grupo de mujeres hispanas; sin embargo, sí lo hicieron para el grupo de mujeres blancas.

Zwick (2007) ofrece dos posibles explicaciones para el fenómeno de subpredicción de las mujeres. La primera de ellas es que las investigaciones tienden a indicar que las féminas, por lo general, toman cursos menos rigurosos y optan por áreas de estudio menos rigurosas que los varones. La subpredicción es menos

marcada en instituciones selectivas en las que los estudiantes, tanto mujeres como varones, toman cursos más rigurosos. Una segunda explicación sostiene que puede ser que las mujeres estén mejor preparadas que los varones o que sean más responsables y aplicadas en sus estudios.

De otra parte, Menéndez (1996) calculó los coeficientes de correlación respecto al promedio de primer año y al promedio de graduación, por género, utilizando una muestra de estudiantes del Recinto de Río Piedras de la UPR. La Tabla 2, recoge los coeficientes de correlación para promedio de primer año y la Tabla 3, los coeficientes de correlación para promedio de graduación.

Tabla 2

Coefficientes de correlación para promedio de primer año, por género, en el Recinto de Río Piedras de la UPR

Predictor	Género	
	Femenino	Masculino
HSGPA	0.3518	0.4030
PAPTV	0.2663	0.2454
PAPTM	0.2026	0.2456
PAPRI	0.2066	0.1553
PAPRM	0.2718	0.2799
PAPRE	0.3311	0.3201

Notas. HSGPA = Promedio de escuela superior
 PAPT V = Prueba de aptitud verbal
 PAPTM = Prueba de aptitud matemática
 PAPRI = Prueba de aprovechamiento en inglés como segundo idioma
 PAPRM = Prueba de aprovechamiento en matemáticas
 PAPRE = Prueba de aprovechamiento en español

A partir de los coeficientes de correlación presentados en la Tabla 2, se puede concluir:

1. Los primeros tres coeficientes de correlación correspondieron, para ambos géneros, al promedio de escuela superior, la prueba de aprovechamiento en español y la prueba de aprovechamiento en matemáticas, respectivamente.
2. Los dos coeficientes de correlación menores, para ambos géneros, correspondieron a la prueba de aptitud

matemática y la prueba de aprovechamiento en inglés, respectivamente.

3. Los coeficientes de correlación de las pruebas de aprovechamiento en español y matemáticas, para ambos géneros, fueron mayores que los coeficientes de correlación correspondientes para las pruebas de aptitud en español y en matemáticas.
4. Los coeficientes de correlación para las pruebas de aptitud y aprovechamiento en matemáticas para los varones fueron mayores que los coeficientes de correlación correspondientes para las mujeres.
5. Por el contrario, los coeficientes de correlación para las pruebas de aptitud verbal, aprovechamiento en español y aprovechamiento en inglés fueron mayores para las mujeres que los coeficientes de correlación correspondientes para los varones.

Los resultados de la Tabla 2 concuerdan con las tendencias reportadas en los Estados Unidos sobre las diferencias de los coeficientes de correlación entre mujeres y varones, según lo confirman los puntos 4 y 5 arriba enumerados.

Tabla 3

Coefficientes de correlación para promedio de graduación, por género, en el Recinto de Río Piedras de la UPR

Predictor	Género	
	Femenino	Masculino
HSGPA	0.3565	0.3934
PAPTV	0.3264	0.3274
PAPTM	0.1424	0.1812
PAPRI	0.2435	0.1355
PAPRM	0.1607	0.1737
PAPRE	0.3170	0.3372

A partir de los coeficientes de correlación presentados en la Tabla 3, se puede concluir:

1. El coeficiente de correlación mayor, para ambos géneros, correspondió al promedio de escuela superior.

2. El segundo y tercer coeficiente de correlación mayor, para ambos géneros y en orden inverso, correspondieron a las pruebas de aptitud verbal y de aprovechamiento en español.
3. Los coeficientes de correlación menores, para ambos géneros, correspondieron a las pruebas de aptitud matemática y de aprovechamiento en matemáticas.
4. Excepto para las pruebas de aprovechamiento en inglés como segundo idioma, los coeficientes de correlación en todas las pruebas fueron mayores para los varones que para las mujeres.

Los resultados de la Tabla 3, contrario a los resultados de la Tabla 2, no reflejan las tendencias reportadas en los Estados Unidos sobre las diferencias en los coeficientes entre las mujeres y los varones según lo indica el punto 4.

Menéndez (1996) también comparó modelos de regresión para predecir promedio de primer año y promedio de graduación para varones y mujeres. La Tabla 4 recoge los modelos de regresión obtenidos por Menéndez con los mejores tres predictores.

Tabla 4

Modelos de regresión, por género, con los mejores tres predictores para promedio de primer año y promedio de graduación

Promedio de Primer Año			
Femenino		Masculino	
Predictor	Coefficiente de Validez	Predictor	Coefficiente de Validez
HSGPA		HSGPA	
PAPRE	0.4246	PAPRE	0.4522
PAPRI		PAPTV	
Promedio de Graduación			
HSGPA		HSGPA	
PAPTV	0.4678	PAPTV	0.4990
PAPRI		PAPRE	

A partir de los resultados de la Tabla 4 se puede concluir:

1. El promedio de escuela superior resultó ser el mejor predictor para ambos criterios y para ambos géneros.
2. Para promedio de primer año, para ambos géneros, el segundo mejor predictor resultó ser la prueba de aprovechamiento en español.
3. Para promedio de graduación, para ambos géneros, la prueba de aptitud verbal sustituyó como segundo mejor predictor a la prueba de aprovechamiento en español.
4. Ni la prueba de aptitud matemática ni la de aprovechamiento en matemáticas figuraron entre los primeros tres predictores para ninguno de los dos criterios para varones o mujeres.
5. Para las mujeres, el tercer mejor predictor, tanto para promedio de primer año como para promedio de graduación, resultó ser la prueba de aprovechamiento en inglés como segundo idioma.
6. Para ambos criterios, los varones exhibieron un coeficiente de validez mayor que las mujeres.

Siguiendo la línea de investigación de Menéndez (1996), este estudio persiguió contestar las siguientes preguntas de investigación:

1. ¿Qué diferencia existe, si alguna, entre el desempeño de los varones y las mujeres en las pruebas de aptitud (verbal y matemática) y en las pruebas de aprovechamiento (español, matemáticas e inglés como segundo idioma) que se han utilizado como criterios de admisión a la Universidad de Puerto Rico?
2. ¿Qué diferencia existe, si alguna, entre los coeficientes de correlación y de validez entre predictores y criterios (promedio de primer año y de graduación) para mujeres y varones?
3. ¿Qué diferencia existe, si alguna, en el poder de predicción de éxito universitario (promedio de primer año y promedio de graduación) de los diferentes predictores utilizados como criterios de admisión respecto al género del solicitante?

Método

Los criterios de admisión seleccionados como variables independientes o predictores en este estudio fueron: el promedio académico de escuela superior, las puntuaciones en las pruebas de aptitud verbal y aptitud matemática, y las puntuaciones en las pruebas de aprovechamiento en español, matemáticas e inglés como segundo idioma. A pesar de que la fórmula de admisión actual no considera las puntuaciones en las pruebas de aprovechamiento, se incluyeron como predictores en el ánimo de determinar su validez de predicción respecto a los demás componentes de la fórmula. Las puntuaciones en las pruebas de aptitud y de aprovechamiento fluctúan entre una puntuación mínima de 200 y una máxima de 800.

Los criterios de éxito universitario seleccionados fueron: el promedio de primer año de estudios universitarios y el promedio de graduación. El promedio académico (GPA) fluctúa entre un mínimo de 0.0 y un máximo de 4.0. Este estudio utilizó datos recolectados de la cohorte del año 2000 de estudiantes admitidos a la Universidad de Puerto Rico, Recinto de Río Piedras. El conjunto original de datos consistía de un total de 3,116 estudiantes.

Como es natural, no todos los estudiantes admitidos a cursar estudios universitarios completan su primer año de estudios. De igual manera, no todos aprueban el mismo número de créditos al completar su primer año de estudios universitarios, por lo que promedios iguales no necesariamente representan igual esfuerzo o aprovechamiento. Tampoco el mismo número de créditos aprobados necesariamente representa igual esfuerzo o aprovechamiento, ya que los cursos difieren en la dificultad de sus requisitos, las exigencias del profesor y los criterios de evaluación. Para minimizar el posible efecto de estas diferencias, solamente aquellos estudiantes que aprobaron por lo menos 24 créditos al completar su primer año fueron considerados en el cómputo del promedio de primer año. Esto representa, en promedio, 12 créditos por semestre, lo que equivale a una carga académica completa para el estudiante. Una vez removidos de la cohorte original aquellos alumnos que no completaron su primer año o que aprobaron menos de 24 créditos, la cantidad de estudiantes se redujo a 2,124.

Similarmente, al momento de llevar a cabo el estudio, algunos estudiantes se habían dado de baja de la universidad, se habían transferido a otra institución universitaria o no habían completado su grado de bachillerato. Una vez removidos estos estudiantes, un total de 1,653 estudiantes de la cohorte original completó su grado de bachillerato, por lo que fueron incluidos en el cómputo del promedio académico de graduación.

Los resultados del estudio fueron generados en el siguiente orden: estadísticas descriptivas de cada predictor y de cada uno de los dos criterios, coeficientes de correlación lineal entre cada predictor y cada uno de los dos criterios, y ecuaciones de regresión lineal múltiple entre la totalidad de los predictores y cada uno de los dos criterios.

Resultados

Para contestar la primera pregunta de investigación, sobre posibles diferencias por género en las puntuaciones en las pruebas de aptitud y de aprovechamiento, la Oficina del CEEB en Puerto Rico proveyó información de estas puntuaciones en los últimos ocho años (2005-2012). A continuación se incluyen dos diagramas de línea que ilustran las diferencias por género entre las puntuaciones promedio de las pruebas de aptitud (verbal y matemática) en estos años.

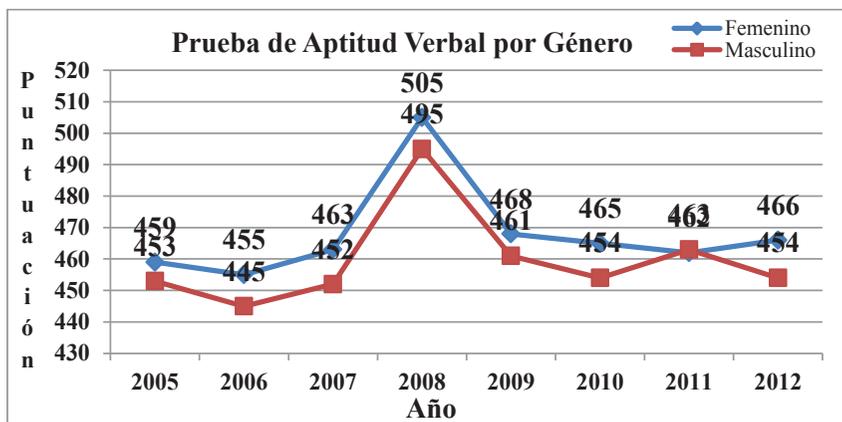


Figura 1. Puntuación promedio, por género, en la prueba de aptitud verbal

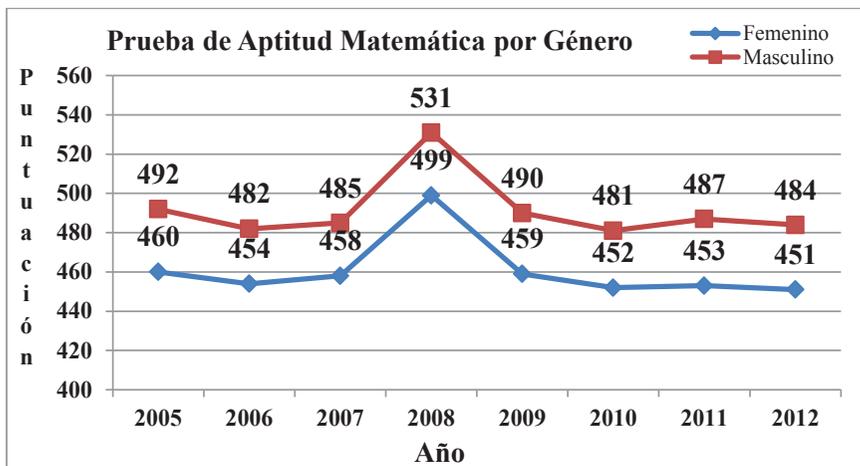


Figura 2. Puntuación promedio, por género, en la prueba de aptitud matemática

Los resultados ilustrados en la Figura 1 y la Figura 2 corroboran el patrón registrado en los Estados Unidos: las mujeres exhiben una puntuación promedio mayor en las pruebas de aptitud verbal, mientras que, por el contrario, los varones registran una puntuación promedio mayor en las pruebas de aptitud matemática. Nótese que las diferencias entre las puntuaciones promedio en las pruebas de aptitud verbal son menores, fluctuando entre 7 y 12, mientras que las diferencias entre las puntuaciones promedio en las pruebas de aptitud matemática son mayores, fluctuando entre 27 y 33.

Seguidamente, se incluyen tres diagramas de líneas con las puntuaciones promedio, por género, obtenidas en las tres pruebas de aprovechamiento administradas en los últimos ocho años (2005-2012).

Los resultados en las pruebas de aprovechamiento ilustrados en la Figura 3 y la Figura 4 corroboraron el patrón registrado en las pruebas de aptitud. Las mujeres registraron una puntuación promedio mayor en la prueba de aprovechamiento en español, mientras que, por el contrario, los varones registraron una puntuación promedio mayor en la prueba de aprovechamiento en matemáticas. Nótese, sin embargo, que las diferencias entre las puntuaciones promedio en las pruebas de aprovechamiento en español son

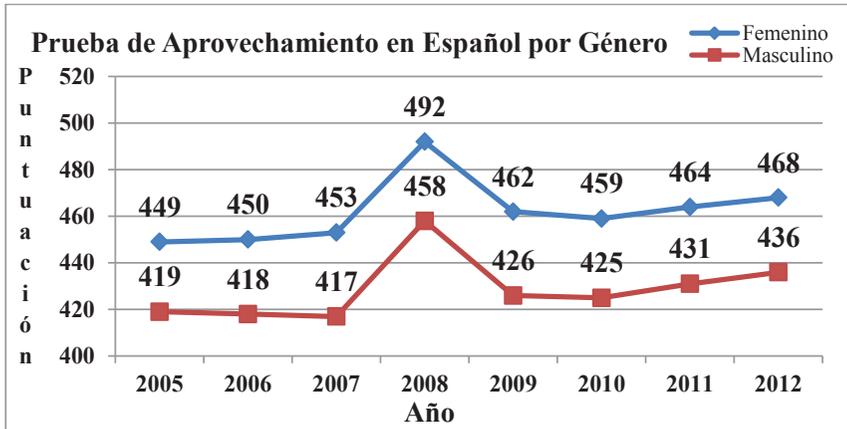


Figura 3. Puntuación promedio, por género, en la prueba de aprovechamiento en español

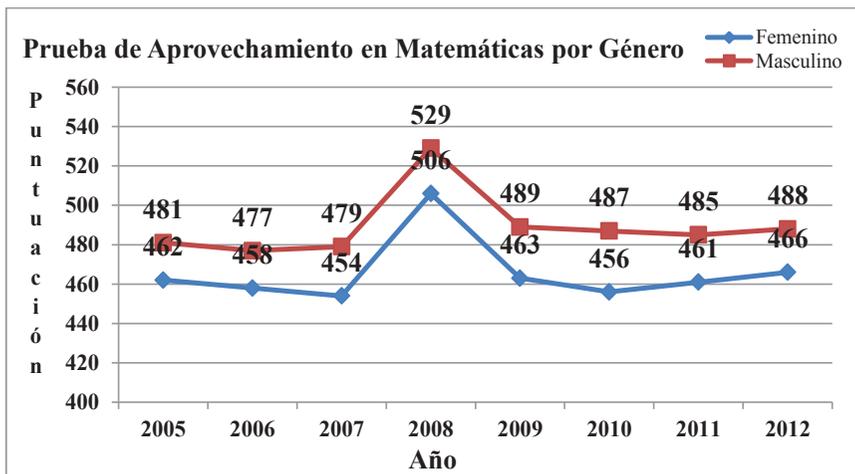


Figura 4. Puntuación promedio, por género, en la prueba de aprovechamiento en matemáticas

mayores, fluctuando entre 30 y 36, mientras que las diferencias entre los promedios en las pruebas de aprovechamiento en matemáticas fueron menores, fluctuando entre 19 y 31.

En lo que concierne a las pruebas de aprovechamiento en inglés como segundo idioma, ilustradas en la Figura 3, contrario

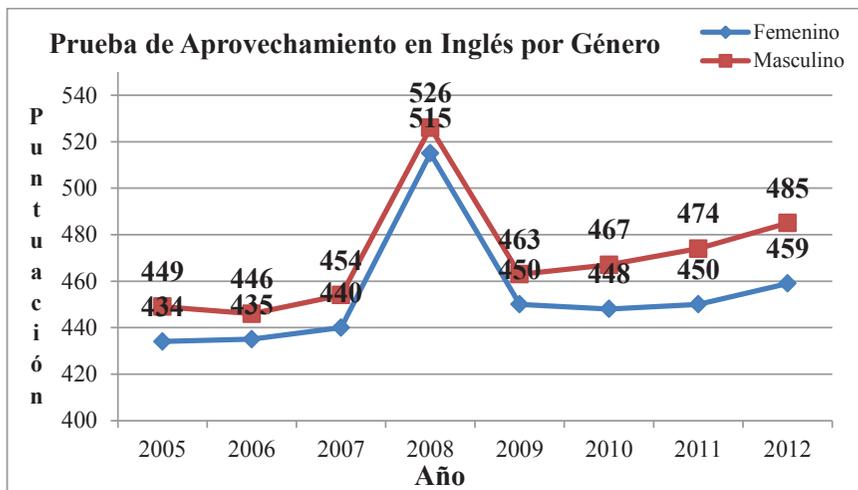


Figura 5. Puntuación promedio, por género, en la prueba de aprovechamiento en inglés como segundo idioma

a lo esperado, los varones registraron una puntuación promedio mayor que las mujeres. Las diferencias fluctuaron entre 11 y 26.

A continuación, la Tabla 5 y la Tabla 6 presentan las estadísticas descriptivas de la cohorte de este estudio con respecto a los predictores de promedio de primer año universitario y de promedio de graduación, respectivamente. La última columna de cada tabla corresponde al coeficiente de correlación de cada predictor con el criterio correspondiente.

1. Las estadísticas presentadas en la Tabla 5 y la Tabla 6 pueden resumirse de la siguiente manera:
2. Las mujeres exhibieron un mayor promedio académico que los varones en escuela superior, al completar su primer año de estudios universitarios y al completar su grado de bachillerato.
3. La puntuación promedio en la prueba de aptitud verbal de las mujeres fue muy similar a la de los varones. Sin embargo, la puntuación promedio en la prueba de aprovechamiento en español de las mujeres fue significativamente mayor que la de los varones.
4. La puntuación promedio de los varones en la prueba de aptitud matemática y en la prueba de aprovechamiento

Tabla 5

Estadísticas descriptivas, por género, de los predictores con respecto al promedio de primer año

Variable	Media Aritmética		Desviación Estándar		Coeficiente con FGPA	
	F	M	F	M	F	M
FSGPA	3.175	3.024	0.502	0.519	1.000	1.000
HSGPA	3.600	3.402	0.348	0.437	0.390	0.390
PAPTV	590.18	582.33	70.023	77.841	0.344	0.322
PAPTM	598.67	632.82	88.609	87.581	0.311	0.306
PAPRE	563.61	535.21	63.787	69.382	0.324	0.320
PAPRM	593.40	621.24	88.380	87.353	0.330	0.338
PAPRI	555.02	559.10	103.994	109.994	0.298	0.258

Notas. Total de mujeres = 1,518

Total de varones = 606

Tabla 6
Estadísticas descriptivas, por género, de los predictores con respecto al promedio de graduación

Variable	Media Aritmética		Desviación Estándar		Coeficiente con FGPA		
	Género	F	M	F	M	F	M
GGPA		3.370	3.260	0.332	0.350	1.000	1.000
HSGPA		3.624	3.411	0.354	0.460	0.298	0.365
PAPTV		594.33	583.80	70.041	80.811	0.312	0.319
PAPTM		601.21	636.52	90.023	92.250	0.255	0.286
PAPRE		566.17	535.91	65.624	72.235	0.277	0.333
PAPRM		598.05	623.89	89.472	92.811	0.259	0.289
PAPRI		560.38	566.09	104.468	112.562	0.277	0.216

Notas. Total de mujeres = 1,181

Total de varones = 472

matemático fue significativamente mayor que las de las mujeres. Estos resultados coinciden con la tendencia registrada en los Estados Unidos.

5. Las puntuaciones promedio en la prueba de aprovechamiento en inglés fueron muy similares. Sin embargo, el coeficiente de correlación de esta prueba con los dos criterios de éxito fue significativamente mayor para mujeres que para varones.
6. Los coeficientes de correlación de los predictores respecto al promedio de primer año son muy similares. Resalta, sin embargo, la diferencia en la correlación respecto a la puntuación en la prueba de aprovechamiento en inglés (para mujeres $r_{xy} = 0.298$ y para varones $r_{xy} = 0.258$).
7. Los coeficientes de correlación de los predictores respecto al promedio de graduación para varones resultaron ser mayores que las de las mujeres, con la excepción de la prueba de aprovechamiento en inglés que, al igual que ocurriera con el promedio de primer año, el coeficiente de correlación para las mujeres fue significativamente mayor que para los varones (para mujeres $r_{xy} = 0.277$ y para varones $r_{xy} = 0.216$).

Como era de esperar, la magnitud de los coeficientes de correlación disminuye, en menor o mayor grado, cuando se utiliza como criterio el promedio de graduación pues los predictores están temporalmente más distantes de este criterio que del promedio de primer año.

La Tabla 7 incluye los resultados de los modelos de regresión, por género, incorporando todos los predictores, para cada uno de los dos criterios de éxito universitario utilizado: promedio de primer año y promedio de graduación. Los *coeficientes beta* que se resaltan representan un nivel de significación $\alpha \leq 0.5$.

Los coeficientes beta de los modelos de regresión ilustrados en la Tabla 7 pueden resumirse de la siguiente manera:

1. El mejor predictor de éxito universitario, a corto y a largo plazos, tanto para mujeres como para varones, resultó ser el promedio de escuela superior.
2. El segundo mejor predictor de éxito universitario, a corto y a largo plazos, tanto para mujeres como para varones, resultó ser la puntuación en la prueba de aptitud verbal.

3. Los únicos dos predictores para varones y para ambos criterios con coeficientes beta estadísticamente significativos resultaron ser el promedio de escuela superior y la puntuación en la prueba de aptitud verbal.
4. Para mujeres, respecto a ambos criterios, el tercer mejor predictor resultó ser la puntuación en la prueba de aprovechamiento en inglés como segundo idioma.
5. Solamente para mujeres y para promedio de primer año hubo un cuarto mejor predictor, que resultó ser la puntuación en la prueba de aprovechamiento en español.
6. Ni la prueba de aprovechamiento en español (excepto en el caso anterior), ni ninguna de las dos pruebas matemáticas (aptitud y aprovechamiento) exhibió un coeficiente beta estadísticamente significativo.

La Tabla 8 recoge los modelos de regresión, por género, con los mejores tres predictores para promedio de primer año y promedio de graduación. Los resultados de la Tabla 8 pueden ser resumidos de la siguiente manera:

1. El mejor predictor para ambos géneros y ambos criterios fue el promedio de escuela superior.
2. El segundo mejor predictor para ambos géneros y ambos criterios fue la puntuación en la prueba de aptitud verbal.
3. Para mujeres y para ambos criterios, el tercer mejor predictor fue la puntuación en la prueba de aprovechamiento en inglés como segundo idioma.
4. Para hombres, el tercer mejor predictor para promedio de primer año fue la puntuación en la prueba de aprovechamiento en matemáticas, y para promedio de graduación, la prueba de aprovechamiento en español.
5. La puntuación en la prueba de aptitud matemática no aparece entre los mejores tres predictores de ninguno de los modelos.
6. Los modelos de regresión, tanto los de las mujeres como los de los varones, reflejan un decremento anticipado en el poder de predicción para criterios a largo plazo. Para mujeres, el coeficiente de validez R disminuye de 0.490 a 0.409, y para varones, de 0.490 a 0.442. La disminución fue menor para varones.

Tabla 7

Modelos de regresión, por género, con los coeficientes beta de cada uno de los predictores respecto al promedio de primer año de estudios universitarios y al promedio de graduación de bachillerato

Criterio	Género	HSGPA	PAPTV	PAPTM	PAPRE	PAPRM	PAPRI	R	R ²	N
FGPA	F	0.304	0.137	0.024	0.061	0.068	0.112	0.499	0.246	1518
	M	0.318	0.130	0.026	0.077	0.092	0.077	0.499	0.242	606
GGPA	F	0.223	0.152	0.026	0.029	0.016	0.131	0.411	0.165	1181
	M	0.277	0.126	0.089	0.111	-0.025	0.012	0.446	0.189	472

Tabla 8

Modelos de regresión, por género, con los mejores tres predictores para promedio de primer año y promedio de graduación, con sus respectivos coeficientes beta

Criterio	Género	Predictor	Coefficiente Beta	R	R ² Ajustado	N
FGPA	F	HSGPA	0.330	0.490	0.239	1,518
		PAPTV	0.187			
		PAPRI	0.159			
	M	HSGPA	0.313	0.490	0.236	606
		PAPTV	0.201			
		PAPRM	0.156			
GGPA	F	HSGPA	0.236	0.409	0.165	1,181
		PAPTV	0.174			
		PAPRI	0.154			
	M	HSGPA	0.281	0.442	0.190	472
		PAPTV	0.150			
		PAPRE	0.135			

■ Discusión

En lo que atañe a la primera pregunta de este estudio, las puntuaciones promedio obtenidas en las pruebas de aptitud y de aprovechamiento durante los últimos ocho años (2005-2012), informadas por la Oficina del CEEB en Puerto Rico, corroboran el patrón reportado en los Estados Unidos (Zwick, 2007). Los varones obtuvieron puntuaciones promedio mayores en la prueba de aptitud matemática y de aprovechamiento matemático, mientras que las mujeres obtuvieron puntuaciones promedio mayores en la prueba de aptitud verbal y en la prueba de aprovechamiento en español.

Respecto a la prueba de aprovechamiento en inglés como segundo idioma, los varones obtuvieron una puntuación promedio mayor que las mujeres. No obstante, la correlación entre esta prueba y los criterios de éxito, promedios de primer año y de graduación, fue significativamente mayor para mujeres que para varones. Además, los modelos de regresión con los mejores tres

predictores para mujeres, para ambos criterios de éxito universitario, incluyeron como tercer mejor predictor la puntuación en la prueba de aprovechamiento en inglés. Por el contrario, la puntuación en esta prueba no apareció entre los mejores tres predictores de los modelos de regresión correspondientes a los varones.

Como se señalara anteriormente, se han invocado múltiples razones para explicar las diferencias por género en las puntuaciones en las pruebas, siendo las principales: sesgo en las pruebas, diferencias de interés y aspiraciones, factores socioeconómicos y hasta diferencias biológicas. Cada uno de estos factores debe ser analizado para tomar las acciones correspondientes que garanticen que los resultados en estas pruebas tengan la mayor validez posible. En primera instancia, corresponde a las organizaciones que confeccionan y administran estas pruebas evaluar periódicamente su validez y confiabilidad, de forma tal que, en la medida de lo posible, estén libres de sesgo por género.

Corresponde a la sociedad en general modificar sus prejuicios y estereotipos, que tradicionalmente tienden a ubicar a los varones en oficios y profesiones que suponen una sólida formación matemática y científica, y a desalentar a las mujeres a hacer lo mismo. Por su parte, el sistema educativo del país debe estimular y propiciar que las mujeres se interesen por carreras de este tipo para que se sientan inclinadas a tomar cursos preparatorios, a nivel de escuela superior, conducentes a estudios postsecundarios que requieran una sólida formación matemática y científica.

Respecto a la segunda pregunta de investigación, sobre las diferencias entre los coeficientes de correlación entre los predictores y el promedio de primer año para mujeres y varones, los resultados de este estudio no confirmaron las tendencias registradas en los Estados Unidos. Las diferencias entre los coeficientes no fueron significativas, excepto por la diferencia entre el coeficiente de correlación de la puntuación en la prueba de aprovechamiento en inglés, que resultó ser significativamente mayor para mujeres que para varones.

En lo que concierne a los coeficientes de correlación de los predictores con respecto al promedio de graduación, los resultados de este estudio confirmaron la tendencia de los varones a obtener una correlación más alta para la prueba de aptitud matemática y

para la prueba de aprovechamiento en matemáticas. Sin embargo, las mujeres no registraron un coeficiente de correlación mayor para la prueba de aptitud verbal y de aprovechamiento en español, sino, más bien, ocurrió todo lo contrario. Por tanto, podemos decir que solo se confirmó parcialmente la tendencia registrada en los Estados Unidos. Finalmente, los coeficientes de correlación de la prueba de aprovechamiento en inglés, para ambos criterios, fueron significativamente mayores para las mujeres. Esto confirma la tendencia de las mujeres a obtener mayores coeficientes de correlación en pruebas de aprovechamiento verbal.

Relativo a la tercera pregunta de investigación, los resultados de este estudio dejan meridianamente claro que el mejor predictor de éxito universitario, tanto a corto como a largo plazos, y para mujeres como para varones, es el promedio de escuela superior. Estos resultados también han sido confirmados por diversos estudios de validez en los Estados Unidos. Es por esto que, desde 1995, se le ha asignado el doble del peso (50%) en la fórmula de admisión a la Universidad de Puerto Rico que el peso que se le ha asignado a cada una de las pruebas de aptitud (25%). Todo parece indicar que un esfuerzo académico sostenido por varios años, como evidencia el promedio académico acumulativo del estudiante, es mejor predictor de éxito universitario que un esfuerzo momentáneo de un examen estandarizado. Por lo tanto, debe ser prioridad de todo sistema educativo que aspire a que un porcentaje considerable de sus estudiantes continúe estudios postsecundarios garantizar una educación de excelencia, proveyendo, en la medida de lo posible, los mejores recursos docentes, instalaciones, materiales educativos y servicios de apoyo al estudiante.

El segundo mejor predictor de éxito universitario, tanto a corto como a largo plazos, y para ambos géneros resultó ser la prueba de aptitud verbal. Para los varones y para promedio de primer año como criterio, el tercer mejor predictor correspondió a la prueba de aprovechamiento en matemáticas; para promedio de graduación, el tercer mejor predictor resultó ser la prueba de aprovechamiento en español. Para mujeres, sin embargo, el tercer mejor predictor de éxito universitario para ambos criterios correspondió a la prueba de aprovechamiento en inglés. Resulta interesante, y debe ser tema de investigaciones futuras, examinar

y buscar explicaciones a la prominencia que tuvo en este estudio, para mujeres, la puntuación en la prueba de aprovechamiento en inglés como predictor de éxito universitario, tanto a corto como a largo plazos.

Resalta la ausencia de la puntuación en la prueba de aptitud matemática como predictor de éxito universitario, tanto a corto como a largo plazos, especialmente cuando se le ha asignado una cuarta parte (25%) del peso en la fórmula de admisión actual. Podemos especular que si la variable moderadora estudiada hubiera sido área o disciplina de estudios universitarios, la validez de predicción de las pruebas de matemáticas, tanto la de aptitud como la de aprovechamiento, hubiera sido mayor para disciplinas que requieren una sólida formación matemática, como, por ejemplo, ciencias naturales e ingeniería. Un futuro estudio de validez podría considerar área de estudios como variable moderadora.

Finalmente, este estudio no generó resultados para examinar el fenómeno de subpredicción de éxito universitario para los modelos de regresión para mujeres, ni su contrapartida, el fenómeno de sobrepredicción de éxito universitario para los modelos de regresión para varones, que se han registrado en los Estados Unidos. Sería pertinente que un estudio futuro aborde este fenómeno para determinar si el mismo también se manifiesta en los alumnos de instituciones universitarias en Puerto Rico.

REFERENCIAS

- Borg, R. B. & Gall, M. D. (1983). *Educational research: An introduction* (4ta. ed.). New York: Longman.
- Bridgeman, W., McCamley-Jenkins, L. & Ervin, N. (2000). *Prediction of freshman grade point average from the revised and recentered SAT I: Reasoning Test*. College Board Research Report No. 2000-1. New York: College Entrance Examination Board. Recuperado de: <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2000-1-predictions-freshman-gpa-revised-recentered-sat-reasoning.pdf>
- Burton, N. W. & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980*. College Board Research Report No. 2001-2. New York: College Entrance Examination Board. Recuperado de: <http://research.collegeboard.org/sites/>

- default/files/publications/2012/7/researchreport-2001-2-predicting-college-success-sat-studies.pdf
- Camara, W. J. & Echternacht, G. (2000). *The SAT and high school grades: Utility in predicting success in college*. The College Board, RN-10, July 2000. Recuperado de: <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-2000-10-sat-high-school-grades-predicting-success.pdf>
- College Board. (1998). SAT and gender differences. College Board Research Summary, No. 4. New York: College Entrance Examination Board. Recuperado de: <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchinreview-1998-4-sat-gender-differences.pdf>
- Hawkins, D.A. & Clinedinst, M. (2006). *State of College Admission*. Alexandria: Virginia: National Association for College Admission Counseling. Recuperado de http://www.immagic.com/eLibrary/ARCHIVES/GENERAL/NACAC_US/N060508H.pdf
- Menéndez, A. (1995). Estudio sobre el proceso de admisiones a la Universidad de Puerto Rico. Informe presentado a la Administración Central de la Universidad de Puerto Rico.
- Menéndez, A. (1996). *The Predictive Validity of Admissions Measures at the University of Puerto Rico*. Dissertation University of Rutgers.
- Punch, K. F. (2004). *Introduction to social research: Quantitative and qualitative approaches*. Londres: SAGE Publications.
- Ramist, L., Lewis, C. & McCamley-Jenkins, L. (1993). *Student group differences in predicting college grades: Sex, language and ethnic groups*. College Board Research Report No. 93-1. New York: College Entrance Examination Board. Recuperado de: <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-1993-1-student-group-differences-predicting-college-grades.pdf>
- Zwick, R. (2007). *College Admissions Testing*. Report for the National Association for College Admission Counseling (NACAC). Recuperado de: <http://www.nacacnet.org/research/PublicationsResources/Marketplace/Documents/TestingWhitePaper.pdf>