

EL SISTEMA DE PRUEBAS DE COLLEGE BOARD LATAM: NUEVAS EVIDENCIAS DE VALIDEZ DE LA PAA EN LATINOAMÉRICA

College Board-LATAM Assessment Suite: New PAA Validity Evidence in Latin American

 <https://doi.org/10.47554/cii.vol15.2024.pp409-417>

Manuel González González

College Board LATAM, Estados Unidos de América

 <https://orcid.org/0000-0002-9956-3748>

 mgonzalez@gmail.com

Carla Mojica de León

College Board LATAM, Puerto Rico

 <https://orcid.org/0009-0009-9536-319X>

 cmojicadeleon@collegeboard.org



Resumen

La PAA es desarrollada y mantenida con el apoyo continuo de expertos de Puerto Rico y de América Latina. Los resultados de las versiones anteriores de la PAA están relacionados con indicadores de éxito académico de alumnos universitarios tanto en Puerto Rico (Colorado-Laguna y Corcino-Marrero, 2014, 2013) como en varios países en América Latina (Estépar García y Pedrosa Roche, 2014; Hernández Laboy, Rivera Ocasio y Pedrosa Roche, 2017; Rivera Ocasio y Hernández Laboy, 2017, González-González y Estépar García, 2023). Este estudio cuantitativo emplea un diseño investigativo correlacional de enfoque predictivo. El objetivo del estudio es examinar la relación entre las puntuaciones de la versión más reciente de la PAA y el índice de escuela de procedencia (PEP) con el índice académico del primer año de estudios universitarios en el contexto latinoamericano. La muestra del estudio corresponde a datos provistos por una institución pública y altamente selectiva de México con 120 732 alumnos aspirantes a la institución y para los análisis concernientes a la predicción del índice académico del primer año de estudios universitarios un total de 57 037 alumnos admitidos. Los resultados reflejan que los estudiantes con puntajes más altos en la PAA obtienen calificaciones más altas en su primer año de estudio en la universidad, aún con los ajustes por el efecto del PEP. Los hallazgos del estudio reafirman lo observado en estudios previos y validan el uso de la PAA como un predictor del desempeño académico de los alumnos en el contexto latinoamericano.

Palabras clave: PAA suite, validez, admisión universitaria, desempeño académico.

Abstract

The PAA is developed and maintained with the continuous support of professionals from Puerto Rico and Latin America. The results of previous versions of the PAA have been related to indicators of academic success of university students both in Puerto Rico (Colorado-Laguna and Corcino-Marrero, 2014, 2013) and in several countries in Latin America (Estépar García and Pedrosa Roche, 2014; Hernández Laboy, Rivera Ocasio, & Pedrosa Roche, 2017; Rivera Ocasio & Hernández Laboy, 2017, González-González y Estépar García, 2023). This study is a quantitative study with a correlational research design with a predictive approach. The objective of this study is to examine the relationship between the scores of the most recent version of the PAA and high school GPA (PEP) with the first year GPA in the context of a Latin American university. The sample for this study corresponds to data provided by a highly selective public university in Mexico, with a sample of 120,732 aspiring students and for the analyses concerning the prediction of the first year gpa a total of 57,037 admitted students. The results show that students with higher scores on the PAA obtain higher grades in their first year of study at the institution, even when controlling for high school gpa. In conclusion, the findings of the study reaffirm what has been observed in previous studies and validates the use of the PAA as a predictor of students' academic performance in the Latin American context.

Keywords: PAA suite, validity, college admission, academic achievement.

1. Introducción

En el contexto de la educación superior, es crucial asegurar la precisión, relevancia y pertinencia de la evaluación para mantener la calidad y salvaguardar la equidad educativa. Con esto como norte, College Board LATAM (CB LATAM) cuenta con un sistema de evaluación estandarizado compuesto de varios instrumentos que permiten, por un lado, evaluar los conocimientos, las habilidades y las destrezas de los alumnos de 6 a 12 años de escolaridad (PIENSE I, PIENSE II y PAA) y conocer, explorar y planificar el futuro académico y los intereses ocupacionales de los estudiantes (CEPA). Este sistema se cimienta en el acercamiento de varios marcos teóricos con los cuales se establecen la importancia del saber, saber hacer, saber ser y saber convivir.

Estos instrumentos se revisan y ratifican con un comité internacional de expertos representativo de las instituciones y los países donde se administra este sistema para que continuamente se desarrollen formas de pruebas paralelas que sean justas para los alumnos y que midan las habilidades, las destrezas y los conocimientos necesarios para la población examinada. De esta manera, se garantizan su alineación con los estándares y las mejores prácticas establecidas por College Board, American Educational Research Association (AERA), American Psychological Association (APA) y Council on Measurement in Education (NCME). Este enfoque riguroso permite que la interpretación de los datos, derivados de las evidencias de validez (Kane 1992; 2006, Messick 1986), se utilice adecuadamente para que las instituciones, en el ejercicio pleno de su autonomía, tome decisiones fundamentadas.

En Estados Unidos, por ejemplo, varios estudios han evidenciado la utilidad de pruebas estandarizadas, específicamente el SAT o el ACT, para predecir el desempeño académico, la retención y graduación de alumnos universitarios (College Board 2019; Westrick et al., 2019). Además, las pruebas estandarizadas potencian la admisión de grupos tradicionalmente desventajados (University of California, Standardized Testing Task Force, 2020).

Por su parte, en Puerto Rico y en varias instituciones hispanoamericanas, estudios similares demostraron que tanto para la Prueba de Evaluación y Admisión Universitaria (PEAU), como su sucesora, la PAA, existen correlaciones moderadas entre los componentes y el promedio de primer año de estudios universitarios; lo que demostró la utilidad de estos instrumentos como predictor del índice académico del primer año de estudios universitarios (Colorado-Laguna & Corcino-Marrero, 2014, 2013; Estépar García & Pedrosa Roche, 2013; Hernández Laboy, Rivera Ocasio & Pedrosa Roche, 2017; Rivera Ocasio & Hernández Laboy, 2017; González-González & Estépar García, 2023).

Por tanto, este estudio tuvo como objetivo evaluar la validez predictiva de la versión más reciente de la PAA en el contexto universitario mexicano y determinar su capacidad para predecir el índice académico del primer año. Este análisis no solo

contribuye a la validación continua de las herramientas de evaluación de CB LATAM, sino que también ofrece datos valiosos para mejorar los procesos de selección, admisión, ubicación y apoyo académico en las instituciones educativas.

2. Metodología

La metodología utilizada es cuantitativa con un diseño de investigación correlacional y de enfoque predictivo.

Muestra

La muestra se configuró con los datos académicos de 120,732 aspirantes y posteriormente admitidos entre 2018 y 2020 en una universidad pública con carreras altamente selectivas en el contexto mexicano. Para los análisis correspondientes al índice académico del primer año de estudios universitarios, se depuró la muestra según los siguientes criterios:

- Cursar estudios universitarios a nivel de licenciatura.
- Poseer una puntuación total de la versión más reciente de la PAA.
- Tener un índice académico de la escuela de procedencia.
- Tener un índice académico del primer año universitario.
- Haber tomado entre 30 y 180 créditos.
- La muestra analítica constó de 57,037 alumnos.

Medidas

Las medidas utilizadas fueron las puntuaciones de los componentes principales de la versión más reciente de la PAA (Lectura y Redacción, Matemáticas y la puntuación total), el índice académico de la escuela de procedencia (PEP) y el índice académico al finalizar el primer año de estudios universitarios.

Análisis de datos

Los resultados del estudio se generaron con el programa estadístico R (*R core team*, 2022). Se emplearon estadísticos descriptivos, análisis de correlación de Pearson y regresiones múltiples. Estos procedimientos permitieron responder a los objetivos de esta investigación.

3. Resultados y discusión

A continuación, se presentan los resultados del estudio.

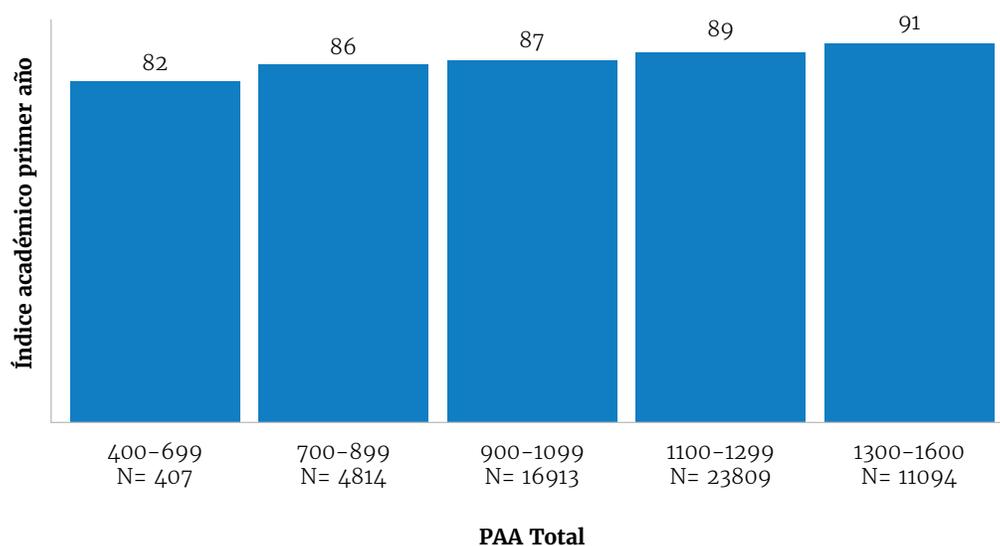
La Tabla 1 presenta las diferencias entre la muestra analítica y la población de aspirantes. Se observa que los promedios en la muestra analítica (alumnos admitidos) son mayores para todas las medidas que en la población de aspirantes. Sin embargo, las desviaciones estándar para todas las medidas son inferiores en la muestra analítica que en la de los aspirantes.

Tabla 1. Estadísticas descriptivas para las medidas de interés

Componentes	Muestra analítica (n = 57 037)				Aspirantes (n = 120 732)			
	M	D.E.	Min	Max	M	D.E.	Min	Max
PEP (0 – 100)	89.67	6.38	60	100	87.83	6.76	60	100
PAA								
Lectura y Redacción	560.08	89.17	200	800	526.05	96.02	200	800
Matemáticas	582.63	101.98	200	800	542.72	107.77	200	800
Puntuación total	1142.71	173.81	435	1577	1068.77	187.90	400	1577
Índice académico								
Primer año	88.53	6.68	52.18	100				

A continuación, la Figura 1 presenta la relación entre la puntuación total en la PAA con el índice académico de primer año.

Figura 1. Promedio del índice académico de primer año por categorías de la puntuación total de la PAA



Las Tablas 2 y 3 presentan los modelos de regresión configurados para predecir el índice académico del primer año. La restricción de rango afecta el valor observado de la correlación múltiple en la muestra analítica, por esta razón, una práctica aceptada es realizar correcciones estadísticas a los resultados de las correlaciones (AERA, APA & NCME, 2014). En situaciones donde se utiliza más de un predictor para la selección, se recomienda aplicar la corrección multivariada de Aitken-Lawley (Sackett & Yang, 2000; Lawley 1943). El presente estudio adopta esta práctica.

A continuación, la Tabla 2 presenta los resultados de los modelos de regresión con los valores de R (correlación simple o múltiple no corregida) y R_c (correlación corregida con la fórmula de Aitken-Lawley) y los resultados de análisis de varianza (ANOVA) para evaluar el modelo global (F , $Sig.$).

Tabla 2. Resultados de las regresiones para predecir el índice académico de primer año

Modelos	R	R_c	F	Sig.
PEP	.43	.46	12 795	< .001
PAA Lectura y Redacción	.28	.32	4864	< .001
PAA Matemáticas	.23	.27	3154	< .001
PAA Total	.28	.32	4782	< .001
PEP, PAA Total	.45	.48	7253	< .001

La Tabla 3 presenta los valores del coeficiente de regresión no estandarizados (B), el coeficiente de regresión estandarizado (β), los intervalos de confianza al 95 % sobre del coeficiente de regresión no estandarizado y los resultados de la prueba t (t , $Sig.$), que evalúan la contribución única del predictor al modelo.

Tabla 3. Coeficientes de los predictores de los modelos de regresión para predecir el índice académico de primer año

Modelo	Componente	B	β	t	Sig.	IC Inferior	IC Superior
PEP	Intercepto	48.319				47.620	49.018
	PEP	0.448	0.428	113	< .001	0.441	0.456
PAA Lectura y Redacción	Intercepto	76.770				76.435	77.105
	PAA Lectura y Redacción	0.021	0.280	70	< .001	0.020	0.022
PAA Matemáticas	Intercepto	79.796				79.487	80.106
	PAA Matemáticas	0.150	0.230	56	< .001	0.014	0.016
PAA Total	Intercepto	76.317				75.967	76.667
	PAA Total	0.010	0.278	69	< .001	0.010	0.011

Modelo	Componente	B	β	t	Sig.	IC Inferior	IC Superior
PEP, PAA Total	Intercepto	46.581				45.885	47.277
	PEP	0.395	0.377	95	< .001	0.387	0.403
	PAA Total	0.005	0.150	37	< .001	0.005	0.006

Se observa que todos los modelos y los predictores son estadísticamente significativos. Particularmente, el modelo que combina los resultados del PEP y la puntuación de la PAA Total ($= 0.48$) ofrece un mayor valor predictivo, pues presenta una correlación múltiple con aumento de 4.3 % en comparación con utilizar solo el modelo con el predictor de PEP.

Conclusiones

Los estudios de validez predictiva ayudan a comprender cómo diferentes variables pueden estar asociadas con un criterio de interés recopilado en una instancia posterior. El presente estudio contribuyó a ampliar el conocimiento de la relación entre las puntuaciones de la versión más reciente de la PAA y el índice de escuela de procedencia (PEP) con desempeño posterior, específicamente con el promedio académico del primer año de estudios universitarios en una universidad pública con carreras altamente selectivas en el contexto mexicano.

Los resultados reafirman lo observado en estudios similares y validan el uso de la PAA como un predictor del desempeño académico de los alumnos en esta institución. La PAA, junto a otros indicadores relevantes, como los promedios de la escuela de procedencia, es una herramienta que las instituciones universitarias pueden utilizar para informar decisiones relacionadas con la selección, admisión y ubicación; para fortalecer y complementar procesos de investigación, y para establecer mecanismos conducentes a la implementación de intervenciones educativas.

Con este estudio, así como otros llevados a cabo para PIENSE, se demuestra cómo el sistema de pruebas de CB-LATAM cuenta con documentación sólida que demuestra cómo los datos que se obtienen de estos instrumentos son un buen indicador para predecir, junto a otros criterios, el nivel de desempeño de los alumnos en períodos importantes de su desarrollo. Es decir, las evidencias de validez y sus respectivos análisis ratifican cómo las puntuaciones de la PAA son un buen predictor de desempeño académico universitario, mientras que los estudios de PIENSE y su respectiva relación entre las puntuaciones, $PIENSE\ I > PIENSE\ II > PAA$, sostienen, con cierto grado de precisión, el desempeño estudiantil en su desarrollo educativo, si no se implementan intervenciones académicas de rigor a tiempo. Así, este estudio debe replicarse en otras instituciones latinoamericanas para proveer a los usuarios un marco de referencia relevante para ponderar adecuadamente los resultados de las pruebas estandarizadas.

Referencias

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*.
- College Board. (2019). *Predictive validity of the SAT for higher education systems and consortia: The value of validating admissions and retention predictors*. <https://r.issu.edu.do/54>
- Colorado-Laguna, R. J., & Corcino-Marrero, L. (2013). Validez de predicción de los criterios de admisión a la Universidad de Puerto Rico. *Pedagogía*, 46(1), 66–92. <https://revistas.upr.edu/index.php/educacion/article/view/16416>
- Colorado Laguna, R. J., & Corcino Marrero, L. (2014). Diferencias en el desempeño académico y en los predictores de éxito universitario por escuela de procedencia. *Pedagogía*, 47(1), 159–191.
- Estépar García, W., & Pedrosa Roche, O. (2013). *Estudio de validez predictiva Universidad Tecnológica de Panamá* [Informe interno]. College Board LATAM.
- González-González, M., & Estépar-García, W. (2023). *Validez de la PAA para predecir el promedio de primer año en la Universidad Nacional Autónoma de Honduras (UNAH): Análisis complementarios de admisión usando las pruebas de conocimientos* [Informe interno]. College Board LATAM.
- Hernández Laboy, J., Rivera Ocasio, R., & Pedrosa Roche, O. (2017). *Estudio de la validez predictiva de los resultados de la PAA en la Universidad Nacional Autónoma de Honduras* [Informe interno]. College Board LATAM.
- Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh, LX/I-Part 1*(1), 28–30.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). American Council on Education.
- Kane, M. T. (1992). An argument-based approach to validation. *Psychological Bulletin*, 112, 527–535.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed), pp. 13–103. Macmillan.
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

- Rivera Ocasio, R., & Hernández Laboy, J. (2017). *Estudio de validez predictiva para la Universidad del Valle de Guatemala* [Informe interno]. College Board LATAM.
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85(1), 112–118.
<https://doi.org/10.1037/0021-9010.85.1.112>
- University of California, Standardized Testing Task Force. (2020). *Report of the UC Academic Council Standardized Testing Task Force (STTF)*. University of California, Academic Senate. <https://r.issu.edu.do/4BC>
- Westrick, P. A., Marini, J. P., Young, L., NG, H. Shmueli, & Shaw, E. (2019). *Validity of the SAT for predicting first-year grades and retention to the second year*. The College Board. <https://r.issu.edu.do/8mT>