

Desempeño de Personas con la Adecuación de Tiempo Adicional en una Prueba Estandarizada

Performance of People with Extra Time for Solving a Standardized Test

Lucrecia Alfaro-Rojas *
Luis Rojas Torres
Universidad de Costa Rica

Se presenta un análisis comparativo del desempeño de un grupo de personas que realizaron una prueba estandarizada con un límite de tiempo uniforme, con el de diferentes subgrupos de personas que tuvieron la posibilidad de utilizar tiempo adicional como adecuación en virtud de una situación de discapacidad o necesidad educativa especial. Se utilizó como muestra la población que aplicó la Prueba de Aptitud Académica de la Universidad de Costa Rica en 2011 y se realizaron comparaciones de los puntajes promedio entre los grupos y subgrupos establecidos, con respecto a los puntajes obtenidos en la prueba. Se encontró que todos los subgrupos que no utilizaron la adecuación de tiempo adicional presentaron diferencias significativas con la población regular; mientras que tres de las cuatro comparaciones de subgrupos que sí utilizaron la adecuación, contra la población regular, presentaron diferencias no significativas o tamaños del efecto pequeños. Únicamente en el subgrupo del área denominada "aprendizaje" se presentó un tamaño del efecto grande.

Descriptor: Adecuaciones, Discapacidad, Medición, Evaluación.

Analysis of the performance of a group of people who took a standardized test, with a uniform time limit, compared to the performance of different subgroups of people who had the possibility of using additional time as an accommodation is presented. Population that applied the Scholastic Aptitude Test (*Prueba de Aptitud Académica*) at the University of Costa Rica in 2011 was used as a sample. Comparisons of mean scores on the Scholastic Aptitude Test were performed by groups and defined subgroups. It was found that all the subgroups who did not use the additional time showed significant differences in comparison to the regular population, while three of the four comparisons of subgroups that did use the additional time showed or not significant differences or small effect sizes. Only the subgroup called "learning" showed a large effect size.

Keywords: Accommodations, Disability, Measurement, Assessment.

*Contacto: lucrecia.alfaro@ucr.ac.cr

1. Revisión de la literatura

La Prueba de Aptitud Académica es un instrumento psicométrico desarrollado por la Universidad de Costa Rica para la medición de habilidades generales de razonamiento en contextos verbales y matemáticos, asociadas al rendimiento académico en el ámbito de la educación superior (Instituto de Investigaciones Psicológicas, 2007). Esta prueba se ha aplicado y depurado por más de 50 años, con el fin de seleccionar estudiantes con aptitudes para la educación superior, y cumple con criterios psicométricos que garantizan su calidad como instrumento actual de selección para el ingreso a dos instituciones universitarias del país.

De acuerdo con la clasificación de Martínez, Hernández y Hernández (2006), se tiene que la Prueba de Aptitud Académica se clasifica como una prueba de altas consecuencias para los examinados, dado que sus resultados son el elemento que define el ingreso a dos de las instituciones de educación superior más importantes del país: la Universidad de Costa Rica y la Universidad de Nacional; por lo que tiene importantes implicaciones para los proyectos de vida de las personas examinadas.

Siguiendo con esta clasificación, esta prueba se describe también como una prueba de papel y lápiz, de respuesta seleccionada; ya que para contestar los ítems las personas examinadas deben elegir una opción dentro de un conjunto de posibles respuestas. Es una prueba de aplicación colectiva, ya que evalúa simultáneamente a una gran cantidad de personas; y de potencia, ya que se aplica en un contexto en el que el tiempo es suficiente para intentar contestar correctamente todos los ítems, que tienen diferentes niveles de dificultad. Finalmente, se considera una prueba referida a normas, ya que las puntuaciones obtenidas por los(as) examinados(as) en la prueba se utilizan como elemento de comparación entre todas las personas que la contestan.

Un concepto clave para la construcción e interpretación de resultados obtenidos mediante la aplicación de una prueba como la Prueba de Aptitud Académica es el de validez, entendida como el grado en que la evidencia y la teoría respaldan las interpretaciones hechas a partir de los puntajes obtenidos en una prueba (AERA, APA y NCME, 2014). La validez es la consideración más importante en el desarrollo y la evaluación de una prueba, y el proceso de validación conlleva acumular evidencia para proveer una base científica sólida a las interpretaciones propuestas de los puntajes.

En general, de acuerdo con Villarreal, Alfaro-Rojas y Brizuela (2015) se deben tomar en cuenta dos aspectos indispensables en cualquier proceso de validación: la validez de las inferencias basadas en los puntajes obtenidos en una prueba y la justificación de los usos de esta. Ambos aspectos no son compensatorios, es decir, poseer evidencia sobre la presencia de uno no compensa las debilidades o carencias en cuanto al otro. Los autores también afirman que la validez puede ser adecuada o no en la medida en que los puntajes obtenidos mediante la correcta aplicación de una prueba permiten fundamentar las inferencias en cuanto a lo que esta evalúa. De esta forma, el proceso de validación sería aquel mediante el cual se recaba, resume y evalúa la evidencia requerida para justificar las inferencias a partir de los resultados en una prueba.

De acuerdo con Villarreal, Alfaro-Rojas y Brizuela (2015), la validación de las inferencias hechas a partir de una prueba requiere iniciar con una propuesta clara sobre las posibles interpretaciones que se pueden realizar a partir de esta, así como sus posibles usos. En este sentido, tanto las interpretaciones como los respectivos usos de una prueba deben

basarse en un marco conceptual claro y explícito, ya que no es conveniente asumir que los potenciales usuarios de esta harán las interpretaciones correctas. Estos autores (2015) también señalan que lo anterior es de vital importancia si se toma en cuenta que la validez es una cuestión de grado y que puede cambiar con el tiempo en función de los avances teóricos, tecnológicos y metodológicos relacionados con el rasgo o característica psicológica (habilidad, actitud, conocimiento, patología, etc.) a medir, por lo que siempre es necesario actualizar la validez de las interpretaciones y usos que se realizan con esta, así como incorporar la evidencia nueva que se vaya generando sobre la validez. El tema de la validez es especialmente importante para la definición de adecuaciones a otorgar a las personas que aplican una prueba, ya que estas no deben, bajo ninguna circunstancia, alterar los resultados obtenidos por los examinados, lo que puede llevar a interpretaciones erradas sobre el nivel de una persona examinada en el constructo medido.

En cuanto a la provisión de adecuaciones para la Prueba de Aptitud Académica, se tiene que desde 1981 se empezaron a documentar los esfuerzos de la institución para garantizar el acceso a la institución a personas con ciertos tipos de discapacidades. Específicamente, estos esfuerzos consistían en la eliminación del requisito de realizar la Prueba de Aptitud Académica para el ingreso de personas con discapacidad visual. Esta medida se mantuvo por un corto periodo de tiempo y, posteriormente, se inició con el proceso de provisión de adecuaciones para personas con otros tipos de discapacidad, brindándose oportunidades de ingreso a cada vez más personas interesadas en aplicar por un puesto como estudiante de esta casa de estudios.

La Prueba de Aptitud Académica es una prueba estandarizada, es decir, que presenta uniformidad en los materiales utilizados (reactivos o folletos de prueba) y en los procedimientos de aplicación (formato de los materiales, tiempo máximo permitido para la emisión de las respuestas, modo de respuesta, entre otros), con el fin de garantizar la precisión en la medición del constructo meta y la oportunidad para que cada persona evaluada muestre su nivel “real” en dicho constructo. La estandarización apunta a la recopilación de información específica que permita hacer inferencias con respecto al nivel de cada persona en el constructo medido y comparar en desempeño entre individuos (Georgia Department of Education, 2008); sin embargo, ante situaciones de discapacidad o necesidades educativas especiales de las personas a quienes se les aplican las pruebas, dicha estandarización debe alterarse, en cierta medida, ante la necesidad de otorgar a estas personas adecuaciones que permitan su acceso a la misma. En estos casos, para garantizar la validez del instrumento utilizado, las adecuaciones ofrecidas no deben, en ningún caso, comprometer la calidad técnica de la medición del constructo, lo que implica que no deben afectar la dificultad de las prueba, ni generar ventajas o desventajas para las personas que las reciben, en comparación con el resto de la población, que realiza la prueba bajo condiciones de estandarización. Solo si se cumplen estas condiciones es posible comparar los puntajes obtenidos y hacer interpretaciones legítimas acerca del nivel de cada persona examinada en el constructo medido (Messick, 1995).

Las adecuaciones en contextos de medición mediante pruebas estandarizadas se definen como cambios en las instrucciones o en las prácticas de medición que reducen el impacto de la discapacidad de una persona en su interacción con el material de la prueba utilizada (Ketterlin-Geller y Johnstone, 2006) y consisten en cambios en el contexto en el que las instrucciones son presentadas o en el que se hace la medición, la provisión de tiempo

adicional para completar una tarea, así como cambios en el método de respuesta, los materiales o el equipo tecnológico o profesional que apoya la interacción con el material de la prueba.

La función primordial de las adecuaciones en pruebas estandarizadas es reducir la varianza irrelevante al constructo relacionada con la condición de discapacidad que pueden presentar las personas evaluadas, sin que se afecte la medición del constructo meta (Georgia Department of Education, 2008). Para esto, las adecuaciones deben contribuir a garantizar la validez de la interpretación de los puntajes derivados de la aplicación. En síntesis, las adecuaciones apuntan a “nivelar el terreno de juego” para que todas las personas examinadas tengan las mismas posibilidades de mostrar su nivel de habilidad (Thurlow y Wiener, 2000; Abedi et al., 2012) sin que se cambien, alteren o reduzcan las expectativas de lo que se mide mediante la prueba aplicada (Georgia Department of Education, 2008; Stone, Cook, Laitusis y Cline, 2010).

La atención brindada a las necesidades de la población con discapacidad o con “necesidades educativas especiales” en el contexto de la aplicación de la Prueba de Aptitud Académica, responde a la evolución en los enfoques, modelos o paradigmas que han moldeado la forma de conceptualizar a esta población (Alfaro-Rojas, 2015), dejando cada vez más de lado las prácticas tendientes a la caridad y la exclusión de la población con discapacidad para dar paso a un enfoque que integra el llamado Modelo Social de la discapacidad, desde el cual se aspira a una participación real y efectiva de esta población en todas las esferas de la vida en la sociedad (Padilla-Muñoz, 2010), lo que incluye el ámbito educativo. Este abordaje sostiene que la exclusión de la población con discapacidad o necesidades educativas especiales obedecen a un entorno que no está diseñado para la diversidad (Strauss y Sales, 2010), y pone de manifiesto que las barreras presentes en el entorno son las que limitan o impiden la participación real de las personas con discapacidad, no así la deficiencia orgánica presente en la persona.

De acuerdo con Alfaro-Rojas (2015), la evolución de estos modelos ha impactado de forma evidente las prácticas de medición en contextos educativos, ya que, en ámbitos como el costarricense, se ha pasado gradualmente de un modelo que descartaba por completo la participación de niños niñas y jóvenes con discapacidad en sistemas educativos regulares, a uno en el que se busca su participación plena y efectiva, lo que incluye la igualdad de oportunidades de acceso a las pruebas estandarizadas que se aplican al finalizar ciertos niveles educativos o para concursar para el ingreso a instituciones de educación superior.

Volviendo a la Prueba de Aptitud Académica, se tiene que la provisión de tiempo adicional es la adecuación más común para personas con discapacidad o necesidades educativas especiales que la realizan en la actualidad. El tiempo adicional que se concede puede ser de 30 o 60 minutos, dependiendo de la funcionalidad de la persona en cuanto al desempeño en el tiempo, la cual debe ser demostrada previamente por el examinado o la examinada mediante documentación aportada o una entrevista realizada durante el proceso de asignación de adecuaciones.

A pesar de la frecuencia con se otorga la adecuación de tiempo adicional para la Prueba de Aptitud Académica, no se han realizado hasta el momento análisis que comparen los resultados de las personas que reciben esta adecuación con los obtenidos por el resto de la población. Es debido a esto que se ha propuesto la realización de un análisis de corte estadístico que busca comparar el desempeño (entendido como el porcentaje de

respuestas correctas en la prueba) de tres grupos de personas que realizaron la Prueba de Aptitud Académica en el 2011: personas que no recibieron ninguna adecuación; personas a quienes se les aprobó la provisión de tiempo adicional y que lo utilizaron de forma total o parcial; y personas a quienes se les aprobó la provisión de tiempo adicional pero que no lo utilizaron. Estas comparaciones se realizaron tanto para la totalidad de la población que recibió la adecuación, como para las subpoblaciones correspondientes a cada área de clasificación de las personas solicitantes.

En la literatura es posible encontrar resultados de investigaciones recientes relativas a los resultados de la provisión de tiempo adicional en el desempeño en pruebas estandarizadas para personas con diferentes discapacidades. Algunas de estas arrojan argumentos que apoyan la provisión de tiempo adicional como una adecuación que “nivela el terreno de juego” para personas con y sin discapacidad (Cahalan-Laitusis, King, Cline y Bridgeman, 2006; Lee, Osborne y Carpenter, 2010; Lesaux, Pearson y Siegel, 2006), convirtiendo a esta adecuación en una herramienta que equipara las oportunidades de personas con discapacidad o necesidades educativas especiales con respecto al del resto de la población.

También se puede ubicar estudios que muestran que la adecuación de tiempo adicional puede ser una fuente de clara ventaja para las personas con discapacidad que las reciban, ya que hace que las personas con discapacidad tengan puntajes más altos que las personas sin discapacidad que no han recibido esta adecuación (Lewandowski, Cohen y Lovett, 2013), así como otros que sostienen que la provisión de tiempo adicional no tiene ningún efecto en los puntajes obtenidos en la prueba aplicada (Cohen, Gregg y Deng, 2005).

Si bien, los resultados de las investigaciones consultadas son contradictorios en cuanto a los resultados que la adecuación de tiempo adicional tiene para la población con discapacidad, en la práctica se tiende a apostar por su papel como garante de la equidad de oportunidades en el acceso a pruebas estandarizadas. Estos resultados tampoco eximen a las personas e instituciones que construyen y validan pruebas estandarizadas, como es el caso de la Prueba de Aptitud Académica, de la necesidad de buscar indicadores que constaten o refuten la pertinencia de esta adecuación. Es por esto que se ha propuesto investigar acerca de las diferencias entre grupos, asociadas al uso del tiempo adicional para la Prueba de Aptitud Académica, con la finalidad de indagar acerca de la pertinencia que la provisión y uso de tiempo adicional tiene para poblaciones específicas, así como plantear recomendaciones para una adecuada atención a la población solicitante de adecuaciones para esta prueba.

2. Metodología

Muestra

Para este análisis se trabajó con la población que aplicó la Prueba de Aptitud Académica en el 2011 ($N=38241$) la cual se dividió en dos grandes grupos: personas que no recibieron ninguna adecuación de tiempo adicional para la prueba ($N_R=37170$), grupo denominado para efectos de este documento como “población regular”, y personas a quienes se les aprobaron adecuaciones ($N_A=1071$); asimismo, este último grupo se dividió en subgrupos de acuerdo con el área de clasificación de la persona (aprendizaje, déficit atencional, visual, auditiva, motora, emocional, sistémica y múltiple) y estos

subgrupos se dividieron, a su vez, de acuerdo con la cantidad de tiempo adicional aprobado (30 o 60 minutos). Lo que dio como resultado los subgrupos que se muestran en la tabla 1.

Tabla 1. Subgrupos en los que se dividió la población con adecuaciones

ÁREA	MINUTOS ADICIONALES	SIMBOLOGÍA	N
Aprendizaje	30	Ap30	688
	60	Ap60	58
Déficit atencional	30	DA30	137
	60	DA60	4
Emocional	30	Em30	18
	60	Em60	8
Visual	30	Vis30	19
	60	Vis60	49
Auditiva	30	Aud30	4
	60	Aud60	20
Motora	30	Mot30	24
	60	Mot60	17
Sistémica	30	Sis30	13
	60	Sis60	7
Múltiple	30	Mul30	3
	60	Mul60	2

Fuente: Elaboración propia.

Con respecto a la cantidad de examinados por área que recibieron la adecuación de tiempo adicional 30 minutos, se tiene que el grupo de Aprendizaje presentó el mayor tamaño, seguido por el área de Déficit Atencional. Por el contrario, los tamaños de muestra de los subgrupos con la adecuación tiempo adicional 60 minutos fueron más pequeños que los observados en el caso anterior, de hecho cuatro subgrupos no superaron las diez personas.

Método

Una vez definidos los grupos descritos en el apartado anterior, se propuso la realización de dos tipos de análisis dependiendo del tamaño de los grupos. Así, para cada uno de los subgrupos compuestos por al menos 50 personas, se realizó una comparación del promedio en las notas de la Prueba de Aptitud Académica de las personas de cada subgrupo que utilizaron el tiempo adicional (denotado como 1) con el promedio obtenido por las personas que no utilizaron el tiempo adicional (denotado como 0), además de comparar cada uno de estos grupos con el promedio de la población regular (denotado como R). Esta comparación se llevó a cabo de forma gráfica e inferencial.

Los análisis estadísticos realizados para estos subgrupos consistieron en comparación de las medias en el puntaje de la Prueba de Aptitud Académica, por medio del análisis de varianza (ANOVA), esto indica si existen diferencias significativas entre las medias obtenidas por los grupos. En los casos en los que se encontraron diferencias significativas, se realizaron pruebas de Tukey; las cuales indican entre que pares de grupos ocurren las diferencias estadísticamente distintas de cero. Además, como análisis extra al test de Tukey, se utilizó el estadístico “d de Cohen”, para determinar el tamaño del efecto.

En el caso de los subgrupos con tamaños menores a 50 pero mayores a 10, se realizó la misma comparación, pero únicamente de manera observacional, ya que se consideró que los tamaños de muestra eran muy bajos, por lo cual era preferible no realizar análisis

inferenciales. Los subgrupos con menos de 10 personas, se excluyeron de los análisis, ya que se estimó que no contienen un número de observaciones adecuado para ilustrar el comportamiento de los puntajes. Los subgrupos eliminados por este criterio fueron: DA60, Aud30, Em60, Sis30, Sis60, Mul30 y Mul60.

En la tabla 2 se muestra la distribución de las personas según los grupos finales con los que se realizaron los análisis descriptos.

Tabla 2. Distribución de la población en los grupos de análisis finales

GRUPO	N		PROMEDIO	
	No Usan TA	Usan TA	No usan TA	Usan TA
Ap30	337	351	30,12	39,8
DA30	60	77	34,66	44,15
Ap60	35	23	16,5	20,82
Vis60	24	25	32,83	38,32
Em30*	8	10	34,73	32,71
Vis30*	7	12	38,58	37,26
Mot30*	8	16	54,14	46,03
Aud60*	6	14	21,96	21,33
Mot60*	9	8	31,44	38,89
Regular	54.854		42,42	

Fuente: Elaboración propia.

3. Resultados

Para iniciar la sección de resultados, se presenta mediante la figura 1 una descripción gráfica de los promedios de cada subgrupo de adecuación. En esta gráfica se observa que para los subgrupos Ap30, DA30, Ap60 y vis60, el promedio obtenido por los grupos que hicieron uso del tiempo adicional fue superior al promedio global del subgrupo y este, a su vez, fue superior al del grupo que no utilizó la adecuación.

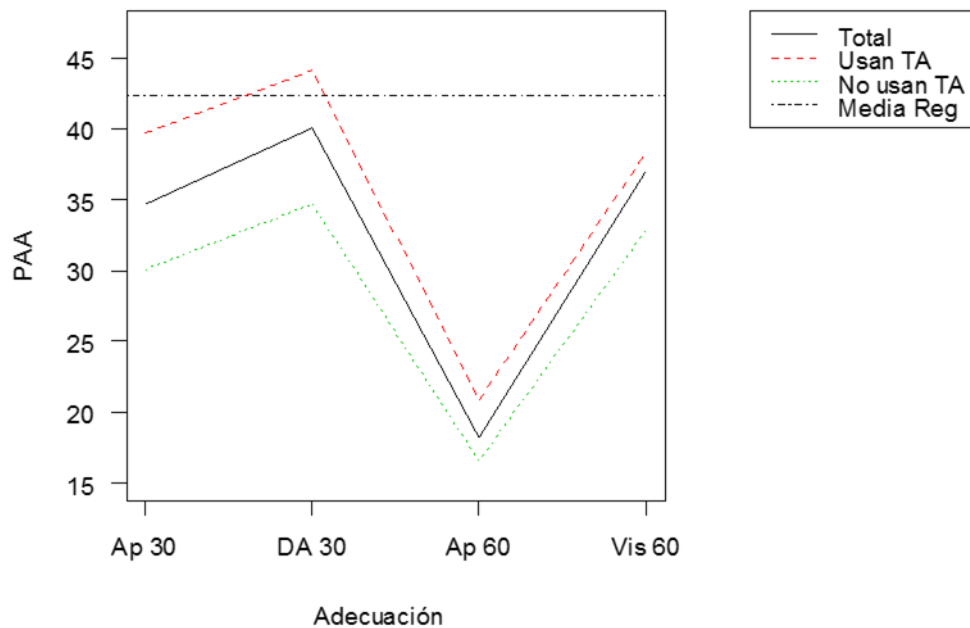


Figura 1. Promedios en la prueba de aptitud académica en los subgrupos con tamaños de muestra grandes, según uso de la adecuación de tiempo

Fuente: Elaboración propia.

Por otra parte, en todos los subgrupos, el promedio de los examinados que utilizaron el tiempo adicional fue muy cercano al promedio obtenido por la población regular; únicamente los examinados del grupo Visual 60 se desvían de esta tendencia, dado que su promedio es considerablemente inferior al obtenido por la población regular.

En la tabla 3 se presentan los resultados de los ANOVAs asociados a las comparaciones anteriores. En los cuatro análisis, se puede observar que las diferencias son significativas al 5%.

Tabla 3. Resultados del ANOVA sobre la Prueba de Aptitud Académica entre la población regular, la que usó el TA y la que no usó el TA, según área de adecuación

GRUPO	DIFERENCIA DE MEDIAS			P ANOVA	PRUEBA DE TUKEY (D DE COHEN)		
	R - 1	R - 0	1 - 0		R - 1	R - 0	1 - 0
Ap 30	2,63	12,30	9,68	0,00	0,02(0,15)	0,00(0,69)	0,00(0,54)
DA 30	-1,73	7,76	9,49	0,00	0,67(-0,10)	0,00(0,43)	0,01(0,53)
Ap 60	21,60	25,92	4,32	0,00	0,00(1,21)	0,00(1,45)	0,64(0,24)
Vis 60	4,09	9,59	5,49	0,02	0,48(0,23)	0,02(0,54)	0,52(0,31)

Fuente: Elaboración propia.

Al realizar la prueba de Tukey para determinar qué pares de contrastes tienen significancia estadística, se obtuvo que en el subgrupo Ap30 las tres comparaciones posibles resultaron estadísticamente significativas, no obstante, el tamaño del efecto (d de Cohen) es pequeño para la comparación entre la población regular y la que utilizó el tiempo adicional, mientras que en los otros dos contrastes el tamaño del efecto se ubica entre mediano y grande (Cohen, 1977).

En el caso del grupo DA30, se obtuvo que la diferencia significativa global fue inducida por las comparaciones contra el grupo que no utiliza el tiempo adicional. El único contraste que no resultó significativo fue el de la población regular con el que utiliza el tiempo adicional.

Para el subgrupo Ap60, la prueba de Tukey indicó que las diferencias significativas se obtienen en las comparaciones con la población regular, la cual alcanza promedios significativamente más altos que los otros dos grupos, y cuyo tamaño del efecto es “muy grande” (Rosenthal, 1996). En esta comparación no se presentan diferencias estadísticamente distintas de cero entre los estudiantes que hacen uso de la adecuación, y los que no.

Para el subgrupo Vis60, al igual que en Ap30 y DA30, se obtuvo que el tamaño del efecto entre la población regular y la que usa la adecuación fue es pequeño. La significancia del ANOVA se debe a la diferencia entre la población regular y la que no hace uso de la adecuación; se puede observar que este contraste es significativo en las cuatro pruebas de Tukey.

Finalmente, en la figura 2 se muestra un gráfico con datos de los subgrupos Em30, Vis30, Mot30, Aud60 y Mot60. Como ya mencionó, este análisis es meramente descriptivo debido a que estos subgrupos cuentan con tamaños de muestra muy pequeños (entre 17 y 24 personas).

En los subgrupos Em30, Vis30 y Mot60, se observa que los promedios de los estudiantes que utilizan la adecuación se acercan al obtenido por la población regular, además, en estas comparaciones, el promedio regular supera al de la población que utiliza la adecuación, y este supera al del grupo que no utiliza la adecuación.

Por otra parte, en la comparación basada en Mot30 se observa que el grupo que utiliza el tiempo adicional presenta un promedio superior en más de 10 puntos al obtenido por la población regular. Además, la población que no utiliza la adecuación obtiene un promedio similar al de la población regular.

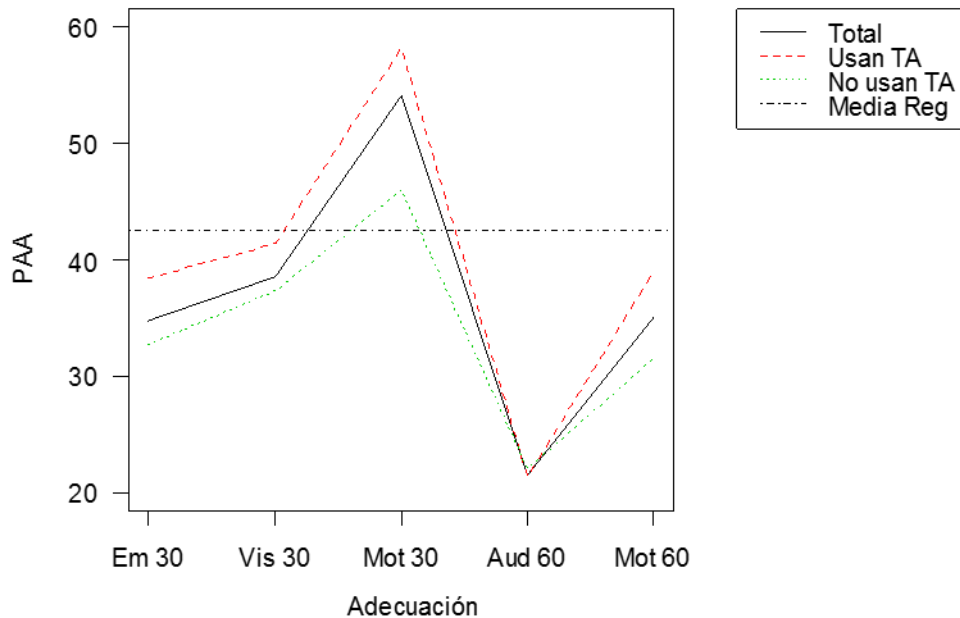


Figura 2. Promedios en la prueba de aptitud académica del subgrupo con tamaños de muestra pequeños, según uso de la adecuación de tiempo adicional
Fuente: Elaboración propia.

En el caso de Aud30, se observa que no hay diferencias entre la población que utiliza o no, la adecuación. También se puede apreciar que la población que tiene acceso a esta adecuación presenta un promedio considerablemente más bajo al alcanzado por la población regular.

4. Conclusiones

4.1. Conclusiones inferenciales

Según los análisis realizados, fue posible constatar que las puntuaciones promedio obtenidas en la Prueba de Aptitud Académica por parte de la población regular fueron significativamente superiores a las de las personas que no hicieron uso del tiempo adicional en los subgrupos de Aprendizaje y Déficit Atencional, ambos con la adecuación de tiempo adicional de 30 minutos, y en los subgrupos de Aprendizaje y Visual con tiempo adicional de 60 minutos.

Para los subgrupos Aprendizaje y DA con 30 minutos adicionales, se evidencia que los examinados y las examinados que hacen uso de la adecuación presentan mejores puntajes, y se puede observar que para estos, así como para el subgrupo de Visual con tiempo adicional de 60 minutos, el uso de la adecuación muestra una tendencia a “nivelar el terreno de juego” para esta población con respecto a la población regular. Esto coincide a cabalidad con lo propuesto por la literatura relevante al tema (Abedi et al.,

2012; Georgia Department of Education, 2008; Stone, Cook, Laitusis y Cline, 2010; Thurlow y Wiener, 2000), donde se señala que la función primordial de las adecuaciones en pruebas estandarizadas es reducir la varianza irrelevante al constructo causada por la discapacidad, sin que se afecte la medición del constructo meta y sin que se cambien, alteren o reduzcan las expectativas de lo que se mide mediante la misma.

En el caso específico del subgrupo de Aprendizaje con tiempo adicional de 60 minutos, se tiene que el uso o no de la adecuación no presenta ninguna diferencia en los promedios obtenidos por las personas en la prueba, lo que puede explicarse por el hecho de que las personas clasificadas dentro de esta área y que reciben la adecuación de mayor tiempo adicional, corresponden a personas con problemas de aprendizaje severos o, incluso, con diagnóstico de retraso mental, de quienes es esperable que tengan un nivel bajo en la variable medida por la Prueba de Aptitud Académica (habilidades generales de razonamiento). Para estas personas, el uso de tiempo adicional no haría diferencia alguna en su nivel de ejecución, sin que esto implique un problema de validez de las interpretaciones hechas a partir de los resultados obtenidos.

4.2. Conclusiones observacionales

Para el caso de los subgrupos con menos de 50 pero más de 10 personas (Emocional y Visual y Motora con 30 minutos adicionales, así como Auditiva y Motora con 60 minutos adicionales), se realizaron algunos hallazgos que, a pesar de ser meramente observacionales, arrojan información descriptiva importante sobre la posible utilidad de la adecuación de tiempo adicional para los mismos.

En primer lugar, se encontró que para los subgrupos Emocional y Visual, con 30 minutos adicionales, así como el de Motora con 60 minutos adicionales el uso de la adecuación parece nivelar el terreno de juego para estas personas en comparación con la población regular, lo que sugiere que la adecuación reduce la varianza irrelevante al constructo en estos casos específicos.

Para el caso del subgrupo de Motora con 30 minutos adicionales, se observa que el grupo de personas que no hicieron uso del tiempo adicional tienen, en promedio, un desempeño igual al de la población regular, mientras que las personas que utilizaron el tiempo adicional tienen un puntaje promedio superior al de esta población.

Este resultado indica que es necesario revisar el criterio para la asignación de tiempo adicional durante la aplicación de la prueba, para personas del área motora que requieren solamente accesibilidad al espacio físico o mobiliario adaptado, ya que para estas no se justificaría la asignación de la adecuación de tiempo adicional si no presentan alguna alteración cognitiva o problemas motores en los miembros superiores. Es probable que la asignación de tiempo adicional a personas de esta área que no lo necesiten, explique los resultados superiores del subgrupo de Motora que sí hizo uso de la adecuación de tiempo adicional 30 minutos, en comparación con los de la población regular.

Finalmente, se tiene que para el subgrupo del área Auditiva con 60 minutos adicionales, el hacer uso o no del tiempo adicional, no hace ninguna diferencia con respecto a su desempeño, que en ambos casos es inferior al de la población regular. Esto puede deberse a que las dificultades específicas con el español oral y escrito en personas sordas, las cuales llevan a problemas para la evaluación y medición del funcionamiento intelectual mediante el uso de tests y mediciones tradicionales de papel y lápiz (Marschark, 2006)

como es el caso de la Prueba de Aptitud Académica, lo que, al igual que con otras poblaciones, hace que el tiempo adicional no haga una diferencia en su desempeño.

Finalmente, queda por explorar la razón por la cual algunas personas no hacen uso del tiempo adicional, a pesar de contar con esta adecuación. Una posible explicación, que ya se mencionó en este apartado, es que las personas con gran compromiso cognitivo no utilicen el tiempo adicional debido a que no profundizan en el análisis de cada uno de los ítems presentados, por lo que responder a ellos les tome poco tiempo. Otra posible explicación es que estas mismas personas decidan abandonar la prueba en corto tiempo por fatiga, o al percatarse de su falta de habilidad para la resolución de los ítems. Sin embargo, es necesario realizar nuevos estudios de corte cualitativo que permitan dilucidar la situación particular de estas personas y, así, proponer mejores estrategias para atender a esta población sin que la medición del constructo mediante la Prueba de Aptitud Académica se vea afectado.

Referencias

- Abedi, J., Bayley, R., Ewers, N., Mundhenk, K., Leon, S., Kao, J. y Herman, J. (2012). Accessible reading assessments for students with disabilities. *International Journal of Disability, Development and Education*, 59(1), 81–95. doi:10.1037/e642002011-001
- Alba, C. y Zubillaga, A. (2012). La utilización de las TICs en la actividad académica de los estudiantes universitarios con discapacidad. *Revista Complutense de Educación*, 23(1), 23-50. doi:10.5209/rev_rced.2012.v23.n1.39100
- Alfaro-Rojas, L. (2015). Nuevos enfoques en materia de medición y evaluación con un enfoque inclusivo: el caso de estudiantes con discapacidad y necesidades educativas especiales. En CIEM (Orgs.), *Educación e investigación. Modelos Educativos e innovaciones curriculares* (pp. 26-31). México: Horson Ediciones Escolares.
- Ballesteros, C. y Vega, M. (2001) Estrategias que construyen y aprovechan las personas con discapacidad física para incorporarse a procesos productivos (Tesis de Licenciatura). Universidad de Costa Rica. San José, Costa Rica.
- Cawthon, S. W., Winton, S. M., Garberoglio, C. L. y Gobble, M. E. (2011). The effects of American sign language as an assessment accommodation for students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, 16(2), 198-211. doi:10.1093/deafed/enq053
- Cohen, A., Gregg, N. y Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research & Practice*, 20(4), 225-233. doi:10.1111/j.1540-5826.2005.00138.x
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Nueva York: Academic Press.
- Georgia Department of Education. (2008). *Accommodations manual: a guide to selecting, administering, and evaluating the use of test administration accommodations for students with disabilities*. Georgia: Georgia Department of Education.
- Instituto de Investigaciones Psicológicas. (2007). *La prueba de aptitud académica*. San José: Lil.
- Ketterlin-Geller, L. R. y Johnstone, C. (2006). Accommodations and universal design: supporting access to assessments in higher education. *Journal of Postsecondary Education and Disability*, 19(2), 163-172.

- Lee, K., Osborne, R. y Carpenter, D. (2010). Testing accommodations for university students with ad/hd: computerized vs. paper-pencil/regular vs. extended time. *Journal of Educational Computing Research*, 42(4), 443-458. doi:10.2190/ec.42.4.e
- Lesaux, N. K., Pearson, M. R. y Siegel, L. S. (2006). The effects of timed and untimed testing conditions on the reading comprehension performance of adults with reading disabilities. *Reading and Writing*, 19, 21-48. doi:10.1007/s11145-005-4714-5
- Lewandowski, L., Cohen, J. y Lovett, B. (2013). Effects of extended time allotments on reading comprehension performance of college students with and without learning disabilities. *Journal of Psychoeducational Assessment*, 31(3), 326-336. doi:10.1177/0734282912462693
- Lovett, B., Lewandowski, L., Berger, C. y Gathie, R. (2010). Effects of response mode and time allotment on college students writing. *Journal of College Reading and Learning*, 40(2), 64-79. doi:10.1080/10790195.2010.10850331
- Mainieri, A. (2010). *Reconstrucción teórica e histórica de los fundamentos de la PAA-UCR. Informe de investigación*. Costa Rica: Instituto de Investigaciones Psicológicas.
- Marschark, M. (2006). Intellectual functioning of deaf adults and children: answers and questions. *European Journal of Cognitive Psychology*, 18(1), 70-89. doi:10.1080/09541440500216028
- Martínez, M., Hernández, M. y Hernández, M. (2006). *Psicometría*. Madrid: Alianza Editorial.
- Montero, E. y Villalobos, J. (2004). *Estudio comparativo del promedio de admisión a la Universidad de Costa Rica*. San José: SIEDIN.
- Montoya, M. (2003) *Evolución histórica de los modelos en discapacidad. Área de capacitación*. Heredia: Consejo Nacional de Rehabilitación y Educación Especial.
- Moeschberger, M. y Klein, J. (2003). *Survival analysis: techniques for censored and truncated data*. Nueva York: Springer.
- Messick, S. (1995). Validity of psychological assessment: validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749. doi:10.1037/0003-066x.50.9.741
- Puig de la Bellacasa, R. (1990). *Concepciones, paradigmas y evolución de las mentalidades sobre la discapacidad*. Madrid: Real Patronato de Prevención y Atención a Personas con Minusvalía.
- Rojas, L. (2013). Validez predictiva de los componentes del promedio de admisión a la Universidad de Costa Rica utilizando el género y el tipo de sexo como variables control. *Actualidades Investigativas en Educación*, 13(1), 24-48. doi:10.15517/aie.v13i1.11707
- Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, 21(4) 37-59. doi:10.1300/j079v21n04_02
- Stone, E., Cook, L., Laitusis, C. C. y Cline, F. (2010). Using differential item functioning to investigate the impact of testing accommodations on an english-language arts assessment for students who are blind or visually impaired. *Applied Measurement in Education*, 23, 132-152. doi:10.1080/08957341003673773
- Thurlow, M. y Wiener, D. (2000). *Non-approved accommodations: recommendations for use and reporting*. Minneapolis, MN: University of Minnesota, National Center for Educational Outcomes.

Agradecimientos

Al Programa Permanente Prueba de Aptitud Académica (PPPAA) por facilitarnos los datos utilizados en esta investigación.

Breve CV de los autores

Lucrecia Alfaro-Rojas

Licenciada en Psicología. Investigadora del Instituto de Investigaciones en Educación de la Universidad de Costa Rica. Encargada de la provisión de adecuaciones para la Prueba de Habilidades Cuantitativas. Ha publicados artículos y presentado ponencias relacionados con los temas de diversidad, equidad y educación. Email: lucrecia.alfaro@ucr.ac.cr

Luis Rojas Torres

Licenciado en Enseñanza de la Matemática y Magister en Estadística. Labora para la Universidad de Costa Rica como docente de la Escuela de Matemática, Investigador principal del proyecto Prueba de habilidades cuantitativas del instituto de investigaciones psicológicas y Docente investigador del programa permanente Prueba de Aptitud Académica. Email: luismiguel.rojas@ucr.ac.cr