

Análisis de sobrevivencia para la estimación de tiempo adicional como adecuación para la aplicación en una prueba estandarizada

Survival Analysis for the Estimation of Additional Time as Accommodation on the Application of a Standardized Test

Luis Rojas Torres y Lucrecia Alfaro-Rojas

Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica

Resumen

Se presenta un método para la predicción del tiempo necesario para resolver pruebas estandarizadas por parte de una población que recibió adecuaciones, según variables relacionadas con el área de clasificación y las adecuaciones otorgadas. Se utilizó como muestra la población que rindió la Prueba de Aptitud Académica (PAA) de la Universidad de Costa Rica con adecuaciones entre el 2009 y el 2011. Se estimaron modelos de sobrevivencia paramétricos donde las variables fueron todas las adecuaciones ofrecidas en cada área, las cuales fueron reducidas mediante un proceso de *stepwise*. Se obtuvieron estimaciones de tiempo requerido en las áreas *aprendizaje y déficit atencional*, así como para ciertos subgrupos de las otras áreas.

Palabras clave: adecuaciones, discapacidad, medición, evaluación

Correspondencia a:

Luis Rojas Torres
Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica
Apdo: 11501-2060, San Pedro, San José, Costa Rica
Correo electrónico: luismiguel.rojas@ucr.ac.cr

© 2014 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi:10.7764/PEL.51.1.2014.23

Abstract

This study presents a method for predicting the time required to complete standardized tests for populations that received accommodations, time is calculated according to variables related to the classification area and the approved accommodations. The population that took the Academic Aptitude Test (PAA) of the University of Costa Rica with accommodations between 2009 and 2011 was used as a sample. Parametric survival models, where the variables were the set of accommodations offered in each area, were estimated, and these variables were reduced by a stepwise process. Estimations of the time required for people classified into *learning* and *ADHD* areas, as well as certain subsets of the other areas, were obtained.

Keywords: accommodations, disability, measurement, assessment

La Prueba de Aptitud Académica (PAA) de la Universidad de Costa Rica (UCR) es un instrumento psicométrico que busca medir habilidades generales de razonamiento asociadas al rendimiento académico en el ámbito de la educación superior. Los ítems de la PAA miden aspectos relacionados con destrezas de inducción, deducción, categorización, relaciones de causalidad, análisis, comprensión de lectura, análisis cuantitativo, pensamiento analógico e interpretación, entre otros (Instituto de Investigaciones Psicológicas, 2007).

Los orígenes de esta prueba se remontan al año 1957, cuando se identificó la necesidad de contar con un indicador de aptitud de los examinados aspirantes a ser admitidos en la UCR, dado el alto grado de deserción que se presentaba por parte del estudiantado de aquella época. Esta prueba, que se ha aplicado anualmente desde 1970 como instrumento de selección, tomó como base, en sus inicios, el modelo del *Scholastic Aptitude Test* (SAT) (Mainieri, 2010); sin embargo, ha evolucionado de forma independiente a este.

La PAA mide estas habilidades generales de razonamiento en contextos verbales y matemáticos. En su componente matemático, utiliza un solo tipo de ítem, mientras que en el componente verbal se presentan dos tipos de ítems: comprensión de lectura y completar oraciones. Se pretende que en cada uno de estos, la población examinada aplique sus habilidades generales para la resolución de los ítems planteados y que necesite para esto conocimientos muy básicos (a nivel de séptimo año de Educación General Básica), por lo que este tipo de medición no comparte las características de una evaluación en un contexto de enseñanza y aprendizaje (Montero y Villalobos, 2004).

En la prueba se utiliza el tipo de ítem denominado *de selección única*. Este consiste en proponer un enunciado base referente a una situación y varias opciones de respuestas entre las cuales solamente una es correcta (Montero y Villalobos, 2004). Estos ítems han sido desarrollados por un equipo de especialistas y depurados mediante la aplicación de criterios científicos para garantizar un grado adecuado de confiabilidad y validez de la prueba (Instituto de Investigaciones Psicológicas, 2007).

El puntaje en la PAA corresponde al porcentaje de respuestas correctas (escala 0–100), que resulta de la calificación de cada ítem. Este puntaje, en combinación con el promedio de notas obtenidas por la persona en educación diversificada (correspondiente a secundaria), permite calcular el llamado *promedio de admisión*, que permite ubicar a las personas en una posición de acuerdo con su desempeño y, según este indicador, ubicarlas en categorías de *elegible* o *no elegible* para su ingreso a la universidad (Montero y Villalobos, 2004).

El proceso de aplicación de la PAA de la Universidad de Costa Rica debe responder a la legislación vigente en materia de discapacidad y, por tanto, brindar a las personas que así lo requieran adecuaciones para su aplicación, en el marco del principio de igualdad y de equiparación de oportunidades. Esto responde específicamente a lo establecido por la Ley N°7600 de Igualdad de Oportunidades para las Personas con Discapacidad (Asamblea Legislativa de la República de Costa Rica, 1996), que establece el deber de las instituciones educativas de realizar las adaptaciones y brindar los servicios de apoyo requeridos para garantizar que las personas con discapacidad puedan ejercer el derecho a la educación. Esta ley remarca, además, la obligatoriedad del Estado de “eliminar las acciones y disposiciones que, directa o

indirectamente, promueven la discriminación o impiden a las personas con discapacidad tener acceso a los programas y servicios” (Artículo 4 inciso c), así como de garantizar el acceso oportuno a la educación independientemente de la discapacidad (Artículo 14).

Otro instrumento jurídico al que responde la asignación de adecuaciones en el contexto de la aplicación de la PAA es la Convención sobre los Derechos de las Personas con Discapacidad, ratificada en Costa Rica como Ley N° 8661 (Asamblea Legislativa de la República de Costa Rica, 2008), que establece el derecho a la igualdad y no discriminación de las personas con discapacidad (Artículo 5) y el derecho a la educación (Artículo 24). Además, define la realización de ajustes razonables que favorezcan la inclusión efectiva de las personas con discapacidad en todos los ámbitos de la vida social.

Breve referente teórico

La estandarización en el contexto de la medición se refiere a la adherencia a un proceso de administración uniforme, en cuanto a procedimientos y condiciones, durante la aplicación de una prueba. Esta es necesaria para recopilar información comparable acerca de las personas evaluadas (Georgia Department of Education, 2008). Sin embargo, la estandarización puede, en algunas ocasiones, intervenir en la correcta medición del nivel de una persona en el constructo de interés debido a que su formato no es accesible a personas con condiciones determinadas. Por ejemplo, una persona con baja visión que no puede acceder adecuadamente al texto de los ítems en el tamaño de letra ofrecido para la mayoría de la población tendría un bajo desempeño que no tendría que ver con sus habilidades de razonamiento.

Las causas del bajo desempeño en pruebas estandarizadas por parte de personas con discapacidad han sido objeto de cuestionamiento y discusión (Abedi et al., 2012; Thurlow & Wiener, 2000). Desde un enfoque tradicional, se partiría de que el desempeño estaría asociado exclusivamente a las características que, a nivel orgánico, presenta la persona con discapacidad. Sin embargo, desde un enfoque más actual se consideraría que este desempeño también está asociado a la falta de oportunidades educativas previas de estas personas, así como a las barreras relacionadas a la obligatoriedad de acceder a una prueba en un formato de aplicación estandarizado (Cawthon, Winton, Garberoglio, & Gobble, 2011). Esto último se remarca cuando las características propias de la discapacidad que, como en el ejemplo citado, no necesariamente tienen que ver con el constructo a medir y pueden llevar a interpretaciones erróneas sobre el nivel de habilidad de la persona evaluada (Cawthon et al., 2011).

Para garantizar la validez de un instrumento psicométrico, se debe asegurar, entre otras cosas, que en caso de brindarse adecuaciones, ninguna de estas comprometa la calidad técnica de la medición del constructo medido, es decir, que las mismas no incidan en la dificultad de la prueba, de manera que no se generen ventajas ni desventajas para las personas que reciben adecuaciones en comparación con el resto de la población. Esta condición permite que los puntajes de la población examinada sean comparables y que se puedan hacer interpretaciones legítimas acerca del nivel de cada persona examinada en el atributo que se desea medir (Messick, 1995). En síntesis, esto implica que si a un examinado o examinada se le aplica una PAA a la que se le han eliminado contenidos o que tiene un nivel inferior de dificultad, esta persona obtendría una puntuación que no reflejaría su habilidad general de razonamiento, por lo que el promedio de admisión que obtendría no permitiría predecir adecuadamente sus probabilidades de éxito académico a nivel universitario.

Tal como se ha explicado en los párrafos anteriores, la atención a la diversidad es un punto de interés relevante en la actualidad y, como afirman Alba y Zubillaga (2012), este se ha visto incrementado en los últimos 30 años dado el proceso gradual de reconocimiento de la diversidad como característica constitutiva básica del entramado social, lo que ha impactado, entre otros, al sector educativo, considerado clave para lograr la autodeterminación e independencia de las personas con discapacidad, así como para el cambio en la mentalidad de las personas sin discapacidad hacia este tema.

El proceso de admisión a la UCR se ha visto impactado por este nuevo modelo y ha ofrecido, desde 1980, la posibilidad de que las personas con discapacidad o necesidades educativas especiales cuenten con adecuaciones que les permitan acceder a la PAA (Universidad de Costa Rica, 1980). Sin embargo, no se cuenta hasta el momento con investigación relativa a la idoneidad de las adecuaciones asignadas en las condiciones de aplicación de la PAA o al desempeño de las personas a quienes se les otorgan, lo que deja

un vacío importante en materia de criterios para una apropiada asignación de las mismas y en cuanto al conocimiento de su efecto en las evidencias de validez para interpretación de los puntajes obtenidos.

Debido a lo expuesto en el párrafo anterior, se ha propuesto la realización de un análisis de corte estadístico que tiene como objetivo la predicción del tiempo requerido para la resolución de la PAA, de acuerdo con el área de clasificación del o de la aspirante¹ y de algunas adecuaciones aprobadas para cada persona; esto, debido a que el tiempo adicional para la aplicación es la adecuación más frecuente concedida en el contexto de la PAA. El tiempo adicional que se concede puede ser de 30 o 60 minutos, dependiendo de la funcionalidad que el examinado o la examinada demuestre mediante la documentación aportada o durante una evaluación realizada por la entidad universitaria pertinente. La cantidad de tiempo adicional que se ofrece a los examinados se ha determinado mediante estimaciones previas relativas al tiempo de desempeño en el tiempo de la población que ha recibido adecuaciones.

Es posible encontrar investigaciones recientes relativas al otorgamiento de la adecuación de tiempo adicional y su efecto en el desempeño en la resolución de problemas. Sin embargo, sus resultados son contradictorios en cuanto al efecto “equiparador” que dicha adecuación tiene para la población con discapacidad en relación con la población sin esta condición. En algunas de estas investigaciones se concluye que esta adecuación, si bien redundante en un mejor desempeño de las personas con discapacidad, también beneficia a la población sin discapacidad que realiza las mismas tareas en las mismas condiciones (para quienes la adecuación no debería hacer diferencia alguna).

Un ejemplo de esto es el trabajo de Lewandowski, Cohen y Lovett (2013), quienes examinaron el efecto de la extensión en un 50% y un 100% del tiempo concedido para una tarea de comprensión de lectura en estudiantes de educación superior con y sin discapacidades en el aprendizaje. Los resultados indicaron que los estudiantes sin discapacidad se beneficiaron más del tiempo extra que los del grupo con discapacidad, lo que los autores interpretan como una indicación de que dicha adecuación no es apta para este tipo de discapacidad. Sin embargo, en comparación con estudiantes sin discapacidad que no recibieron tiempo adicional, los estudiantes con discapacidad que contaron con tiempo adicional, sobre todo aquellos con un 100% de tiempo extra, superaron con creces las puntuaciones obtenidas por este grupo.

Estos autores concluyeron que, dada la importancia del tiempo adicional en los puntajes obtenidos por toda la población de estudiantes, debe concederse tiempo ilimitado para pruebas en las que la velocidad no es parte del constructo a medir, mientras que si la velocidad es relevante para la medición del constructo, las adecuaciones de tiempo solo deberían otorgarse en condiciones muy especiales, de manera que se evite el beneficio que arbitrariamente podrían recibir quienes cuenten con esta adecuación sin necesitarla.

En el mismo sentido, Mandinach, Bridgeman, Cahalan-Laitusis y Trapani (2005) examinaron los efectos del tiempo adicional en el desempeño de estudiantes con y sin discapacidad que realizaron las secciones verbal y matemática del test de razonamiento del SAT. Específicamente, se exploró el impacto de conceder tiempo estandarizado a un grupo, de proveer a un segundo grupo de un 50% más de tiempo, combinado con cortes definidos durante la aplicación, y de dar el 100% de tiempo adicional sin cortes especificados durante la aplicación.

Si bien los autores citados encontraron que los resultados obtenidos para los estudiantes con discapacidad no fueron significativos debido al escaso tamaño de la muestra, estos detectaron que los estudiantes con discapacidad (de aprendizaje o con déficit atencional con hiperactividad) tuvieron un desempeño inferior al obtenido por estudiantes sin discapacidad, independientemente del tiempo adicional concedido. El tiempo adicional pareció afectar al desempeño en la sección de matemáticas más que en la parte verbal. Para estudiantes sin discapacidad con niveles de habilidad media y alta, el mejor desempeño se alcanzó bajo las condiciones de un 50% más de tiempo para completar la prueba, mientras que el desempeño más bajo se obtuvo en condiciones de tiempo estándar. Para estudiantes con baja habilidad, el tiempo adicional no se tradujo en ningún tipo de ventaja. Por último, el 50% de tiempo adicional dio beneficios

¹ Actualmente, las áreas de clasificación de la población solicitante de adecuaciones son: *aprendizaje, déficit atencional, visual, auditiva, motora, emocional, sistémica y múltiple*. Estas áreas de clasificación han sido establecidas por el Centro de Asesoría y Servicios a Estudiantes con Discapacidad (CASED), la entidad de la Universidad de Costa Rica encargada de determinar los requerimientos de adecuaciones de las personas a evaluar mediante la PAA.

para la sección verbal del test en todos los grupos de habilidad, pero los efectos no fueron tan marcados como para la parte matemática.

Resultados contrarios encontraron Lesaux, Pearson y Siegel (2006), quienes examinaron los efectos de la adecuación de tiempo adicional (tiempo indefinido para completar la prueba) en el desempeño en un test de comprensión de lectura aplicado a un grupo de adultos con discapacidades que afectan a esta área, en comparación con un grupo de personas adultas con un nivel de lectura apropiado. Encontraron que la totalidad de participantes con discapacidad se vieron beneficiados por el tiempo extra, mientras que los participantes sin discapacidad tuvieron un desempeño en la prueba, similar al que obtendrían sin recibir la adecuación. En el caso de los individuos con discapacidades que afectaban a la lectura de una manera más leve, no tuvieron un desempeño significativamente distinto a las personas sin discapacidad.

Un estudio similar fue realizado por Cahalan-Laitusis, King, Cline y Bridgeman (2006) para aportar información sobre el tiempo real usado por estudiantes con discapacidades en el aprendizaje y/o con déficit atencional con hiperactividad para resolver el SAT. Esta información se comparó con la obtenida de una muestra de estudiantes sin discapacidad que completaron el mismo test.

La muestra de estudiantes con discapacidad, a quienes se les concedió un 50% más de tiempo de aplicación, utilizó más tiempo para resolver cada sección de la prueba en comparación con los estudiantes sin discapacidad que aplicaron en condiciones estandarizadas (entre un 4% y un 14% para la población con ambas discapacidades y en solo unos pocos casos se utilizó todo el tiempo adicional). Esto indicó que, si bien el total de tiempo adicional no fue necesario para la mayoría de la muestra con discapacidad, para la mayoría de estudiantes sí fue necesario contar con tiempo adicional para que sus puntajes en la prueba no se vieran afectados.

De acuerdo con el estudio citado, los estudiantes con déficit atencional e hiperactividad, pero sin discapacidades de aprendizaje, generalmente no necesitaron tiempo adicional al utilizado por la muestra sin discapacidad. Sin embargo, se reportó un mayor consumo de tiempo en los ítems de lectura crítica y de matemáticas. Por otra parte, algunos estudiantes con discapacidades de aprendizaje y déficit atencional con hiperactividad no alcanzaron a concluir el subtest de lectura crítica, para cuyo primer pasaje los estudiantes con discapacidades que afectaban a sus capacidades de lectura utilizaron un 25% más de tiempo para responder.

En otro estudio, Lee, Osborne y Carpenter (2010) compararon los efectos de un test computarizado con los de un test en formato de papel y lápiz, así como los efectos del tiempo adicional para cada modalidad, en el desempeño de los estudiantes con déficit atencional e hiperactividad.

Los estudiantes que tomaron la versión computarizada del test puntuaron significativamente más alto que aquellos que tomaron la versión de papel y lápiz del mismo test. Mientras que la cantidad de tiempo que los participantes tuvieron disponible no afectó significativamente su desempeño en ninguna de las dos modalidades de presentación del test, sí se relacionó con un mayor puntaje en la prueba.

Finalmente, Cahalan-Laitusis, Morgan, Bridgeman, Zanna y Stone (2007) buscaron determinar si los estudiantes examinados bajo condiciones de tiempo adicional en el test de razonamiento del SAT estarían sufriendo de fatiga excesiva (calculada mediante el incremento en el Funcionamiento Diferencial del Ítem [DIF] y el decrecimiento en la tasa de ítems contestados) en comparación con los estudiantes examinados bajo condiciones estándar. La muestra incluyó estudiantes con discapacidades de aprendizaje y/o con déficit atencional e hiperactividad, así como estudiantes sin discapacidades medidos bajo condiciones de tiempo estandarizadas.

Los resultados indicaron pocos cambios en los niveles de DIF para la población con tiempo adicional. Además, las tasas de ítems contestados para estudiantes que recibieron tiempo extra fue comparable con la de los estudiantes sin discapacidad y sin tiempo adicional, lo que no constituye evidencia de un efecto asociado a la fatiga excesiva por parte de las personas con discapacidad que recibieron tiempo adicional para esta prueba.

Los resultados de las investigaciones citadas muestran que el efecto del otorgamiento de tiempo adicional como adecuación en pruebas estandarizadas o en tareas de resolución de problemas puede verse

influenciado por factores aun no determinados, relacionados con la funcionalidad de las personas y la naturaleza de la tarea. Por ejemplo, mientras que el otorgamiento de tiempo adicional puede beneficiar a poblaciones que presentan condiciones que interfieren en el procesamiento de textos, también puede perjudicar a personas que tiene periodos cortos de atención. De esto se desprende la necesidad de investigar acerca de mecanismos que permitan definir la cantidad de tiempo necesaria para personas de cada área de discapacidad, de acuerdo con las particularidades de cada grupo.

En cuanto a la elaboración conceptual y teórica sobre el tema de las adecuaciones en pruebas estandarizadas, el tema de la discapacidad ha sido abordado históricamente desde diversos paradigmas y modelos que han dictado el modo de brindar atención a las personas en esta situación. En la actualidad se pretende que las acciones para la atención de las personas que presentan discapacidad o necesidades especiales se basen en el llamado *Paradigma Social y de Derechos Humanos*, el cual sostiene que las desventajas, la segregación y la inaccesibilidad a recursos no están determinadas por la deficiencia biológica que pueden presentar las personas, sino en la discriminación social que estas experimentan (Puig de la Bellacasa, 1990), dando paso a una nueva definición de la discapacidad como:

(...) un concepto que evoluciona y que resulta de la interacción entre las personas con deficiencias y las barreras debidas a la actitud y al entorno que evitan su participación plena y efectiva en la sociedad, en igualdad de condiciones con las demás (Asamblea Legislativa de la República de Costa Rica, 2008).

En consonancia con esto, las personas con discapacidad son definidas como:

(...) aquellas que tengan deficiencias físicas, mentales, intelectuales o sensoriales a largo plazo que, al interactuar con diversas barreras, puedan impedir su participación plena y efectiva en la sociedad, en igualdad de condiciones con las demás (Asamblea Legislativa de la República de Costa Rica, 2008).

El abordaje desde el paradigma social señala que los problemas y desventajas que enfrentan las personas con discapacidad se originan no en las dificultades inherentes a su discapacidad, sino en las barreras encontradas en su entorno (Alba y Zubillaga, 2012). Asimismo, reconoce que la cultura y las reglas sociales son las que determinan cuáles necesidades se valorarán o desestimarán según el momento de atención imperante. Siguiendo este enfoque, el entorno social tiene la obligación de eliminar las barreras al acceso y a la comunicación, de manera que ningún tipo de deficiencia a nivel orgánico tenga implicaciones más allá del ámbito fisiológico de la persona (Ballesteros y Vega, 2001; Montoya, 2003). Esta eliminación de barreras implica el ajuste del entorno de manera que haya accesibilidad a la infraestructura y productos y servicios para la totalidad de las personas que integran la sociedad.

En el contexto específico de la PAA, este enfoque implica que no son las personas con discapacidad las que deben ajustarse a los requerimientos en materia de acceso al espacio físico y al material de la prueba, sino que deben modificarse las disposiciones para la aplicación del instrumento y se deben variar los formatos de presentación y las condiciones de espacio físico para su aplicación de manera que se ajusten a los requerimientos específicos de las personas con discapacidad o necesidades especiales sin afectar la medición del constructo.

Siguiendo la idea anterior, las adecuaciones se definen, en el contexto de la medición, 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 & Johnstone, 2006). Pueden incluir cambios en el contexto en el que las instrucciones son presentadas o en el que se hace la medición, la cantidad de tiempo que se concede para completar una tarea, el método de respuesta, o bien los materiales o el equipo que apoya la interacción con el material de la prueba. Para que se consideren efectivas, las adecuaciones deben reducir la varianza irrelevante para el constructo causada por la discapacidad de la persona sin afectar a la medición del constructo meta, es decir, no deben atentar contra las evidencias de validez en la interpretación de los puntajes derivados de la aplicación (Georgia Department of Education, 2008). Así, las adecuaciones vienen a ser intentos por “nivelar el terreno de juego”, de manera que todas las personas examinadas tengan las mismas posibilidades de mostrar su nivel de habilidad. Estas apuntan a la equidad, no a la ventaja para el grupo que las recibe, mediante la remoción de las fuentes de varianza irrelevante para el constructo, de manera que la medición de habilidades se realice con la debida precisión (Abedi et al., 2012; Thurlow & Wiener, 2000).

De acuerdo con el Departamento de Educación del Georgia (Georgia Department of Education, 2008), solamente se deben otorgar las adecuaciones estrictamente necesarias para garantizar el acceso de cada persona a la evaluación, ya que proveer de adecuaciones no requeridas puede interferir e impactar de forma negativa el desempeño y la medición. Además de lo anterior, las adecuaciones brindadas en procesos de evaluación deben adherirse a los siguientes principios:

- Deben permitir a la persona evaluada participar de forma completa en la evaluación de manera que esta pueda demostrar de mejor manera su conocimiento o habilidades.
- Deben basarse en las necesidades individuales de cada persona y no en una categoría o en un tipo de discapacidad, en el nivel de instrucción o en un programa de estudios.
- Deben ser justificadas y debidamente documentadas para cada persona.
- Deben ir en consonancia con las que la persona recibe en su proceso educativo y no deben ser introducidas por primera vez en el contexto de la evaluación.
- Deben facilitar la independencia de las personas examinadas.

El concepto de adecuaciones en las condiciones de aplicación de una prueba puede confundirse con otro que sí tiene implicaciones para la validez de las inferencias realizadas a partir de los resultados obtenidos por una persona: el concepto de modificaciones. Estas consisten en la realización de prácticas que cambian, alteran o reducen las expectativas de lo que se mide mediante la prueba aplicada. Estas pueden aumentar la brecha entre el logro obtenido por personas con discapacidad y las expectativas sobre el desempeño del grupo total de personas examinadas.

Las modificaciones necesariamente implican algún impacto en la medición del constructo y, a diferencia de las adecuaciones, implican más que un cambio en el escenario de medición (Stone, Cook, Laitusis, & Cline, 2010). Algunos ejemplos de modificaciones en contexto de evaluación incluyen: bajar los objetivos de evaluación, presentar una prueba con menos ítems o problemas, permitir a una persona completar solamente los ítems más fáciles, presentar una prueba más fácil, reducir las opciones de respuesta en pruebas de selección única o dar pistas de las respuestas correctas (Georgia Department of Education, 2008).

Metodología

Los datos utilizados en este artículo han sido recogidos por el Programa Permanente Prueba de Aptitud Académica durante el periodo comprendido entre el año 2009 y el 2011 mediante la solicitud a cada uno de los aplicadores de prueba de anotar las horas de inicio y de finalización de cada examinado. Posteriormente, esta información se integró en una base que contenía las adecuaciones asignadas a cada aspirante. Valga aclarar que la hora de finalización se define como el momento en el que la persona regresa la prueba al aplicador, y que, por la naturaleza confidencial del contenido de la prueba, los aplicadores no comprueban que la persona haya resuelto todos los ítems en el folleto de la prueba, sino solamente verifican que la hoja de respuestas esté completa. Debido a esto, es imposible determinar si la persona resolvió toda la prueba o si marcó respuestas al azar.

La unión en una sola base de las poblaciones que rindieron la prueba en el período 2009-2011 se justifica en la forma en que se diseña la PAA, ya que las pruebas se ensamblan con condiciones que se mantienen constantes año tras año, por ejemplo: la dificultad general y por componente, el orden de los ítems y la cantidad de ítems por área; lo cual garantiza que las pruebas de los años en cuestión sean muy similares entre sí (Programa Permanente Prueba de Aptitud Académica, 2013).

Descripción del proceso estadístico

Para estimar el tiempo necesario en cada una de las áreas de adecuación, se realizó un análisis de sobrevivencia con el software STATA 11.0, donde el evento fue la finalización de la prueba y el tiempo correspondió a la cantidad de minutos que tardó cada aspirante en resolver la prueba. Siguiendo la terminología estadística, a los sujetos que no terminaron la prueba en el plazo fijado (que puede ser de tres horas y media, cuatro horas o cuatro horas y media, dependiendo de la adecuación) se les denominó *censuras*.

Se utilizó un modelo paramétrico de sobrevivencia en cada una de las áreas, donde los predictores fueron las adecuaciones ofrecidas en cada una de ellas. Este modelo fue depurado mediante un proceso de *stepwise*, cuyo criterio de entrada fue que el *p-value* tuviera un valor menor a 0,20. En cada una de las áreas se ajustaron varios modelos completos (con todas las variables) con diferentes distribuciones (*Weibull*, *loglogística*, *lognormal* y *gompertz*). Luego, mediante el criterio del BIC y el AIC, se obtuvo que la distribución que mejor ajustó a los datos en todas las áreas fue la Weibull, a excepción del área *auditiva* donde la distribución que presentó el mejor ajuste fue la loglogística (ver Apéndice B).

Luego, se verificó gráficamente la pertinencia de las distribuciones seleccionadas. En el caso de las áreas donde se seleccionó la distribución Weibull se observó una relación lineal entre el logaritmo del tiempo y el logaritmo de la tasa de *hazards* acumulada (H), calculada por el estimador de Nelson-Aalen. En el caso del área *auditiva*, donde se utilizó la distribución loglogística, se observó una relación lineal entre el logaritmo del tiempo y el logaritmo de $\exp(H)-1$, donde H nuevamente es aproximada con el estimador de Nelson-Aalen. Estas relaciones lineales son supuestos necesarios para usar dichas distribuciones (Moeshberger & Klein, 2003).

Para las áreas de clasificación *sistémica* y *múltiple* se realizó un *stepwise* conjunto, cuyas variables fueron las adecuaciones (ambas áreas presentan las mismas adecuaciones) y una interacción de estas con la variable *múltiple*, la cual es 1 si la persona está clasificada dentro del área *múltiple* ó 0 si es del área *sistémica*. Este proceso se realiza cuando hay dos grupos con muestras pequeñas y se quieren obtener resultados de cada área por separado, pero considerando muestras más grandes para la estimación.

El *stepwise* para estas áreas se realizó por pares de variables (la adecuación y la interacción), donde una pareja se excluía únicamente si ambos predictores presentaban $p > = 0,20$. Además, la unión de ambas poblaciones presentó, al igual que en la mayoría de áreas de adecuación, un mejor ajuste a la distribución Weibull que a las otras cuatro distribuciones.

Finalmente, se decidió utilizar un modelo de tiempo acelerado de falla (AFT) debido a que las interpretaciones de sus coeficientes presentan una facilidad de comprensión para lectores no especializados en estadística. En el Apéndice A se muestra que los coeficientes obtenidos en los modelos finales concuerdan con los obtenidos en un modelo de Cox, un modelo muy utilizado en análisis de sobrevivencia, con el inconveniente de que no permite realizar predicciones de tiempo.

En todos los modelos se satisfizo el supuesto básico de los análisis de sobrevivencia: los riesgos proporcionales. Esto se verificó mediante el *test global de hazards* proporcionales y por medio de la comparación gráfica de los *hazards* de la población dividida según sexo (esta comparación se consideró pertinente debido a que el género es una de las variables con mayor influencia en los resultados obtenidos en la PAA, Rojas, 2013). En el primer análisis se obtuvieron *p-values* mayores a 0,05 en todos los modelos, mientras que en el estudio gráfico se observó que el cociente de los *hazards* de ambos grupos describía una relación constante.

Para cada modelo finalmente ajustado se realizó un análisis gráfico de los residuos de Cox-Snell contra la tasa de *hazards* acumulados y se observó que estos residuos no se desvían de la recta identidad, lo cual indica que el modelo presenta un buen ajuste (Moeshberger & Klein, 2003). Por otro lado, en los gráficos de residuos de deviancia contra el tiempo, se comprobó que la mayoría de estos residuos no se alejan drásticamente del eje horizontal, lo cual brinda otra evidencia del buen ajuste de los modelos acá planteados.

Predictores

Como se dijo anteriormente, las variables que se utilizan como predictores en cada área son las adecuaciones brindadas en estas. Todas están codificadas como variables dicotómicas, donde 1 significa presencia de la adecuación y 0, ausencia.

Las adecuaciones que están a disposición de todas las áreas son: letra ampliada (letra), calculadora con operaciones básicas (calc), ubicación especial (ubic), marcar con equis en el folleto (MarcaX), aplicación en grupo de 15 o 5 personas (G_peq), aplicación individual (Indiv) y tablas de multiplicar (Tabla).

Por otro lado, en las áreas *motora*, *sistémica* y *múltiple* hay varias adecuaciones en común, tales como aula accesible (Acceso), mobiliario adaptado (Mob) y lector-escribiente (L_Esc).

En el área *visual*, además de las adecuaciones generales también se hace uso del lector-escribiente (L-Esc) y de otras adecuaciones específicas del área como aplicación en Braille (Braille), uso de lupa (Lupa), iluminación especial (Luz), prueba en audio (audio), uso de calculadora parlante (Calc_p) y formulario adaptado para personas con deficiencias visuales (F5).

Para el área *auditiva*, se cuenta con varias adecuaciones específicas que consisten en son: uso de diccionario (Dicc), presencia de interprete oral (I_Oral), interpretación del folleto de prueba en Lengua de Señas Costarricense (LESCO) y formulario adecuado a examinados con adecuación en el área *auditiva* (F6).

En las áreas *aprendizaje*, *déficit atencional* y *emocional* no hay adecuaciones específicas además de las generales que se ofrecen a todas las áreas.

Ecuaciones

Luego de realizar el proceso de stepwise con los predictores mencionados anteriormente, se obtiene que las ecuaciones para estimar el tiempo son las siguientes:

$$\text{Aprendizaje: } t(S|Z) = \{\ln(1/S)\}^{1/p} * \exp\{\beta_0 + \beta_1\text{Letra} + \beta_2\text{Indiv}\}$$

$$\text{Déficit atencional: } t(S|Z) = \{\ln(1/S)\}^{1/p} * \exp\{\beta_0\}$$

$$\text{Motora: } t(S|Z) = \{\ln(1/S)\}^{1/p} * \exp\{\beta_0 + \beta_1\text{L_Esc} + \beta_2\text{G_peq}\}$$

$$\text{Visual: } t(S|Z) = \{\ln(1/S)\}^{1/p} * \exp\{\beta_0 + \beta_1\text{Calc_p} + \beta_2\text{MarcaX} + \beta_2\text{Indiv}\}$$

$$\text{Auditiva: } t(S|Z) = (1/S-1)^\gamma \exp\{\beta_0 + \beta_1\text{LESCO}\}$$

$$\text{Emocional: } t(S|Z) = \{\ln(1/S)\}^{1/p} * \exp\{\beta_0 + \beta_1\text{MarcaX} + \beta_2\text{Tabla}\}$$

Sistémica-múltiple:

$$t(S|Z) = \{\ln(1/S)\}^{1/p}$$

$$* \exp\{\beta_0 + \beta_1\text{Mob0} + \beta_2\text{Mob1} + \beta_3\text{Letra0} + \beta_4\text{Letra1} + \beta_5\text{G_peq0} + \beta_6\text{G_peq1}\}$$

Aquí, Z es un vector de valores en las variables observadas, S es la proporción de personas a las que no les ha ocurrido el evento, p es el parámetro de la distribución Weibull ajustada y γ , el de la distribución loglogística.

Muestra

Este estudio presenta un diseño transversal debido a que se utilizan las poblaciones de tres años distintos como una sola población, compuesta por los examinados a los que se les aprobó alguna adecuación entre los años 2009 y 2011. En las áreas *visual* y *auditiva* se optó por eliminar el año 2009, debido a que en ese año varias adecuaciones importantes de las áreas no estaban implementadas (prueba en LESCO o uso de calculadora parlante). Además, cabe mencionar que se eliminaron 1000 observaciones del análisis por dos razones: tiempos no registrados o tiempos “inflados” debido a que algunos aspirantes tienen aprobado el uso de descansos durante su aplicación.

En la Tabla 1 se puede observar la frecuencia absoluta de cada una de las áreas por año. Este cuadro refleja que en cada año el área con más adecuaciones es la de *aprendizaje*, la cual representa poco más del 70% del total de adecuaciones asignadas. Esta es seguida por las de *déficit atencional*, luego por *visual*, *emocional*, *motora*, *auditiva*, *sistémica* y, por último, *múltiple*.

Tabla 1
Frecuencia absoluta de adecuaciones según área y año

Año	<i>Apren</i>	<i>Déf. At.</i>	<i>Motora</i>	<i>Visual</i>	<i>Audit.</i>	<i>Emocio.</i>	<i>Sistém.</i>	<i>Múlti.</i>	Total
2009	619	120	22	-	-	40	19	5	825
2010	746	141	26	68	21	41	20	5	1,068
2011	848	123	23	64	30	51	19	8	1,166
Total	2,213	384	71	132	51	132	58	18	3059

Nota: Fuente: elaboración propia de los autores (2013).

Resultados y análisis

En la Tabla 2 se presentan los resultados de los modelos utilizados para estimar el tiempo de duración en resolver la PAA según área de adecuación, a excepción de las áreas *sistémica* y *múltiple*, que se detallan en la Tabla 3.

Como ya se vio en las ecuaciones de los modelos, en el área de *aprendizaje* solamente las adecuaciones que implicaron un cambio en el tipo o tamaño de letra (tipo letra) y aplicación individual cumplieron con el criterio de ingreso al modelo.

Las personas que utilizan un tipo o tamaño de letra distinto del de la población con adecuación por *aprendizaje* presentan diferencias significativas a un nivel del 5% y en promedio duran un 10% menos de tiempo que el resto de aspirantes en esta área. Por otro lado, los examinados que requieren aplicación individual, en promedio duran un 34% más de tiempo que la población evaluada en grupos. Sin embargo, esta variable no es significativa al 5%.

Para la población con adecuación por *déficit atencional* se reveló que ninguna de las variables incorporadas en el análisis cumplió con los criterios de ingreso al modelo. Esto parece indicar que para estos sujetos ninguna de las adecuaciones presenta un efecto significativo en la conclusión de la PAA.

Respecto al área *motora*, se obtuvo que la adecuación de grupo reducido presenta una asociación significativa al 1% con el tiempo de resolución de la PAA, mientras que el uso de lector escribiente es significativa al 10%. La presencia de cualquiera de estas adecuaciones se asocia a un mayor tiempo en la ejecución de la prueba.

Tabla 2
Modelos resultantes del stepwise según área de adecuación

<i>Aprendizaje</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
Letr.	-0,108	-10,232	0,042	-2,570	0,010	-0,190	-0,026
Indiv.	0,298	34,758	0,182	1,640	0,101	-0,059	0,655
Intercepto	5,283		0,004	1183,240	0,000	5,274	5,291
<i>Déficit Atencional</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
Intercepto	5,280		0,010	528,840	0,000	5,260	5,299
<i>Motora</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
L_Esc	0,153	16,532	0,081	1,890	0,059	-0,006	0,312
G_peq	0,278	32,080	0,083	3,360	0,001	0,116	0,441
Intercepto	5,266		0,022	234,820	0,000	5,222	5,310
<i>Visual</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
Calc_p	0,153	16,527	0,096	1,600	0,110	-0,035	0,341
MarcaX	-0,461	-36,929	0,160	-2,880	0,004	-0,775	-0,147
Indiv.	-0,373	-31,108	0,147	-2,530	0,011	-0,661	-0,084
Intercepto	5,798		0,162	35,800	0,000	5,481	6,116
<i>Auditiva</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
LESCO	0,341	40,635	0,062	5,480	0,000	0,219	0,463
Intercepto	5,150		0,047	108,440	0,000	5,056	5,243
<i>Emocional</i>							
Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
MarcaX	0,146	15,719	0,071	2,050	0,040	0,007	0,285
Tabla	0,305	13,566	0,222	1,380	0,168	-0,129	0,740
Intercepto	5,340		0,026	206,320	0,000	5,289	5,391

Nota: Coef.: Coeficiente, Porc_t: aumento porcentual en el tiempo, EE: error estándar, z: Coef/EE, P>|z|: probabilidad asociada de que un valor en la distribución normal estándar sea mayor o igual que el valor absoluto de z, Intervalo 95 %: intervalo de confianza al 95 % para el coeficiente. Fuente: Elaboración propia de los autores (2013).

En cuanto a la población con necesidades de adecuación en el área *visual*, se reveló que las variables con efectos significativos al 5% fueron: marca en el folleto y aplicación individual. Además, debido al criterio de selección, también se incluyó en el modelo la variable de uso de calculadora parlante. Los coeficientes de las variables de marca en el folleto y aplicación individual indican que, en promedio, los examinados que utilizaron alguna de estas adecuaciones requieren menos tiempo que el resto de la población de esa área.

Por su parte, para la población con discapacidad en el área *auditiva*, la variable que marca diferencias significativas a un nivel del 1% es la aplicación en LESCO. Los examinados de esta área que hacen uso de la prueba en LESCO duran en promedio un 40% más que la población del área *auditiva* que no usa esta adecuación.

Finalmente, para el área *emocional* únicamente la adecuación de marcar en el folleto presenta un efecto significativo, mientras que el uso de tablas de multiplicar cumple con el criterio de inclusión en el modelo pero no tiene un efecto significativo. Además, el uso de ambas adecuaciones está asociado a mayores tiempos en la culminación de la prueba.

Tabla 3
Modelo resultante del stepwise en las áreas *sistémica* y *múltiple*

Tiempo	Coef.	Porc_t	EE	z	P> z	Intervalo 95%	
Mob0	0,149	16,054	0,082	1,820	0,068	-0,011	0,309
Mob1	-0,260	-22,877	0,105	-2,470	0,013	-0,466	-0,054
Letra0	-0,234	-20,883	0,100	-2,330	0,020	-0,431	-0,037
Letra1	0,609	83,794	0,160	3,800	0,000	0,295	0,922
G_peq0	-0,070	-6,770	0,049	-1,420	0,155	-0,167	0,026
G_peq1	-0,169	-15,541	0,093	-1,810	0,070	-0,351	0,014
Intercepto	5,335		0,037	146,020	0,000	5,263	5,406

Nota: Coef.: Coeficiente, Porc_t: aumento porcentual en el tiempo, EE: error estándar, z: Coef/EE, P>|z|: probabilidad asociada de que un valor en la distribución normal estándar sea mayor o igual que el valor absoluto de z, Intervalo 95%: intervalo de confianza al 95 % para el coeficiente. Fuente: Elaboración propia de los autores (2013).

En la Tabla 3 aparece el modelo resultante para las poblaciones del área *sistémica* y *múltiple*. En esta tabla, debido a la definición del modelo, las variables con terminación 0 comparan la población del área *sistémica* que tiene la adecuación especificada en la primera columna con la población *sistémica* sin la misma adecuación. De manera análoga, las variables con terminación 1 comparan la población del área *múltiple* con esa adecuación y la restante de *múltiple*.

Para el área *sistémica* las variables significativas fueron el uso de mobiliario especial y diferente tipo de letra; la primera, a un nivel del 10% y la segunda a un 5%. Respecto del área *múltiple*, se obtuvieron las siguientes variables significativas: uso de mobiliario especial, diferente tipo o tamaño de letra y aplicación en grupo pequeño. Estas presentaron una significancia de 5%, 1% y 10%, respectivamente.

El uso de mobiliario especial en el área *sistémica* se asocia a más tiempo, mientras que en el área *múltiple*, a menos tiempo. Inversamente, la adecuación de distinto tipo de letra se asocia a menos tiempo en la población del área *sistémica* y a más tiempo en el área *múltiple*. En ambos grupos de población el grupo pequeño se asocia a menos tiempo en la resolución de la prueba en comparación con los individuos de la misma área que no utilizan esa adecuación.

En la Tabla 4 se presenta una estimación del tiempo en el cual el 70% de la población habrá finalizado la prueba según su área y las adecuaciones seleccionadas por el stepwise. En algunas subcolumnas de la columna de adecuación aparecen puntos debido a que los modelos seleccionaron menos de tres tipos de adecuaciones.

Tabla 4
Estimación del tiempo en que el 70% de la población habrá concluido la prueba según área y sus adecuaciones seleccionadas en el stepwise

<i>Aprendizaje</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Letra	Indiv.				
23	1	0	.	182,803	168,436	198,395
2.186	0	0	.	203,639	201,865	205,429
4	0	1	.	274,419	192,078	392,060
<i>Déficit atencional</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
384	.	.	.	202,719	198,791	206,724
<i>Motora</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	G_peq	L_Esc				
48	0	0	.	198,456	189,922	207,373
5	0	1	.	231,265	198,578	269,333
18	1	0	.	262,121	224,187	306,475
<i>Visual</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Calc_p	MarcaX	Indiv			
2	0	1	1	148,841	111.,852	198,062
110	0	1	0	216,050	207,094	225,394
11	0	0	1	235,992	207,127	268,879
5	1	1	0	251,756	209,239	302,913
4	1	0	1	274,993	223,759	337,957
<i>Auditiva</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	LESCO					
26	0	.	.	192,182	175,102	210,928
25	1	.	.	270,275	249,800	292,427
<i>Emocional</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Tabla	MarcaX				
102	0	0	.	217,183	206,439	228,485
26	0	1	.	251,310	220,553	286,358
4	1	0	.	294,730	191,316	454,044
<i>Sistémica</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Mob	G-Peq	Letra			
1	0	0	1	169,026	139,289	205,112
2	1	0	1	196,161	160,589	239,612
30	0	1	0	199,177	186,681	212,509
6	1	0	0	247,937	213,651	287,726
<i>Múltiple</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Mob	G-Peq	Letra			
3	1	0	1	278,053	221,744	348,663
1	0	0	1	310,660	240,915	400,598
1	1	1	1	218,944	179,746	266,689
1	0	1	1	244,619	199,391	300,105
4	1	0	0	191,217	165,942	220,341
4	0	1	0	168,224	141,814	199,552
<i>Sistémica y múltiple</i>						
N	Adecuaciones			Tiempo	Intervalo 95%	
	Mob	G-Peq	Letra			
23	0	0	0	213,640	198,877	229,500

Nota: N: Número de observaciones en la muestra que presentaron cada combinación de adecuaciones. Fuente: Elaboración propia de los autores (2013).

A continuación, se mencionan aquellos tiempos que presentan intervalos de confianza al 95% cuya extensión no supera los 50 minutos. Además, se denomina *tiempo óptimo* al tiempo promedio en que el 70% de la población ha finalizado la prueba (porcentaje deseable de población que finalice la prueba) y *tiempo óptimo máximo* al límite superior del intervalo de confianza al 95% del tiempo óptimo.

Para que el 70% de la población del área de *aprendizaje* que no tenga asignada adecuación de letra ni individual culmine la prueba se necesitan, como máximo, 205 minutos (extremo superior del intervalo de 95% de confianza), esto es, menos de tres horas y media. Por su parte, para los individuos del área de *aprendizaje* que tienen asignada la adecuación de letra pero no individual, el tiempo óptimo máximo es de 198 minutos. En el área de *déficit atencional* en promedio, el porcentaje deseable de población que finaliza la prueba lo hace a los 202 minutos, con una duración máxima de 207 minutos.

En el área *motora*, los individuos a quienes no se les asignó una adecuación de grupo pequeño ni de lector escribiente presentan un tiempo óptimo máximo de 207 minutos (inferior a las tres horas y media, que es el tiempo extra mínimo que se concede). Luego, los individuos del área *visual* que marcan en el folleto y que no utilizan calculadora parlante ni aplicación individual presentan un tiempo óptimo máximo de 225 minutos, o sea, menos de cuatro horas (tiempo extra mínimo que se concede). Igualmente, la población del área *emocional* que no utiliza tablas de multiplicar y marca en el folleto presenta un tiempo óptimo máximo inferior a las cuatro horas (229 minutos).

Con respecto a la población del área *auditiva*, el tiempo óptimo máximo para los que no tienen asignada la aplicación en LESCO es de 211 minutos, mientras que para los que sí requieren esta adecuación es de 292 minutos. Esto indica que el tiempo óptimo máximo para el primer grupo es de aproximadamente tres horas y media, mientras que para el segundo es de casi cinco horas.

Por último, la población del área *sistémica y múltiple* que no utilizó distinto tipo de letra, ni mobiliario especial, ni grupo pequeño, presenta un tiempo óptimo máximo de 213 minutos. Si la población es exclusiva del área *sistémica* y no utiliza distinto tipo de letra ni mobiliario especial, pero sí grupo pequeño, el tiempo óptimo máximo es de 230 minutos. En los dos casos anteriores el tiempo es inferior a las cuatro horas. Para el área *múltiple* por sí sola, no hay ningún tiempo óptimo con un intervalo de longitud menor a 50 minutos, pero se pueden mencionar dos casos en que el intervalo es de 55 minutos aproximadamente: cuando se utiliza mobiliario pero no grupo pequeño ni tipo de letra, o cuando se utiliza grupo pequeño pero no mobiliario ni tipo de letra. Para el primer caso el tiempo óptimo máximo es inferior a las cuatro horas, mientras que para el segundo es inferior a las tres horas y media.

Otros casos que se pueden considerar son aquellos que exceden la condición puesta a los intervalos pero que tienen más de 10 individuos. Estos son el de la población del área *visual* que rinde la prueba individualmente y que no tiene asignada marcar en el folleto ni calculadora parlante, el de los individuos del área *emocional* que no tienen asignadas tablas de multiplicar pero sí marcar en el folleto y el de *motora*, al que se asigna un grupo pequeño. Para estos grupos se tienen los siguientes tiempos óptimos máximos: para el de *visual*, 268 minutos (casi cuatro horas y media), para el de *emocional*, 287 minutos (poco menos de cinco horas) y para el de *motora*, 306 minutos (poco más de cinco horas). En todos los casos, el intervalo de confianza es de aproximadamente una hora.

Discusión

Los resultados obtenidos en este estudio arrojan un primer indicio científico acerca de la cantidad de tiempo requerido en la solución de la PAA por parte de un con una adecuación específica, lo cual es un avance para el programa, ya que hasta el momento la estimación del tiempo se ha hecho únicamente a juicio de expertos y mediante análisis descriptivos previos de los datos relativos al desempeño en el tiempo de la población con adecuaciones.

En el Programa PAA se considera óptimo que al cumplirse el tiempo máximo otorgado para la prueba, al menos el 70% de la población logre terminarla. Los resultados presentados en este estudio se limitan a revisar si la adecuación de tiempo adicional realmente logra equiparar las poblaciones con o sin discapacidad en cuanto a la posibilidad de finalizar y entregar la PAA. La consideración del efecto que las adecuaciones otorgadas puedan tener en los puntajes obtenidos en la PAA por la población con discapacidad está pendiente de ser abordada en un estudio posterior.

A los examinados del área *aprendizaje* generalmente se les otorgan tres horas y media para realizar la prueba, lo cual es muy similar a la predicción realizada para el percentil 70 de esta población que no requiere de una adecuación con un efecto considerable en el tiempo (letra o aplicación individual). Además, los examinados de esta área que utilizan adecuación de letra pero no individual también presentan un tiempo óptimo menor a las tres horas y media. Esto apunta a que la apreciación realizada por los especialistas en cuanto al tiempo necesario para estos sujetos fue bastante atinada.

De igual manera que en el área *aprendizaje*, la predicción realizada del tiempo óptimo máximo para el área de *déficit atencional* fue bastante similar al tiempo otorgado generalmente a esta población, que es de tres horas y media. Los resultados indican que para esta área no hay ninguna adecuación que requiera una variación en el tiempo otorgado a los examinados.

Los examinados de *motora* que no requieren aplicaciones en grupo pequeño o con lector escribiente y los de *sistémica* que aplican en grupo pequeño y no requieren mobiliario especial ni modificaciones en la letra, según los tiempos óptimos obtenidos necesitan solamente tres horas y media para realizar la prueba (en la PAA se acostumbra a redondear a la media hora superior más cercana). Por otro lado, a los examinados del área *emocional* que no requieren tablas de multiplicar ni marcar en el folleto, a los del área *visual* que marcan en el folleto y no requieren una aplicación individual o uso de calculadora parlante y a los de *sistémica* y *múltiple* que no tienen asignado un grupo pequeño ni mobiliario especial ni modificaciones de letra se les debería asignar al menos cuatro horas. Los tiempos que se han asignado a estas poblaciones no son estándares y dependen de las condiciones del examinado; sin embargo, el análisis realizado en esta investigación brinda criterios para la construcción de estos estándares.

Las estimaciones indican que a los examinados del área *auditiva* que no requieren LESCO se les puede haber sobreestimado el tiempo, ya que el tiempo óptimo máximo de estos es exactamente tres horas y media y muchas veces a estos se les dan cuatro horas. Lo inverso sucede con los que sí tienen asignada la PAA en LESCO, ya que a estos se les dan únicamente cuatro horas y media de tiempo, cuando su tiempo óptimo máximo es de aproximadamente cinco horas.

A pesar de que las estimaciones realizadas en este estudio justifican que a los individuos del área *auditiva* con adecuación de LESCO se les den cinco horas para realizar la prueba, puede que no sea esta una adecuación favorable para esta población, ya que dos horas adicionales al tiempo normal de aplicación introducen factores que perjudican al examinado, como lo son el cansancio físico y mental o la distracción, por lo cual no se estaría reduciendo el impacto de la discapacidad y no se tendría una adecuación válida (Ketterlin-Geller & Johnstone, 2006). Lo anterior lleva al Programa PAA a considerar otras opciones para evaluar a esta población, como lo es la reducción de la prueba o la aplicación en dos sesiones.

Por otro lado, se obtuvieron estimaciones de tiempo para la mayoría de la población, pero para ciertos grupos con muy pocos sujetos no se logró obtener una estimación confiable, lo cual implica la necesidad de reunir poblaciones de una mayor cantidad de aplicaciones, ya que con solamente tres aplicaciones algunos grupos no obtuvieron tiempos estables.

En cuanto a la influencia de las adecuaciones en el tiempo necesario para resolver la PAA, en los

modelos finales se escogieron como máximo tres de estas, lo cual parece indicar que muchas no influyen en el tiempo de ejecución de la prueba. Las adecuaciones de tipo de letra, aplicación individual y grupo pequeño fueron seleccionadas en dos modelos, pero las tres presentan direcciones contrarias en cada modelo. Así, por ejemplo, el tipo de letra se asocia a menores tiempos que el resto de la población en el área *aprendizaje* y en el área *múltiple* se asocia a mayores tiempos.

Por último, esta investigación proporciona un sustento teórico para la estimación de los tiempos para las poblaciones con adecuación y un antecedente de las adecuaciones que tienen una asociación considerable con los tiempos necesarios para resolver la PAA.

El artículo original fue recibido el 6 de mayo de 2013
El artículo revisado fue recibido el 19 de noviembre de 2013
El artículo fue aceptado el 20 de diciembre de 2013

Referencias

- Abedi, J., Bayley, R., Ewers, N., Mundhenk, K., Leon, S., Kao, J., & Herman, J. (2012). Accessible reading assessments for students with disabilities. *International Journal of Disability, Development and Education*, 59(1), 81-95. doi: 10.1080/1034912X.2012.654965
- 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
- Asamblea Legislativa de la República de Costa Rica (29 de mayo de 1996). Ley N° 7600: Ley de Igualdad de Oportunidades para las Personas con Discapacidad. *Diario Oficial La Gaceta*, 102.
- Asamblea Legislativa de la República de Costa Rica (29 de septiembre de 2008). Ley N° 8661: Convención Interamericana para la Eliminación de todas las Formas de Discriminación contra las Personas con Discapacidad. *Diario Oficial La Gaceta*, 187.
- Ballesteros, C. y Vega, M. (2001). *Estrategias que construyen y aprovechan las personas con discapacidad física para incorporarse a procesos productivos* (Tesis de licenciatura). Trabajo Social, Universidad de Costa Rica, San José, Costa Rica.
- Cahalan-Laitusis C., King T., Cline F., & Bridgeman, B. (2006). *Observational timing study on the SAT Reasoning Test for test-takers with learning disabilities and/or AD/HD* (College Board Research Report No. 2006-4 ETS RR-06-23). Nueva York: The College Board.
- Cahalan-Laitusis, C., Morgan, D. L., Bridgeman, B., Zanna J., & Stone, E. (2007). *Examination of fatigue effects from extended-time accommodations on the SAT Reasoning Test™* (College Board Research Report No. 2007-1). Nueva York: The College Board.
- Cawthon, S. W., Winton, S. M., Garberoglio, C. L., & 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
- 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., & 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., & 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., & 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., & 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
- Mainieri, A. (2010). *Reconstrucción teórica e histórica de los fundamentos de la PAA-UCR*. (Informe de investigación). San José: Instituto de Investigaciones Psicológicas, Universidad de Costa Rica.
- Mandinach, E., Bridgeman, B., Cahalan-Laitusis, C., & Trapani, C. (2005). *The impact of extended time on SAT® test performance* (College Board Research Report No. 2005-8 ETS RR-05-20). Nueva York: The College Board.
- 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
- Moeschberger, M., & Klein, J. (2003). *Survival analysis: techniques for censored and truncated data*. Nueva York: Springer-Verlag. doi: 10.1080/00401706.1998.10485206
- 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*. Heredia, Costa Rica: Área de capacitación, Consejo Nacional de Rehabilitación y Educación Especial.
- Programa Permanente Prueba de Aptitud Académica (2013). *Manual de procesos de la Prueba de Aptitud Académica* (Sin publicar). San José: Universidad de Costa Rica.

- 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), 1-24.
- Stone, E., Cook, L. L., Laitusis, C. C., & 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., & Wiener, D. (2000). *Non-approved accommodations: Recommendations for use and reporting* (Policy Direction No. 11). Minneapolis, MN: University of Minnesota, National Center for Educational Outcomes.
- Universidad de Costa Rica (1980). *Oficio R-510-80* (Sin publicar). San José, Costa Rica: Universidad de Costa Rica.

Apéndice A

Comparación de los modelos obtenidos con el modelo de Cox

Al comparar los coeficientes obtenidos en los modelos paramétricos con los obtenidos en el modelo semiparamétrico de Cox, se evidencia que en ambos modelos los valores estimados siguen iguales direcciones. Esta comparación entre modelos se presenta en la Tabla A1. El supuesto de *odds* proporcionales requerido para la regresión de Cox se cumplió en todos los modelos y variables con un 95% de confianza; este se verificó con el `stphtest` de STATA.

Por ejemplo, el porcentaje de cambio en el tiempo del predictor *letra* del modelo paramétrico en el área *aprendizaje* indica que las personas de dicha área que reciben la adecuación de letra resuelven la prueba con un 10% menos de tiempo que el resto de la población de *aprendizaje*. Lo anterior concuerda con el *hazard ratio* obtenido en el modelo de Cox, el cual indica que los individuos de *aprendizaje* con la adecuación de letra tienen unos *hazards* de terminar la prueba 82,6% mayores que los del resto de sujetos.

En función de la interpretación anterior se puede deducir que los coeficientes del modelo paramétrico y los del modelo de Cox concuerdan si los porcentajes de cambio en el tiempo son negativos y los *hazard ratio* son mayores a 1, o si los porcentajes de cambio en el tiempo son positivos y las razones de *hazards* son menores a 1. En la Tabla A1 se puede apreciar que en ambos modelos todos los predictores utilizados brindan resultados concordantes.

Además, la significancia de todos los coeficientes en ambos modelos es muy similar, a excepción del predictor *lector-escribiente* en el área *motora*, que es significativo al 10% en el modelo paramétrico pero no en el modelo de Cox, aunque se puede decir que en ninguno de los dos modelos es significativo al 5%.

Los argumentos anteriores indican que el modelo de Cox brinda evidencia de que las estimaciones realizadas con los modelos paramétricos son consistentes.

Tabla A1
Comparación entre los coeficientes del modelo paramétrico y el modelo de Cox

<i>Aprendizaje</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
Letra	-0,108	-10,232	0,010	1,826	0,009
Indiv	0,298	34,758	0,101	0,204	0,113
<i>Motora</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
L-Esc	0,153	16,532	0,059	0,574	0,397
G_peq	0,278	32,080	0,001	0,186	0,006
<i>Visual</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
Calc_p	0,153	16,527	0,110	0,491	0,125
MarcaX	-0,461	-36,929	0,004	8,447	0,008
Indiv	-0,373	-31,108	0,011	5,551	0,020
<i>Auditiva</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
LESCO	0,561	75,238	0,000	0,147	0,000
<i>Emocional</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
MarcaX	0,146	15,714	0,040	0,544	0,065
Tabla	0,305	35,706	0,168	0,239	0,156
<i>Sistémica y múltiple</i>					
Tiempo	Modelo paramétrico			Modelo de Cox	
	Coef.	Porc_t	P> z	H. ratio	P> z
Mob0	0,149	16,054	0,068	0,418	0,099
Mob1	-0,260	-22,877	0,013	4,453	0,029
Letra0	-0,234	-20,883	0,020	4,528	0,022
Letra1	0,609	83,794	0,000	0,023	0,000
G_peq0	-0,070	-6,770	0,155	1,492	0,206
G_peq1	-0,169	-15,541	0,070	3,002	0,072

Nota: Coef.: Coeficiente, Porc_t: aumento porcentual en el tiempo, P>|z|: significancia estadística del coef. o el ln (H.Ratio), H. ratio: hazard ratio. Fuente: Elaboración propia de los autores (2013).

Apéndice B

AIC y BIC de los modelos con todas las variables según área y distribución

En la Tabla B1 se puede observar que en todas las áreas los modelos con menor AIC y BIC son los ajustados con la distribución Weibull, a excepción del área *auditiva*, donde la distribución con menores valores en estos indicadores es la loglogística.

Tabla B1

AIC y BIC de los modelos con todas las variables según área y distribución

Modelo	AIC						
	<i>Aprend.</i>	<i>Déf. at.</i>	<i>Motora</i>	<i>Visual</i>	<i>Audit.</i>	<i>Emocio.</i>	<i>Sist.-múlt.</i>
Exponencial	3515,469	617,242	114,953	299,126	125,496	210,934	164,351
Weibull	829,846	121,666	11,552	88,034	20,344	115,837	38,812
Gompertz	873,015	127,261	14,026	96,617	26,407	119,964	39,315
LogLogística	894,776	139,639	14,673	88,100	14,715	117,1639	46,724
LogNormal	953,470	145,807	13,907	90,326	30,084	122,933	51,845
Modelo	BIC						
	<i>Aprend.</i>	<i>Déf. at.</i>	<i>Motora</i>	<i>Visual</i>	<i>Audit.</i>	<i>Emocio.</i>	<i>Sist.-múlt.</i>
Exponencial	3559,032	635,693	131,161	339,485	144,814	231,695	209,357
Weibull	878,862	143,808	29,782	131,276	41,594	139,193	85,961
Gompertz	922,021	149,403	32,253	139,859	47,657	143,319	86,464
LogLogística	943,787	161,781	32,908	131,342	35,965	140,521	93,873
LogNormal	1002,488	167,950	32,134	133,568	51,334	146,289	98,994