Actualidades en Psicología, 29(119), 2015, 115-130 ISSN 2215-3535
http://revistas.ucr.ac.cr/index.php/actualidades DOI: http://dx.doi.org/10.15517/ap.v29i119.19283

Efecto de una capacitación sobre los puntajes de la prueba de admisión de la Universidad de Costa Rica: una aproximación bayesiana*

Effects of a Training Program on the Scores from the Admission Test at the University of Costa Rica: a Bayesian Approximation

Eiliana Montero Rojas[1]
Guaner Rojas-Rojas[2]
Susan Francis-Salazar[4]
Universidad de Costa Rica, Costa Rica

Miguel Negrín-Hernández[3]
Universidad de Las Palmas de Gran Canaria, España

Resumen. Se usó un diseño cuasi-experimental con pre y post-test para estimar el efecto de una capacitación para la prueba de admisión de la Universidad Costa Rica, un test estandarizado que mide habilidades de razonamiento en contextos verbales y matemáticos. Cuatro colegios públicos del área metropolitana central del país participaron en el estudio, asignándose dos de ellos aleatoriamente al grupo de intervención y los otros dos al grupo de control, con 61 estudiantes en el primer grupo y 80 en el segundo. La intervención consistió de 5 sesiones de capacitación de 3 horas, utilizando como guía un manual desarrollado por una experta pedagoga, con enfoque constructivista. Las medidas antes y después fueron formas reducidas de la prueba de admisión 2014. La variable dependiente fue la diferencia entre ambas mediciones. El efecto de la capacitación fue de 3.5 puntos porcentuales y significativo, y se estimó utilizando un modelo bayesiano de regresión multinivel.

Palabras clave. Evaluación de impacto, modelos jerárquicos bayesianos, efectos de capacitación, prueba de admisión a la Universidad, modelos multinivel.

Abstract. A quasi-experimental design with pre and post- test was used to estimate training effects for the University of Costa Rica's admission test, a standardized exam that measures reasoning abilities in mathematical and verbal contexts. Four secondary public schools from the metropolitan central area of the country participated in the study; two of them were randomly assigned to the intervention group and the other two to the control group, with 61 students in the first group and 80 in the second. The intervention consisted of five three-hour training sessions, using a written guide developed by a pedagogy expert with a constructivist approach. Before and after measures were reduced test forms of the real admission test from the year 2014. The dependent variable was the difference between the two measures. The effect of the training was estimated using a multilevel Bayesian regression model with a significant magnitude of 3.5 percentage points.

Keywords. Impact evaluation, Bayesian hierarchical models, effects of training, university admission test, multilevel models.

[1]Eiliana Montero Rojas. Instituto de Investigaciones Psicológicas, Universidad de Costa Rica. Dirección Postal: 11501-2060, San José , Costa Rica. E-mail: eiliana.montero@ucr.ac.cr
[2]Guaner Rojas-Rojas. Instituto de Investigaciones Psicológicas, Universidad de Costa Rica. E-mail: guaner.rojas@ucr.ac.cr
[3]Miguel Negrín-Hernández, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España. E-mail: : miguel.negrin@ulpgc.es
[4]Susan Francis-Salazar. Universidad de Costa Rica. E-mail: susan.francis@ucr.ac.cr
*Agradecimientos: A Floribeth Amador, Erick Montoya, María Isabel Sánchez, Gustavo Garita, Kenner Ordóñez, Mirania Astorga, Carlos Solera, Jenny Bolaños y Vanessa Smith, funcionarios de la Universidad de Costa Rica.

Introducción

Los y las aspirantes a cursar una carrera de grado en la Universidad de Costa Rica (UCR) deben tomar una prueba estandarizada de admisión para concursar por un cupo en alguno de sus programas académicos. El proceso de ingreso a la UCR es altamente competitivo, ya que cada año hay cerca de 50000 aspirantes, de los cuales un máximo de 8000 logran obtener cupo en alguna de las aproximadamente 100 carreras de grado que ofrece esa casa de estudios.

Dentro de los proyectos de investigación realizados con la prueba de admisión se perfiló una carencia de estudios enfocados a la estimación de los posibles efectos producidos por la preparación o la capacitación sobre los puntajes de este examen. Dado que la prueba intenta medir habilidades generales de razonamiento en contextos verbales y matemáticos, los desarrolladores e investigadores de la prueba asumieron, tradicionalmente, y sin evidencia empírica, que la preparación previa o capacitación tendría un efecto desdeñable sobre los puntajes de la prueba. Entonces, a pesar de que la prueba de admisión y su correspondientes bancos de ítems tienen ya más de 50 años de existencia, no existe a la fecha ningún estudio científico que arroje luz sobre esta temática.

A nivel internacional diversos estudios afirman que factores relacionados con la situación operativa del examen permiten generar mejoras en los puntajes. Tales factores incluyen el conocimiento del contexto, familiaridad con la tarea, manejo de la ansiedad y del tiempo, la repetición, la práctica, así como también el uso deliberado de estrategias auto- regulatorias y meta-cognitivas. (Cohen, 2006; Flippo, Becker & Wark, 2000; Crocker, 2005; Koriat & Bjork, 2006). Varias de estas dimensiones se engloban dentro del constructo que se denomina en inglés "test-wiseness" (Millman, Bishop & Ebel, 1965; Harmon, Morse & Morse, 1996; Morse, 1998; Sarnacki, 1979).

De hecho, a nivel mundial desde hace varias décadas, existe un floreciente negocio comercial alrededor del tema de preparación para las pruebas de selección a la Universidad, y Costa Rica no es la excepción.

No se trata de inferir que los entrenamientos cortos podrían incrementar las habilidades de razonamiento, sino que la preparación en destrezas y estrategias para realizar estas pruebas podría producir efectos positivos en los puntajes de las mismas (Cohen, 2006; Shulman, 2013; Burns, Siers & Christiansen, 2008; Martínez-Cardenoso, Muniz & García Cueto, 2000; Powers & Rock, 1999). Esta discusión adquiere mayor relevancia para Costa Rica, bajo un contexto de inequidad creciente en el acceso a las oportunidades educativas (Fernández & Del Valle, 2013; Montero, Rojas, Rodino & Zamora, 2012; Trejos, 2010), y de ahí que sea evidente la necesidad de documentar científicamente la posibilidad de efectos de capacitación potenciales, no directamente relacionados al constructo que se pretende medir con la prueba de admisión.

El marco de referencia conceptual que sustenta este estudio nos remite al concepto moderno de validez de Samuel Messick (1989) y ligado específicamente a su definición de varianza irrelevante al constructo. Bajo este marco conceptual el estudio se propone generar evidencia del posible efecto de una capacitación sobre el desempeño de los examinados en la prueba de admisión de la UCR. La posibilidad de encontrar evidencia de un efecto relevante no necesariamente implica que la prueba enfrente una amenaza de validez per se, sino que implica para algunos autores y para estos investigadores, tomar acciones para que en la medida de lo posible, se pueda garantizar que todos los examinados tengan similar acceso a esta capacitación, de forma tal que las diferencias en las puntuaciones de la prueba no reflejen diferencias en el acceso a oportunidades para desarrollar destrezas de toma de tests estandarizados de razonamiento, sino diferencias reales en el constructo bajo medición (Crocker, 2005; Shulman, 2013).

También es necesario indicar que aunque muchos estudios han encontrado que las destrezas y estrategias que se enseñan para tomar exámenes provocan efectos positivos en las puntuaciones de las pruebas, su magnitud parece ser bastante variable y depender de muchos factores contextuales, tales como el rasgo bajo

medición y las características de la población a la que va dirigido el instrumento (Samson, 1985; Cohen, 2006; Shulman, 2013).

En cuanto al tamaño específico de los efectos de la capacitación o entrenamiento en pruebas de selección para la Universidad, nos remitimos inicialmente a Kulik et al. (1984) quienes realizaron un meta análisis con 38 estudios. En 14 de ellos que se referían a la prueba Scholastic Aptitude Test (SAT) de los Estados Unidos, la capacitación aumentó los puntajes en 0.15 desviaciones estándar, en promedio. En otros 24 estudios con otras pruebas de aptitud e inteligencia el entrenamiento elevó el puntaje promedio en 0.43 desviaciones estándar.

Por su parte Powers (1993) expresa que de los estudios meta-analíticos que se han realizado en el Educational Testing Service para estimar los efectos del entrenamiento en la prueba SAT, se puede concluir que la capacitación previa tiene un pequeño y consistente efecto que varía de 25 a 32 puntos de ganancia en la prueba, en promedio. Sabiendo que el SAT tiene una desviación estándar de 100 puntos, entonces los tamaños del efecto estarán precisamente entre 0.25 y 0.32 desviaciones estándar, en promedio. Finalmente Baydar (1990) utilizando estudios de simulación, encontró un efecto de alrededor de un quinto de una desviación estándar (0.20 DE) para la prueba SAT.

En el caso de la prueba de admisión de la UCR y bajo el escenario de inequidad creciente en el acceso a oportunidades educativas de las diversas poblaciones estudiantiles (Fernández & Del Valle, 2013; Montero et al., 2012; Trejos, 2010), surge la necesidad de documentar científicamente la posibilidad de estos efectos potenciales, no asociados directamente al constructo objeto de la medición, pero que, en ciertas condiciones podrían generar cambios en las puntuaciones.

Método

La meta de este estudio fue estimar los efectos de una capacitación para la prueba de admisión de la UCR. El diseño del estudio fue cuasi-experimental con medidas pre y post-test para los grupos de control e intervención. La posibilidad de usar un diseño experimental fue rápidamente descartada debido a la amenaza a la validez interna que presenta su implementación en un contexto escolar, por los llamados efectos de difusión o contaminación (Campbell & Stanley, 1963; Shadish, Cook & Campbell, 2002).

Participantes

En total participaron en el estudio 141 estudiantes de décimo nivel de cuatro instituciones de educación secundaria pública, con tamaños y características socioeconómicas similares. Estas instituciones se ubican en el área metropolitana central del país. Otro criterio que se tomó en cuenta para la selección de estos centros educativos fue el hecho de que tienen una proporción relativamente baja de estudiantes admitidos en la UCR.

Es importante aclarar que la secundaria académica en Costa Rica termina en el undécimo nivel, de tal manera que la población que toma la prueba de admisión todos los años debe haber cursado o estar cursando ese nivel. Esta investigación incluyó, por el contrario, estudiantes de décimo, el nivel anterior, que no tomarán la prueba sino hasta el año siguiente. La principal razón que justifica este proceder es que regulaciones internas de la UCR advierten que la capacitación otorgada por la misma Universidad a grupos seleccionados de estudiantes de último año de secundaria genera a su vez una ventaja no legítima y mayor inequidad, comparados con aquellos grupos que no reciben el entrenamiento.

Dos de las instituciones fueron asignadas aleatoriamente al grupo de intervención y las otras dos al grupo de control; la cantidad de estudiantes del grupo de intervención fue de 61 mientras que el grupo de control tuvo 80. Todos los estudiantes de décimo año de cada uno de los cuatro centros educativos fueron invitados a participar en la capacitación, la cual involucraba asistencia a sesiones de aproximadamente 3 horas de duración durante 5 sábados consecutivos. Entre un 20 y un 25% de todos los estudiantes de décimo año de los 4 colegios se registraron para

participar. Todos ellos firmaron un consentimiento informado y también una carta de compromiso donde manifestaban su propósito de asistir a todas las sesiones de capacitación.

En cuanto al grado de representatividad estadística de estas muestras de estudiantes y de centros educativos, ciertamente no son muestras aleatorias de la población de décimo año de los colegios públicos de Costa Rica, sin embargo debe recordarse que en un diseño experimental o cuasi-experimental lo crucial es tratar de asegurar validez interna (la rigurosidad de la evidencia para el argumento causal) y no la representatividad estadística de los sujetos (Shadish, Cook & Campbell, 2002).

Procedimiento

En este estudio se usaron dos estrategias para controlar los posibles efectos de variables confusoras, que amenazan la deseada interpretación causal de los resultados. La primera estrategia fue la selección de los grupos participantes y su composición, emparejando globalmente algunas de las variables de los centros educativos. La segunda estrategia implicó la medición de posibles variables confusoras y su control vía análisis estadístico usándolas como covariables.

En los grupos de control e intervención se administraron como medidas de pre y post-test formas reducidas del examen real empleado en el proceso de admisión del año 2014. El pre-test incluyó 35 ítems y el post-test 46; los 11 ítems adicionales del post-test son anclajes, es decir ítems que también se incluyeron en el pre-test. Ambas formas presentaron la misma dificultad y exhibieron buenas propiedades psicométricas (Alfa de Cronbach > .75). Los 61 estudiantes en el grupo intervención completaron la capacitación total de aproximadamente 15 horas. Debido a un tema de ética en la investigación con personas, el grupo de control también recibió la capacitación inmediatamente después de haber tomado el pos-test.

Las facilitadoras fueron jóvenes profesoras graduadas recientes de la carrera de Enseñanza de la Matemática. Se definió que el perfil de las facilitadoras debía incluir formación en Enseñanza de la Matemática debido a que en entrevistas previas con algunos estudiantes manifestaron, de entrada, y sin conocer aún el test, aprehensiones muy evidentes para resolver los ítems de contexto Matemático en la prueba, los cuales representan el 40% del puntaje.

Una experta en pedagogía desarrolló un manual escrito para orientar el proceso de capacitación. También las capacitadoras recibieron una inducción previa sobre sus roles en la intervención y el uso del manual. Se usó un enfoque de aprendizaje activo y participativo, con elementos constructivistas, y enfocándose en la comprensión y el uso de estrategias auto-regulatorias incluyendo la evaluación de la dificultad de la tarea, y la planificación de la respuesta. El trabajo en grupos fue un componente fundamental para el desarrollo de las sesiones. Durante su implementación se enfrentaron problemas propios de los procesos educativos vinculados con el manejo del grupo y estrategias de trabajo con grupos numerosos.

En uno de los colegios de control y en uno de los colegios de intervención se presentaron problemas con estudiantes que parecían desmotivados y que provocaron distracciones en varias de las sesiones de capacitación. En estos casos el tema del manejo apropiado del grupo fue relevante y las capacitadoras tuvieron que usar estrategias extraídas de su arsenal como docentes de secundaria para tratar de mantener a la mayoría focalizada en la tarea. Igualmente se vieron enfrentadas a la necesidad de utilizar estrategias Ad Hoc para trabajar estas dinámicas con grupos relativamente grandes (mayores a 30 estudiantes en los 4 casos), cuando lo ideal es que sean más pequeños.

Una parte importante del proceso consistió en generar una situación de examen lo más auténtica posible para la administración del pre y post-test en ambos grupos, simulando el contexto real operativo de aplicación que enfrentan los estudiantes con la prueba de admisión.

Junto con el pre-test se administró un cuestionario a los estudiantes de los grupos control e intervención para medir algunas variables socio demográficas, actitudes hacia la lectura y destrezas de lectura.

Modelo multinivel bayesiano. Para analizar el impacto de la intervención se propuso el uso de los modelos multinivel (Goldstein, 1995; Snijders & Bosker, 1999). Los modelos multinivel se han mostrado como adecuados cuando los datos poseen una estructura jerárquica natural, consistente en múltiples unidades micro, estudiantes, anidadas en unidades macro, colegios. Las primeras aplicaciones de estos modelos se realizaron en el ámbito de la educación (Goldstein, 1987; Book, 1988; Raudenbush & Willms, 1991), aunque se ha extendido su uso en otras áreas como la salud (Rice & Leyland, 1996; Rice & Jones, 1997; Carey, 2000; Leyland & Goldstein, 2001, Goldstein et al., 2002).

Se propuso la utilización de un enfoque bayesiano debido a las ventajas que varios autores (Raudenbush & Bryk, 2002; Browne & Draper, 2006; Gelman & Hill, 2007) han destacado frente a la estimación clásica: asuntos relacionados con los problemas de convergencia de los modelos clásicos, resultados más precisos aún con muestras relativamente pequeñas y la estimación real de los efectos aleatorios para cada conglomerado (colegios en este caso), en vez de estimar solamente la matriz de covariancias de los efectos aleatorios.

La especificación del modelo sería la siguiente. Dada una muestra de *n* estudiantes, se tiene información para cada alumno *i* perteneciente al centro *j* sobre la mejora en las calificaciones de la prueba de admisión a la UCR como variable a explicar (y_{ij}), y una serie de *k* variables explicativas, $x_{ij}=(x_{1,ij}, x_{2,ij},\ldots, x_{k,ij})$, que incluyen una variable indicadora sobre la participación de dicho alumno en el programa de capacitación.

$$y_{ij} = \beta_0 + \beta_1 x_{1,ij} + \beta_2 x_{2,ij} + \cdots + \beta_k x_{k,ij} + u_{ij} + v_j$$

$$u_{ij} \sim N\left(0, \sigma_u^2\right), v_j \sim N\left(0, \sigma_v^2\right), cov\left(u_{ij}, v_j\right) = 0$$

El término de perturbación v_j modeliza la variabilidad existente entre los centros educativos.

En el análisis bayesiano es necesario determinar la distribución a priori sobre los parámetros del modelo. Se utiliza un modelo de distribución a priori condicionalmente conjugado que permite la aplicación del algoritmo Gibbs sampling para la estimación de la distribución a posteriori (Gilks et al., 1996; Goldstein et al., 2002). Aunque las distribuciones propuestas permitirían la incorporación de información a priori, en este artículo se realiza un análisis bayesiano objetivo con distribuciones a priori no informativas. En este caso la no-información se expresa a través de distribuciones con varianza alta, tanto para los parámetros del modelo, como para las varianzas (Gelman, 2006).

$$\beta_0, \beta_1, \ldots, \beta_k \sim N\left(0, 1000\right), \sigma_u^2 \sim IG(1, 0.001), \sigma_v^2 \sim IG(1, 0.001),$$

donde *IG* se refiere a la distribución inversa-gamma.

La distribución Inversa-Gamma ha sido ampliamente utilizada en la literatura bayesiana para modelizar la varianza (Gelman, 2006). Como distribuciones gamma no informativas existe un amplio abanico de ellas que cumplen en su mayoría la propiedad de tener varianzas grandes que se corresponderían con distribuciones gamma con parámetro de escala cercano a 0. Por supuesto, esta no es la única opción posible y en la literatura se han propuesto otras distribuciones desinformativas como Gamma (0.1,0.1), Gamma(0.001,0.001), Gamma(1,1), Gamma(0.5,0.001), etc. aunque la sensibilidad de las estimaciones a las distintas distribuciones a priori es pequeña (Lambert et al., 2005).

Las estimaciones se realizaron utilizando el software OpenBUGS en su versión 3.2.3. Se estimaron dos cadenas simultáneamente y se utilizó una muestra de calentamiento de 40000 iteraciones. Las estimaciones de los coeficientes se realizaron con las siguientes 40000 iteraciones. La convergencia fue verificada a partir del análisis gráfico de las autocorrelaciones, la traza y la función de densidad, además del estadístico Gelman-Rubin. Se incluye como Anexo los códigos utilizados en este análisis.

Variables

A continuación se presenta la definición operativa de las variables utilizadas en el análisis:

Variable dependiente Diferencia_Segunda_menos_Primera, es la diferencia entre el porcentaje de respuestas correctas en la primera medición y el porcentaje de respuestas correctas en la segunda medición.

Variables independientes. Intervención: 1 si el estudiante es parte del grupo de intervención, 0 si es del grupo de control.

Covariables (todas medidas a nivel individual):

-Sexo: 1 si el estudiante es hombre, 0 si es mujer

-Pad_Uni: toma el valor de 1 si al menos uno de los padres estuvo/está en la Universidad, 0 en otro caso

-Indicador_Capital_Cultural: Indicador proxy de Capital Cultural, es una escala de 10 ítems, presenta un Alfa de Cronbach de .614. Está compuesta por los siguientes reactivos, calificados cada uno con 1 si el estudiante indica que lo posee en su hogar y con 0 si indica que no lo posee: 1-línea telefónica fija, 2-lector de DVD, 3-computadora de escritorio, 4-acceso a internet (desde la casa), 5-escritorio para estudiar, 6-una computadora que pueda usar para hacer los trabajos del colegio, 7-programas educativos para computadora, 8-libros de literatura (poesía, cuento, novela, ensayo, etc., 9-libros de ayuda para su trabajo escolar y 10-un diccionario.

El rango de la medida va de 0 a 1, pues se calcula, para cada estudiante, como el promedio de las respuestas a los 10 ítems.

-Indicador_Acceso_BienesYServ: Indicador de Acceso a Bienes y Servicios. Medida compuesta por los siguientes ítems, calificados cada uno con 1 si el estudiante indica que lo posee en su hogar y con 0 si indica que no lo posee: 1-consolas de video juego, 2-pantalla de TV de plasma, LCD o LED, 3-televisión por cable o por satélite, 4-sistema de agua caliente para toda la casa, 5-computadora portátil, 6-asociación con algún club privado (vacaciones, deportivo, etc.), 7-Cuarto propio y 8-un lugar tranquilo para estudiar. Esta medida presenta un rango de 0 a 1, pues es, para cada estudiante, el promedio de las respuestas a los 8 ítems.

Como evidencia de validez divergente reportamos la correlación entre la escala proxy de Capital Cultural y el indicador compuesto de acceso a bienes y servicios, según un análisis factorial exploratorio. El valor para esta medida de asociación fue de .145.

-Escala_Actitud_Lectura: Escala de actitud hacia la lectura, compuesta por 11 ítems y con un Alfa de Cronbach de .885, brinda evidencia de una elevada consistencia interna. Los ítems se responden en un formato tipo Likert de 4 categorías ordinales, desde Totalmente en desacuerdo hasta Totalmente de acuerdo, valores altos indican actitud favorable. Incluye los siguientes reactivos: 1- Solamente leo si lo tengo que hacer, 2-La lectura es uno de mis pasatiempos preferidos, 3-A mí me gusta hablar con la gente sobre libros, 4-Me cuesta terminar de leer un libro, 5- Me gusta que me regalen libros, 6-Para mí, la lectura es una pérdida de tiempo, 7-Me gusta ir a una librería o a una biblioteca, 8-Yo solo leo para buscar la información que necesito, 9-No logro permanecer sentado(a) leyendo tranquilamente más de unos pocos minutos, 10-Me gusta dar mi opinión sobre los libros que he leído, 11-Me gusta intercambiar libros con mis amigos. El rango de esta medida va de 1 a 4, pues es el promedio de las respuestas a los 11 ítems.

-Escala_Comprender_Texto: Escala de conocimiento de estrategias eficaces para comprender un texto. Incluye los siguientes ítems, en cada uno de ellos se asigna un 1 si identifica correctamente si la estrategia es o no eficaz, 0 en otro caso: 1-Me concentro en las partes del texto que son fáciles de entender, 2-Leo rápidamente todo el texto dos veces, 3- Después de haber leído el texto, discuto sobre su contenido con otras personas, 4- Subrayo los pasajes importantes del texto, 5-Resumo el texto con mis propias palabras, 6-Le leo el texto en voz alta a otra persona. Esta medida varía entre 0 y 1, pues es el porcentaje de respuestas correctas.

-Primera medición: Porcentaje de respuestas correctas obtenidas por el estudiante en la primera medición, antes de iniciarse la intervención.

Es importante notar que los ítems para la construcción de todas las medidas compuestas utilizadas como covariables (excepto por supuesto el puntaje en la primera medición de la prueba de admisión) fueron

extraídos del cuestionario del estudiante de la aplicación 2009 en Costa Rica de las pruebas PISA (Programme for International Student Assessment) (Montero et al., 2012), si bien algunos sufrieron modificaciones menores para adaptarlos al contexto de este estudio. Se decidió utilizar este conjunto de covariables, junto con la dummy que identifica a los grupos, dado que en estudios recientes desarrollados para predecir los puntajes de PISA en Costa Rica y los puntajes en una aplicación piloto de la prueba de admisión de la UCR, estas variables resultaron predictoras relevantes (Montero et al., 2012; Montero, 2014).

Resultados

La tabla 1 presenta las estadísticas descriptivas y pruebas de comparación de promedios con el estadístico t de Student para los grupos intervención y control en las covariables. Se evidencian diferencias estadísticamente significativas entre grupo de control y grupo de intervención para el porcentaje de correctas en la primera medición (pre-test), siendo más altos los puntajes promedio en el grupo de control y también en el indicador de acceso a bienes y servicios, en donde también el promedio en el grupo de control es superior. Por su parte la Figura 1 contiene los diagramas de cajas correspondientes a los datos de los cuatro colegios en la variable dependiente: diferencia post-test menos pre-test, o puntaje de ganancia.

Es relevante mencionar que se descartaron de los análisis multinivel dos valores muy extremos correspondientes a residuos estandarizados superiores a 3 en valor absoluto, pues se desviaban claramente de la tendencia general de los datos y tenían una influencia exagerada en las estimaciones. Se especula que las razones para estos valores extremos tendrían que ver con circunstancias muy particulares de los dos estudiantes que los originaron.

Como primera aproximación al modelo multinivel se estima el modelo nulo o vacío que no contiene variable explicativa alguna. El modelo vacío permite obtener una estimación puntual de la media poblacional, que se sitúa en -0.1958 con un intervalo bayesiano al 95% de (-1.866, 1.446). Este resultado indica un empeoramiento de los resultados del test en media, debido principalmente al grupo de no-intervención,

Tabla 1

Estadísticas descriptivas y comparación de promedios para los grupos de intervención y control en las covariables del modelo de regresión

Covariables	Sin intervención n=80		Con intervención n=61		Valor p diferencia de promedios (prueba t de Student)
	Promedio	Desviación Estándar	Promedio	Desviación Estándar	
Sexo	.40	.49	.27	.448	.116
Pad_Uni: Al menos un padre en la Universidad	.34	.486	.20	.406	.083
Primera medición	30.75	13.49	26.49	10.48	.046
Indicador_Capital_Cultural	.68	.18	.66	.23	.546
Indicador_Acceso_BienesYServ	.62	.20	.52	.22	.005
Escala_Actitud_Lectura	2.85	.63	2.82	.64	.785
Escala_Comprender_Texto	.80	.21	.75	.27	.296
Escala_Resumir_Texto	.81	.24	.86	.21	.179

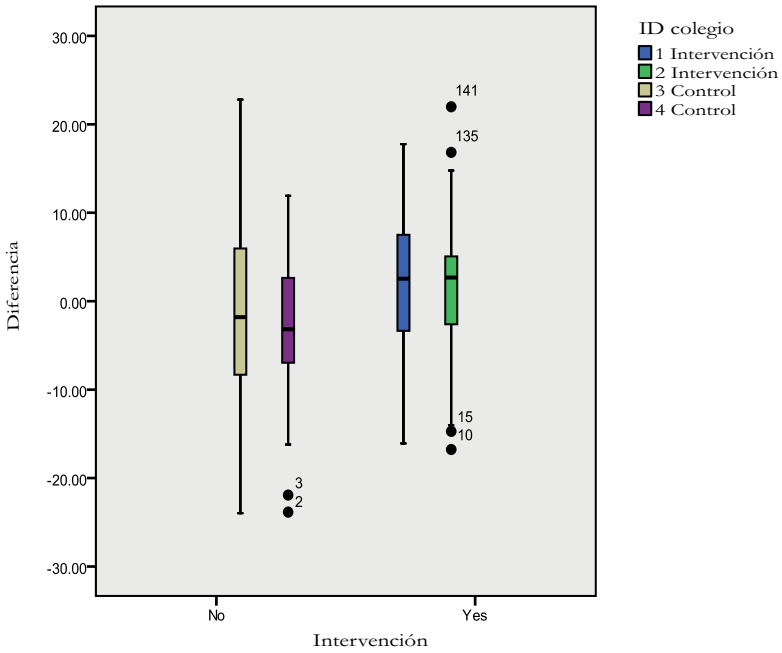

Figura 1. Diagramas de cajas para la variable dependiente (Post menos Pre-test) en los cuatro colegios de la muestra

que registraron un empeoramiento medio de 2 puntos entre ambas pruebas. Además, el modelo vacío ofrece información útil sobre la variabilidad del resultado en cada uno de los dos niveles. En este caso la varianza entre colegios se estima en 0.0101 y la varianza entre individuos en 98.0487, indicando que la mayor proporción de la variabilidad total es atribuible a diferencias entre individuos.

La tabla 2 muestra los resultados del modelo multinivel bayesiano completo. Según se indica aquí, el coeficiente de la variable de intervención es positivo y relevante. En concreto, el efecto de la capacitación se estima en una media en 3.495 puntos porcentuales, con un intervalo de credibilidad bayesiano al 95% de (0.421, 6.582).

En relación al resto de variables incluidas como variables de control, y asumiendo como variables relevantes aquellas cuya probabilidad de ser positivas (o negativas) supera al 95%, se muestran explicativas para el cambio en el puntaje de la prueba únicamente la variable de intervención, como ya indicamos antes y la medición antes o primera medición. Si consideramos una probabilidad superior al 90%, además de las anteriores se muestran explicativas las variables sexo y escala de estrategias para comprender un texto. Los hombres muestran mejores resultados que las mujeres, en promedio superiores en 2.688 puntos porcentuales. Los estudiantes con mejores resultados en la escala de estrategias para comprender un texto muestran mejoras mayores en el puntaje de ganancia, en promedio, por cada punto de aumento en esta la escala el puntaje de ganancia en la prueba (diferencia post-pretest) aumenta 4.40 puntos porcentuales, manteniendo constantes o controlando las otras variables en el modelo. Por último, los estudiantes con peores resultados en la primera medición muestran mayores mejoras en el puntaje de ganancia.

El análisis de la estimación de la varianza de los términos de perturbación muestra que la mayor parte de la variabilidad no explicada por el modelo se debe a las características de los estudiantes y no al centro de estudio. La variabilidad entre colegios únicamente explicaría el 0.01% de la varianza del error. La escasa variabilidad entre colegios se confirma al estimar

Tabla 2

Estimación bayesiana del modelo de regresión multinivel

Variable dependiente: (Post menos Pre-test)

Modelo	Promedio	Desviación Estándar	IB95%	Prob.
Constante	1.837	5.236	(-8.341, 12.030)	0.361
Sexo	2.688	1.828	(-0.907, 6.299)	0.070
Al menos un padre universitario	1.449	1.723	(-1.923, 4.840)	0.199
Indicador Capital Cultural	3.443	3.942	(-4.264, 11.160)	0.191
Indicador Acceso Bienes y Serv	0.508	3.799	(-6.949, 7.964)	0.447
Escala Actitud Lectura	1.696	1.369	(-0.996, 4.355)	0.106
Escala Comprender Texto	4.403	3.380	(-2.248, 11.060)	0.096
Escala Resumir Texto	-3.391	3.382	(-10.043, 3.331)	0.158
Primera medición	-0.445	0.071	(-0.581, -0.307)	0.000
Intervención	3.495	1.577	(0.421, 6.582)	0.014
σ_u^2	74.521	9.377	(58.432, 95,234)	
σ_v^2	0.006	0.038	(0.000, 0.036)	

Tabla 3

Estimación del efecto aleatorio para cada uno de los cuatro colegios de la muestra

Estimación bayesiana del modelo de regresión multinivel.Variable dependiente: (Post menos Pre-test)

Colegio	Promedio	Desviación Estándar	IB95%
Intervención 1	0.0001119	0.08199	(-0.1343, 0.1354)
Intervención 2	-0.000669	0.08191	(-0.137, 0.1331)
Control 1	0.0005418	0.07952	(-0.1315, 0.1419)
Control 2	-0.000595	0.08565	(-0.1398, 0.1353)

Tabla 4
Estimación clásica (frecuentista) del modelo de regresión multinivel
Variable dependiente: (Post menos Pre-test)

Modelo	Promedio	Desviación Estándar	IB95%	Prob.
Constante	1.854	5.222	(-8.380, 12.088)	0.723
Sexo	2.696	1.828	(-0.886, 6.278)	0.140
Al menos un padre universitario	1.455	1.722	(-1.921, 4.832)	0.398
Indicador Capital Cultural	3.420	3.951	(-4.325, 11.164)	0.387
Indicador Acceso Bienes y Serv	0.484	3.797	(-6.958, 7.927)	0.899
Escala Actitud Lectura	1.700	1.372	(-0.990, 4.390)	0.215
Escala Comprender Texto	4.410	3.374	(-2.202, 11.022)	0.191
Escala Resumir Texto	-3.422	3.360	(-10.007, 3.163)	0.308
Primera medición	-0.445	0.070	(-0.581, -0308)	0.000
Intervención	3.500	1.566	(0.430, 6.570)	0.025
σ_u^2	74.482	9.274	(58.353, 95.068)	
σ_v^2	0.000	0.000		

los efectos aleatorios para cada uno de los 4 centros educativos, los cuales son todos muy cercanos a cero, según indica la tabla 3.

La tabla 4, muestra por su parte, y con propósitos comparativos, los resultados de la estimación clásica o frecuentista, con estimación máximo verosímil restringida. No se observan diferencias destacadas en la estimación de los coeficientes del modelo. El efecto capacitación se estima en 3.5 puntos, frente a los 3.495 puntos obtenidos en la estimación bayesiana.

Discusión

En primer lugar, es relevante notar que este estudio es el primero de su tipo en Costa Rica, y debe considerarse una primera aproximación a una temática compleja. Es de esperar que en el futuro se sigan proponiendo investigaciones en esta línea, con muestras de mayor tamaño y explorando de manera más específica factores particulares que, de acuerdo con la teoría, se establezcan como posibles causas para explicar la variabilidad de los puntajes en la prueba de admisión a la UCR.

Ahora bien, en el contexto de este estudio cuasi-experimental y con la naturaleza particular de la variable de criterio (diferencia entre post-test y pre-test en formas reducidas de la prueba de admisión de la UCR) la evidencia presentada apunta a que la estructura multinivel de los datos no tiene una influencia importante en la estimación puntual del efecto de la capacitación, dada que la magnitud

estimada de los efectos aleatorios con el modelo bayesiano es prácticamente nula para los cuatro colegios del estudio. En otras palabras, no hay evidencia de efectos contextuales en la variable de criterio o variable dependiente generados por el "ambiente" del colegio. Este hallazgo, explicado expost facto puede tener bastante sentido si se considera la naturaleza del constructo bajo medición (habilidades de razonamiento en contextos verbales y en contextos matemáticos).

Sin embargo, también puede darse explicación alternativa en términos de que estos efectos contextuales fueron desdeñables en este caso, precisamente por tratarse de un diseño cuasi-experimental, en donde se hizo un esfuerzo de emparejamiento de centros educativos públicos con características similares para lograr mayor validez interna. En este contexto de homogeneidad entre colegios, la modelización multinivel no aporta una variación significativa a los resultados que se obtendrían de un modelo de regresión lineal. Sin embargo, hemos preferido mantener esta modelización ya que es la que se ajusta mejor a la naturaleza de los datos.

La carencia de efectos contextuales encontrada en nuestro estudio no puede generalizarse a otras secundarias del país con características más diversas. En ese caso cabría esperar que variables contextuales asociadas al centro educativo (clima de clase, liderazgo del director, dependencia (pública o privada), zona geográfica, dotación y calidad del personal docente, etc.) puedan afectar significativamente a los resultados. En diversos estudios observacionales con muestras representativas de secundarias en Costa Rica se han evidenciado estos efectos contextuales cuando la variable dependiente es el resultado en una prueba estandarizada o un indicador de rendimiento escolar (Montero, 2014; Montero et al., 2012; MEP, 2012; Moreira, 2009; Rojas, 2004).

Los dos resultados más contundentes, y en donde coinciden los modelos bayesiano y frecuentista, se refieren al papel fundamental de la primera medición como variable de control y al efecto de la intervención, cuya estimación puntual es 3.5 puntos porcentuales. Refiriéndonos a este último hallazgo, el hecho de que, en promedio, y controlando las otras variables en el modelo, un estudiante que participó en el grupo de intervención tenga 3.5 puntos porcentuales más en su puntaje de ganancia (diferencia post-test menos pre-test) que un estudiante del grupo de control, reviste una gran importancia práctica en el contexto de esta investigación, por las consecuencias que pueden derivarse a partir de ese resultado. Y es que al tratarse de una prueba tan competitiva como la prueba de admisión de la UCR, 3.5 puntos porcentuales de diferencia en los puntajes del examen pueden significar la diferencia entre ingresar o no la Universidad, o ingresar o no a la carrera deseada.

Aún más, este puntaje de ganancia representa un aumento de 0.26 desviaciones estándar de la medida antes (3.5/13.49 = 0.26), tamaño del efecto que coincide de manera muy cercana con las estimaciones de tamaño del efecto que se han reportado en estudios internacionales y que se describen en la Introducción (Kulik et al., 1984; Powers, 1993; Baydar, 1990).

Por otra parte, el papel esencial de la medida antes (pre-test) como variable de control en el modelo queda también claramente evidenciado, al recoger esta medida todas aquellas dimensiones idiosincráticas de cada individuo que no pueden ser explicadas por otras variables y que justifican la siempre vigente recomendación de "usar a los sujetos como sus propios controles". De ahí la importancia de contar con al menos dos mediciones antes y después del tratamiento en ambos grupos (intervención y control).

El signo negativo para el coeficiente asociado a esta variable se interpreta en términos de un posible efecto "techo" (ceiling effect) para la ganancia entre pre-test y post-test en los estudiantes con puntajes más altos, y también como que los estudiantes con peores resultados en la primera medición tienen mayor espacio para mejorar que aquellos con puntajes más altos.

En cuanto a las otras covariables utilizadas en los modelos multinivel, hay evidencia menos contundente de relación, para dos de ellas con el puntaje de ganancia. Estas son las variables que exhibieron una probabilidad

superior al 90% en el modelo bayesiano: el sexo del estudiante y la escala de conocimiento de estrategias para comprender un texto.

En el caso del primer resultado, y si el estudiante es varón, en promedio y controlando todas las otras variables en el modelo, tendrá 2.69 puntos porcentuales más en su puntaje de ganancia comparado con una estudiante mujer. En este momento no podemos aventurar ninguna explicación sólida que justifique este comportamiento, se debe indicar eso sí que el posible sesgo de autoselección producido al trabajar únicamente con estudiantes voluntarios, pudo ser mayor en el caso de los hombres, pues la mayoría de las participantes fueron mujeres, como muestran las estadísticas descriptivas de la tabla 1.

Respecto a la relación positiva encontrada entre la escala de estrategias para comprender un texto y el puntaje de ganancia, sí se puede generar una explicación plausible y con fundamentación teórica, en el sentido de que un buen nivel de comprensión de lectura potencia o incrementa la capacidad de mejorar en los puntajes. Cabe resaltar que este resultado coindice con los hallados en otros estudios realizados en Costa Rica para predecir los puntajes en la prueba de admisión y los puntajes en la pruebas de PISA con datos transversales observacionales (Montero, 2014).

Finalmente, en términos metodológicos y a nivel de análisis de datos, se evidenciaron las bondades de la estimación bayesiana en cuanto a mayor precisión en los análisis inferenciales y la posibilidad de estimar directamente los efectos aleatorios para cada conglomerado. Además, el Cuadro 4 confirma las limitaciones de la estimación clásica en la pobre estimación de la varianza entre colegios. Tal y como se afirma en Raudenbush y Bryk (2002), "cuando el número de grupos es pequeño puede no haber suficiente información para estimar la varianza con precisión en los métodos frecuentistas".

Por último y a modo de reflexión final, podemos indicar que estos resultados merecen discutirse desde la dimensión de equidad en el acceso para los diversos grupos de estudiantes que aspiran a ingresar a la UCR, pues es claro que los de menos recursos económicos tienen menos posibilidades de participar en cursos y talleres de preparación para tomar la prueba de admisión, al involucrar la gran mayoría de estas actividades un costo monetario.

Dado que uno de los productos de este proyecto fue precisamente el diseño de la capacitación, incluyendo la guía escrita para su implementación y otros materiales de apoyo, se espera poder tomar acciones para distribuir estos materiales en instituciones secundarias públicas, con la recomendación de que alguno de sus docentes pueda utilizarlos con los estudiantes que así lo deseen, como preparación para la prueba de admisión a la UCR.

Referencias

Baydar, N. (1990). Effects of coaching on the validity of the SAT: Results of a simulation study. In W. W. Wilingham, C. Lewis, R. Morgan, and L. Ramist (Eds.), *Predicting college grades: An analysis of institutional trends over two decades*. Princeton, New Jersey: ETS.

Bock, R.D. (Ed.) (1989). *Multilevel Analysis of Educational Data*. San Diego: Academic Press.

Browne, W.J., & Draper D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis, 1*(3), 473-514.

Burns, G. N., Siers, B. P., & Christiansen, N. D. (2008). Effects of providing pre-test information and preparation materials on applicant reactions to selection procedures. *International Journal of Selection and Assessment, 16*(1), 73-77.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago, IL: Rand McNally.

Carey, K. (2000). A multilevel modeling approach to analysis of patient costs under managed care. *Health Economics, 9,* 435-446.

Cohen, A. D. (2006). The coming of age of research on test taking strategies. *Language Assessment Quarterly, 3*(4) 307-331.

Crocker, L. (2005). Teaching for the test: How and why test preparation is appropriate. In R. P. Phelps (Ed.), *Defending standardized testing* (pp. 159–174). Mahwah, NJ: Lawrence Erlbaum Associates.

Fernández, A. & Del Valle, R. (2013). Desigualdad educativa en Costa Rica: La brecha entre estudiantes de colegios públicos y privados. Análisis con los resultados de la evaluación internacional PISA. *Revista CEPAL, 111*.

Flippo, R. F., Becker, M. J., & Wark, D. M. (2000). *Preparing for and taking tests.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1,* 515-533.

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

Gilks, W., Richardson, S., & Spiegelhalter, D. (1996). *Markov Chain Monte Carlo in Practice.* Londres: Chapman and Hall.

Goldstein, H. (1987). *Multilevel models in education and social research.* New York:Oxford University Press.

Goldstein, H. (1995). *Multilevel Statistical Models.* Londres: Edward Arnold.

Goldstein, H., Browne, W., & Rasbash, J. (2002). Multilevel modelling of medical data. *Statistics in Medicine, 15*, 3291-3315.

Harmon, M. G., Morse, D. T., & Morse, L. W. (1996). Confirmatory factor analysis of the Gibb Experimental Test of Testwiseness. *Educational and Psychological Measurement, 56*(2), 276-286.

Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition,34*(5),959-972.

Kulik, J. A., Bangert-Drowns, R. L., & Kulik, C. (1984). Effectiveness of coaching for aptitude tests. *Psychological Bulletin, 95*(2), 179-188.

Lambert, P., Sutton, A., Burton, P., Abrams, R., Jones, D. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine, 24*, 2401-2428.

Leyland, A., & Goldstein, H. (2001). *Multilevel Modeling of Health Statistics.* Willey: Chichester.

Martínez-Cardenoso, J., Muniz, J., & García Cueto, E. (2000). Efecto del entrenamiento sobre las propiedades psicométricas de los tests. *Psicothema, 12*(Suppl2) 2000, 363-367.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18* (2), 5-11.

Millman, J., Bishop, C. H., & Ebel, R. (1965). Analysis of test wiseness in the cognitive domain. *Educational and Psychological Measurement, 18*, 787–790.

Ministerio de Educación Pública de Costa Rica (MEP) (2012). *Informe Nacional de Factores Asociados al Rendimiento Académico en las Pruebas Nacionales Diagnósticas, III Ciclo de la Educación General Básica, 2010.* San José, Costa Rica: Dirección de Gestión y Evaluación de la Calidad, Departamento de Evaluación Académica y Certificación, Ministerio de Educación Pública.

Montero, E., Rojas, S., Rodino, A. & Zamora, E. (2012). *Costa Rica en las pruebas PISA 2009 de competencia lectora y alfabetización matemática.* Cuarto Informe Estado de la Educación. San José: Programa Estado de la Nación.

Montero, E. (2014). *Dimensiones de lectura como predictoras de los puntajes en las pruebas PISA-Costa Rica-2009 y PAA de la UCR: Evidencias de regresiones corregidas por estructura multinivel.* Ponencia presenta en las Jornadas de Investigación 2014 del Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica.

Moreira, T. E. (2009). Relación entre factores individuales e institucionales con el rendimiento en matemática: Un análisis multivariado. *Avances en Medición, 7*, 115-128.

Morse, D.T. (1998). The Relative Difficulty of Selected Test-Wiseness Skills among College Students. *Educational and Psychological Measurement, 58*(3), 399-408.

Powers, D. E. (1993). Coaching for the SAT: A summary of the summaries and an update. *Educational Measurement: Issues and Practice, 12*, 24-39.

Powers, D. E., & Rock, D. A. (1999). Effects of coaching on SAT I: Reasoning Test scores. *Journal of Educational Measurement, 36*(2), 93-118.

Raudenbush, S. W., & Bryk, A. S. (2002) *Hierarchical Linear Models. Applications and Data Analysis Methods.* Second Edition. Londres: Sage Publications.

Raudenbush, S.W., & Willms, J.D. (Eds.). (1991). S*chools, Classrooms, and Pupils: International Studies of Schooling from a Multilevel Perspective*. San Diego: Academic Press.

Rice, N., & Jones, A. (1997). Multilevel models and health economics. *Health Economics, 6*, 561-575.

Rice, N., & Leyland, A. (1996). Multilevel models: Applications to health data. *Journal of Health Services Research and Policy, 1*, 154-164.

Rojas, L. (2004). *Factores Asociados a la Repitencia de los y las Estudiantes que Cursan Sétimo Año en Colegios Académicos, Diurnos y Públicos: Un Modelo de Análisis de Niveles Múltiples*. Tesis doctoral. San José, Costa Rica: Universidad Estatal a Distancia.

Samson, G.E. (1985). Effects of Training in Test-Taking Skills on Achievement Test Performance: A Quantitative Synthesis. *The Journal of Educational Research, 78*(5), 261-266.

Sarnacki, R.E. (1979). An Examination of Test-Wiseness in the Cognitive Test Domain. *Review of Educational Research, 49*(2), . 252-279.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin Company.

Shulman, L. S. (2013). *When Coaching and Testing Collide*. Carnegie Perspectives: Current news views and links form the Carnegie Foundation http://www.carnegiefoundation.org/perspectives/when-coaching-and-testing-collide

Snijders, T., & Bosker, R. (1999). *Multilevel Analysis*. Londres: Sage Publications.

Trejos, J.D. (2010). *Indicadores sobre equidad en la educación para Costa Rica. Tercer Informe Estado de la Educación.* San José: Programa Estado de la Nación.

Recibido: 20 de mayo de 2015
Aceptado: 9 de setiembre de 2015

Anexos

Códigos OpenBUGs

Modelo multinivel vacío

```
model{

#Verosimilitud
for(i in 1:n){
diferencia[i]~dnorm(mu[i],tau.u)
mu[i]<-alpha+v[colegio[i]]*cte[i]}
for(j in 1:4){
v[j]~dnorm(0,tau.v)}

#Distribuciones a priori
alpha~dnorm(0,0.001)
tau.u~dgamma(1,0.001)
tau.v~dgamma(1,0.001)

#Cálculo de varianzas y desviaciones típicas y probabilidades
var.u<-1/tau.u
var.v<-1/tau.v
st.u<-sqrt(var.u)
st.v<-sqrt(var.v)
ICC<-var.v/(var.v+var.u)
prob.alpha<-step(alpha)}

#Datos
list(n=139)

diferencia[ ]      colegio[ ]cte[ ]
-6.96      3       1
-1.24      2       1
…
17.89      2       1
END

#Valores iniciales
list(tau.u=1, tau.v=1, v=c(0,0,0,0), alpha=0)
list(tau.u=0.1, tau.v=0.1, v=c(0,0,0,0), alpha=2)
```

Modelo multinivel incluyendo covariables

```
model{
#Verosimilitud
for(i in 1:n){
diferencia[i]~dnorm(mu[i],tau.u)
mu[i]<-
alpha[1]+alpha[2]*sexo[i]+alpha[3]*pad_uni[i]+alpha[4]*Icapital[i]+alpha[5]*Iacceso[i]+alpha[6]*Eactitud[i]+alph
a[7]*Ecomprender[i]+alpha[8]*Eresumir[i]+alpha[9]*primera[i]+alpha[10]*intervencion[i]+v[colegio[i]]*cte[i]}
for(j in 1:4){
v[j]~dnorm(0,tau.v)}

#Distribuciones a priori
alpha[1:10]~dmnorm(a[1:10],A[1:10,1:10])
tau.u~dgamma(1,0.001)
tau.v~dgamma(1,0.001)

#Cálculo de varianzas y desviaciones típicas y probabilidades
var.u<-1/tau.u
var.v<-1/tau.v
st.u<-sqrt(var.u)
st.v<-sqrt(var.v)
for(h in 1:10){
prob.alpha[h]<-step(alpha[h])
}}

#Datos y parámetros a priori
list(n=139,a=c(0,0,0,0,0,0,0,0,0,0),A=structure(.Data c(0.001,0,0,0,0,0,0,0,0,0,0.001,0,0
,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0,0,0,0,0,0,
0.001,0,0,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0,0,0,0,0,0,0.001,0,0,0,0,0
,0,0,0,0,0,0.001), .Dim=c(10,10)))

diferencia[ ] primera[ ] sexo[ ] pad_uni[ ] intervencion[ ] Icapital[ ] Iacceso[ ]
Eactitud[ ] Ecomprender[ ] Eresumir[ ] Colegio[ ] cte[ ]
-6.96   20   1   0   0   0.6   0.63   1.36   0.25   1   3   1
-1.24   14.29   0   1   0   0.4   0.25   1.73   0.5   1   2   1
…
17.89   34.29   1   0   0   0.3   0.75   1.82   0.75   1   2   1
END

#Valores iniciales
list(tau.u=1, tau.v=1, v=c(0,0,0,0), alpha= c(0,0,0,0,0,0,0,0,0,0))
list(tau.u=0.1, tau.v=0.1, v=c(0,0,0,0), alpha= c(2,2,2,2,2,2,2,2,2,2))
```