



CUARTO INFORME DEL ESTADO DE LA EDUCACIÓN

Aporte Especial

Clasificación de los colegios con base en indicadores de la calidad educativa

Investigador:

Luis Rojas Torres

2013



CONSEJO NACIONAL DE RECTORES

Nota: Las cifras de las ponencias pueden no coincidir con las consignadas por el IV Informe Estado de la Educación en el tema respectivo, debido a revisiones posteriores. En caso de encontrarse diferencia entre ambas fuentes, prevalecen las publicadas en el Informe.

Contenido

Resumen Ejecutivo	3
Introducción.....	4
Materiales y método	5
Objetivos y metodología	5
Datos	6
Método	8
Resultados	9
Construcción de los indicadores	9
Agrupación de los colegios.....	11
Validación	15
Conclusiones	18
Bibliografía.....	19
Anexos.....	21
Notas	24

Resumen Ejecutivo

Este estudio fue realizado en el 2013 con el objetivo de clasificar los colegios de Costa Rica con base en indicadores de las dimensiones de calidad educativa. Mediante un análisis de conglomerados con indicadores de las dimensiones de contexto, eficacia y procesos se logró determinar 5 grupos en que se clasifican los colegios costarricenses: *rezagados*, *promedios*, *eficaces con condiciones deficientes*, *emprendedores* y *eficaces con condiciones óptimas*. La variable conglomerado de clasificación del colegio presentó una correlación eta con el índice de calidad educativa de 0.85, dando evidencias de la validez de esta clasificación para determinar los grupos en que se dividen las instituciones educativas de secundaria. Se observó que en el conglomerado *eficaces con condiciones óptimas* está compuesto en casi un 99% por colegios privados y científicos, en contraparte con el grupo *rezagados* que no presenta ni un 4% de instituciones privadas. Esta clasificación permite identificar situaciones específicas asociadas a la calidad educativa en grupos de colegios, lo cual le da al Ministerio de Educación Pública un panorama de qué es prioritario trabajar en cada uno de los conglomerados.

Descriptores

Indicadores de calidad educativa, colegios públicos, colegios privados, análisis de conglomerados.

Introducción

Son muchos los resultados que apuntan a que hay una diferencia en la calidad educativa de los colegios de Costa Rica, por ejemplo en el Tercer Estado de la Educación se indica que los colegios técnicos presentan mayor porcentaje de aprobación que los académicos, que las telesecundarias, los liceos rurales y los colegios nocturnos tienden a tener menor rendimiento que el resto de secundarias, que las instituciones educativas de la zona central presentan mejor desempeño que los de la periferia, que la deserción y la repitencia son mucho mayores en los colegios nocturnos que en los demás tipos de secundaria o que el acceso a recursos tecnológicos es desigual entre los colegios de la zona central con respecto a los de otras zonas (Programa Estado de la Nación, 2011).

Los resultados anteriores nos hablan de varios aspectos que al verlos por separado dan señales de múltiples problemas en los colegios costarricenses, pero no muestran por dónde empezar a trabajarlos. Quizás una de las formas más eficientes para determinar hacia dónde dirigir las políticas educativas, es establecer una clasificación de los colegios que indique cuales tienen más deficiencias en la calidad educativa.

En este trabajo se pretende clasificar los colegios de Costa Rica según los indicadores de la calidad educativa, lo cual brinda al Ministerio de Educación Pública una visión sintetizada de la situación de las instituciones educativas de secundaria y le da una herramienta sobre la cual diseñar estrategias que ayuden a mejorar la educación costarricense.

En cuanto a las definiciones de calidad educativa, Toranzos (2010) dice que un sistema educativo de calidad es “aquel que garantiza oportunidades de aprendizaje equivalentes para todos los alumnos, es decir aquel que se preocupa por generar una conjunción favorable entre los factores externos e internos a la escuela para el desarrollo de resultados exitosos en términos de aprendizaje” (p.47).

Por su parte, el considerando C, punto 1D de la Política Costarricense Hacia el Siglo XXI define proceso educativo de calidad a aquel que “involucra el proceso y los resultados de acuerdo con los Fines [de la Educación]. Además, ofrece igualdad de oportunidades para el éxito a quienes participan en él y propone una oferta coherente con las necesidades, problemas y aspiraciones de los participantes en relación con su entorno” (MEP, 1994, p.5).

Otra definición de calidad de la educación dice que esta es “la vinculación que debe existir entre el qué, el cómo, y el para qué, con base en la coherencia interna de un eje vertical constituido por lo filosófico, lo teórico y lo práctico, en donde lo uno sirve de sustrato a lo otro, y por la articulación de ese eje vertical con un horizontal conformado por la coherencia interna entre planes y programas; recursos didácticos, la mediación del docente (didáctica) y la intermediación gerencial” (Política Educativa Hacia el Siglo XXI, 2010, p.83).

Finalmente, MEP (2012) define calidad de la educación como la “satisfacción de las necesidades educativas de la persona, la comunidad y la sociedad, mediante un proceso que potencie con equidad el desarrollo humano y la identidad nacional” (p.12).

A excepción de la tercera definición dada por Política Educativa Hacia el Siglo XXI (2010), todas las definiciones son muy similares entre sí, haciendo énfasis principalmente en el alcance de un aprendizaje óptimo por parte del estudiante y tomando en cuenta la igualdad de oportunidades. La última definición tiene como aspecto sobresaliente, que lleva la calidad de la educación a un nivel de importancia para el País, ya que no solo debe satisfacer las necesidades del individuo, sino también las de la comunidad y la sociedad; es por esto que en este trabajo se decidió adoptar esta, como definición de la calidad de la educación.

Por otro lado, basándose en MEP (2011) y Toranzos (2010) se pueden distinguir cuatro dimensiones de la calidad educativa:

- Eficacia: Asociada a resultados de la institución como deserción, aprobación y sobriedad.
- Procesos: Relacionada con los medios que el sistema proporciona para el proceso educativo como adecuaciones, formación de los docentes e infraestructura.
- Papel del estudiante, el docente y la comunidad: Trata sobre las funciones de cada uno de estos actores en la educación.
- Contexto: Abarca las características sociales, económica y culturales del contexto donde está ubicada la institución.

En cuanto a mediciones de la calidad educativa que se han sido realizado anteriormente, se puede mencionar el índice SNED utilizado en Chile (SNED, 2011), los indicadores educativos de los Estados Unidos Mexicanos (Tamez, Zúñiga y Martínez, 2006) y el índice de situación educativa (Murillo, 2011). El primero es poco aplicable a Costa Rica debido a la ventaja que tiene Chile sobre Costa Rica en el campo de la recolección de estadísticas de colegios (SNED, 2011), y los otros dos son desarrollados para evaluar regiones, no colegios.

Materiales y método

Objetivos y metodología

El objetivo principal de este trabajo es clasificar a los colegios de Costa Rica según indicadores asociados a las dimensiones de la calidad educativa. Para lograr este objetivo es necesario considerar una serie de objetivos específicos que permitan la consecución del principal.

Como primer objetivo específico se plantea el desarrollar un indicador para cada una de las dimensiones de la calidad educativa. Para su alcance se realiza un análisis factorial

exploratorio con las variables más representativas de cada dimensión de la calidad educativa, a partir del cual se identifican los factores asociados a las dimensiones de la calidad educativa y se les estima las puntuaciones factoriales, las cuales funcionan como indicadores de cada una de las dimensiones (para la dimensión *papel del estudiante, el docente y la comunidad* no se incluyeron variables en el análisis factorial, debido a que no se tienen variables asociadas a esta, por tanto esta dimensión no se trabaja en este estudio).

El segundo objetivo específico es establecer la cantidad óptima de grupos en que se deben clasificar los colegios. Este objetivo se trabaja analizando cuál es la cantidad de grupos que brinda más información al realizar varios análisis de conglomerados con distintas cantidades de grupos.

Como tercer objetivo específico se plantea diseñar y caracterizar los grupos de colegios según los indicadores de la calidad educativa, para esto se realiza un análisis de conglomerados con la cantidad óptima de grupos, posteriormente se analizan las medias de los indicadores de la calidad educativa en cada uno de estos grupos y por último, se estudia la cantidad de colegios en los conglomerados según zona (urbana o rural) y dependencia (pública o privada).

Finalmente, el cuarto objetivo específico es hallar evidencias sobre la validez de los resultados, para esto se calculan las correlaciones eta de la variable *conglomerado de pertenencia* con los indicadores de la calidad educativa y un índice de calidad educativa y además, con el fin de corroborar los resultados, se vuelve a realizar todos los análisis con la mitad de la muestra.

Datos

En este estudio se trabajó con los datos de colegios para el año 2009, que se presentan en la Megabase de datos georreferenciados de secundaria elaborada por el Estado de la Nación, basadas en las estadísticas recolectadas por el MEP al finalizar los ciclos lectivos. Además, se utilizaron datos asociados al índice de desarrollo social (IDS) del 2007 y a la matrícula en la Universidad de Costa Rica en el proceso 2009-2010.

En el año 2009 la Megabase presenta datos para 810 colegios, los cuales según Murillo (2011b) eran el total de colegios de Costa Rica para ese año, de estos 596 eran públicos y 210, privados; por otro lado, 396 eran rurales y 414, urbanos. Dentro de este grupo de colegios no se incluyen instituciones educativas que brindan Educación Abierta u otros tipos de educación no convencional, como es el caso de los CINDEA o los CONED. Además, hubo varios colegios con valores perdidos en ciertas variables, los cuales fueron recodificados mediante el uso de procesos estadísticos que se presentan en el anexo 1, con esta recodificación se logró tener 786 colegios sin valores perdidos en las variables de interés (más del 97% del total).

Con base en los resultados obtenidos en Rojas (2013), se decidió utilizar como variables asociadas a la dimensión procesos a las cantidades de: aulas de asignaturas

especiales, bibliotecas, comedores, laboratorios de cómputo, pizarras, computadoras para personal administrativo, administrativos, docentes administrativos y profesores titulados. Todas estas se presentan como razones con respecto a la cantidad de estudiantes matriculados en el colegio.

Los docentes titulados son aquellos que cuentan con un título para dar clases en secundaria, el personal docente-administrativo incluye al director, asistente de dirección y labores afines, mientras que el personal administrativo incluye personas que trabajan en labores que no requieren un título profesional, como conserjes, cocineros, secretarios y jardineros.

Con respecto a variables asociadas a la dimensión eficacia se utilizaron las siguientes: a) proporción de aprobación en bachillerato con respecto a la cantidad de estudiantes que presentan esas pruebas (bachillerato), b) proporción de personas que no están repitiendo el año lectivo con respecto a la matrícula total (flujo), c) proporción de personas matriculadas inicialmente que no abandonaron el colegio (retención) y d) proporción de personas matriculadas inicialmente que aprobaron el ciclo lectivo (aprobación).

Para la dimensión contexto se utilizaron los indicadores de salud, economía y educación del 2007, utilizados por el MIDEPLAN para elaborar el IDS del distrito de ubicación del colegio. Las estadísticas descriptivas asociadas a las variables utilizadas en este estudio se presentan en el cuadro 1.

Por último, debido a que las variables utilizadas en este estudio están medidas de diversas formas, se decidió estandarizarlas con el fin de tener una unidad de medida común en todas las variables.

La estandarización se realizó mediante la conversión de las variables a otras con media igual a 0 y desviación estándar igual a 1, lo cual resultó de aplicar la fórmula

$$Z = \frac{X - \bar{X}}{s}$$

donde \bar{X} es la media de la variable X y s su desviación estándar.

Cuadro 1
Estadísticas descriptivas de las variables usadas en el estudio. Año 2009

Variable	N	Media	desviación estándar	mínimo	máximo
Aulas de asign. espec.	796	0.00	0.02	0.00	0.33
Biblioteca	796	0.00	0.01	0.00	0.33
Comedor	796	0.00	0.02	0.00	0.33
Lab. de cómputo	796	0.00	0.02	0.00	0.33
Pizarras	796	0.07	0.08	0.00	1.00
Comp. para admin.	796	0.02	0.03	0.00	0.67
Docente-admin.	796	0.02	0.03	0.00	0.67
Administrativos	796	0.02	0.04	0.00	0.67
Titulados	796	0.11	0.15	0.00	3.00
IDS-económico	810	26.44	18.00	0.00	100.00
IDS-salud	810	60.58	21.06	0.00	100.00
IDS-educación	810	65.46	21.54	0.00	100.00
Bachillerato	689	0.66	0.25	0.00	1.00
Flujo	796	0.77	0.08	0.52	1.00
Retención	753	0.90	0.11	0.36	1.00
Aprobación	760	0.82	0.14	0.32	1.00

N=Cantidad de datos omitiendo los valores perdidos.

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria.

MIDEPLAN. (2007). Índice de Desarrollo Social.

Método

La técnica estadística sobre la cual se basa este estudio se denomina análisis de conglomerados, esta tiene como objetivo principal la agrupación de datos en un número reducido de grupos, a partir de sus datos en ciertas variables, mediante un algoritmo que permite conformar grupos con los casos más similares entre sí y a la vez, buscando que estos sean diferentes con respecto a los casos que conforman los otros grupos (Cea, 2002).

Para este trabajo se utilizó el método denominado k-medias, el cual según Cea (2002) es el más utilizado de los métodos de análisis conglomerados no jerárquicos (caracterizados principalmente por la reasignación de casos en el proceso de estimación de los grupos, esto es, que si conforme se van formando los grupos resulta que un caso ya incorporado en un grupo A presenta mejor identificación con el grupo B, este puede ser movido al grupo B).

El proceso de k medias básicamente sigue el siguiente algoritmo: a) definir la cantidad K de grupos que se desean formar, b) definir K centroides (medias de los grupos en las

variables usadas para crear los conglomerados) iniciales (si no se especifican, el programa puede utilizar el conjunto de medias de los primeros K datos), c) calcular la distancia de los casos a los centroides, d) determinar la distancia más pequeña presentada entre un caso y un centroide, que no sean del mismo grupo, e) asignar ese caso al grupo con ese centroide, f) reestimar el centroide de ese grupo, g) repetir los pasos c) al f) hasta lograr que los casos de cada conglomerado presenten una distancia al centroide de este, más pequeña que a los otros centroides (Cea, 2002).

Además, en este estudio se utilizó el análisis factorial exploratorio para estimar los indicadores de cada dimensión. En esta técnica, se utilizó el método de componentes principales con rotación VARIMAX. Como criterio de selección de los factores se utilizó que las raíces características de estos fueran mayores a 1 y como regla para considerar que una variable cargaba en un factor se estableció que la carga en el factor de identificación fuera mayor a 0.60 y en los demás factores su carga fuera menor a 0.35.

Resultados

Construcción de los indicadores

Para la construcción de los indicadores se elaboró un análisis factorial exploratorio con las variables presentadas en la sección de materiales y método. Este análisis dio como resultado, la agrupación de las variables en tres factores y una explicación de la variancia de los datos del 71.2%; además presentó estadísticos de bondad de ajuste buenos, como un índice KMO de .913 y un valor-p asociado a la prueba de esfericidad de Bartlett de .000 (Hernández, 2002). En el cuadro 2 se presentan las cargas factoriales de cada una de las variables en los factores obtenidos.

Se puede observar que las variables asociadas previamente a una dimensión específica de la calidad educativa, se agruparon entre sí determinando un factor asociado a esa dimensión.

Los indicadores de las dimensiones fueron elaborados estimando las puntuaciones factoriales de cada uno de los factores. En el indicador de la dimensión procesos, la variable presentó una distribución bastante asimétrica, por tanto se decidió aplicarle la transformación logaritmo natural, con la cual se logró normalizar el indicador.

Posteriormente, estas puntuaciones fueron estandarizadas con la fórmula

$$Z = \frac{X - \min(X)}{\max(X) - \min(X)}$$

donde X representa el indicador, min significa mínimo y max significa máximo.

En el cuadro 3 se presentan las estadísticas descriptivas de los indicadores asociados a las dimensiones de la calidad educativa. Se observa que en los tres indicadores el mínimo está dado por el valor cero y el máximo por el uno, lo cual es debido a la estandarización utilizada.

En el anexo 2 se presentan varias evidencias de la validez de estos indicadores.

Cuadro 2
Cargas factoriales de las variables en los componentes rotados

Variables	Factor			Componente con carga > 0.60
	Procesos	Contexto	Eficacia	
Aulas de asign. esp.	0.799	0.137	0.017	1
Biblioteca	0.916	0.015	0.014	1
Comedor	0.885	-0.077	0.021	1
Lab. de cómputo	0.905	0.056	0.066	1
Comp. para admin.	0.890	0.147	0.164	1
Pizarras	0.740	0.028	0.234	1
Titulados	0.885	-0.047	0.164	1
Administrativos	0.926	0.088	0.132	1
Docente-admin.	0.796	0.268	0.198	1
IDS-económico	0.099	0.799	0.147	2
IDS-salud	0.068	0.909	0.066	2
IDS-educación	0.051	0.882	0.076	2
Flujo	0.115	-0.040	0.761	3
Retención	0.074	0.206	0.718	3
Aprobación	0.115	-0.036	0.673	3
Bachillerato	0.106	0.258	0.685	3

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Cuadro 3
Estadísticas descriptivas de los indicadores utilizados en el Índice de Calidad Educativa Año 2009

	Media	Mínimo	Máximo	Desviación estándar
Contexto	.533	.000	1.000	.179
Eficacia	.657	.000	1.000	.187
Procesos	.455	.000	1.000	.136

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Agrupación de los colegios

Primeramente se analizaron las características de los grupos determinados por la división de los colegios en dos conglomerados, este análisis produjo dos grupos diferentes entre sí, principalmente en el promedio del indicador de la dimensión procesos. Luego, se dividió la muestra en tres conglomerados con el fin de obtener más grupos diferentes entre sí que los hallados en la separación de dos conglomerados, y así sucesivamente, se hizo hasta analizar la división en seis conglomerados. En esta última división se presentaron dos grupos muy similares entre sí, por lo que se decidió utilizar un análisis de conglomerados de cinco grupos. En el cuadro 4 se presentan las medias obtenidas por los colegios de cada uno de los conglomerados, en los indicadores de las dimensiones de la calidad educativa y la cantidad de colegios presentes en cada grupo.

Cuadro 4
Medias de los indicadores de la calidad educativa y total de colegios según conglomerado. Año 2009

Conglomerado	Indicador			Cantidad de colegios
	Contexto	Eficacia	Procesos	
1	0.643	0.386	0.440	133
2	0.406	0.538	0.476	161
3	0.621	0.717	0.324	186
4	0.337	0.775	0.453	161
5	0.690	0.831	0.616	145

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Con el fin de caracterizar de una manera más sencilla cada uno de los grupos, se decidió asignar nombres a los rangos de valores tomados por los promedios de los indicadores. Al rango de 0 a .20 se le denominó muy bajo; al de .20 a .40, bajo; al de .40 a .60, medio; al de .60 a .80, alto y al de de .80 a 1, muy alto. Los nombres asignados a los rangos transforman los promedios en variables ordinales de cinco categorías, que son muchas sencillas de interpretar que las variables continuas. En el cuadro 5 se presentan los nombres de los rangos a los que pertenecen los promedios obtenidos por los conglomerados en cada uno de los indicadores.

El primer grupo es el único grupo con un promedio bajo en la dimensión eficacia, lo cual indica que los colegios de este grupo presentan problemas en sus resultados, es por esto que a este grupo se le denomina *rezagados*. Por otro lado, los colegios de este conglomerado presentan un promedio alto en el contexto y uno medio en los recursos, lo cual no es concordante con su baja eficacia, por tanto no son solamente rezagados con el resto de colegios, sino también con las expectativas creadas por las condiciones en las que se encuentran.

Con respecto al segundo conglomerado, se obtuvo que los promedios de sus tres indicadores presentaron valores dentro del rango medio, por lo cual se le asignó el nombre *promedios*.

Luego, al tercer grupo se le denominó *eficaces con condiciones deficientes*, ya que presentan una eficacia promedio alta, pero son el único grupo que presenta un promedio bajo en la dimensión procesos.

Al cuarto grupo se le asignó el nombre de *emprendedores* porque a pesar de las situaciones adversas logran resultados favorables, esto se concluye al ver que son el único grupo con una puntuación baja en el promedio del indicador de la dimensión contexto, pero a pesar de esto, presenta un promedio alto en el indicador de la dimensión eficacia.

Finalmente, al quinto grupo se le denominó *eficaces con condiciones óptimas*, ya que en todos sus indicadores se presentan promedios altos, de hecho son el único colegio que obtiene un promedio alto en procesos y uno muy alto en eficacia.

Cuadro 5
Valor ordinal de los promedios de los conglomerados en los indicadores de la calidad educativa. Año 2009

Conglomerado	Indicador		
	Contexto	Eficacia	Procesos
Rezagados	Alto	Bajo	medio
Promedios	Medio	Medio	medio
Eficaces con condiciones deficientes	Alto	Alto	bajo
Emprendedores	Bajo	Alto	medio
Eficaces con condiciones óptimas	Alto	muy alto	alto

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Para analizar las características de los colegios que componen cada uno de los grupos formados por el análisis de conglomerados, se estudió la cantidad de colegios según zona urbana o rural, que componen cada uno de estos grupos. Se obtuvo que los colegios urbanos se encuentran principalmente en los grupos denominados *rezagados*, *eficaces con condiciones deficientes* y *eficaces con condiciones óptimas*, mientras que los colegios de las zonas rurales se agrupan mayoritariamente en los grupos *promedios* y *emprendedores*. En el cuadro 6 se presenta la cantidad de colegios en cada conglomerado según zona de ubicación.

Cuadro 6
Cantidad de colegios por conglomerado según zona de ubicación. Año 2009

Zona	Conglomerado					Total
	Rezagados	Promedios	Eficaces con condiciones deficientes	Emprendedores	Eficaces con condiciones óptimas	
Rural	32	141	53	148	11	385
Urbano	101	20	133	13	134	401
Total	133	161	186	161	145	786

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

En cuanto a la cantidad de colegios según su dependencia, se obtuvo que más del 65% de colegios privados se ubica en el conglomerado *eficaces con condiciones óptimas*, mientras que ni un 3% de los colegios públicos se encuentra en esta categoría. Por otro lado, el 22% de los colegios privados se ubica en la categoría *eficaces con condiciones deficientes*.

En los grupos *rezagados*, *promedios* y *emprendedores* más del 90% de los colegios son públicos, mientras que en la dimensión *eficaces con condiciones óptimas* más del 90% son privados. En el conglomerado *eficaces con condiciones deficientes* se tiene que más del 75% de los colegios son públicos. En el cuadro 7 se presentan la cantidad de colegios por conglomerado según dependencia.

Cuadro 7
Cantidad de colegios por conglomerado según dependencia. Año 2009

Zona	Conglomerado					Total
	Rezagados	Promedios	Eficaces con condiciones deficientes	Emprendedores	Eficaces con condiciones óptimas	
Privados	5	3	44	15	129	196
Públicos	128	158	142	146	16	590
Total	133	161	186	161	145	786

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Finalmente, con el fin de analizar la relación de esta clasificación con la calidad educativa, se decidió construir un índice asociado a este constructo. El índice de calidad educativa (ICE) fue definido como el promedio de los indicadores de las dimensiones contexto, eficacia y procesos, y en el anexo 2 se presentan evidencias de su validez.

En el cuadro 8 se presentan unos estadísticos de asociación entre los indicadores del ICE y la variable categórica compuesta por los grupos creados por el análisis de conglomerados. En esta tabla se puede ver que hay una alta correlación eta entre los conglomerados y el ICE, la cual es de .846. Por otro lado, los indicadores de cada dimensión presentan correlaciones eta con los conglomerados mayores a 0.6, mostrando la asociación esperable entre los conglomerados y las variables utilizadas para formarlos.

Cuadro 8
Estadísticos de asociación entre los indicadores del ICE y los conglomerados. Año 2009

Indicador	Eta	Eta Cuadrado	P ANOVA
ICE	.846	.715	.000
Contexto	.792	.628	.000
Eficacia	.843	.710	.000
Procesos	.696	.485	.000

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

La prueba ANOVA indica que hay diferencias significativas al 5% entre los cinco conglomerados, en los promedios de los indicadores asociados a las dimensiones del ICE y en los promedios del ICE.

Por medio del análisis post-hoc de la prueba de Tuckey (ver cuadro 9), se evidenció que el conglomerado Eficaces con condiciones óptimas presenta ventajas significativas al 1% en todos los indicadores del ICE, con respecto a los demás conglomerados. Luego, el conglomerado Emprendedores al ser comparado con los conglomerados Eficaces con condiciones deficientes, Promedios y Rezagados, presenta ventajas significativas al 1% en el ICE y la Eficacia, pero presenta desventajas significativas al 1% en el indicador de Contexto. Por otro lado, los colegios del conglomerado Eficaces con condiciones deficientes presentan desventajas en el indicador de procesos, significativas al 1% con los otros cuatro conglomerados y los rezagados presentan la misma tendencia en el indicador de Eficacia.

Cuadro 9
Resultados de la Prueba de Tuckey asociados a la diferencia de medias de los indicadores del ICE entre conglomerados

Conglomerado i	Conglomerado j	Diferencia de medias i - j			
		Contexto	Eficacia	Procesos	ICE
Efic. con cond. ópt.	Emprendedores	.353**	.056**	.163**	.191**
	Efic. con cond. def.	.069**	.115**	.293**	.159**
	Promedio	.294**	.293**	.140**	.243**
	Rezagados	.047**	.446**	.176**	.223**
Emprendedores	Efic. con cond. def.	-.284**	.059**	.130**	-.032**
	Promedios	-.059**	.237**	-.023	.052**
	Rezagados	-.306**	.390**	.013	.032**
Efic. con cond. def.	Promedios	.225**	.179**	-.152**	.084**
	Rezagados	-.022	.331**	-.117**	.064**
Promedios	Rezagados	-.248**	.152**	.035*	-.020*

** significativo al 1%, * significativo al 5%

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria.

MIDEPLAN. (2007). Índice de Desarrollo Social.

Por otro lado, el eta cuadrado indica que porcentaje de la variancia de los indicadores es explicada por los conglomerados, se puede concluir que el 71.5% de la variancia del índice de calidad educativa es explicada por los conglomerados conformados en este análisis. El indicador de eficacia es explicado en un 71.0%, el de contexto en un 62.8% y el de procesos en un 48.5%.

Validación

Con el fin de dar evidencias de la validez de los resultados se realizó el análisis de conglomerados con una selección aleatoria de aproximadamente el 50% de los casos (386 colegios). Para este análisis se trabajó con 5 conglomerados, los cuales presentaron estructuras muy similares a las de los hallados con toda la población. En el cuadro 10 se presentan las medias de los indicadores en cada uno de los conglomerados y en el cuadro 11 se presentan estos valores recodificados en la escala ordinal utilizada en el análisis de conglomerados previo.

Cuadro 10
Medias de los indicadores de la calidad educativa según conglomerados de la submuestra. Año 2009

Conglomerado	Indicador		
	Contexto	Eficacia	Procesos
1	.607	.366	.455
2	.411	.811	.385
3	.661	.664	.338
4	.353	.603	.501
5	.672	.844	.635

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria.
 MIDEPLAN. (2007). Índice de Desarrollo Social.

Cuadro 11
Valor ordinal de las medias de los indicadores de la calidad educativa según conglomerados de la submuestra. Año 2009

Conglomerado	Indicador		
	Contexto	Eficacia	Procesos
1	alto	bajo	medio
2	medio	muy alto	bajo
3	alto	alto	bajo
4	bajo	alto	medio
5	alto	muy alto	alto

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria.
 MIDEPLAN. (2007). Índice de Desarrollo Social.

Los conglomerados 1, 3, 4 y 5 respectivamente, presentaron valores ordinales de sus medias, exactamente iguales a los de los conglomerados denominados rezagados, eficaces con condiciones deficientes, emprendedores y eficaces con condiciones óptimas. Únicamente el conglomerado 2, no presentó un conjunto de medias similar a la de algún grupo obtenido en el análisis con toda la muestra.

Si se compara el conglomerado 2 con el conglomerado *promedios* obtenido en el análisis con toda la muestra se obtiene que solamente coinciden en que el contexto tiene un promedio medio. Siendo un poco laxo, se puede decir que en el conglomerado 2 el indicador de procesos tiene un promedio cercano a la magnitud media (le falta .015 para ser considerado en esa categoría), por lo cual este grupo ya tendría dos semejanzas con el conglomerado *promedios*. La gran diferencia entre ambos conglomerados está dada por el indicador de eficacia, ya que en esta submuestra presenta un promedio muy alto, mientras que con toda la muestra el promedio es de magnitud media.

Finalmente, en el cuadro 12 se presenta la cantidad de colegios en los conglomerados según dependencia y el cuadro 13 se presenta según zona. En estos cuadros se observa que la distribución es muy similar a la observada en la muestra total, ya que en el conglomerado 5 (análogo al conglomerado *eficaces con condiciones óptimas*) la mayoría de colegios son privados y urbanos.

Por otro lado, en el conglomerado 3 se agrupa la segunda cantidad más grande de colegios privados y en los conglomerados 1, 2, 3 y 4 la mayoría de colegios son públicos. Además, los colegios rurales se agrupan en los grupos 2 y 4, mientras que los urbanos se agrupan en el 1, 3 y 5. Lo anterior, cambiando de nombre a los conglomerados de la submuestra por sus asociados en los obtenidos con toda la muestra, resulta muy similar a lo obtenido en el análisis de conglomerados original.

Cuadro 12

Cantidad de colegios por conglomerados de la submuestra según dependencia. Año 2009

Dependencia	Conglomerado					Total
	1	2	3	4	5	
Privado	2	13	21	3	46	85
Público	55	71	73	96	6	301
Total	57	84	94	99	52	386

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Cuadro 13

Cantidad de colegios por conglomerados de la submuestra según zona de ubicación. Año 2009

Zona	Conglomerado					Total
	1	2	3	4	5	
Rural	22	60	24	89	5	200
Urbana	35	24	70	10	47	186
Total	57	84	94	99	52	386

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Otra forma con la cual se validaron los resultados, fue por medio de la aplicación de un análisis discriminante, el cual es una técnica estadística que permite asignar casos a grupos establecidos a partir de características dadas. En el caso de este trabajo, se evaluó si el análisis discriminante basado en la información de los indicadores del ICE asignó los colegios a los conglomerados que han sido trabajados a lo largo de este estudio; se obtuvo que el análisis discriminante clasificó un 97.2% de los casos en los mismos grupos que lo hizo el análisis de conglomerados, proveyendo una nueva evidencia de validez de los resultados de este estudio.

Conclusiones

Primeramente, mediante el análisis de conglomerados se logró determinar cinco grupos que presentaron diferencias marcadas entre sí, lo anterior permite clasificar a los colegios costarricenses de una manera más informativa que sencillamente hacer una diferenciación por ubicación en los quintiles del ICE.

Es importante destacar que estos grupos parecen tener como criterio de formación de los conglomerados a las variables zona y dependencia, ya que cuando se analiza conglomerado por conglomerado hay una notable mayoría de alguna de las categorías de cada una de estas dos variables. Esto parece indicar que la clasificación está bien hecha, ya que variables no utilizadas para agrupar, presentan notables diferencias entre los grupos.

Por otro lado, el análisis de conglomerados realizado con la mitad de la muestra brindó aún más evidencias de la validez de los resultados, ya que el análisis de la submuestra fue bastante similar al obtenido con toda la muestra.

Como evidencia final de la validez de los resultados se destaca la correlación entre el conglomerado de pertenencia con el ICE, la cual fue de .846. Esta correlación presenta una magnitud bastante alta, dando una idea de que esta clasificación representa muy bien la calidad educativa.

Ahora, con respecto a los grupos formados se presentan resultados importantes, por ejemplo los colegios *eficaces con condiciones óptimas* no están repartidos de manera aleatoria en el País, sino que son instituciones que en su gran mayoría son de zonas urbanas y de dependencia privada, de hecho los colegios públicos que se ubican en esta categoría son en su mayoría científicos.

Por otro lado, resultó interesante la conformación del grupo *emprendedores*, el cual es un grupo de colegios que a pesar de sus condiciones adversas presentaron una eficacia alta. Este grupo es muy similar al conglomerado *promedios*, ya que ambos presentan una mayoría de colegios públicos y de colegios rurales, su diferencia principal radica en su eficacia. Queda pendiente para otro estudio, determinar a qué se debe la diferencia en el promedio de su eficacia.

El grupo *eficaces con condiciones deficientes* muestra que hay un grupo de colegios de Costa Rica que cuentan con servicios no deseables para el desarrollo del proceso educativo, acá se pueden mencionar situaciones asociadas a las variables de la dimensión procesos, como falta de profesores titulados o exceso de población en comparación con la cantidad de materiales presentes en la institución. La segunda situación se conjetura como la hipótesis más fuerte, cuando se observa que la mayoría de colegios de este grupo son de la zona urbana y de dependencia pública.

El otro grupo formado fue el de colegios *rezagados*, el cual contrariamente a lo esperado estuvo conformado por colegios del área urbana. Este grupo presenta problemas serios de eficacia a pesar de ubicarse en zonas con un indicador alto de contexto y un indicador medio de procesos. Con el fin de analizar la situación de este

grupo se observó que los cantones con más colegios en este conglomerado son Desamparados y el central de San José, con 13 y 14 instituciones, respectivamente y luego sigue el cantón central de Limón con 6 colegios. En estas zonas a pesar de que el indicador de contexto es alto, su nivel de desarrollo humano para el 2009 resultó débil o desfavorable (PNUD, 2011).

Se puede establecer como hipótesis que los problemas con el desarrollo humano que enfrentan estas regiones pueden estar ligados a su calidad educativa, este nivel de desarrollo humano contempla aspectos ligados con la pobreza, el género y la seguridad.

Finalmente, se recomienda utilizar esta clasificación de los colegios para atender las necesidades específicas de cada uno de estos grupos con el fin de mejorar su calidad educativa. Queda pendiente realizar este trabajo tomando en cuenta un indicador que mida la dimensión *papel del estudiante, el docente y la comunidad*, para la cual se debe iniciar con la tarea de recolectar estadísticas asociadas a esta.

Bibliografía

- Cea, M. A. (2002). Análisis multivariable. Teoría y práctica en la investigación social. Madrid: Editorial Síntesis. [230-260]
- Da Silveira, P. (1995). La segunda reforma. Montevideo: Fundación Banco de Boston. [82]
- Hernández, O. (1998). Temas de análisis estadístico multivariado. San José: Editorial Universidad de Costa Rica. [capítulo 5]
- MEP. (1994). Política Educativa Hacia el Siglo XXI. San José: MEP. [5]
- MEP. (2011). Modelo de evaluación de la calidad de la educación. Costa Rica: Ministerio de Educación Pública. [4]
- MEP. (2012). La calidad de la educación es un asunto de todos. Costa Rica: Ministerio de Educación Pública. [12]
- Murillo, D. (2011). Índice de Situación Educativa. En Estado de la Nación (Ed.). III Estado de la Educación (293-296). San José: Programa Estado de la Nación. [293-296]
- Murillo, D. (2011b). Megabases de datos georreferenciados para la educación primaria y secundaria de los centros educativos de Costa Rica. (Informe Final de Investigación). San José: Estado de la Nación. [10]
- PNUD. (2011). Atlas del desarrollo humano cantonal 2011. San José: PNUD. [Capítulo 7]

- Política Educativa Hacia el Siglo XXI. (2010). Definición de calidad de la educación dentro de la política educativa aprobada por el Consejo Superior de Educación. En M. Marín (Ed.). Antología de calidad (83-85). Costa Rica: Ministerio de Educación Pública. [83]
- Programa Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria [base de datos]. Costa Rica: PEN.
- Programa Estado de la Nación. (2011). Tercer Informe Estado de la Educación. Costa Rica. PEN. [117-136]
- Rojas, L. (2013). Construcción de un índice de calidad educativa de colegios. Manuscrito sin publicar. [1-15]
- SNED. (2011). Sistema Nacional de Evaluación del Desempeño de los Establecimientos Educativos Subvencionados y de los Regidos por el Decreto Ley N° 3166. Chile: SNED y Departamento de Estudios y Desarrollo de la División de Planificación y Presupuesto del Ministerio de Educación de Chile. [Todo el documento]
- Tamez, R., Zúñiga, L. y Martínez, F. (2006). Sistema de indicadores educativos de los Estados Unidos Mexicanos. México: SEP-INEE. [Todo el documento]
- Toranzos, L. (2010). En la búsqueda de estándares de calidad. En M. Marín (Ed.). Antología de calidad (46-53). Costa Rica: Ministerio de Educación Pública. [247]
- UNESCO. (1999). Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados en Tercero y Cuarto Grado. Disponible en <http://unesdoc.unesco.org/images/0012/001231/123143s.pdf>, recuperado el 8 de marzo del 2013. [32-33]

Anexos

ANEXO 1: Operacionalización de los valores perdidos

Debido a que aproximadamente el 15% de los registros no presenta registrada el porcentaje de aprobación en bachillerato, se procedió a estimar un modelo de regresión lineal para calcular una aproximación a esos valores perdidos. Este modelo tiene como variable dependiente a bachillerato y como independientes al horario de atención del colegio (diurno o nocturno), la dependencia (privada o pública), el flujo, la retención y la aprobación. Los coeficientes obtenidos por este modelo se presentan en el cuadro 14.

Todas las variables resultaron significativas al 5%, a excepción del horario de la institución, por otro lado las variables con mayor importancia relativa (coeficiente beta) en la estimación del porcentaje de aprobación en bachillerato fueron la dependencia del colegio y el porcentaje de aprobación. Este modelo explicó un 30,53% de la variancia de bachillerato, por lo cual se consideró apropiado para estimar los valores perdidos en la variable dependiente.

Con la estimación de los valores perdidos del porcentaje de aprobación en bachillerato, se logró tener valores de esta variable para 795 colegios. Luego, se procedió a ordenar los datos en orden ascendente según la variable bachillerato, para asignar a casos con valores perdidos en retención y aprobación, la media de los diez casos más cercanos.

Cuadro 14

Modelo de regresión para estimar el porcentaje de aprobación en bachillerato

Variable	Coeficiente	Error estándar	Coeficiente beta	t	p
Horario	-2.878	2.478	-0.040	-1.160	0.246
Dependencia	-15.596	2.317	-0.268	-6.730	0.000
Flujo	0.290	0.119	0.099	2.430	0.015
Retención	0.594	0.095	0.250	6.240	0.000
Aprobación	0.183	0.067	0.099	2.720	0.007
Constante	-17.551	12.025		-1.460	0.145

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria.

Estas recodificaciones permitieron obtener 786 colegios con valores en todas las variables en estudio, lo cual representa más del 97% de los colegios registrados para el año 2009. En el cuadro 15, se presenta la media de las variables antes de asignar valores a los valores perdidos y la media posterior a la asignación, estas dos mediciones prácticamente no varían ni en la población general, ni en las subcategorías público y privado.

Cuadro 15

Medias en las variables originales que fueron recodificadas y en las recodificadas, según dependencia del colegio. Año 2009.

Variable	N	Media			razón priv/púb
		General	Privado	Público	
Bachillerato (orig.)	689	0.658	0.849	0.586	1.449
Bachillerato (rec.)	795	0.655	0.849	0.588	1.444
Retención (orig.)	753	0.899	0.987	0.873	1.131
Retención (rec.)	790	0.900	0.981	0.873	1.124
Aprobación (orig.)	760	0.817	0.890	0.794	1.121
Aprobación (rec.)	790	0.817	0.887	0.794	1.117

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. Originales (orig.), Recodificadas (rec.)

ANEXO 2: Evidencias de validez del ICE y sus indicadores

Con el fin de dar evidencias de validez de los indicadores de la calidad educativa y del ICE en sí, en el cuadro 16 se presenta la matriz de correlaciones de estas variables y una variable que se considera asociada a la calidad educativa, la cual es el porcentaje de estudiantes de cada colegio que fueron elegibles¹ a la Universidad de Costa Rica (Eleg_UCR) en el proceso 2009-2010², con respecto a los estudiantes de ese colegio que solicitaron ingresar a esa Universidad.

El ICE presentó una correlación con la variable Eleg_UCR de .64, lo cual mostró evidencias de validez del ICE. Por otro lado, los indicadores del ICE presentaron correlaciones con este índice positivas y mayores 0.48, lo cual indica que los componentes del ICE presentan iguales direcciones que este, esto brinda una evidencia de que el índice está bien diseñado.

Cuadro 16

Matriz de correlaciones del ICE, sus componentes y el porcentaje de elegibles a la UCR. Año 2009

	ICE	Procesos	Eficacia	Contexto	Eleg_UCR
ICE	1.000	.488	.661	.594	.640
Procesos	.488	1.000	.067	-.025	.159
Eficacia	.661	.067	1.000	.000	.572
Contexto	.594	-.025	.000	1.000	.384
Eleg_UCR	.640	.159	.572	.384	1.000

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social. UCR. (2010). Base de datos de matrícula.

En el cuadro 16, también se brinda evidencias de validez para los indicadores de las dimensiones eficacia y contexto, ya que se presentan correlaciones con la variable Eleg_UCR de magnitudes medias.

Finalmente, en el cuadro 17 se presentan estadísticas de asociación entre la dependencia del colegio y cada uno de los indicadores de la calidad educativa. Se observa que las correlaciones eta son medias y altas, lo cual da evidencias de la validez del ICE y sus componentes, ya que la dependencia del colegio está asociada a la calidad educativa (Da Silveira, 1995; UNESCO, 1999). Además, se aprecian diferencias significativas al 5% entre los promedios de los indicadores, según dependencia del colegio y explicaciones de la variabilidad de los indicadores por parte de la dependencia del colegio, superiores al 15% (eta cuadrado).

Cuadro 17

Estadísticos de asociación entre los componentes del ICE y los conglomerados. Año 2009.

Indicador	Eta	Eta Cuadrado	P ANOVA
ICE	.753	.567	.000
Contexto	.445	.198	.000
Eficacia	.484	.234	.000
Procesos	.389	.151	.000

Fuente: Estado de la Nación. (2010). Megabase de datos georreferenciados de secundaria. MIDEPLAN. (2007). Índice de Desarrollo Social.

Notas

¹ Se dice que un estudiante es elegible a la Universidad de Costa Rica si obtiene un promedio de admisión superior a 442. El promedio de admisión es un indicador elaborado con las notas de los últimos años de secundaria y una prueba de selección de estudiantes, el rango de valores de esta variable es de 200 a 800.

² Esta población estaba compuesta en su mayoría por los egresados del 2009 de cada colegio.