

Prediction and Analysis of the Cognitive Level of 15 Year Olds in Ecuador using Linear Regression of Multiple Variables

R. Cueva, MSc.¹, J. Calderón, MBA¹, N. Medina, PhD.², J.C. Almachi, MSc.¹, E. Bone, MSc.¹, M. Medina, MSc.¹ y J. Montenegro, MSc.¹

¹Escuela Politécnica Nacional, Departamento de Formación Básica, Ecuador

²Escuela Politécnica Nacional, Departamento de Física, Ecuador

ruth.cueva@epn.edu.ec, edwin.bonee@epn.edu.ec

Abstract— The cognitive level is measured in the PISA tests that are carried out every 3 years in the countries associated with the OECD. Ecuador participated for the first time in 2017 in these tests. It is proposed to adjust a mathematical model based on linear regression of multiple variables involving previously investigated factors that have a relationship with the average cognitive level of a country. The factors used are: level of education of parents, minimum wage, fertility rate, infant mortality, access to early education, unemployment rate, government expenditure on education and teenage pregnancy. With the results obtained, it is analyzed which factors have the most influence to give general guidelines where one must work to increase the level of education in Ecuador.

Keywords—education, statistics, cognitive level, PISA test, Linear Regression of Multiple Variables

Digital Object Identifier (DOI):<http://dx.doi.org/10.18687/LACCEI2018.1.1.93>

ISBN: 978-0-9993443-1-6

ISSN: 2414-6390

Predicción y Análisis del Nivel Cognitivo de Jóvenes de 15 Años del Ecuador utilizando Regresión Lineal de Múltiples Variables

R. Cueva, MSc.¹, J. Calderón, MBA¹, N. Medina, PhD.², J.C. Almachi, MSc.¹, E. Bone, MSc.¹, M. Medina, MSc.¹ y J. Montenegro, MSc.¹

¹Escuela Politécnica Nacional, Departamento de Formación Básica, Ecuador

²Escuela Politécnica Nacional, Departamento de Física, Ecuador

ruth.cueva@epn.edu.ec, edwin.bonee@epn.edu.ec

Resumen— *El nivel cognitivo de los estudiantes de 15 años es medido en las pruebas PISA que se realizan cada 3 años en los países asociados a la OCDE. Ecuador participó por primera vez en el 2017 en estas pruebas, cuyos resultados serán publicados a inicios de 2019. Se propone ajustar un modelo matemático en base de regresión lineal de múltiples variables involucrando factores previamente investigados que tienen una relación con el nivel cognitivo medio de un país. Los factores utilizados son: nivel de educación de los padres, ingresos mensuales, tasa de fertilidad, mortalidad infantil, acceso a la educación inicial, tasa de desempleo, gasto público en educación y embarazo adolescente. Con los resultados obtenidos, se analiza cual o cuales factores tienen mayor influencia para dar lineamientos generales en donde se debe trabajar para incrementar el nivel de educación en el Ecuador.*

Palabras clave— *educación, estadística, nivel cognitivo, pruebas PISA, regresión lineal de múltiples variables.*

I. INTRODUCCIÓN

La adquisición y desarrollo de habilidades intelectuales de las sociedades están vinculados a una serie de factores que interactúan de una manera estrecha [1]. Identificar, corregir y fortalecer estos factores ayudan al avance científico y tecnológico de las sociedades. Está totalmente comprobado que países cuyos estudiantes de educación media tienen un alto coeficiente intelectual contribuyen de manera significativa al desarrollo científico y tecnológico en comparación con estudiantes de países que no han trabajado en el desarrollo de su coeficiente intelectual [2].

Numerosas organizaciones han dedicado tiempo y recursos para determinar el coeficiente intelectual promedio por país. El programa para la Evaluación Internacional de Alumnos PISA (Programme for International Student Assessment) de la Organización para la Cooperación y el Desarrollo Económicos OCDE, determina el nivel cognitivo medio de los países participantes con los resultados obtenidos de evaluaciones en tres áreas de competencia: comprensión lectora, matemática y ciencias naturales tomadas a jóvenes de 15 años de todo el mundo cada tres años. Los resultados obtenidos proporcionan un indicador para mejorar sus políticas de educación. El nivel cognitivo es la capacidad de los estudiantes de poder entender y resolver problemas reales a partir de la aplicación de conocimientos de cada una de las áreas. Evidencia bibliográfica respalda el hecho de que países con bajos salarios, altos índices de embarazos en adolescentes y la falta de educación de los

padres impactan negativamente al coeficiente intelectual de la población [3] [4]. Según datos de la UNESCO en los países Latinoamericanos uno de cada tres jóvenes entre 15 y 19 años no culmina la enseñanza primaria, además dentro de la región se encuentran los niveles cognitivos más bajos del ranking mundial [4].

El análisis de regresión lineal de múltiples variables es ampliamente aplicable en la investigación en ciencias del comportamiento humano, salud, educación, negocios, entre otros [5]. En ciencias de la salud, Rahe, Mahan y Arthur [6] trabajaron en encontrar la relación entre la cantidad de estrés en un individuo con el número de días de enfermedad que experimentará esa persona por los siguiente seis meses. En políticas de la educación, Coleman, Hoffer y Kilgore [8], Lee y Bryk [9] comparan el rendimiento de los estudiantes que asisten a escuelas públicas con los estudiantes de escuelas privadas. Utilizan variables características de los alumnos, tales como: rendimiento académico previo, coeficiente intelectual, descendencia étnica, ingreso familiar. Goldberger y Cain [10] analiza el nivel cognitivo en áreas como: vocabulario, matemática y lectura comparando estudiantes de escuelas católicas, públicas y privadas a partir de trabajos de Coleman, Hoffer y Kilgore. En la psicología experimental, Yerkes y Dodson [11] proponen una ley clásica de que existe una relación entre la excitación fisiológica y el rendimiento de la tarea conductual, de modo que se puede determinar un nivel óptimo de excitación para un rendimiento óptimo. La excitación excesiva o insuficiente reduce el rendimiento de la tarea. El punto donde se produce el máximo rendimiento está determinado por la dificultad de la tarea emprendida. En ciencias de la salud, Aiken, West, Woodward y Reno [12] desarrollaron un modelo predictivo para comparar el número de mujeres que se realizan periódicamente exámenes de mamografía con mujeres que no lo hacen. Se analizaron un conjunto de creencias de salud, tales como: gravedad percibida del cáncer de mama, susceptibilidad a cáncer de mama, beneficios percibidos al realizarse las mamografías, dificultades percibidas en la realización de los exámenes por encima de otros grupos de variables como: demografía, historial familiar médico, información médica y conocimiento previo de la enfermedad.

Cada una de estas hipótesis propone algún tipo de relación entre uno o más factores de interés (regresores) y una variable de

Digital Object Identifier (DOI): <http://dx.doi.org/10.18687/LACCEI2018.1.1.93>

ISBN: 978-0-9993443-1-6

ISSN: 2414-6390

16th LACCEI International Multi-Conference for Engineering, Education, and Technology: “Innovation in Education and Inclusion”, 19-21 July 2018, Lima, Peru.

resultado (regresando). El uso de regresión de múltiples variables, es una función lineal de n números de regresores correspondientes a las variables explicativas con una perturbación aleatoria o error. Si designamos por Y al regresando, por $X_1, X_2, X_3, \dots, X_n$ a los regresores y por U al error o perturbación aleatoria, el modelo poblacional de regresión lineal múltiple vendrá dado por (1) que es conocida como función de regresión poblacional (FRP) o hiperplano poblacional [13].

$$Y = C_0 + C_1X_1 + C_2X_2 + C_3X_3 + \dots + C_nX_n + U \quad (1)$$

Los parámetros $C_0, C_1, C_2, C_3, \dots, C_n$ son fijos y desconocidos, indican matemáticamente el grado de asociación entre los regresores.

II. METODOLOGÍA DE LA INVESTIGACIÓN

El diagrama de flujo con la metodología utilizada se muestra en la Fig 1. Se realiza la verificación de la correlación de factores que tienen una influencia tanto positiva como negativa empleadas en el estudio. El estudio de sensibilidad determina los factores que tienen una relación significativa con el nivel cognitivo de los estudiantes de 15 años. Se establece una base robusta de datos de 42 países miembros de la OCDE obtenidas desde el 2006 hasta el 2015 [14] [15]. La variable dependiente o regresando es el nivel cognitivo y las variables independientes seleccionadas o regresor se presentan en la Tabla 1. Con el método de regresión múltiple, se ajusta un modelo matemático que en función de los factores estudiados anteriormente puede predecir el nivel cognitivo del Ecuador para los años utilizados en la base de datos. Con el análisis de resultados de los coeficientes fijos del modelo matemático, se determina el factor que tiene una mayor influencia en el incremento del nivel cognitivo. La validación del modelo numérico se realiza comparando los resultados disponibles en la literatura de libre dominio del nivel cognitivo de países latinoamericanos para los años analizados. Con el modelo validado se obtendrá una proyección de la tendencia media del puntaje de nivel cognitivo para el año 2017 en donde Ecuador participó por primera vez y cuyos resultados finales serán publicados por la OCDE a finales de 2018.

Tabla I
FACTORES SELECCIONADOS COMO REGRESORES

Factor	Unidad	Variable
Promedio de educación de padres	Años	X_1
Salario mínimo mensual	Dólares	X_2
Tasa de fertilidad	Personas	X_3
Mortalidad infantil	Por cada 1000 nacimientos	X_4
Niños en edad escolar sin estudiar	Personas	X_5
Tasa de desempleo	Porcentaje	X_6
Gasto público en educación	Porcentaje del PIB	X_7
Embarazo en adolescentes	Por cada 1000 nacimientos	X_8

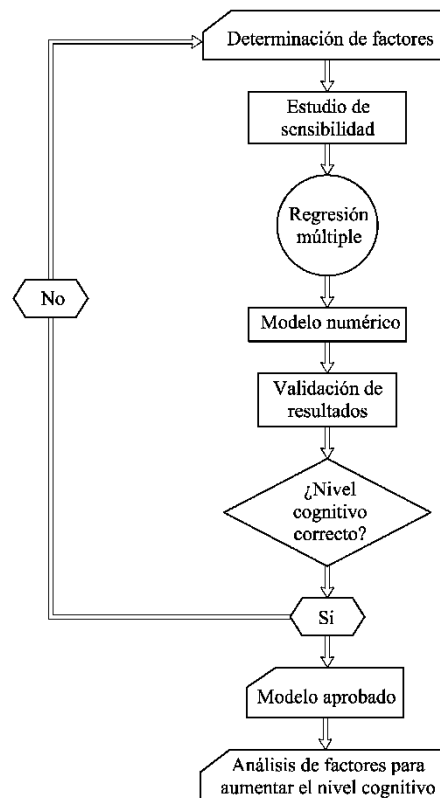


Fig. 1. Diagrama de flujo de la metodología utilizada.

III. RESULTADOS OBTENIDOS

La Tabla 2 muestra los resultados estadísticos por año de la regresión múltiple realizado en el programa de Matlab.

Tabla II
RESULTADOS ESTADÍSTICOS DE LA REGRESIÓN MÚLTIPLE

Año	2006	2009	2012	2015
Coefficiente de correlación múltiple	0.8300	0.8794	0.8983	0.9162
Coefficiente de determinación R^2	0.6889	0.7733	0.8069	0.8394
R^2 ajustado	0.6136	0.7184	0.7586	0.7992
Valor crítico de F	1.59E-6	1.18E-8	1.984E-9	1.16E-10
Error típico	34.8626	23.6859	21.4744	16.975

En la Tabla 3, se presentan los resultados de los parámetros fijos por año de las variables independientes o regresores empleados en el ajuste matemático.

Tabla III
COEFICIENTES OBTENIDOS EN EL MODELADO

Año	2006	2009	2012	2015
C_0	544.516	499.636	484.239	474,748
C_1	-2.989	-0.123	4.467	3,691
C_2	-0.006	-0.002	0.0003	-0,002
C_3	1.277	-6.530	-17.913	-11,353

C_4	-0.752	-0.633	-0.146	-1,643
C_5	4.900E-07	1.086E-6	-1.112E-6	6,428E-07
C_6	-2.017	-0.586	-1.121	-1,158
C_7	8.138	8,262	4,432	5,265
C_8	-1.558	-1.021	-0.987	-0,483

En la Fig. 2 a la Fig. 5, se presenta la validación mediante la comparación de los resultados de las pruebas PISA por año acerca del nivel cognitivo en los países de Latinoamérica con la proyección utilizando el ajuste matemático propuesto. Con el ajuste matemático se encontró el valor medio del nivel cognitivo para Ecuador, si hubiera participado en los años utilizados en la base de datos.

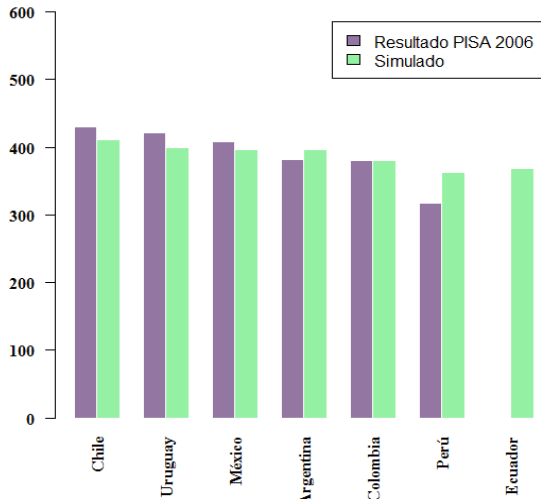


Fig 2. Comparación de los resultados medios de PISA 2006 con los resultados del ajuste matemático.

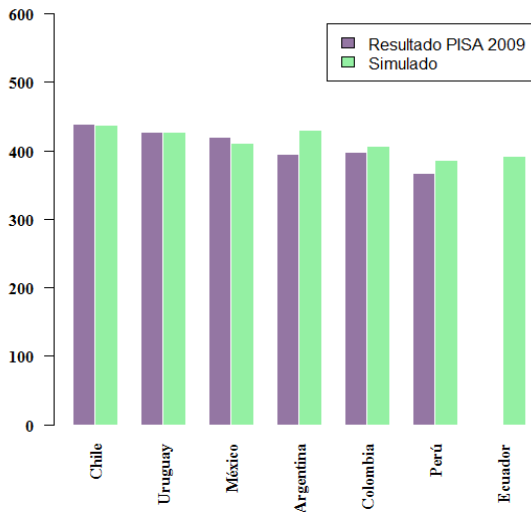


Fig. 3. Comparación de los resultados medios de PISA 2009 con los resultados del ajuste matemático.

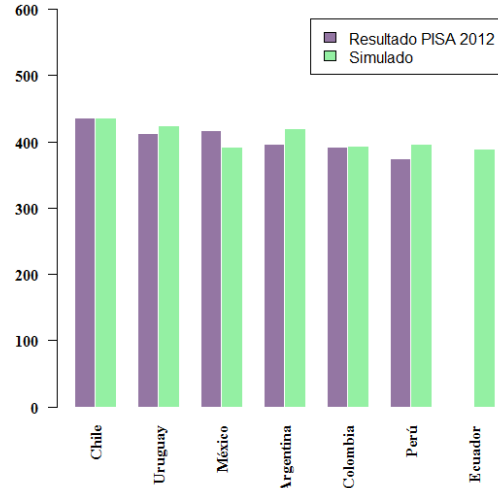


Fig. 4. Comparación de los resultados medios de PISA 2012 con los resultados del ajuste matemático.

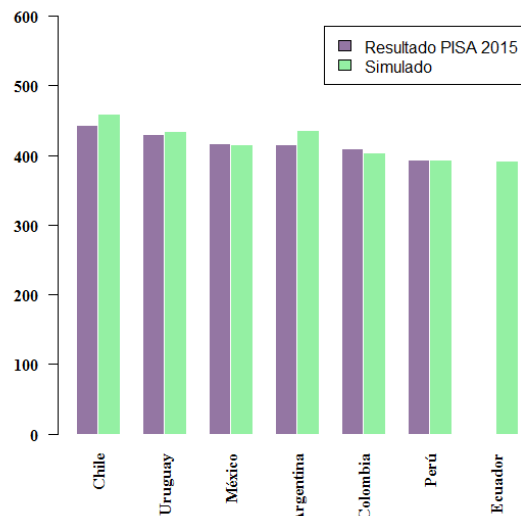


Fig. 5. Comparación de los resultados medios de PISA 2015 con los resultados del ajuste matemático.

IV. ANÁLISIS DE RESULTADOS

El análisis por regresión lineal múltiple probó la heterocedasticidad, la multicolinealidad y la especificación de los factores de mayor peso en el cambio del nivel cognitivo [16]. El coeficiente de determinación tiene una media de 0.7744 para los años considerados. Este coeficiente proporciona una idea del porcentaje de variabilidad de las variables a modelizar, lo cual asegura que el ajuste matemático utilizado presenta resultados confiables con un error típico medio de 24.24 puntos de nivel cognitivo.

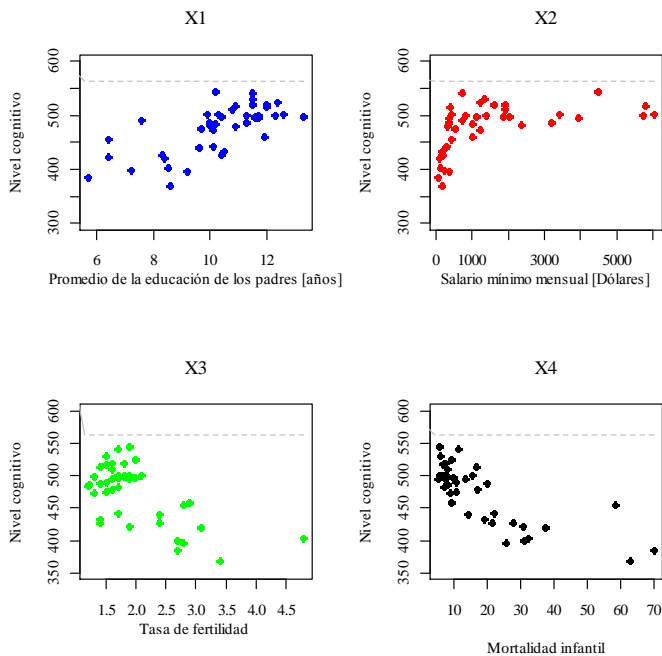


Fig. 6. Dispersión de los regresores X_1, X_2, X_3 y X_4 para el año 2012.

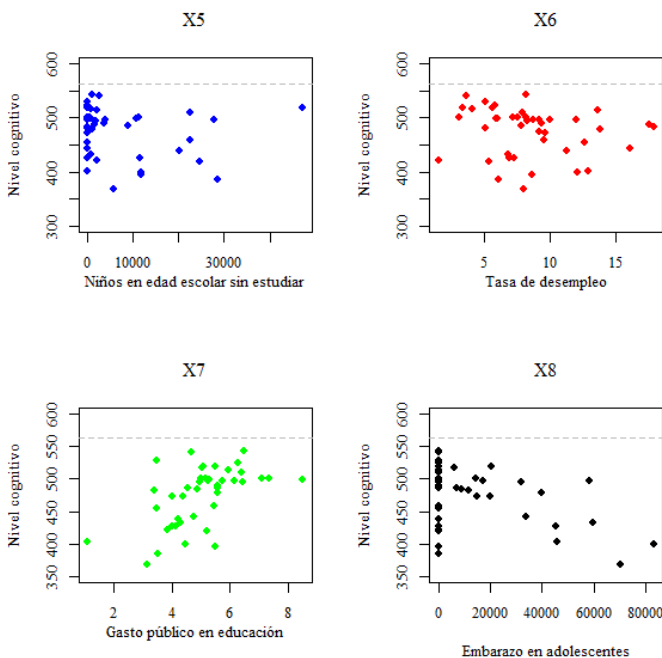


Fig. 7. Dispersión de los regresores X_5, X_6, X_7 y X_8 para el año 2012.

Para los años 2012 y 2015, las variables independientes “educación de los padres” y “gasto público en la educación” se convierten en el pilar fundamental para el incremento del nivel cognitivo de los jóvenes. La Fig. 6 (variable X_1) muestra que los padres, al no tener una formación adecuada de estudios en el ámbito de las matemáticas, lectura y ciencia, no pueden actuar como tutores o guía en el crecimiento intelectual del joven, provocando en él un declive significativo de su nivel cognitivo.

Programas gubernamentales de ayuda para los padres existen muy pocos, la mayoría se centra en cómo mejorar la calidad de la educación alumno profesor y olvidan por completo los factores intrafamiliares. La importancia de aumentar el índice cognitivo es muy alta debido a que según Antoine Fischbach y Ulrich Keller [17][18] muestran que mayores niveles cognitivos están relacionados con una mayor probabilidad de ser promovido a los curso posteriores, en especial para los alumnos de bajo rendimiento, los puntajes de competencia PISA indican una creciente discriminación de los estudiantes en sus entornos con el tiempo. Así también los estudiantes con mayores niveles cognitivos tendrán más probabilidades de acceder a estudios superiores y menos posibilidad de deserción escolar.

Según la UNESCO el acceso a la educación superior se expandió aceleradamente en la región durante la década del 2000, acumulando un promedio de crecimiento de aproximadamente 40%. El incremento de ingresos a la educación superior se refleja por el aumento del valor cognitivo de la región “como conjunto” en el promedio de la tendencia internacional. El patrón de crecimiento de la educación superior ha sido, sin embargo, muy inequitativo, favoreciendo principalmente a los sectores de mayores ingresos y de zonas urbanas, excluyendo a jóvenes de sectores marginados [19].

Los ingresos económicos en países de América Latina son bajos en comparación con los ingresos económicos de los países desarrollados [19]. Esto supone que en términos financieros y de políticas compensatorias, el gobierno debe enfocarse en, fortalecer y mejorar las instituciones de educación media, tanto a nivel de infraestructura como pedagógica. El resultado de estas políticas se evidenciará en el incremento del nivel cognitivo de los estudiantes de 15 años de acuerdo al ajuste del modelo matemático propuesto. Al incrementar el nivel cognitivo de los jóvenes, las instituciones de educación superior podrán disponer de estudiantes con una mayor capacidad para la producción científica y tecnológica, aspecto en el que la región se encuentra crónicamente retrasada.

V. CONCLUSIONES

El ajuste matemático desarrollado, fue validado con la comparación de los resultados de la prueba Pisa con los resultados obtenidos en el modelo de acuerdo a las Fig. 2 al 5, y permitió trazar un perfil cognitivo aproximado para el caso del Ecuador. Se tendrá que esperar hasta el 2019 el informe oficial de PISA para tener el resultado real, sin embargo, los resultados obtenidos viabilizan a cuantificar de manera aproximada cómo el nivel cognitivo ha evolucionado desde el 2006 al 2015 para nuestro país. Si Ecuador hubiera participado en las pruebas PISA en los años considerados, el valor obtenido de nivel cognitivo tiene la misma tendencia de los resultados alcanzados por los países de Latinoamérica, siempre y cuando los determinantes sean los regresores considerados en la investigación.

La variable de resultado, que es el nivel cognitivo de los jóvenes de 15 años, depende de las variables regresores consideradas en la investigación. De la Tabla 3, el coeficiente C_5 , peso de la variable “niños en edad escolar sin estudiar”, es la que menos influye en el modelo de predicción, analizándola individualmente; esta podría ser considerada una variable confusora, sin embargo, no se puede descartar del modelo, debido a que, al examinar una serie de gráficos de residuos, se evidenció que un aumento porcentual, alrededor del 20% de la variable “niños en edad escolar sin estudiar”, provoca cambios significativos en el índice cognitivo promedio por país. La normalidad y equidistribución de los residuos en el modelo predictivo resultó ser confiable, así lo demuestran las Fig. 2 y Fig. 3, donde se observan los errores típicos entre los valores reales y los valores del modelo. Al analizar en detalle el caso Ecuador, se observa que el nivel cognitivo se encuentra entre los más bajos de la región de Latinoamérica. Para los años tomados en cuenta en la investigación, el nivel cognitivo de los jóvenes ecuatorianos de 15 años tiene una tendencia a ser igual a los resultados obtenidos por Perú en las pruebas PISA.

Los sistemas de educación de América Latina, de acuerdo a la Fig. 7 invierten en educación hasta una media del 4% del producto interno bruto, la inversión de los países que se encuentran en los primeros lugares en el ranking mundial del resultado de la prueba PISA, mantiene o superan el valor medio de inversión de Latinoamérica, demostrando que, el dinero puede estar siendo desperdiciado o invertido en factores secundarios que no ayudan al mejoramiento de la calidad de la educación. El estudio presentado muestra a Chile como el país sudamericano con mayor índice cognitivo, pero sigue estando muy por debajo de la media establecida por la OCDE, un resultado desalentador, considerando que países como Argentina, Colombia Uruguay y Perú están por debajo de los resultados obtenidos del nivel cognitivo de los jóvenes chilenos.

Las comparaciones internacionales aportan valiosas lecciones, por ejemplo, no solo la inversión de dinero puede resolver los problemas de educación, disminuir los índices como el embarazo adolescente, disminuir la tasa de fertilidad de las mujeres y aumentar los años de educación de los padres, representan cambios significativos en el nivel cognitivo de los jóvenes. Al variar estos factores con datos de países desarrollados adaptándoles a nuestra realidad, en simulaciones realizadas con el modelo de regresión propuesta, mostraron un aumento del nivel cognitivo hasta un valor de 437 para el caso del Ecuador.

Para incrementar el nivel cognitivo de los jóvenes de 15 años, es preciso promover políticas de mejoramiento en la calidad de la educación media y asegurar una educación continua de los padres de familia. La tasa de fertilidad es un factor negativo que mayor peso tiene en la disminución del nivel cognitivo de los jóvenes. Es necesario hacer un llamado de atención a este factor

social, pues viene a ser un problema que involucra a toda la región de Latinoamérica

Se debe conceptualizar la coordinación de la educación de manera integral entre la educación básica, media y superior, para que se tenga una participación activa con la sociedad. Los proyectos PISA ayudan a promover políticas en donde se debe actuar para mejorar la educación. Los proyectos de vinculación de la universidad con la sociedad, deben enfocarse en el desarrollo del nivel cognitivo proponiendo metodologías innovadoras activas y participativas en donde el estudiante es el protagonista de su aprendizaje.

REFERENCIAS

- [1] N. Ahmad, M. Yunus, “Why are cognitive abilities of children so different across countries? The link between major socioeconomic factor and PISA test scores”, 105, 95-106, 2017.
- [2] A. Fischbach, U. Keller, “PISA proficiency scores predict educational outcomes”, 24, 63-72, 2013.
- [3] G. Jones, W. Schneider, “IQ in the production function: Evidence from immigrant earnings”, 48, 743-755, 2003.
- [4] R. Lynn, T. Vanhanen, “IQ and the wealth of nations”, Westport, CT: Praeger., 2002
- [5] J. Cohen y otros, “Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences”, tercera edición, 2013.
- [6] Rahe, R. H., Mahan, J. L., & Arthur, R. J., Prediction of near-future health change from subjects' preceding life changes. *Journal of Psychosomatic Research*, 14(4), 401-406, 1970.
- [7] England, P., Farkas, G., Kilbourne, B. S., & Dou, T. Explaining occupational sex segregation and wages: Findings from a model with fixed effects. *American Sociological Review*, 544-558, 1988.
- [8] Coleman, J., Hoffer, T., & Kilgore, S., Achievement and segregation in secondary schools: A further look at public and private school differences. *Sociology of education*, 55(2), 162-182, 1982.
- [9] Lee, V. E., & Bryk, A. S., A multilevel model of the social distribution of high school achievement. *Sociology of education*, 172-192, 1989.
- [10] Goldberger, A. S., & Cain, G. G., The causal analysis of cognitive outcomes in the Coleman, Hoffer and Kilgore report. *Sociology of education*, 55(2), 103-122, 1982.
- [11] Yerkes, R. M., & Dodson, J. D., The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology*, 18(5), 459-482, 1908.
- [12] Aiken, L. S., West, S. G., Woodward, C. K., & Reno, R. R., Health beliefs and compliance with mammography-screening recommendations in asymptomatic women. *Health Psychology*, 13(2), 122, 1994.
- [13] E. Uriel, “Regresión lineal múltiple: estimación y propiedades”, Universidad de Valencia, 2013.
- [14] OCDE, Estadísticas, <http://www.oecd.org/estadisticas/>
- [15] Grupo Banco Mundial, Datos de libre acceso del Banco Mundial, <https://datos.bancomundial.org/>
- [16] R. Montero, “Modelos de regresión lineal múltiple”, Universidad de Granada, Granada, 2016.
- [17] C. Marchioni, “Habilidades no cognitivas en América Latina una medición desde pruebas estandarizadas”: Banco del desarrollo de América Latina, 2016.
- [18] C. Hitt, J. Trivitt, “When you say nothing at all: The predictive power of student effort on surveys.”, 55, 105-119, 2016.
- [19] UNESCO, Situación Educativa de América Latina y el Caribe, Santiago de Chile, 2013.