

# Memorias



CONGRESO INTERNACIONAL  
DE INNOVACIÓN EDUCATIVA

**16-18 de diciembre, 2019**  
Campus Monterrey, N. L., México



Tecnológico  
de Monterrey

# Formas alternativas de evaluación a las pruebas de alto impacto para un currículo por habilidades

## *Alternative Evaluation Forms of High Impact Test for an Skill Based Curriculum*

Mauricio Cristhian Portillo Torres, Universidad Estatal a Distancia, Costa Rica, portillozeledon@gmail.com

### Resumen

A partir del año 2015 se implementa en el sistema educativo costarricense programas de estudio basados en el desarrollo de habilidades para el ciudadano del siglo XXI, lo cual conlleva un cambio en el modelo evaluativo de las pruebas nacionales. Esta investigación tuvo como propósito indagar sobre formas alternativas de evaluación de los aprendizajes que se pueden implementar en las pruebas nacionales a partir de una revisión de la literatura internacional, y el modelo evaluación criterial del Programa Bachillerato Internacional.

A partir de la revisión documental sobre modelos evaluativas para pruebas nacionales, la visita de un colegio que aplica el Programa de Bachillerato Internacional, así como la realización de entrevistas a docentes y expertos en evaluación se observó que: la evaluación educativa debe ser sistémica, requiere del diseño de un marco referencial alineado a la política educativa, el modelo evaluativo debe tener un valor educativo, y se requieren medios de evaluación más auténticos e integrales. Se concluye que la evaluación por ejecución (*performance assessment*) o evaluación auténtica es un referente importante para una evaluación criterial de las habilidades requeridas por el ciudadano en el siglo XXI.

### Abstract

*The Costa Rican educational system implements study programs based on the development of skills for the 21st century citizens since 2016, which involve a change in the evaluation model of national tests. The purpose of this research was to investigate alternative forms of educational evaluation that can be implemented in national tests based on a review of the international literature and the criteria evaluation model of the International Bachelor Degree Program.*

*Documental review of evaluative models for national tests, the visit of a school that applies the International Bachelor Degree Program, and interviews with teachers and experts in evaluation allow establish that educational evaluation must be systematic and it requires a design of a aligned referential framework to the educational policy. The evaluation model must have an educational value, and requires more authentic and comprehensive means of evaluation. Concluding, the evaluation by execution (*performance assessment*) or authentic evaluation is an important referent for a criteria evaluation of the skills required in the XXI century.*

**Palabras clave:** evaluación educativa, evaluación criterial, currículo por habilidades, habilidades del siglo XXI

**Keywords:** *assessment, criterial evaluation, skills based curriculum, 21st century skills*

## 1. Introducción

Desde 1988 el sistema educativo costarricense ha realizado pruebas nacionales para los estudiantes que concluyen el plan de estudio de la educación media superior. Esta prueba ha sido de carácter censal y estandarizada según los principios de pruebas psicométricas, es decir, construidas a partir de la teoría clásica de construcción de ítems y la evaluación referida a normas.

A partir de la aprobación en el año 2016 de una política curricular y la elaboración de programas de estudio orientados al desarrollo de habilidades del siglo XXI, se vuelve imperativo repensar la validez de la prueba nacional vigente. Pues el proceso evaluativo no es independiente de las estrategias y técnicas de enseñanza, del perfil final del estudiantado o los saberes contenidos en los programas de estudio, al contrario, estos deben ser su principal fuente de referencia. El modelo evaluativo debe fundamentarse en la misma conceptualización del aprendizaje, solo así los juicios de valor podrán mejorar el proceso de enseñanza y dar una realimentación oportuna. El propósito de la investigación fue indagar sobre formas alternativas de evaluación de los aprendizajes acordes con el desarrollo de habilidades en el currículo nacional para la prueba nacional. Lo cual supone analizar los modelos evaluativos y medios más auténticos de evaluación de las habilidades.

## 2. Desarrollo

### 2.1 Marco teórico

Las tendencias actuales de reforma educativa se orientan hacia modelos curriculares construidos a partir de estándares educativos o de resultados de aprendizaje (Sahlberg, 2006). Estos estándares o resultados del currículo se expresan a través de competencias o habilidades considerados relevantes para la ciudadanía del siglo XXI.

### Currículo por habilidades

Una revisión de los currículos de varios países evidencia una variedad de términos sobre este particular como “competencias básicas”, “habilidades para la vida”, “habilidades del siglo 21” o “competencias del siglo 21” (MECD, 2013). En el caso de la política curricular costarricense las trece habilidades definidas tienen como fuente principal el proyecto *Assessment & Teaching of 21st Century Skills* (ATC21s), liderado por la Universidad de Melbourne entre 2009-2012, cuyo propósito fue validar

instrumentos de medición de la resolución colaborativa de problemas, una habilidad compleja.

Las trece habilidades de la política curricular son: pensamiento sistémico, pensamiento crítico, aprender a aprender, resolución de problemas, creatividad e innovación, ciudadanía global y local, responsabilidad personal y social, estilos de vida saludables, vida y carrera, colaboración, comunicación, apropiación de tecnologías digitales y manejo de la información (MEP, 2016).

Según Patrick Griffin, líder del proyecto ATC21s, las habilidades son acciones que las personas pueden realizar, acciones que incluyen características actitudinales. Por su parte, la competencia abarcaría la calidad y capacidad de transferencia de una acción en el tiempo y en el contexto, por consiguiente la competencia es la capacidad de ajustar el rendimiento de la habilidad según la demanda del contexto (Griffin y Care, 2014). Esta definición permite comprender la habilidad como una acción situada en un contexto específico con diferentes niveles de competencia.

En un currículo por habilidades los niveles de competencia se expresan en los mapas de progreso. Se denomina mapa de progreso a la representación gráfica de una habilidad que permite observar su nivel de desarrollo competencial de forma creciente según avance el estudiantado por el sistema educativo (Portillo-Torres, 2017). Este mapa facilita la construcción de perfiles de salida del estudiantado, por ejemplo, por ciclos educativos (ver Figura 1).

	Preescolar	I ciclo	II ciclo	III ciclo	Ciclo Diversificado y Técnico
Colaboración	Identifica a los compañeros y los compañeros con quienes debe trabajar de forma grupal en la realización de una tarea.	Reconoce que las personas tienen diferentes estilos de trabajo según las habilidades y la motivación que tienen.	Comunica de forma clara a los demás compañeros su expectativa sobre lo que cada uno aportará para la tarea.	Propone cómo mejorar el desempeño grupal desde los aportes individuales.	Aprovecha los aportes de sus compañeros y compañeros, sin discriminación, fortaleciendo la cohesión del grupo y el éxito esperado.
	Describe de forma clara la tarea que debe realizarse con los demás.	Explica lo que piensa y lo que piensan otros acerca de lo que debe hacer el grupo.	Resume las diferentes ideas para facilitar la toma de decisión del grupo.	Expone de las diferentes posiciones las convergencias y las divergencias con el fin de llegar a acuerdos.	Formula objetivos grupales individuales a partir de las ideas de los demás compañeros sobre los contenidos de aprendizaje.
	Muestra actitudes que favorecen el trabajo colaborativo en torno a proyectos o tareas.	Reconoce que sus responsabilidades contribuyen a un trabajo exitoso.	Sugiere constantemente cómo mejorar el progreso para alcanzar la meta común.	Asume diferentes roles de trabajo durante el progreso de las actividades que se realizan.	Vela por la eficiencia y la eficacia del trabajo grupal asistiendo a otros en la solución de problemas y cumpliendo sus propias responsabilidades.

Figura 1. Mapa de progreso de la habilidad de colaboración, tomado de MEP, 2016a, p.45

Otro elemento clave de la habilidad es el contexto específico en el cual se ejecuta, en el caso del currículo nacional, corresponde a cada uno de los programas de estudio, es decir, cada habilidad se encuentra contextualizada por los saberes de una disciplina, así como por las tareas

procedimentales y los valores actitudinales de dicha disciplina (Portillo-Torres, 2017).

Un currículo basado en habilidades es un paradigma emergente que supone un reto para la medicación y la evaluación de los aprendizajes. Algunas habilidades tienen una larga tradición de investigación, y otros son más recientes. Al respecto, el *National Research Council* (2012) de las Academias Nacionales de Estados Unidos agrupa las habilidades en tres dominios: el dominio cognitivo, el dominio interpersonal y el dominio personal. De estos tres dominios, el cognitivo ha sido históricamente más estudiado. Este dato es importante de cara a la

evaluación de los aprendizajes, pues representa un reto evaluar y medir habilidades del dominio interpersonal e intrapersonal.

Las trece habilidades de la política curricular pueden agruparse en los dominios anteriormente dichos, si bien los límites de una habilidad no siempre están bien definidos:

Tabla 1. Dominios de las trece habilidades

Dominio cognitivo	Dominio intrapersonal	Dominio interpersonal
Pensamiento crítico	Estilos de vida saludable	Colaboración
Pensamiento sistémico	Vida y carrera	Comunicación
Resolución de problemas	Ciudadanía global y local	Apropiación de las tecnologías digitales
Manejo de la información	Responsabilidad personal y social	
Aprender a aprender		
Creatividad e Innovación		

Nota: elaboración propia para este trabajo a partir de la clasificación de *National Research Council*, 2012.

Comprender la naturaleza de las habilidades permite fundamentar el objeto de la evaluación educativa, así como los modelos y los instrumentos que se deben implementar. En el caso de las pruebas nacionales estandarizadas, estas solo han sido utilizadas para medir el dominio cognitivo del currículo nacional. Hecho que los informes nacionales costarricenses de los resultados de la prueba así reconocen en sus títulos, por ejemplo: *Informe Pruebas Nacionales Diagnósticas de II Ciclo de la Educación General Básica 2012 Prueba Cognitiva*.

### Modelos de evaluativos de los aprendizajes

Martinic (2010) identifica tres ciclos de reforma educativa en Latinoamérica en los últimos años, los cuales asocia a tres modelos de evaluación: referido a normas, referido a criterios y referido a estándares. Cada uno de estos modelos coincide con cambios en las teorías y desarrollo metodológico de la evaluación. Además, existe dos grandes métodos utilizados para la elaboración de las pruebas nacionales: el método estadístico y el método no estadístico.

### Método estadístico

Concibe la evaluación como medición a través de

pruebas estandarizadas. Utiliza modelos psicométricos de evaluación. Para la elaboración de pruebas utiliza la Teoría de Respuesta al Ítem (TRI). En el caso de la prueba nacional costarricense, la Prueba Nacional de Bachillerato (PNB) es una prueba de evaluación referente a un comportamiento de distribución normal estadístico.

### Modelo de evaluación referente a la norma

Este modelo evaluativo se enfoca en la evaluación de contenidos educativos según los objetivos de aprendizajes de los programas de estudio. La respuesta correcta de los ítems de la prueba nacional supone el dominio por parte de los estudiantes del contenido. El comportamiento de las respuestas al ítem se visualiza como una curva de campana, lo que permite el cálculo de medidas de tendencia central. A partir de estas medidas se compara el resultado de cada estudiante en relación a todas las personas que tomaron la prueba. La norma es según el desempeño del grupo y la posición relativa de cada uno de los estudiantes (Martinic, 2010).

Otra de las características de este tipo de prueba es su utilidad para establecer un ranking de puntaje, así como la facilidad de procesar los datos de la prueba a través de máquinas y no a través de seres humanos.

### **Modelo referente al estándar**

Según Scriven (1981) un estándar es un nivel de desempeño asociado con una calificación, o un criterio dado, o una dimensión de logros. Los puntajes de corte suelen definir el estándar para el objeto evaluado. Ya que los estándares pueden estar asociados a criterios, algunos se refieren a este tipo de prueba como criterial, sin embargo, el criterio puede no ser es el referente principal de la construcción de la prueba. El establecimiento de los puntajes de cortes es la parte más relevante e influyente en la interpretación de la prueba (Ministerio de Educación y Deportes, 2017).

La evaluación referida al estándar es ampliamente utilizada en pruebas internacionales como PISA, TIMMS y PIRLS. Es un modelo asociado a enfoques educativos basados en competencias. El estándar es una descripción de lo que el estudiante debería saber (contenido del estándar) y lo que debería lograr (desempeño del estándar) en diferentes niveles momentos del proceso educativo, en asignaturas claves como ciencias, matemáticas y lectura. Los ítems de las pruebas se elaboran según diferentes niveles de dominio previamente establecidos. PISA, por ejemplo, utiliza seis niveles de dominio (OCDE, s.f.). Los cortes de puntaje para cada nivel de dominio son determinados de forma estadística. Los estándares son de naturaleza prescriptiva pues se redactan según lo que el estudiante debe ser capaz de hacer, y no de lo que es capaz de hacer.

### **Método no estadístico**

Se concibe la evaluación como un proceso de indagación, de recolección de múltiples evidencias a partir de los procesos y los productos de aprendizaje. Es un enfoque evaluativo cuyo propósito no es hacer *ranking* de los alumnos o los centros educativos, sino usar la información para la mejora del currículo. Su principal referente es la interpretación de las evidencias a partir de unos criterios establecidos y que permiten realizar un juicio valorativo de los procesos y los productos.

### **Evaluación referida al criterio**

De acuerdo con la Organización para la Cooperación y el Desarrollo Económico (OCDE) (2013 y 2015) en varios países miembros existen instrumentos y enfoques de evaluación alternativos a las pruebas estandarizadas y pruebas de opción múltiple. Este enfoque evaluativo se le conoce como *performance assessment* o basados en

el rendimiento (OECD, 2013 y OCDE, 2015). La principal característica del *performance assessment* es evaluar de forma integrada el conocimiento, las habilidades y las actitudes del estudiantado al realizar una tarea, en lugar de una respuesta a través de un ítem. Las tareas pueden incluir presentaciones orales, portafolios, experimentos, y trabajo en grupo.

Al centrarse la evaluación en la producción y realización de tareas, se promueve la ejecución de habilidades o competencias complejas o difíciles de medir en una prueba estandarizada. Mientras que las pruebas estandarizadas utilizan máquinas para procesar la información, este tipo de prueba se fundamenta en el criterio de personas expertas. Esto supone un costo mayor que las pruebas estandarizadas, así como la capacitación de evaluadores expertos.

Un referente de la evaluación del desempeño o rendimiento (*performance assessment*) que valora las acciones y productos que realiza el estudiante durante el proceso de aprendizaje, es el Aprendizaje para el dominio o *Mastery Learning* de Benjamin S. Bloom y John B. Carroll que sugiere que todos pueden adquirir el mismo nivel de aprendizaje con el tiempo y las estrategias instruccionales adecuados (Block, 1971). Otros conceptos relacionados con este enfoque son el concepto de evaluación alternativa, que hace referencia a técnicas y métodos que superan la metodología tradicional de evaluación con el propósito principal de calificar; así como evaluación auténtica, es decir, actividades que evalúan el desempeño en situaciones, actividades y contenidos reales de aprendizaje (López, 2016).

Para inferir e interpretar el nivel de dominio del estudiante se utilizan criterios según la tarea y el nivel de desempeño que se espera lograr. Como instrumentos de evaluación se suelen utilizar rúbricas que contienen los criterios y el nivel de desempeño esperado. En algunos países estos criterios y rúbricas vienen en documentos adjuntos a los programas de estudio para cada asignatura. Por su naturaleza el criterio es dependiente de la tarea y específico para un área del conocimiento.

### **2.2 Planteamiento del problema**

Las pruebas nacionales han adquirido una creciente importancia como forma de medir y monitorear la calidad. Son pruebas de alto impacto, estandarizadas y establecidas a niveles centrales y administrados a nivel nacional. Los procedimientos que rigen su elaboración

y contenido, así como su interpretación y uso de los resultados se establecen a nivel central (Consejo Europeo, 2009).

En el caso costarricense, las primeras pruebas de carácter nacional y de tipo estandarizada se aplicaron de forma continua entre 1954 y 1973. Se restablecieron con fines de certificación en el año 1988. Están diseñadas siguiendo un modelo de medición con referencia a normas, es decir, compara los resultados de un individuo con los resultados de una población o grupo al cual pertenece. Actualmente, el sistema educativo costarricense tiene diversos tipos de pruebas nacionales:

- a) Las Pruebas Nacionales de Bachillerato de Educación Media, con carácter censal y de certificación.
- b) Las pruebas nacionales escritas comprensivas estandarizadas en Educación Técnica, con carácter censal y de certificación.
- c) Las pruebas nacionales diagnósticas: conclusión del II Ciclo de EGB y de conclusión de III Ciclo de EGB (Decreto N°40862).

Las Pruebas Nacionales de Bachillerato (PNB) son de carácter censal, es decir, se aplica a toda la población que concluye el plan de estudio de Educación Media, y certifican el conocimiento adquirido por el estudiantado. Desde el año 1999 se estableció la nota mínima de aprobación en 70, en el caso de la PNB, la cual es el resultado de sumar la nota de presentación (40%) con la nota obtenida en el examen de Bachillerato (60%). Esta condición sigue aún vigente. Sin embargo, las PNB son sólo una de las formas posibles de evaluación de los alumnos, pues la medición es una parte de la evaluación, pero no es toda la evaluación. La evaluación de los aprendizajes es un fenómeno mucho más amplio y valorativo del fenómeno. Lo ideal es que la PNB forme parte de una política más amplia de evaluación de los aprendizajes.

La OCDE (2015) señala que uno de los retos para los países que tienen pruebas nacionales estandarizadas es evitar una excesiva dependencia hacia este tipo de prueba: los países que tienen evaluaciones nacionales estandarizadas, sea a nivel nacional o en regiones específicas, tienen el reto de gestionar la dependencia excesiva de las evaluaciones estandarizadas y establecer salvaguardias, como el uso de varias formas de evaluación de los estudiantes para supervisar el aprendizaje (OECD, 2015, p. 110).

Establecer salvaguardas se podría incluir formas alternativas de evaluar el aprendizaje, por ejemplo, la introducción de evaluaciones basadas en el rendimiento o desempeño que permiten conocer con mayor precisión el logro de los estudiantes y las competencias complejas mediante tareas abiertas como ensayos, presentaciones, tareas colaborativas y portafolios, así como otras prácticas innovadoras.

Al aprobar el Consejo Superior de Educación, órgano máximo rector del sector educativo del país, la *Política Curricular bajo la visión de Educar para una Nueva Ciudadanía* (Acuerdo 07-44-2016) y la *Política Educativa La persona Centro del proceso educativo y sujeto transformador de la sociedad* (Acuerdo 02-64-2017), establecen que los programas de estudio fruto de la transformación curricular 2014-2018 sean según las habilidades que la ciudadanía requiere para el siglo XXI. Esto supone un cambio en el planeamiento, medicación y evaluación de los aprendizajes, los cuales están orientados al desarrollo de habilidades y no de contenidos educativos.

Las trece habilidades identificadas de la *Política Curricular* se desarrollan desde preescolar hasta la educación media superior, incluyendo la educación de personas jóvenes y adultas, lo cual permite establecer perfiles de salida del estudiantado por ciclos educativos. A la vez, estos perfiles son referentes para la elaboración de los perfiles de salida de cada asignatura según el ciclo educativo en el cual se imparte. De esta forma, las habilidades se contextualizan según los conocimientos y procedimientos propios de cada asignatura.

Esta transformación del currículo nacional implica algunos retos como: la construcción de mapas de progresos sólidos de las habilidades, el alineamiento de un sistema de evaluación con un énfasis más en lo formativo y menos en lo sumativo, trayectorias escolares basadas en el dominio demostrado y no sobre una calificación, considerar la flexibilidad del tiempo en los procesos de aprendizaje, y apoyos educativos que enfatizan en el dominio del estudiantado (Portillo-Torres, 2017).

En este contexto una evaluación pobremente diseñada tiene el potencial de entorpecer el aprendizaje y suprimir la transformación curricular. En cambio, un sistema de evaluación bien diseñado puede mejorar el aprendizaje como herramienta formativa para aprender y/ o herramienta sumativa que indica lo ya aprendido. A

partir de estas consideraciones se plantean las siguientes preguntas orientadoras:

¿Cuáles formas alternativas de evaluación de los aprendizajes pueden implementarse en las pruebas nacionales para un currículo basado en habilidades?

Otras preguntas afines son:

¿Existen a nivel nacional formas de evaluación del aprendizaje con fines de certificación que se ajustan a la política curricular vigente?

¿Cuál modelo evaluativo de los aprendizajes conviene como marco referencial para implementar pruebas nacionales de un currículo por habilidades?

### 2.3 Método

Esta investigación se realizó con un diseño de estudio de caso. Un estudio de caso permite estudio de situaciones que no siempre son susceptibles de un análisis numérico. Según Robson (2002), citado por Cohen, Manion y Morrison (2007), los estudios de caso optan por generalizaciones analíticas, más que generalizaciones estadísticas, es decir, desarrollan ideas que pueden ayudar a otras investigaciones a comprender casos, fenómenos o situaciones similares. En este sentido, la investigación utilizó una aproximación cualitativa al objeto de estudio.

Así mismo los estudios de caso se establecen en contextos temporales, geográficos, institucionales que permiten definir los límites alrededor del caso. De acuerdo a lo anterior, la unidad de análisis de esta investigación corresponde a la prueba nacional de bachillerato al finalizar la educación diversificada en el sistema educativo costarricense, en el marco de la política curricular orientado al desarrollo de habilidades para el siglo XXI.

Una vez delimitado el objeto de estudio se siguió un diseño investigativo correspondiente a un estudio de caso (ver Figura 1):

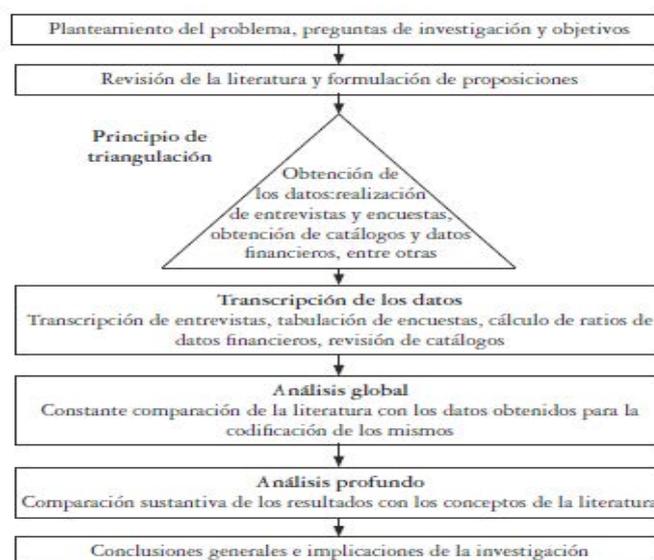


Figura 2. Fases de un diseño de casos, tomado de Martínez 2006, p.186

Para la fase de obtención de los datos, según Yin (2014), se pueden utilizar seis fuentes para la recolección de evidencias en un estudio de caso: documentos, registro de archivos, entrevistas, observación directa, observación participativa y artefactos físicos. En la presente investigación se aplicaron tres técnicas de recolección de los datos: la revisión documental, la entrevista, y el grupo focal.

#### Revisión documental

En primer lugar se revisaron investigaciones nacionales respecto a la Prueba Nacional de Bachillerato, así como documentos oficiales del Ministerio de Educación Pública respecto a las características de esta prueba. Esto permitió definir el objetivo de la prueba, su modelo evaluativo, y el impacto que esta prueba ha tenido en el sistema educativo costarricense. Una segunda revisión documental permitió conocer informes internacionales sobre los tipos de pruebas nacionales que se aplican en otros países. En total, se revisaron cuatro investigaciones nacionales y ocho informes internacionales.

#### Entrevista

Se visitó el Liceo Santo Domingo de Heredia, colegio público que desde el año 2016 ingresó al Programa de Bachillerato Internacional (BI). En este lugar se realizaron dos entrevistas a profundidad: una con el director del centro educativo y otra con el coordinador del programa de Bachillerato Internacional. Se les preguntó sobre el

modelo evaluativo de BI, y la articulación que existen entre la evaluación interna y la evaluación externa.

Posteriormente se entrevistó a Juan Manuel Esquivel Alfaro Ph.D., miembro del Consejo del Sistema Nacional de Acreditación de la Educación Superior (SINAES) y reconocido especialista nacional e internacional en el campo de la evaluación de los aprendizajes. Al especialista se le entrevistó sobre los modelos evaluativos, sobre la naturaleza de las habilidades y cómo *operativizar* su evaluación en el sistema educativo costarricense. Todas las entrevistas fueron grabadas y transcritas.

### **Grupo focal**

El grupo focal se conoce como una entrevista grupal integrada de tres a diez personas en la cual se conversa de uno a varios temas a profundidad (Hernández, Fernández y Baptista, 2014). En el grupo focal participaron dos investigadores: uno de ellos dirigió el diálogo grupal, y la otra tomó notas de aspectos relevantes del desarrollo de la misma. Se realizaron preguntas abiertas para generar la conversación del tema.

El grupo focal se realizó en la sede de San José del Colegio de Licenciados y Profesores en Letras, Filosofía, Ciencias y Artes (COLYPRO) de Costa Rica, con la participación de ocho personas, docentes de diversas especialidades.

### **Análisis de la información**

La información recopilada fue sistematizada y comparada en relación a la literatura revisada para una codificación de la misma. En este sentido se realizó una triangulación de datos, cuyo resultado permitió la delimitación de cuatro grandes temas de análisis:

- La evaluación educativa debe ser sistémica
- El diseño de un marco referencial para la evaluación
- El valor educativo de la prueba
- Medios evaluativos más auténticos e integrales

## **2.4 Resultados**

### **a. La evaluación educativa debe ser sistémica**

Obtener información de los sistemas educativos para tomar decisiones basadas en la evidencia es un imperativo para la mejora continua. No existe una única prueba capaz de recopilar toda la información que se necesita para una valoración integral de lo que ocurre a nivel del aula, a nivel de la escuela, o de una región o a nivel nacional. Según la literatura revisada, los países aplican diversas estrategias, métodos y técnicas de evaluación para una

mayor validez de lo que se desea conocer (Consejo Europeo, 2016).

Dada la complejidad de los sistemas educativos, el primer reto es la articulación de un marco de evaluación y valoración a través de una política nacional de evaluación. Los marcos de evaluación son acuerdos coordinados que buscan apoyar la mejora los resultados de los estudiantes de un sistema escolar. Reúnen la evaluación de los estudiantes, la evaluación de los profesores, la evaluación de los centros escolares y la evaluación del sistema, intentando lograr una armonización coherente hacia los objetivos del aprendizaje de los estudiantes (OCDE, 2015).

Una revisión de los diferentes marcos de evaluación de los países miembros de la OCDE, se clasifica en cuatro perspectivas las estrategias e instrumentos de evaluación:

- Evaluación del sistema: cómo los gobiernos evalúan el progreso de los objetivos educativos y el rendimiento global.
- Evaluación de los centros escolares: evaluar centros escolares individuales como organizaciones.
- Evaluación de los profesores: formas de valorar y evaluar a los profesores.
- Evaluación de los estudiantes: forma en la que se mide y planifica el progreso de los estudiantes de forma sistemática para medir las pruebas de aprendizaje y hacer evaluaciones sobre el aprendizaje del estudiante. (OCDE, 2015, p.37).

Para construir un marco de evaluación, se requiere que la política educativa evaluativa cubra las siguientes áreas (OECD, 2013):

Gobernanza: gestión del marco de evaluación y medición, incluidos los objetivos, distribución de responsabilidades, funciones dentro del marco, el concepto de evaluación y valoración, y la integración del sector no público, financiamiento.

Diseño: Configuración y arquitectura del marco de evaluación y valoración, incluyendo sus componentes principales, la articulación entre ellos, los principios fundamentales en el que se basan los procedimientos de evaluación, alineación con objetivos, y los enlaces a las prácticas de aula.

Construcción de capacidades: Competencias y habilidades para la evaluación y valoración a través del marco,

incluidas las directrices y herramientas para la evaluación y evaluación y las oportunidades de aprendizaje.

En los próximos años el sistema educativo costarricense tiene el reto de construir un sistema de evaluación integral y alineada, algo que no solo compete a la Dirección de Gestión y Evaluación de la Calidad, dependencia del Ministerio de Educación Pública, sino debería concebirse como una política nacional de evaluación con la participación de diversos actores interesados.

En este orden de ideas, la actual PNB no cumplen con los objetivos definidos en el artículo 81 del *Reglamento de Evaluación de los Aprendizaje* debido a las limitaciones propias de la forma de la prueba, una prueba estandarizada basada en norma, y a la falta de otro tipo de instrumentos que se aplique para recopilar mayor cantidad de evidencia (González, 2018). A continuación se detallan los objetivos de la Pruebas Nacionales estipulado en el citado reglamento:

- Contribuir a la formación integral de los estudiantes.
- Coadyuvar en la determinación de la promoción o la certificación de los educandos.
- Conocer los resultados de logro de los objetivos curriculares de la educación costarricense, basado en criterios técnicos de medición y evaluación, de modo que permitan mejorar la calidad del sistema educativo en todos sus niveles y procesos.
- Incorporar con base en los resultados obtenidos por la población estudiantil en las respectivas Pruebas Nacionales, según lo permita esta información; las estrategias y propuestas conducentes al mejoramiento cualitativo de los procesos de la enseñanza y el aprendizaje, en aquellas áreas donde el Sistema Educativo lo requiera.
- Ofrecer a la población estudiantil un desafío académico, que contribuya a mejorar las posibilidades de éxito para su incorporación a los ciclos o niveles educativos inmediatos superiores o al mundo del trabajo.
- Promover una actitud de superación académica en los profesionales de la docencia, motivándolos para que aporten lo mejor de sus conocimientos en la búsqueda de un mayor y mejor aprendizaje de los educandos.
- Hacer de las pruebas nacionales un recurso adecuado para el proceso de evaluación y control del rendimiento escolar (MEP, 2018, pp.71-73).

Los Directores Regionales del MEP y especialistas de evaluación entrevistados para el estudio de la Universidad Nacional, coinciden que los objetivos de las PNB no responden a las expectativas que promueven, ni responde a la realidad de la población estudiantil (González, 2018). Por consiguiente, deberá revisarse otras formas alternativas de evaluación de la prueba nacional, si se mantienen los mismos objetivos.

#### b. **El diseño de un marco referencial para la evaluación**

De acuerdo al análisis documental hecho, el proceso de evaluación de una prueba nacional inicia con la elaboración el marco referencial o conceptual, documento que ordena las ideas y el enfoque de la prueba. Este documento orienta las etapas siguientes de la evaluación (UNESCO, 2009).

La actualización de los marcos referenciales de evaluación es una tarea necesaria para construir políticas y sistemas de evaluación sólidos. Son varios los ejemplos de países que recientemente han actualizados sus políticas en evaluación educativa (OCDE, 2015). En el caso costarricense con la aprobación del la Política Curricular bajo la visión Educar para una nueva ciudadanía aprobada por el Consejo Superior de Educación, se establece que los programas de estudio fruto de la transformación curricular 2014-2018 sean según las habilidades que la ciudadanía requiere para el siglo XXI. Por consiguiente es necesaria una actualización del marco referencial de la evaluación para la prueba nacional.

El referente teórico más próximo para comprender un currículo por habilidades en el contexto costarricense es el proyecto *Assessment and Teaching of 21st Century Skill* (ATC21s), cuyo propósito fue validar un modelo evaluativo de una competencia compleja, a saber, la resolución colaborativa de problemas. Dicho proyecto sirvió como base para la medición de esta competencia en la prueba internacional PISA el año 2015. En total, la Política Curricular contiene una definición de trece habilidades claves para los programas de estudio, agrupadas en cuatro dimensiones, al igual que el proyecto original ATC21S.

En opinión del experto en evaluación entrevistado, Juan Manuel Esquivel Alfaro PhD, existe la duda sobre la naturaleza de las habilidades, objeto de evaluación de las pruebas nacionales, si estas deben entenderse como un constructo en términos psicométricos, o si son algo

observable de forma directa y empírica. Por ejemplo, la habilidad de leer es algo directamente comprobable y observable. En cambio, un constructo es un atributo o característica que no puede observarse directamente, sino que requiere de conceptos abstractos y principios que puedan ser inferidos a partir de la conducta y explicados por una teoría educativo-psicológica. Tal es el caso de la inteligencia. Definir la naturaleza del objeto evaluado es muy importante para elegir el modelo evaluativo a seguir, y por consiguiente la selección de las múltiples fuentes de evidencia para la interpretación del constructo evaluado o la conducta evaluada.

Según la teoría consultada, las pruebas nacionales suelen utilizar metodologías estadísticas para su construcción, pues de preferencias son pruebas estandarizadas. Esto supone una limitación importante para la medición de habilidades pues la prueba se restringe al dominio cognitivo (pensamiento, razonamiento, creatividad), y no se extienden al dominio interpersonal (comunicación, colaboración) y el dominio intrapersonal (responsabilidad personal y social, vida y carrera, motivación) (MEP, 2016a).

#### c. **El valor educativo de la prueba nacional**

Desde 1999, la prueba nacional de bachillerato se elabora a partir del modelo de medición referido a normas, en primera instancia a partir de objetivos establecidos en temarios y, posteriormente, a partir de los programas de estudio de Ciencias (Biología, Física o Química), Español, Matemáticas, Estudios Sociales e Inglés. Para la determinación de la calificación se utiliza una escala de 1 a 100, donde el 40% corresponde a la nota de presentación y un 60% a la nota de obtenida en la respectiva prueba. Aquellos estudiantes que obtienen una calificación final igual o superior a 70, se consideran aprobados en la respectiva asignatura. La Dirección de Gestión y Evaluación de la Calidad tiene la tarea de remitir a cada institución educativa los resultados de esa institución así como el solucionario correcto (MEP, 2016b). Este resultado permite conocer el desempeño promedio institución y regional, pero no se percibe el valor pedagógico para los docentes.

De acuerdo al estudio de González (2018) este modelo evaluativo, en opinión de expertos y ex ministros de educación, presenta las siguientes debilidades: es una prueba descontextualizada, la prueba pierde relación con el proceso educativo, no refleja valores y actitudes, no detecta completamente capacidades, la prueba no

es reflejo de una educación integral, es un filtro social, y el tipo de ítem no favorece la evaluación de habilidades superiores. Por otro lado, el perfil socioeducativo de la población estudiantil influye en los resultados de la PNB en el 2016: quienes obtuvieron mejores resultados son hombres y mujeres menores de 18 años residentes de la zona metropolitana de distritos de índice de desarrollo social medio y alto (González, 2018).

Tanto los docentes consultados en el grupo focal, como las investigaciones revisadas coinciden en el que los resultados de las pruebas nacionales no se aprovechan para mejorar los procesos educativos, tanto a nivel local como regional. No se revela el uso de estos resultados para la construcción y ejecución de política educativa (gestión curricular, capacitación docente, orientaciones pedagógicas) (González, 2018).

La transformación curricular iniciada en el período 2014-2018 para un currículo nacional por habilidades requiere una actualización del marco referencial de la prueba nacional, así como de su valor para mejorar los procesos de aprendizaje. Las investigaciones y las entrevistas coinciden en que un modelo evaluativo criterial es un modelo más apropiado para la evaluación y medición de habilidades, pues en la evaluación criterial los resultados de cada estudiante son comparados con unos criterios previamente establecidos, así como a los niveles de desempeño asociados.

De acuerdo al experto Juan Manuel Esquivel Alfaro PhD, en una prueba criterial cada estudiante debe recibir una realimentación individual de su desempeño expresada en forma descriptiva, lo cual supone un reto logístico y de recursos para el Ministerio de Educación Pública. Por otro lado, en opinión de este experto, la habilidad no debería ser tratada como un constructo y por ende aplicar sólo criterios psicométricos para elaborar las pruebas, más bien, debería prevalecer el valor educativo de la prueba. Concluye que medir habilidades complejas requiere de instrumentos complejos.

#### d. **Medios evaluativos más auténticos e integrales**

La prueba nacional de bachillerato sigue un formato de ítem característico de las pruebas estandarizadas a nivel internacional. El formato de un ítem se refiere a la manera en que se presentan, estructuran o disponen sus partes para plantear la tarea de evaluación en una prueba. Existen dos grandes categorías de formato según lo que debe hacer el evaluado: selección de respuesta y producción de respuesta (Ver Figura 2).



Figura 3. Tipos de formatos de ítem según UNESCO, 2009, p.17

La PNB dado su carácter masivo y racionalización del costo, sobre todo, utiliza un formato de ítem de selección múltiple, que está compuesto por tres elementos: contexto, enunciado y opciones. Un ejemplo de la organización de este tipo ítem sería:

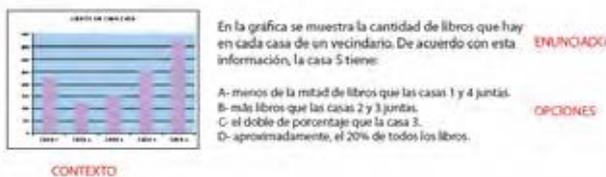


Figura 4. Ejemplo de ítem de selección múltiple, tomado de UNESCO, 2009, p.21

Los informes internacionales, el grupo focal y las entrevistas coinciden en que el formato de selección de respuesta es insuficiente para la evaluación de habilidades, pues existe un creciente interés en formas innovadoras y auténticas de evaluación del aprendizaje socialmente valorado. Según Díaz Barriga (2005) se le denomina auténtica por considerarse alternativa a la cultura evaluativa centrada en la prueba escrita que recoge el conocimiento declarativo y factual. Además, es congruente con el postulado constructivista del aprendizaje centrado en el desempeño en un contexto y situaciones determinados. Este tipo de evaluación también recibe el nombre de *performance assessment*.

El estudio de González (2018) sobre las PNB señala que los expertos consultados coinciden en que la prueba nacional debería ser criterial y no referida a la norma. Dado que el Programa de Bachillerato Internacional (BI) utiliza un modelo criterial de evaluación de los aprendizajes se visitó el Liceo de Santo Domingo de Heredia, para conocer de cerca su implementación. Las entrevistas realizadas al director y al coordinador académico del programa BI coinciden en lo siguiente:

- Bachillerato Internacional entiende cada asignatura dentro de su propia naturaleza, entonces las pruebas varían dependiendo de la asignatura, no se evalúa de la misma forma literatura que historia, por ejemplo.
- Los estudiantes además de las pruebas escritas llamadas “papers” demuestran el aprendizaje con formas alternativas de evaluación: presentan y defienden un ensayo de dos mil palabras, hacen defensas orales, monografía, y acción social.
- La evaluación interna se basa en la confianza hacia el estudiante.
- La evaluación interna está articulada con la evaluación externa internacional, se dice que la evaluación interna es “predictiva” de la evaluación externa.
- Se utiliza una evaluación criterial pues el rendimiento viene interpretado por bandas de desempeño.
- Cada asignatura tiene su propia banda de desempeño y criterios.
- Cada docente y coordinador de BI debe capacitarse para implementar el programa.
- Existe libertad para que el estudiante elija el tema que desea desarrollar en la monografía, pues lo que interesa es la coherencia, estructura, citas y argumentación.
- El BI está pensado para que el estudiante demuestre lo que sabe.
- Cada asignatura tiene su propia banda.
- Los estudiantes conocen lo que se espera de ellos y la forma en que serán evaluados

## 2.5 Discusión

La actualización del marco evaluativo costarricense es una tarea que supera la mera confección de una nueva prueba nacional para bachillerato. No solo implica un cambio en el modelo evaluativo, sino que debe concebirse como parte de un sistema evaluativo nacional. El potencial que tiene la evaluación y la medición sólo se logrará cuando exista una articulación de los distintos tipos de evaluación en el marco de un sistema único y coherente, que contribuyen a apoyar el aprendizaje en todos los niveles: alumnado, centros educativos, autoridades locales y sistema educativo.

De acuerdo a las buenas prácticas en política educativa, para el éxito de la actualización del marco evaluativo se requiere la participación de las partes interesadas como docentes, estudiantes, academia, sociedad civil,

expertos, entre otros (OCDE, 2015). Una vez asegurada la participación se deberá a llegar a consensos para la creación de un sistema de evaluación nacional, el cual integre diversos instrumentos evaluativos.

Un tema crucial que aún no se ha resuelto es la comprensión de un currículo por habilidades en el contexto educativo costarricense. Según la política curricular y los programas de estudio, las habilidades deben comprenderse desde una perspectiva socioconstructivista del aprendizaje. Lo cual implica que todo aprendizaje es situado a las prácticas, lenguaje, instrumentos de una comunidad, según ya fue señalado, cada disciplina del currículo es una comunidad con su propia cultura, lenguaje e instrumentos y modelos (Portillo-Torres, 2017). Esto supone un problema para modelos estadísticos de elaboración de pruebas nacionales. Un modelo no estadístico es más conveniente, sin embargo, también es mucho más caro. Requiere de un personal calificado, pues la prueba se fundamenta en el juicio de expertos.

Para los docentes del grupo focal y el modelo evaluativo del *Programa Bachillerato Internacional* la evaluación debe aspirar a formas de evaluación más auténticas, como la evaluación basada en problemas (*problem-based assessment*) y la evaluación de desempeño. La evaluación auténtica presenta al alumno tareas o desafíos de la vida real para cuya resolución debe desplegar un conjunto integrado de conocimientos, destrezas y actitudes.

La evaluación auténtica presenta al alumno tareas o desafíos de la vida real para cuya resolución debe desplegar un conjunto integrado de conocimientos, destrezas y actitudes. Entre las actividades o tareas comúnmente utilizadas en las evaluaciones auténticas por ejecución se incluyen: presentación oral, ensayos, experimentos, proyectos, presentaciones, tareas colaborativas, casos de la vida real, resolución de problemas y portafolio. La principal característica de estas actividades de ejecución es que se evalúa una variedad de conocimientos y habilidades integradas al pedirles a los estudiantes que realicen una tarea en vez de proporcionar una respuesta correcta. Son más efectivos que los formatos cerrados (tipo prueba escrita) para evidenciar el logro de aprendizajes complejos.

Investigaciones en algunos países muestran que este tipo de evaluación es conveniente para habilidades superiores las cuales son contextualizadas y específicas para una situación. Varios estudios muestran que este tipo de evaluación tiene un impacto positivo en la enseñanza, es

decir, es probable que el docente ajuste sus estrategias para alinearlas a las tareas señaladas en la evaluación del desempeño (OCDE 2015).

### 3. Conclusiones

Las investigaciones a nivel nacional e internacional apuntan al poco valor educativo de las pruebas nacionales diseñadas bajo un modelo evaluativo referidos a normas, como es el caso de la Prueba Nacional de Bachillerato. Entre las debilidades de esta prueba esta su descontextualización, su poca relación con el proceso educativo, su reducción al dominio cognitivo y el uso de ítems de selección de respuesta que no favorece la evaluación de habilidades superiores.

A corto plazo debe ser una prioridad para el sistema educativo costarricense el desarrollo de un marco conceptual sólido para la mediación y evaluación de un currículo por habilidades, lo cual favorecerá una efectiva articulación entre los diferentes componentes de la evaluación, por ejemplo, entre la evaluación interna (aula) y la evaluación externa (nacional). Las ventajas de una adecuada articulación se observó en la visita hecha al colegio Liceo Santo Domingo de Heredia donde se aplica del Programa de Bachillerato Internacional.

Se concluye que modelos evaluativos no estadísticos, como la evaluación criterial a través de *performance assessment* son un medio más idóneo para la evaluación de las habilidades pues el estudiantado debe demostrar sus capacidades en la ejecución de actividades en contextos específicos.

### Referencias

- Block, J. (Ed.) (1971). *Mastery Learning: Theory and Practice*. [Aprendizaje para el dominio: teoría y práctica], New York: Dyden Press.
- Cohen, L., Manion, L., y Morrison, K. (2007). *Research Methods in Education* [Métodos de investigación en educación], USA y Canada: Routledge
- Consejo Europeo (2009). *Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados*. España: EACEA P9 Eurydice
- Díaz Barriga, F. (2005). *Enseñanza situada: vínculo entre la escuela y la vida*. México: McGraw Hill
- Dominique, M., Sluijsmann, F., Prins, y Materns (2006). The design of competency-based performance as-

- assessment in e-learning. *Learning Enviroments Research*, pp.46-66, DOI: 10.1007/s10984-005-9003-3
- Griffin, P. y Care, E. (2014). Developing learner collaborative problem solving skills. [Desarrollando habilidades de resolución de problemas en estudiantes]. Recuperado de <https://sodas.ugdome.lt/bylos/GENERAL/8af7dd98-d82c-4d81-90ed-7f912c0dfcf0.docx>
- González, F. (coord.) (2018). Balance crítico de objetivos, modelo educativo, resultados y utilidad de la información de las Pruebas Nacionales de Bachillerato en Educación Media como instrumento de certificación del conocimiento de las y los estudiantes. Resumen general. Recuperado de <https://drive.google.com/file/d/1rA1xMf3WheK-cfacKq3xmF8SAUFzmpH/view>
- Hernández, R., Fernández, C. y Baptista, M. (2014). Metodología de la Investigación. 6ta ed. México: Mc Graw Hill.
- López, V. (Coord.)(2016). Evaluación formativa y compartida en Educación Superior. México: Alfaomega Gurpo Editor.
- Martínez, P. (2006). El método de estudio de caso. Estrategia metodológica de la investigación científica. *Revista Pensamiento y Gestión*, n°20, 167-193. Recuperado de <http://www.redalyc.org/articulo.oa?id=64602005>
- Martinic, S. (2010). Evaluación y las reformas educativas en América Latina. *Revista Iberoamericana de Evaluación Educativa*, 3(3), 31-43. Recuperado de [https://repositorio.uam.es/bitstream/handle/10486/661623/RIEE\\_3\\_3\\_2.pdf?sequence=1](https://repositorio.uam.es/bitstream/handle/10486/661623/RIEE_3_3_2.pdf?sequence=1)
- Ministerio de Educación y Deporte (2017). Aprender 2016 Bookmark. Establecimiento de puntos de corte. Serie de Documentos Técnicos. Argentina: Secretaría de Evauación Educativa
- MEP (2016a). Fundamentación Pedagógica de la Transformación Curricular 2015. Viceministerio Académico. San José: MEP. Recuperado de : <https://www.mep.go.cr/sites/default/files/documentos/transf-curricular-v-academico-vf.pdf>
- MEP (2016b). Informe Nacional Bachillerato de la Educación Formal 2015. Dirección de Gestión y Evaluación de la Calidad. San José: MEP
- MEP (2018). Reglamento de Evaluación de los Aprendizajes, Decreto N°40862, La Gaceta Diario Oficial, San José: Imprenta Nacional
- MECD (2013). Las competencias básicas: una visión internacional desde el informe OECD “Sinergias para aprender mejor: perspectiva internacional sobre evaluación”. Recuperado de: <http://blog.educalab.es/inee/2013/06/19/las-competencias-basicas-una-vision-internacional-desde-el-informe-ocde-sinergias-para-aprender-mejor-perspectiva-internacional-sobre-evaluacion/>
- National Research Council (2012). Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century [Educación para la Vida y el Trabajo: desarrollando conocimiento y habilidades transferibles en el Siglo 21], Washington, DC: The National Academies Press.
- OCDE (s.f). El programa PISA de la OCDE. Qué es y para qué sirve. París: OCDE Recuperado de <https://www.oecd.org/pisa/39730818.pdf>
- OECD (2013). Synergies for Better Learning: an international perspective on evaluation and assessment. [Sinergias para un mejor aprendizaje: una perspectiva internacional de la evaluación], Paris: OECD Publishing. Recuperado de <http://dx.doi.org/10.1787/9789264190658-en>
- OCDE (2015). Política educativa en perspectiva 2015. Hacer posible las reformas. España: Fundación Santillana. Recuperado de [https://publiadmin.fundacion-telefonica.com/index.php/publicaciones/add\\_descargas?tipo\\_fichero=pdf&idioma\\_fichero=\\_&title=Pol%C3%ADtica+educativa+en+perspectiva+2015&code=456&lang=es&file=politica-educativa-en-perspectiva-2015.pdf](https://publiadmin.fundacion-telefonica.com/index.php/publicaciones/add_descargas?tipo_fichero=pdf&idioma_fichero=_&title=Pol%C3%ADtica+educativa+en+perspectiva+2015&code=456&lang=es&file=politica-educativa-en-perspectiva-2015.pdf)
- Portillo-Torres, M. (2017). Educación por habilidades: Perspectivas y retos para el sistema educativo. *Revista Educación*, 41 (2), 2215-2644, DOI: <http://dx.doi.org/10.15517/revedu.v41i2.21719>
- Rodriguez, G. e Ibarra, S. (2016). E-Evaluación orientada al e-Aprendizaje estratégico en Educación Superior. México: Alfaomega, Narcea
- Sahlberg, P. (2006). *Models of curriculum development: international trends and the way forward* [Modelos de desarrollo curricular: tendencias internacionales y camino a seguir] Trabajo presentado en la Conferencia Internacional sobre reformas curriculares e implementación en el siglo XXI, Estambul, Turquía.
- Yin, R. (2014). *Case Study Research. Design and Methods*. [Investigación de Estudio de Caso. Diseño y Métodos] (5ed). California: Sage

### **Reconocimientos**

Se agradece a la Fiscalía del Colegio de Licenciados y Profesores en Letras, Filosofía, Ciencias y Artes de Costa Rica (COLYPRO) que solicitó y financió la realización de esta investigación. Especial agradecimiento para el Lic. Jorge Quesada Lacayo quien acompañó en todas las fases de este estudio y brindó consejos oportunos.

Un reconocimiento también para el M.Sc. Olman Bolaños Ortíz, Coordinador de las carreras de Bachillerato y Licenciatura en Administración Educativa de la Universidad Estatal a Distancia, por su apoyo en la socialización y divulgación de los resultados de esta investigación.