

[Cierre de edición el 01 de Setiembre del 2020]

<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

Análisis de casos múltiples de prácticas evaluativas en relación con estándares internacionales de evaluación de los aprendizajes

Analysis of Assessment Practices With Multiple Cases Methodology Related to International Learning Assessment Standards

Análise de múltiplos casos de práticas de avaliação em relação aos padrões internacionais de avaliação da aprendizagem



Arcelia Ramos-Monobe
Tecnológico de Monterrey
Zapopan, México

 arcelia.ramos.monobe@outlook.com
<http://orcid.org/0000-0001-8052-7272>

Katherina Gallardo-Córdova
Tecnológico de Monterrey
Monterrey, México

 katherina.gallardo@tec.mx
<http://orcid.org/0000-0001-8343-9518>

Dulce Camacho-Gutiérrez
Tecnológico de Monterrey
Zapopan, México

 dulce@tec.mx
<http://orcid.org/0000-0002-1084-3082>

Recibido • Received • Recebido: 23 / 11 / 2018
Corregido • Revised • Revisado: 19 / 04 / 2020
Aceptado • Accepted • Aprovado: 21 / 07 / 2020

Resumen: El presente artículo es resultado de un estudio de casos múltiples (cuatro casos) en México y Colombia, con estudiantes del curso Evaluación del Aprendizaje (Gallardo Córdova, 2017), en un posgrado en educación ofrecido en modalidad en línea. El objetivo de este estudio fue analizar las decisiones tomadas por las personas participantes a partir de los resultados obtenidos de la aplicación de un cuestionario diagnóstico sobre prácticas de evaluación del aprendizaje. Las decisiones del estudiantado estuvieron focalizadas en la elaboración de materiales o el cambio en prácticas y tuvieron la intención de incidir en la conducción de la evaluación sumativa y formativa. Participaron un total de 18 docentes y personal directivo en los cuatro casos analizados. Los principales resultados fueron: comprender que, en la mayoría de los casos, la toma de decisiones se inclinó más por consolidar los aspectos teóricos que sustentan las prácticas, tales como definición de tipos de evaluación, integración de taxonomías de aprendizaje, instrumentación y retroalimentación, entre las principales. Asimismo, en cuanto a los estándares internacionales propuestos por el Joint Committee on Standards for Educational Evaluation (JCSEE), especialmente los referidos a la utilidad, factibilidad y ética, se integraron en la toma de decisiones



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

para instruir y sensibilizar al personal docente. Se confirmaron la complejidad y la demanda de tiempo que exigen los procesos de evaluación, específicamente aquellos de tipo formativo. Se concluye que el proceso formativo sobre las prácticas evaluativas permitió la reflexión sobre la necesidad de incluir y visualizar los estándares, buscando la mejora continua de los procesos de evaluación del aprendizaje.

Palabras clave: Evaluación del aprendizaje; casos múltiples; estándares internacionales de evaluación; retroalimentación.

Abstract: This article is the result of four case studies in Mexico and Colombia with students from the Learning Evaluation course (Gallardo Córdova, 2017) of an online postgraduate program in education. The study aimed to analyze the participants' decisions based on the results of a diagnostic questionnaire on learning assessment practices. The students' decisions relate to the preparation of materials or the change in practices to influence the conduct of the summative and formative evaluation. Eighteen teachers and management staff members participated in the four cases analyzed. The main results were to understand that, in most cases, decision making was more inclined to consolidate the theoretical aspects that sustain the practices, such as the definition of types of evaluation, integration of learning taxonomies, instrumentation, and feedback. Likewise, in terms of international standards stated by the Joint Committee on Standards for Educational Evaluation (JCSEE), especially those related to the Utility, Feasibility and Ethics were integrated into the decision making to educate and sensitize teachers. The complexity and the demand for the time required by the evaluation processes were confirmed, specifically those of a formative nature. It is concluded that the process carried out allows the participants to reflect on the evaluation practices for a continuous improvement.

Keywords: Evaluation of learning; feedback; international evaluation standards; multiple cases.

Resumo: Este trabalho de investigação científica é o resultado de um estudo de casos múltiplos (quatro casos) em México e Colômbia com estudantes do Curso de Avaliação de Aprendizagem (Gallardo Córdova, 2017) na pós-graduação em ensino oferecido em modo online. O objetivo deste estudo foi analisar as decisões tomadas pelas pessoas participantes com base nos resultados obtidos com a aplicação de um questionário de diagnóstico sobre práticas de avaliação da aprendizagem. As decisões foram feitas por estudantes em relação ao desenvolvimento de materiais ou mudança nas práticas, com a intenção de influenciar o comportamento da avaliação somatória e formativa. Um total de 18 professores e executivos participaram dos quatro casos analisados. Os principais resultados foram: compreender que, na maioria dos casos, a decisão tomada inclinou-se para consolidar os aspectos teóricos que sustentam as práticas, tais como a definição de tipos de avaliação, integração da aprendizagem taxonomias, instrumentação e retroalimentação. Da mesma forma, no que diz respeito aos padrões internacionais proposto pelo Joint Comittee on Standars for Educational Evaluation (JCSEE), especialmente aqueles que se referem à utilidade, viabilidade e ética foram integrados na tomada de decisões para educar e sensibilizar os professores. As demandas de tempo e a complexidade que exigem os processos de avaliação foram confirmados, especificamente, aqueles orientados a formação. Conclui-se que o processo formativo sobre as práticas de avaliação permitiu a reflexão sobre a necessidade de incluir e visualizar os padrões, buscando uma melhoria continua das avaliações de aprendizagem.

Palavras chave: avaliação de aprendizagem; casos múltiplos; retroalimentação; padrões internacionais de avaliação.



Introducción

Con base en un estudio exploratorio sobre prácticas de evaluación vinculadas con estándares de evaluación internacional (Gallardo Córdova et al., 2015), se planteó la necesidad de analizar el proceso de toma de decisiones realizada por estudiantes de posgrado (desde ahora *participantes*) a partir de los resultados obtenidos de la aplicación de un cuestionario diagnóstico en las escuelas donde laboran. Se parte de la idea que del aprendizaje, que de ello se deriva, coadyuva a que identifiquen discrepancias entre lo que los estándares indican y las prácticas que se llevan a cabo.

Por lo anterior, este estudio se centra en la experiencia del estudiantado de posgrado en educación y tecnología educativa, en modalidad en línea, principalmente en escenarios de México y Colombia. Una de las materias que comparten estos programas de posgrado es la denominada *Evaluación del aprendizaje*. El proyecto integrador de la materia se enfoca en llevar a cabo prácticas en la que el personal docente sostiene un proceso formativo auténtico y situado (Hancock, 2007). Se espera que el estudiantado logre apropiarse de aspectos teórico-conceptuales en primera instancia, entre los cuales se encuentra el estudio de los estándares internacionales de evaluación como pilar para el proceso de diagnóstico y activación de soluciones educativas (JCSEE, 2014).

Así, el objetivo de este estudio fue analizar las decisiones tomadas por las personas participantes a partir de los resultados obtenidos de la aplicación de un cuestionario diagnóstico sobre prácticas de evaluación del aprendizaje clasificadas en los cuatro estándares propuestos.

Referentes conceptuales

Los estándares internacionales de evaluación son emitidos por el organismo binacional *Joint Committee on Standards for Educational Evaluation* (JCSEE); dichos estándares se han concebido con el objetivo de establecer prácticas de calidad al momento de llevar a cabo procesos de evaluación educativa. Según este organismo, cada cinco años se realiza una revisión periódica para actualizar la pertinencia de dichos estándares en cuanto a prácticas, tendencias y nuevas problemáticas que se presentan en el medio educativo. Los estándares relacionados con la evaluación del aprendizaje se han clasificado de la siguiente forma: utilidad, factibilidad, precisión y ética (JCSEE, 2014).

En primer lugar, la *utilidad*, es entendida como un conjunto de parámetros que determina en qué medida responden las prácticas de evaluación a las necesidades de información de su público usuario. Escudero (2013) comenta que la utilidad se encuentra generalmente como primer eslabón de la cadena de evaluación del aprendizaje, asimismo, argumenta que el análisis de la utilidad y el uso de la evaluación constituyen unos de los mayores problemas a los que hoy se enfrentan los equipos evaluadores, pues, sin esta condición, el proceso de evaluación no cobra valor.



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

En segundo lugar, la *factibilidad*, comprendida como un conjunto de parámetros que permiten asegurar que el proceso de evaluación podrá ser diseñado y aplicado, cuidando la confiabilidad de su aplicación y sus resultados. En tercero, la *precisión* que postula la evaluación mediante el mecanismo que transmite información que el sujeto evaluado requiere para comprender y asegurar el logro de las metas u objetivos formativos. Por último, el estándar de *ética* está conformado por una serie de parámetros que buscan garantizar que el proceso evaluativo se realice de forma justa y legítima, anteponiendo el bienestar de los sujetos involucrados, así como de los sujetos que se ven afectados por sus resultados. Cabe mencionar que este tipo de estándares se utiliza e investiga en varias disciplinas con fines de evaluación, como lo es, por ejemplo, la educación médica (George, et al., 2014).

Un aspecto importante en torno a la diseminación de estos estándares radica en que, en programas de preparación docente en países latinoamericanos, éstos son escasamente conocidos y practicados (Cabra Torres, 2008; Gallardo Córdova et al., 2015). No obstante, se sabe que una de las competencias pedagógicas que los programas de profesionalización docente en América Latina busca desarrollar la relacionada con la evaluación del aprendizaje, particularmente en quienes se enfocan en el modelo educativo basado en competencias (MEBC) y que implica un cambio de paradigma (Álvarez, 2005; Gallardo Córdova y Gil Rendón, 2016; Tejada Fernández, 2013; Valenzuela, 2005). Desarrollar la competencia de evaluar implica todo un reto, en tanto se deba trabajar en el proceso de evaluación desde una perspectiva diferente a la que se ha realizado por años en el modelo por objetivos (Cano García, 2008; Fernández March, 2010).

En países como México y Colombia, por ejemplo, se están haciendo esfuerzos en aras de ofrecer programas de educación formal para la profesionalización docente acorde con las exigencias educativas contemporáneas, como las que exige el MEBC (Castillo Ortega y Canul de Dios, 2015; Morán et al., 2015). En consecuencia, para cumplir con estas expectativas de formación, el personal docente no solo debe conocer nuevos parámetros y prácticas para la evaluación del aprendizaje, que se refiere a las decisiones tomadas a partir del diagnóstico realizado. De hecho, las expectativas radican en que puedan desempeñarse en situaciones auténticas que permitan dar cuenta de su capacidad para comprender las nuevas problemáticas de la evaluación, y tomar decisiones para modificar y mejorar los procesos existentes (Alarcón y Donoso-Díaz, 2012; Alarcón Leiva et al., 2014; Hancock, 2007; Tobón, 2008).

Por lo anterior, las preguntas del estudio respecto a las prácticas de evaluación son:

¿Cómo toman las decisiones las personas participantes del proceso formativo? ¿Cómo contribuyen los procesos formativos a los que se expusieron al desarrollar su competencia en evaluación del aprendizaje? ¿En qué coinciden las experiencias en los diferentes casos?

A fin de dar respuesta a estas preguntas se plantean los siguientes supuestos o proposiciones:

- 1) Los estándares internacionales de evaluación del aprendizaje emitidos por la JCSEE permiten:
 - a) Diagnosticar la forma en que se llevan a cabo los procesos de evaluación en el aula con la finalidad de entender la situación de evaluación con mayor precisión (JCSEE, 2014).
 - b) Guiar en el proceso formativo para la mejora e implantación de prácticas evaluativas (JCSEE, 2014).
- 2) Comprender los procesos inherentes a la evaluación sumativa y la consideración de su validez y confiabilidad; favorecer un análisis general de las decisiones y mejora continua del proceso (Gallardo Córdova, 2013).
- 3) El uso del modelo de retroalimentación de Hattie y Timperley (2007) permite emitir recomendaciones al estudiantado de manera organizada y en función de los objetivos de aprendizaje, pudiendo acercarlo a los desempeños esperados (Hattie y Timperley, 2007; Shute, 2008; Wiggings, 1998).

Metodología

Esta investigación se realizó con base en la metodología de casos múltiples conforme lo explica Yin (2003). La selección de los casos se basó en la simultaneidad de los eventos en torno a la mejora de prácticas de evaluación del aprendizaje con una misma metodología de trabajo, aunque en diferentes contextos. Se juzgó pertinente esta metodología, dado que el objetivo es analizar en detalle similitudes y diferencias con respecto al proceso de diagnóstico, intervención, mejoras y áreas de oportunidad en procesos de evaluación del aprendizaje desde el marco de los estándares internacionales. Asimismo, como parte del protocolo propio de esta metodología se buscaron patrones comunes en los casos analizados. A su vez, facilitó predecir posibles repeticiones en situaciones similares, así como contrastar los resultados en función de la teoría (Baxter y Jack, 2008).

Este estudio partió de los datos recolectados en el curso en línea denominado *Evaluación del aprendizaje* (Gallardo Córdova, 2017) que se imparte para dos programas de posgrado en educación. Se inscribieron 248 estudiantes de diferentes nacionalidades, quienes formaron 58 grupos de trabajo (casos).

Se seleccionaron cuatro casos para ser analizados a partir de un muestreo intencional, por ser representativos del proceso, y por incorporar diferentes disciplinas y niveles educativos.

Para fines de esta investigación, se utilizaron las diferentes evidencias del proceso generado por el estudiantado: diagnóstico de necesidades en materia de evaluación, entrevistas (ver Apéndice A), reporte final de los proyectos de mejora para las evaluaciones formativa y sumativa, y un video con el testimonio de agentes informantes. Todas las evidencias fueron etiquetadas para su referencia, utilizando las iniciales de la fuente y el número de caso; por



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

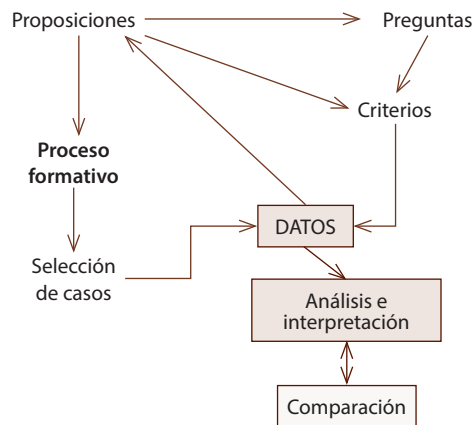
ejemplo, se codificó el mapa de decisiones (MD_#), el proyecto final (PF_#) y el video (V_#); en todos se añade el número de caso en cuestión.

A continuación, se presenta una clasificación de las personas que participaron -según su rol- dentro de la investigación:

- *Participantes*: El estudiantado de posgrado, quienes realizaron diagnóstico, toma de decisiones, prácticas de mejora, elaboración de reporte de resultados y entrevistas a agentes informantes.
- *Agentes informantes*: es aquel personal directivo y docente que participó brindando información para el diagnóstico general y específico sobre las prácticas de evaluación para cada caso. En este rubro también se encuentran las personas beneficiarias de las mejoras realizadas, quienes juzgaron los resultados para cada caso.

Con base en los pasos de la metodología de casos múltiples (Yin, 2003), se *determinaron las proposiciones* o supuestos básicos de la investigación asociándose a las preguntas que guiaron el análisis realizado (Figura 1). Se seleccionaron cuatro casos de acuerdo con los supuestos planteados.

Figura 1: Procedimiento de investigación acorde a la metodología de casos múltiples



Nota: Elaboración propia.

El siguiente paso fue la vinculación de los datos acopiados en las evidencias con las proposiciones que guiaron la investigación; entre ellos se distinguió el proceso formativo como una condición general vivencial de las personas participantes, no así las decisiones tomadas en cada caso, situación que permite declarar los estándares internacionales de evaluación: utilidad, factibilidad, ética y precisión (JCSEE, 2014) como criterios para interpretar los datos.



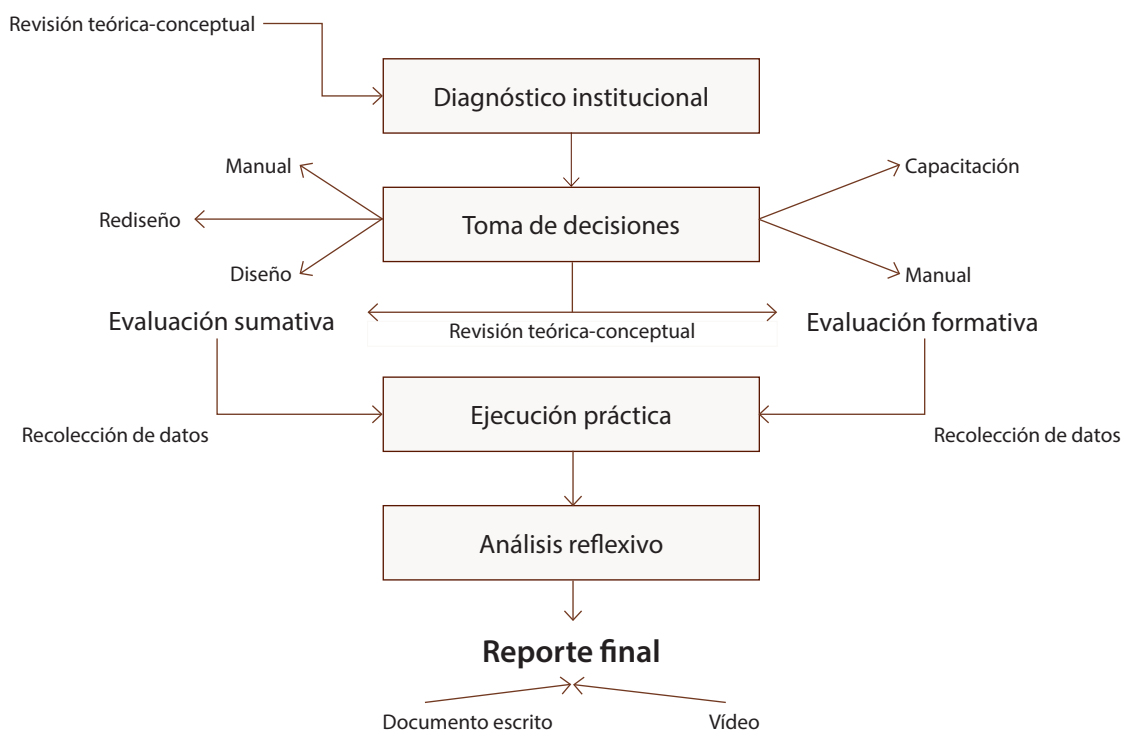
De manera que, partiendo de las semejanzas del proceso formativo al analizar e interpretar los datos recolectados, se busca responder las preguntas de investigación, las similitudes (aún en contextos diferentes) entre las decisiones tomadas, y los beneficios percibidos tanto por participantes como para las personas agentes informantes.

Conforme lo dicta la metodología de casos múltiple, se definieron tres fuentes de información: 1) las decisiones sobre las posibles mejoras a realizar con base en el diagnóstico; 2) los reportes de resultados, y 3) las entrevistas de cierre a las personas beneficiarias. Estas fuentes sirvieron para dar soporte a la credibilidad de los datos, comprender mejor el fenómeno estudiado, saber qué diferencias existen y qué generalizaciones se podrían emitir para los casos de estudio.

Resultados

Los casos tienen una base común que permitió la generación de los reportes finales según se ilustra en la [Figura 2](#). Las generalidades del proceso formativo, por tanto, se refieren a la condición vivencial general y estructurada a fin de que las personas participantes realizaran sus prácticas.

Figura 2: Proceso formativo para las prácticas



Nota: Elaboración propia.

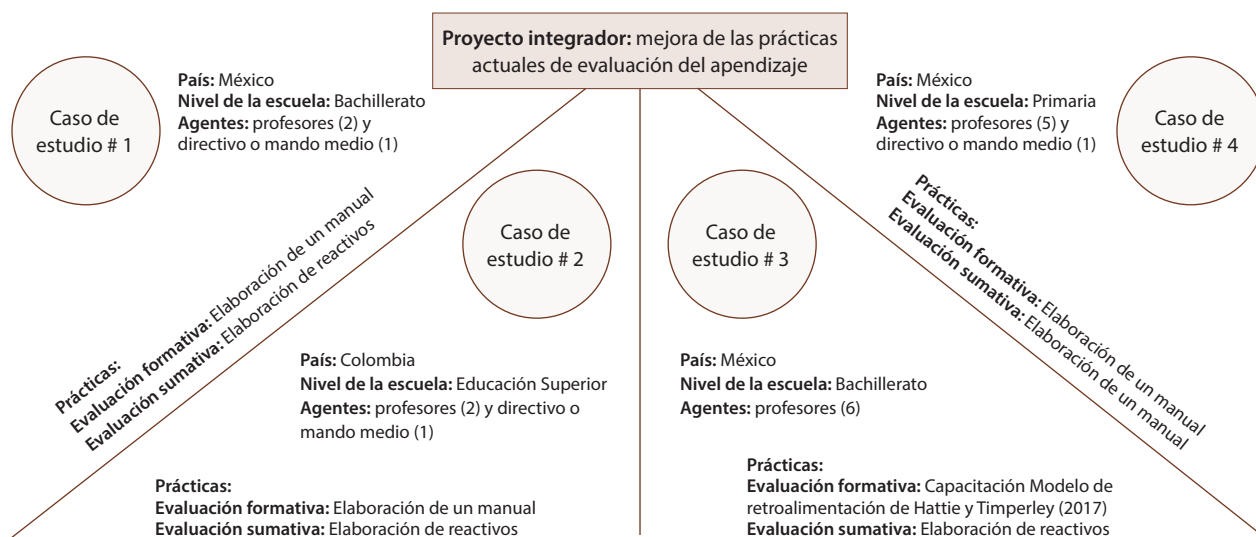
<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

El proceso general marcó un referente común y constó de:

- a. Preparación teórica: las personas participantes estudiaron (ocho semanas) aspectos paradigmáticos de la evaluación y las bases para diseñar un sistema integral de evaluación a partir de estándares, lo que les ayudaría a sustentar los cambios necesarios en las prácticas de evaluación.
- b. Diagnóstico: las personas participantes aplicaron el cuestionario sobre prácticas de evaluación del aprendizaje en función de estándares internacionales (JCSEE, 2014). Este instrumento fue desarrollado y validado por las profesoras investigadoras expertas en el área de evaluación con base en los estándares declarados por la JCSEE. Su nivel de consistencia interna es de 0,91 a partir de la aplicación del alfa de Cronbach con base en 678 aplicaciones en diferentes países (Gallardo Córdova et al., 2015). Este instrumento consta de 35 preguntas con respuesta en escala tipo Likert, cuya intención fue indagar en prácticas de evaluación en institucionales educativas. Con estos datos quienes participaron emitieron un diagnóstico institucional y tomaron decisiones con base en opciones controladas para la evaluación sumativa y formativa.
- c. Preparación para la práctica: las personas participantes accedieron a sesiones sincrónicas y grabadas en las que se trataron diferentes procesos asociados a las opciones controladas. Para las dos opciones de evaluación formativa se requirió prepararse en el modelo de retroalimentación (Hattie y Timperley, 2007) que implica llevar a cabo el proceso considerando tres partes que se complementan entre sí: *Feed Up* (hacia dónde voy) *Feed Back* (cómo lo estoy haciendo) y *Feed Forward* (que sigue después de esto). Para la evaluación sumativa, se revisaron aspectos técnicos básicos para la elaboración de reactivos de opción múltiple a fin de generarlos o revisar bancos previamente creados. Así, se realizaron análisis de resultados para estimar tanto el rendimiento académico del estudiantado, como las características psicométricas del examen; procedimientos que podrían también exponerse al elaborar un manual de procedimientos (Gallardo Córdova, 2013).
- d. Recolección de datos: el estudiantado realizó, a la luz de los estándares, las prácticas sumativa y formativa seleccionadas como parte de su plan de mejora de los procesos de evaluación institucional, con el fin de emitir un reporte final que se acompañaría del testimonio de las personas usuarias beneficiadas.

A partir de este proceso general, cada uno de los casos que se presentan a continuación refleja el resultado de las decisiones tomadas como parte de su proceso formativo. La **Figura 3** presenta el panorama general de las decisiones tomadas en cada uno de los casos seleccionados.

Figura 3: Particularidades de los casos analizados



Nota: Elaboración propia.

Caso 1

El estudiantado seleccionó por conveniencia dos instituciones educativas de nivel medio superior privadas a nivel nacional en locaciones distintas, guiadas por similares regulaciones, para llevar a cabo sus prácticas. No obstante, los procesos de evaluación competen a cada una de las escuelas y sus cuerpos docentes adscritos.

A partir del diagnóstico realizado, se supo que las instituciones seleccionadas comparten ciertas falencias vistas en el marco de los estándares; es decir, no se perciben variaciones de importancia en los datos acopiados de las personas agentes informantes.

El diagnóstico fue aplicado a seis agentes informantes (un directivo y cinco docentes del área de matemáticas en bachillerato). En ambas instituciones se reporta, a la luz de los estándares, lo siguiente:

1. Se carece de un manual de procedimientos de evaluación (ética),
2. El proceso evaluativo no se analiza a la luz de los resultados buscando una mejora continua (precisión) que implica falta de seguimiento y control.

<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

3. Los procesos de evaluación y la evaluación misma se realizan desde una perspectiva unidireccional (desde la perspectiva docente) dejando de lado la integración del resto de los actores en el proceso, que solo son sujetos de la acción y no partícipes del proceso (factibilidad).
4. No es posible determinar si el estudiantado percibe los beneficios con respecto al proceso de evaluación (utilidad).

Lo anterior pone de manifiesto dos áreas de oportunidad: 1) la preparación del personal docente, dado que reportan no saber cómo transformar la evaluación en una oportunidad de aprendizaje para el estudiantado, y 2) se carece de seguimiento al proceso de su aprendizaje. En lo que respecta a los procesos de evaluación sumativa, las personas participantes eligieron generar un banco de reactivos, mientras que para la evaluación formativa eligieron diseñar un manual de procedimientos.

La decisión de generar una base de reactivos se justificó en la necesidad de atender los elementos asociados a los estándares de precisión y utilidad; el estudiantado reportó que cuentan con reactivos que se aplican, pero que no existe un análisis de estos mismos y, por ello, el estudiantado no se puede percibir como protagonista del proceso. Estas personas plantean, de forma textual, que, como resultado de su intervención práctica en el proceso sumativo, serían capaces de *analizar detalladamente las causas que generan estas debilidades en el proceso, cumplir con el reglamento académico, y además detectar fortalezas y áreas de oportunidad como pocas veces se hace con los reactivos existentes* (MD_1).

La decisión tomada de realizar un banco de reactivos implicó un proceso de planeación reflexiva, teniendo en cuenta la taxonomía de [Marzano y Kendall \(2008\)](#) y los saberes asociados para un curso de matemáticas, que implicó el diseño integral de la evaluación. Se aplicaron dos exámenes que incluían todos los ítems, con los resultados obtenidos se realizó un análisis estadístico complejo desde dos perspectivas: 1) desde los resultados de la aplicación en un grupo, y 2) desde un análisis de las características psicométricas que implicó el análisis de las características de los reactivos y su confiabilidad (factibilidad).

El enfoque del análisis y su reporte final (PF_1) condujeron con claridad a incidir en la problemática detectada en los estándares de factibilidad y de precisión, ya que se presentan los datos paramétricos e inferenciales con mucha claridad, justificados, con fórmulas y resultados obtenidos, entendiendo con ello lo complejo de los procesos estadísticos asociados al examen, su confiabilidad, y atendiendo los índices de discriminación y dificultad de cada uno de los ítems construidos. Sin embargo, desde esta elección, no se observa que se haya asociado al estándar de utilidad en cuanto a la percepción del alumnado como el centro del proceso; reiteró una postura unidireccional de los procesos de evaluación donde el personal docente tiene el control del proceso y no está preparado para transformar el paradigma hacia una evaluación para el

aprendizaje. Tampoco se cuenta con datos específicos sobre qué se hizo con estos resultados; no se sabe si fueron utilizados efectivamente para determinar fortalezas y debilidades del estudiantado y así marcar un camino hacia su mejora, como fue planteado en sus intenciones iniciales, aunque no era parte de lo requerido en su proceso formativo.

En lo que respecta a la evaluación formativa, con la decisión de elaborar un manual de procesos para evaluación formativa, quienes participaron reportan que las personas informantes declararon: *las evaluaciones no discriminan las debilidades y fortalezas de los alumnos... así como el desarrollo de habilidades enfatizado en la Precisión y la Utilidad (MD_1)*.

El enfoque de las personas participantes sigue alineado a la decisión que tomaron en la evaluación sumativa: pensar en el alumnado como el centro del proceso. Para ello, se realizaron entrevistas, aunque se sabe que dos de este grupo no participaron en el diagnóstico.

El manual presentado en este caso aludió a los conceptos teóricos revisados sobre la evaluación. No obstante, se observó que participantes y agentes informantes confundieron los procedimientos sumativos con los formativos, situación que se observa claramente en el manual (PF_1) y en las declaraciones sobre beneficios y utilidad del manual, en donde las respuestas aluden a *calificaciones, ponderaciones parciales, uso de rúbricas, y coevaluación (V1)*. El estudiantado no analizó las respuestas de las entrevistas a la luz del marco teórico conceptual de la evaluación formativa. En su manual se pudo apreciar, claramente, que se requiere, desde el punto de vista institucional, del trabajo colegiado y de la emisión y la difusión de criterios de evaluación (JCSEE, 2014) y la importancia de la retroalimentación como parte de un proceso que permitirá que el alumnado se empodere de su aprendizaje; aunque, en suma, como manual que se haya entregado a personal docente y directivo institucional, este pueda no cumplir con el cometido de convertirse en una guía de acciones (manual de procesos), ya que el enfoque es eminentemente teórico e integra información sobre procesos de evaluación sumativa.

Caso 2

Las personas participantes docentes de inglés como segundo idioma eligieron, de forma intencional, dos instituciones de nivel superior privadas, aunque de diferentes grados educativos. La aplicación del cuestionario diagnóstico se realizó a seis informantes (tres con posiciones directivas y tres docentes). Los programas de inglés responden al desarrollo formativo desde un nivel básico hasta uno avanzado.

El diagnóstico reveló diferencias importantes entre ambas instituciones educativas con respecto a los estándares de ética y utilidad. En la universidad, es evidente que existieron discrepancias entre las personas informantes (dos docentes, un directivo). En cuanto a ética, hubo diferencias con respecto a la existencia de manuales de procedimientos y falta de seguridad (confidencialidad) en los procesos de evaluación. Con respecto a la utilidad, también



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

se presentaron discrepancias sobre la preparación docente para llevar a cabo procesos de evaluación. De forma contraria, en la segunda institución, los reportes de las personas informantes (dos directivos y un docente) confluyen en dos grandes categorías: el alumnado como actor importante del proceso (para quien debe resultar útil la información, y a quien debe informarse sobre los criterios y consideraciones de equidad y discreción en las prácticas de evaluación), y el análisis de los procesos de evaluación como parte de la mejora continua.

En tanto, en factibilidad y precisión, ambas instituciones coinciden en que la evaluación se realiza desde la perspectiva docente y no se desarrollan análisis para la mejora continua, particularmente en lo que respecta a la evaluación sumativa (MD_2).

Después de realizar el análisis diagnóstico, las personas participantes eligieron realizar un banco de reactivos y elaborar un manual de procedimientos para la evaluación formativa.

Con la elaboración del banco de reactivos, su intención era incidir en la precisión; argumentan que este se haría *orientado a la validez y la justificación clara de la evaluación para los estudiantes, quienes merecen mayor claridad en las pruebas y sus resultados* (MD_2). La construcción de los reactivos de opción múltiple implicó un proceso de planeación inicial para, de acuerdo con la taxonomía de [Marzano y Kendall \(2008\)](#), identificar los dominios de conocimiento y niveles de procesamiento requeridos para evaluar el nivel básico de inglés en cuanto a gramática, vocabulario y comprensión oral. Se aplicó el examen en línea con estudiantes de tres carreras diferentes en la universidad; no se administró en la segunda escuela por no coincidir con la temática revisada al tiempo de la aplicación. Los resultados fueron analizados con respecto al desempeño del estudiantado y de acuerdo con las características psicométricas del examen en su conjunto.

El análisis del reporte final, en su sección sumativa, denota la suficiencia explicativa en cuanto a los procedimientos realizados, y la propuesta de mejora de los reactivos cuya discriminación era baja o con el grado de dificultad disminuido, en dicho reporte se indica -de forma adicional- la confiabilidad obtenida en el instrumento, al tiempo que se sugiere que estos reactivos podrían formar parte de la base de reactivos institucionales, luego de su proceso analítico crítico (factibilidad y utilidad).

Con respecto a la evaluación formativa, y en aras de elaborar el manual, se realizaron entrevistas a profundidad con las personas informantes (seis docentes y personal directivo, una persona no había participado en el cuestionario diagnóstico) y contrastaron la información con las fuentes teórico-conceptuales. El estudiantado reportó que con esta decisión se atenderán los indicadores de ética y precisión. Textualmente declaran: *enriquecería nuestra formación y apuntaría a la mejora de los procesos* (MD_2). Agregan que ayudaría a clarificar los manuales de procedimientos existentes, aludiendo a las discrepancias de información en las respuestas de las personas informantes.

El manual integra los tipos de evaluación y los momentos para la evaluación; se alude a la necesidad de un proceso de planeación en el que se haya reflexionado sobre las actividades y su relación con el proceso formativo del estudiantado, y se solicita asimismo la incorporación de criterios de evaluación y el uso de instrumentos. Se introduce la participación estudiantil en el proceso, ya sea para la autoevaluación o la coevaluación, lo que haría al estudiantado corresponsable informado de las prácticas de evaluación que implican la utilidad, la factibilidad y la precisión.

Al concluir su proyecto final el estudiantado reporta: *nuestra perspectiva de la evaluación cambia radicalmente. Ésta no está limitada sólo a la comprobación de saberes, sino a una compleja, pero significativa, secuencia de pasos que permite una evaluación del aprendizaje más justa y balanceada que da cuenta de lo alcanzado por los estudiantes, que arroja información útil para la mejora del programa mismo y de las prácticas educativas de docentes e institución por igual* (PF_2) lo que da cuenta de haber alcanzado las expectativas planteadas al tomar decisiones luego del diagnóstico.

Por su parte, las personas beneficiarias reportan, con respecto a la evaluación formativa, *mayor claridad, consciencia* (V_2), así como la posibilidad de estandarización de los procesos de evaluación formativa consensuada y atender, en el corto plazo, el diálogo entre el personal docente para favorecer el aprendizaje e incidir en la *dimensión metacognitiva* (V_2) del estudiantado de licenciatura y en la comunicación con los padres y madres de familia en el centro escolar. Es claro, por los planteamientos que se pueden apreciar en el video, que tras la revisión del manual, tanto personal docente como directivo comprenden las implicaciones y los retos en la realización de evaluación formativa acorde con los estándares internacionales (utilidad, precisión, ética).

Caso 3

Las dos instituciones seleccionadas (3A y 3B) son públicas de nivel medio superior, asentadas en México. En el diagnóstico participaron seis agentes informantes (cinco docentes y un psicopedagogo).

Con respecto a la ética, en escuela 3A se observó ausencia de información documentada en un manual para docentes; en la 3B se pudo apreciar falta de información de cada estudiante (antecedentes, identificación de fortalezas y áreas de oportunidad).

En cuanto al concepto utilidad se consideró hacer cambios para que la evaluación sea de la siguiente manera: *(una evaluación) constructiva U1 y contener valores explícitos U5, cuidando que, en los procesos de planeación y conducción, los profesores identifiquen los valores utilizados para emitir sus juicios de desempeño* (MD_3). En la escuela 3ª, los sujetos entrevistados indicaron que se carece de formación docente para realizar evaluación de tipo sumativa y formativa, además, el alumnado desconoce los valores que sustenta la evaluación, las estrategias para



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

poder mejorar su desempeño, y cuál es, en todo caso, la utilidad de dicha evaluación. En la escuela 3B, los sujetos entrevistados declararon falta de creación de reportes de resultados y de acciones de seguimiento, así como ausencia en la percepción del vínculo de evaluación para la construcción del aprendizaje.

En factibilidad, en ambas instituciones se observó la necesidad de *orientar la evaluación a la práctica a fin de arrojar información necesaria para convocar la participación de otros agentes de la comunidad escolar* (MD_3). Se detectó, asimismo, que hace falta considerar todos los elementos que inciden en la planeación de evaluación y en el proceso para mejorar evaluaciones futuras (precisión). Además, se notó ausencia de vinculación entre metas de aprendizaje y evaluación, lo mismo que la falta de declaración de niveles de desempeño (escuela 3B).

Por todo ello, el estudiantado eligió rediseñar reactivos existentes (evaluación sumativa) y brindar la capacitación del modelo de retroalimentación (evaluación formativa) de [Hattie y Timperley \(2007\)](#).

Para la evaluación sumativa se revisó un examen previamente diseñado por la academia de la escuela 3A. Para este rediseño tomaron en cuenta la realidad que se vive en el plantel, las expectativas, las necesidades y los requerimientos normativos planteados por la institución, además de los requeridos para su ejecución práctica.

Una vez obtenidos los resultados de la evaluación sumativa, quienes participaron llevaron a cabo un análisis de las características del grupo y de las características psicométricas de la prueba, tomaron en cuenta la factibilidad al considerar que la evaluación debe estar orientada hacia la práctica y con el fin de obtener información necesaria para determinar el grado de aprendizaje. La utilidad se integró al considerar que la evaluación debe ser constructiva y también debe proveer valores explícitos, para lo cual se toman en cuenta procesos como planeación, conducción e identificación de valores en aras de emitir juicios de desempeño.

Para la evaluación formativa, se consideró que *la retroalimentación integra aspectos emotivos, procedimentales y cognitivos del alumno* (PF_3); las personas participantes entrevistaron a dos docentes y analizaron diversas tareas para determinar la presencia de elementos de retroalimentación requeridos, de acuerdo con el modelo citado.

Se ratificó que el personal docente (al retroalimentar) no brindaba datos precisos en función de la tarea, el alcance de objetivos, ni para comprender el progreso en función del perfil de egreso (PF_3). Tras una capacitación del modelo de retroalimentación (de 60 minutos), las personas agentes informantes realizaron una práctica que fue analizada por las personas participantes. En ésta se observó *feed up* y el *feed forward* ([Hattie y Timperley, 2007](#)). Sin embargo, se pudo apreciar que la retroalimentación aún carece de los elementos propios del modelo (precisión).

Las personas participantes describieron, a manera de conclusiones, la utilidad de haber integrado el uso de la taxonomía de [Marzano y Kendall \(2008\)](#) para el rediseño de reactivos, y cómo fue que al realizar la propuesta de evaluación sumativa, siguiendo la planeación, se tomó conciencia de su utilidad y su precisión. También concluyeron que la edad de sus estudiantes (15 y 16 años) influye *para sentirse y hacerse responsable de su propio proceso de aprendizaje* (PF_3) (utilidad) y aprovechar la retroalimentación, pero no dejaron evidencia de buscar la manera de incluirlos, en un futuro, en el proceso evaluativo (factibilidad).

Asimismo, las personas informantes expresaron que se elevó el nivel de conciencia (ética) de lo que debe ser un plan de evaluación (V_3); consideraron de gran valor el haber podido realizar un trabajo conjunto con otras personas docentes, tomar en cuenta su opinión y escucharles (utilidad) (V_3).

Caso 4

El estudio se realizó en dos escuelas primarias públicas suburbanas del Estado de México (4A y 4B). El diagnóstico se trabajó con seis agentes informantes (cinco docentes y un directivo), quienes participaron en las entrevistas a profundidad sobre evaluación sumativa y formativa.

En ambas instituciones se determinó:

1. Falta de un manual de operaciones para evaluación (ética). La escuela 4B señaló que existía una ausencia de antecedentes de estudiantes y la falta de enfoque de la evaluación, lo que impedía estimar el logro respecto a los objetivos de aprendizaje; se señaló que suelen *maquillarse* (MD_4) resultados en aras de cumplir con ciertos requerimientos.
2. Ausencia de capacitación docente para llevar a cabo procesos de evaluación formativa y sumativa, lo cual se puede apreciar en el siguiente testimonio: *se desconoce cómo realizar un análisis posterior de resultados... y hay deficiencia en retroalimentación* (MD_4) (utilidad). Asimismo, no existen reportes de evaluación ni para estudiantes ni para autoridades locales.
3. *No existe la participación de todos los actores educativos* (MD_4). Quien diseña los reactivos es la supervisión escolar; el personal docente es un mero administrador de exámenes (factibilidad).
4. No hay congruencia entre los objetivos declarados para el aprendizaje y las actividades con las que se evalúan; el personal docente no se siente en control del proceso de evaluación y no existe evidencia de los procesos de mejora de la evaluación (precisión).



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

Las personas participantes decidieron elaborar un manual de procedimientos para ambos tipos de evaluación, con el propósito de disminuir falencias encontradas e integrar lo indicado por los estándares internacionales.

El manual diseñado para la evaluación sumativa describe *recomendaciones acerca de la elaboración de un banco de reactivos, de la aplicación y la forma de interpretar los exámenes, de la manera de dar a conocer los resultados a las autoridades educativas a los padres de familia y alumnos* (PF_4); no se alude a la taxonomía de [Marzano y Kendall \(2008\)](#); brinda generalidades sobre diseño de reactivos, sobre qué considerar y qué no al momento de administrar exámenes y cómo divulgar resultados; no brinda ninguna explicación sobre trabajo colegiado (factibilidad), tampoco integra qué información se puede extraer del análisis de datos o cómo realizar la mejora del examen analizado; no contempla la consideración acerca de informes sobre estudiantes, aunque brinda aspectos de cómo divulgar los resultados para tres de los grupos usuarios (utilidad), solo indica la difusión pero no qué se puede hacer con dicha información para la mejora continua (precisión). El reporte carece de información que promueva la ética en el proceso, que se vincule cuando sus participantes documenten aspectos institucionales que coartan al personal docente; no se analizan resultados a favor del aprendizaje del estudiantado y su empoderamiento, y se pide coaccionar a través de la evaluación *evaluación igual a herramienta de control* (V_4).

El manual de procedimientos para la evaluación formativa fue elaborado a partir de los hallazgos, e integró *cómo podrían usar la información para el mejoramiento académico y los aspectos éticos que hay que cuidar antes, durante y después del uso de este tipo de evaluación* (PF_4), y consideró también integrar información sobre cómo diseñar un instrumento para evaluar (rúbrica, por ejemplo) y cómo llevar a cabo un proceso de retroalimentación y divulgación de resultados; no se integran aspectos en lo referente a la misión escolar (ética); se omitió la capacitación docente y la emisión de reportes de seguimiento, claramente detectadas como necesidades en su diagnóstico (utilidad) y en cuanto a factibilidad, no se mostró la integración del resto de la comunidad escolar, toda vez que se hace alusión al personal docente como único actor de la evaluación formativa.

En síntesis, ambos manuales incluyeron ciertos aspectos indicados desde la teoría, pero no los aspectos débiles detectados en el diagnóstico. La reflexión final mostró que quienes participaron identificaron la importancia de cuidar los procesos de evaluación de acuerdo con los estándares, aunque en la realidad reconocieron presiones administrativas que los desenfocan: *tiempo, cantidad de evaluaciones por hacer, presión de los padres de familia* (V_4).

Análisis de los cuatro casos

En cada caso, el estudiantado seleccionó las opciones de mejora de acuerdo con los resultados del diagnóstico realizado. En todos ellos, se ratificó que la decisión tomada respondía a necesidades que fueron detectadas en el diagnóstico.

Con respecto a la evaluación sumativa, en tres de los cuatro casos se optó por trabajar con el diseño y el rediseño de reactivos de evaluación. Estos casos tienen en común que corresponden a los niveles educativos medio superior y superior. El caso en el que se optó por la elaboración de un manual de procedimientos que permitiera definir procesos acordes a los estándares (precisión, ética, utilidad), corresponde a educación básica; no obstante, por las condiciones administrativas institucionales se percibió como algo fuera de control docente (factibilidad).

El procesamiento estadístico realizado con los datos de la aplicación de los reactivos (casos 1, 2 y 3) aportó un marco de referencia para el análisis de los resultados de la evaluación del aprendizaje (utilidad y precisión). El análisis de los resultados psicométricos abrió las puertas para la comprensión de los procesos inmersos en la construcción de exámenes masivos y la mejora continua, aunque afirmaron también la dificultad de permearlo al resto de sus pares; solo en un caso (2) se reporta que se integrará a la base de reactivos de evaluación existente.

Por otra parte, en todos los casos se informa que se reconoció la necesidad de diseñar los procesos de evaluación desde las metas de aprendizaje esperadas, con el fin de determinar el desempeño real del estudiantado en su quehacer cotidiano. Las personas participantes tomaron conciencia de la vinculación entre la generación de reactivos y la taxonomía, al reflexionar que se requiere de un proceso de planeación y de diseño integral.

Respecto a la evaluación sumativa, se hallaron expresiones tales como: que los docentes y las docentes la convierten en útil, precisa y ética pero difícilmente factible por el tiempo y la dificultad de trabajar en academia o por estar sujetos a procedimientos administrativos regulatorios; no obstante, no existe proyección o claridad sobre cómo establecer su uso posterior, ya que sólo en un caso se afirma que los reactivos podrán ser considerados para el banco institucional, aunque se desconoce cómo convertirlos en una constante de dichos procesos institucionales.

El estudiantado fue capaz de actuar en una situación real utilizando lo aprendido y propusieron soluciones tras una evaluación diagnóstica; esto puede implicar que la competencia proyectada en el curso quede en el terreno de un ejercicio práctico y contextualizado como evidencia de aprendizaje, y no, necesariamente, que sea transferible a su ejercicio docente cotidiano.

Con respecto a la evaluación formativa, el diagnóstico confirmó que los procesos y procedimientos de retroalimentación son desconocidos o poco comprendidos en las instituciones educativas; se dijo, asimismo, que aunque se reporte interés en el aprendizaje del estudiantado y su empoderamiento, la retroalimentación como proceso de evaluación no ocurre, por lo que en tres casos se planteó elaborar manuales de procedimientos.

En los manuales de procedimientos finales no existe una constante. En un caso (1) aún se confunde la evaluación sumativa con la formativa con un grupo de formación matemática, quizá fue la tendencia al cálculo numérico y al enfoque en los resultados, las calificaciones



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

y las ponderaciones más que en el proceso lo que impidió la apropiación del sentido de la evaluación formativa; mientras que, como manual, la propuesta de docentes de educación básica no aportaría para cumplir con los estándares de evaluación, dadas las imprecisiones de su proyecto final. Solamente en un caso se reporta con claridad el papel de la evaluación formativa en el proceso y cómo puede ser incorporado de forma institucional a fin de abonar a los estándares de utilidad, precisión y ética.

Por otra parte, en el caso en que se eligió realizar un proceso de capacitación en el modelo de retroalimentación con colegas, no se alcanzó el objetivo esperado en su proceso práctico. Asimismo, se aludió a que para su uso futuro se requiere tiempo y que no se puede llevar a cabo por el gran número de estudiantes que se atiende (factibilidad), por lo tanto, se optaría por utilizarse únicamente con alumnado de desempeño bajo o medio.

A pesar de que la necesidad de utilidad, factibilidad, precisión y ética es clara, los manuales concebidos por las personas participantes no dan cuenta de poder auxiliar o conducir al público usuario respecto a cómo llevar a cabo el proceso, solo fueron documentos informativos.

Finalmente, el proceso formativo al que fue expuesto el estudiantado le produjo el compromiso de llevar a cabo el proyecto práctico en su totalidad. Se generaron vínculos con las personas agentes informantes, dado que, de forma consistente (más del 75%), estos fueron quienes participaron en las entrevistas a profundidad para la elaboración de los manuales o participaron con testimonios al cierre del proceso.

Conclusiones

Al comparar los resultados de estos casos se observan –independientemente del contexto socioeconómico, nivel educativo y el país en donde se lleva a cabo el proceso de evaluación del aprendizaje– tres constantes:

- Primera. La carencia de un material de referencia (manual de procedimientos) en las instituciones educativas que les indique qué y cómo hacerlo, así como la importancia de la planeación del proceso, a lo que se suman ecos asociados a falta de capacitación y recursos.
- Segunda. La falta de participación de todos los actores en el proceso de evaluación. Con respecto a la evaluación sumativa, sólo en uno de los casos analizados, el docente no tiene un rol protagónico en la generación de los instrumentos de evaluación sumativa, sino que depende de la autoridad escolar oficial. En el resto, cada docente tiene el control principal de los procesos de evaluación, es quien determina qué y cómo se evalúa, de manera tal que los demás actores del proceso no participan.

- Tercero. No se plantean procesos de mejora a la luz del análisis de los resultados obtenidos. Las personas participantes no proyectan a futuro planes de mejora para la evaluación del aprendizaje.

Tras el análisis de lo observado en los cuatro casos y de las decisiones tomadas por las personas participantes en estos escenarios educativos sobre prácticas de evaluación del aprendizaje, se concluye:

- La realización de un diagnóstico con base en los estándares internacionales de evaluación del aprendizaje (JCSEE) es una manera confiable y práctica para identificar fortalezas y debilidades de las tareas de evaluación en instituciones de diferentes niveles y sectores (público o privado).
- El diagnóstico brinda la posibilidad de deducir el estado actual de la situación institucional con facilidad, al identificar lo que es necesario realizar para acotar la planeación a dichos estándares, por tanto, provee una ruta para los planes de mejora.
- El hecho de proporcionar opciones controladas para las prácticas de evaluación puede subsanar las debilidades encontradas en el diagnóstico. Es una forma de desarrollar en las personas participantes (docentes en proceso de formación y actualización) la comprensión sobre lo que implica el proceso general de diseño de la evaluación del aprendizaje a la luz de los estándares internacionales: utilidad, factibilidad, precisión y ética.
- El personal docente que debe llevar a la práctica las mejoras esperadas argumenta su poca posibilidad de tomar sus propias decisiones sobre las mejoras en la evaluación del aprendizaje, debido a falta de tiempo (por su carga horaria de trabajo) y disposiciones administrativas, ausencia de manuales y poca capacitación como las causas que impiden un trabajo colegiado; es decir, son circunstancias externas sobre las que no tienen control.
- Las personas participantes y agentes informantes expresaron la necesidad de continuar su formación sobre las prácticas de evaluación para desarrollar las competencias necesarias a fin de lograr un mejor desempeño. Tal fue el caso del uso del modelo de retroalimentación de [Hattie y Timperley \(2007\)](#). Las personas participantes indicaron, como ventaja, que a través de este modelo se puede brindar información a los grupos de estudiantes sobre lo realizado adecuadamente para lograr la meta de aprendizaje y la conexión del trabajo con futuros aprendizajes. Cabe señalar que el modelo de retroalimentación era desconocido, por quienes participaron, hasta el proceso de práctica.

A manera de recomendación, se sugiere que los procesos de evaluación del aprendizaje deben estar comprendidos y relacionados con los beneficios que aportan al proceso enseñanza aprendizaje, ya que la forma en que lo realizan actualmente no favorece procesos estándares de



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

ética, utilidad y precisión. Se necesita generar más espacios de divulgación de los modelos para desmitificar la idea de que la actividad consume mucho tiempo y lograr, así, que se perciban sus beneficios como parte del proceso de aprendizaje.

Como limitantes de esta investigación se señala que esta se ha restringido, únicamente, a cuatro casos de los 58 generados en el curso, además de que los datos del proceso están basados en un curso específico que se imparte en línea, con el enfoque de competencias. Otra limitante consiste en que es factible que la información brindada por las personas agentes informantes, particularmente en los videos finales, presenten sesgo, dado que reportaron algunos beneficios asociados a lo que se supondría que los entes responsables de las intervenciones y mejoras esperarían escuchar.

Se sugiere, como futuras investigaciones, proponer ejercicios más profundos de prácticas específicas que permitan consolidar el sentido de la evaluación formativa y evaluar los resultados de esta mejora; además, continuar desarrollando investigación asociada a los estándares de evaluación del aprendizaje, dada la escasa evidencia empírica con la que se cuenta en Latinoamérica.

Declaración de Material complementario

Este artículo tiene disponible como material complementario:

- La versión preprint del artículo en <https://doi.org/10.5281/zenodo.3353707>

Referencias

- Alarcón, J. y Donoso-Díaz, S. (2012). Nuevos desafíos de las universidades y demandas a la formación por competencias; ¿qué camino tomar? *Revista Electrónica de Desarrollo de Competencias*, 5(1), 225-236.
- Alarcón Leiva, J., Hill, B. y Frites, C. (2014). Educación basada en competencias: Hacia una pedagogía sin dicotomías. *Educação & Sociedade*, 35(127), 569-586. <https://doi.org/10.1590/S0101-73302014000200013>
- Álvarez, I. (2005). Evaluación como situación de aprendizaje o evaluación auténtica. *Perspectiva Educacional, Formación de Profesores*, 45, 45-68. <http://www.redalyc.org/articulo.oa?id=333329100004>
- Baxter, P. y Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, 13(4), 544-559. <https://nsuworks.nova.edu/tqr/vol13/iss4/2/>
- Cabra Torres, F. (2008). La calidad de la evaluación de los estudiantes: Un análisis desde



estándares profesionales. *Magis: Revista Internacional de Investigación en Educación*, 1(1), 95-112. <https://revistas.javeriana.edu.co/index.php/MAGIS/issue/view/349>

Cano García, M. E. (2008). La evaluación por competencias en la educación superior. Profesorado. *Revista de Currículum y Formación de Profesorado*, 12(3), 1-16. <http://www.ugr.es/~recfpro/rev123COL1.pdf>

Castillo Ortega, S. E. y Canul de Dios, P. Y. (2015). Docente mediador en la formación de competencias en educación básica. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 2(4), 1-14. <http://www.pag.org.mx/index.php/PAG/article/view/334/376>

Escudero, T. (2013). Utilidad y uso de las evaluaciones. Un asunto relevante. *Revista de evaluación educativa*, 2(1), 184-205. <https://studylib.es/doc/8151005/utilidad-y-uso-de-las-evaluaciones.-un-asunto-relevante>

Fernández March, A. F. (2010). La evaluación orientada al aprendizaje en un modelo de formación por competencias en la educación universitaria. *REDU: Revista de Docencia Universitaria*, 8(1), 11-34. <https://doi.org/10.4995/redu.2010.6216>

Gallardo Córdova, K. E. (2013). *Evaluación del aprendizaje: Retos y mejores prácticas*. Editorial Digital del Tecnológico de Monterrey.

Gallardo Córdova, K. E. (2017, agosto-diciembre). *Evaluación del aprendizaje* (ED5048). [Planes de Estudio]. Tecnológico de Monterrey, Escuela de Humanidades y Educación. <https://samps.itesm.mx/Materias/VistaPreliminarMateria?clave=ED5048&lang=ES>

Gallardo Córdova, K. E. y Gil Rendón, M. E. (2016). Evaluación de desempeño en estudiantes de educación superior: Uso de la herramienta competere. *Revista de Pedagogía*, 37(100), 187-205. http://saber.ucv.ve/ojs/index.php/rev_ped/article/view/11974

Gallardo Córdova, K. E., Ramos Monobe, A., Camacho Gutiérrez, D. y Gil Rendón, M. E. (2015). Evaluación de los estudiantes: Las prácticas institucionales y docentes vistas desde el marco de estándares internacionales. *Apertura*. 7(2), 96-109. <http://www.udgvirtual.udg.mx/apertura/index.php/apertura/article/view/657>

George, B. C., Teitelbaum, E. N., Meyerson, S. L., Schuller, M. C., DaRosa, D. A., Petrusa, E. R., Petito, L. C. y Fryer, J. P. (2014). Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. *Journal of Surgical Education*, 71(6), e90-e96. <https://doi.org/10.1016/j.jsurg.2014.06.018>

Hancock, D. R. (2007). Effects of performance assessment on the achievement and motivation of graduate students. *Active Learning in Higher Education*, 8(3), 219-231. <https://doi.org/10.1177/1469787407081888>

Hattie, J. y Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1),



<http://doi.org/10.15359/ree.24-3.12>
<http://www.una.ac.cr/educare>
educare@una.ac.cr

81-112. <https://doi.org/10.3102/003465430298487>

Joint Committee on Standards for Educational Evaluation (2014). *Estándares para la evaluación de estudiantes*. <http://shorturl.at/vwJ67>

Marzano, R. J. y Kendall, J. S. (2008). *Designing & assessing educational objectives: Applying the new taxonomy*. Corwin Press.

Morán, R., Cardoso, E. O., Cerecedo, M. T. y Ortiz, J. C. (2015). Evaluación de las competencias docentes de profesores formados en instituciones de educación superior: El caso de la asignatura de tecnología en la enseñanza secundaria. *Formación Universitaria*, 8(3), 57-64. <https://doi.org/10.4067/S0718-50062015000300007>

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189. <https://doi.org/10.3102/0034654307313795>

Tejada Fernández, J. (2013). Professionalisation of teaching in universities: implications from a training perspective. *RUSC. Revista de Universidad y Sociedad del Conocimiento*, 10(1), 345-358. <http://dx.doi.org/10.7238/rusc.v10i1.1471>

Tobón, S. (2008). *La formación basada en competencias en la educación superior: El enfoque complejo* [Documento de Curso]. México, Universidad Autónoma de Guadalajara. https://www.researchgate.net/publication/329440312_La_formacion_basada_en_competencias_en_la_educacion_superior_el_enfoque_complejo

Valenzuela, J. R. (2005). El profesor como evaluador. En A. Lozano Rodríguez (Coord.), *El éxito en la enseñanza. Aspectos didácticos de las facetas del profesor* (pp. 227-256). Trillas.

Wiggins, G. (1998). *Educative assessment. Designing assessment to inform and improve student performance*. Jossey-Bass Publisher.

Yin, R. K. (2003). *Case study research: Design and methods* (2ª ed.). Sage.

Apéndice A: Guía de preguntas para entrevistas semiestructuradas

El proceso formativo correspondiente a las prácticas en el curso de Evaluación del aprendizaje, preveía la toma de decisiones a partir de la aplicación del cuestionario diagnóstico. Las entrevistas se realizaron en dos momentos del proyecto: (1) durante la evaluación formativa y (2) al cierre del proyecto integrador, según se muestra.

Etapa del proyecto	Diagnóstico institucional	Selección de escenarios y agentes informantes
		Construcción y validación de un banco de reactivos.
	Evaluación sumativa	Rediseño y validación de un banco de reactivos. Manual de procedimientos de evaluación sumativa.
	Evaluación formativa	Aplicación del modelo de retroalimentación. Manual de procedimientos de evaluación formativa.
	Integración del proyecto	Testimonios de los agentes informantes beneficiarios.

Evaluación formativa opción 1: Análisis del proceso de evaluación formativa existente con los sujetos informantes. Entrevista semiestructurada individual a una persona docente.

1. ¿Qué significa para ti retroalimentar?
2. ¿Qué es lo más importante para ti y para tus estudiantes a la hora de realizar y recibir la retroalimentación?
3. ¿Con qué frecuencia retroalimentas a tus estudiantes de forma escrita?
4. ¿Has estudiado o trabajado con algún modelo de retroalimentación últimamente?

Evaluación formativa opción 2: Análisis del proceso general para elaborar un manual de procesos de evaluación del aprendizaje institucional. Entrevista semiestructurada grupal.

1. ¿Qué se busca a través de la evaluación formativa en tu institución?
2. ¿Qué se le solicita atender o cuidar en los procesos de evaluación formativa?
3. ¿Cómo se planea este tipo de evaluación en tu institución?
4. ¿Los procesos de evaluación formativa están alineados a la visión y modelo educativo de la institución?
5. ¿Cuáles son los mecanismos o ejercicios, desde tu punto de vista, los más utilizados?
6. ¿Toman en cuenta el grado de dificultad y la disciplina para planear la evaluación formativa?
7. ¿Cómo orienta a los maestros y maestras en el aprovechamiento de la información que emerge de la evaluación formativa hacia la sumativa?
8. ¿Cómo plantean procesos de mejora para la planeación, puesta en marcha y colecta de datos que emergen de los procesos de evaluación formativa?

Integración del proyecto: Cierre de proceso de prácticas de evaluación, con entrevista de personas beneficiarias del proyecto

1. ¿Cuáles han sido los principales beneficios de su intervención?
2. ¿Qué aspectos detectan fuertes en sus prácticas de evaluación del aprendizaje?
3. ¿Qué aspectos consideran débiles o con cierta amenaza que deban mejorar en el corto plazo?
4. ¿Qué se realizará y cómo mejorará la comunicación de los resultados a los diferentes públicos de ahora en adelante?

