

Evaluación de programas de motivación y educación emocional en Educación Secundaria: mirada analítica a programas internacionales

Evaluation of motivation and emotional education programs in secondary education: An analytical look at international programs

Gisela Monserrate Lagos-Vélez, gislagos130@gmail.com,
<https://orcid.org/0000-0002-6406-91414>;

Vanessa Liliana Paredes-Tacuri, <https://orcid.org/0000-0003-2063-2374>;
Frank Stalin Poma-Sánchez, <https://orcid.org/0000-0001-8324-557X>

Universidad Técnica Estatal de Quevedo, Ecuador

Resumen

El presente trabajo tiene como objetivo analizar los procesos y resultados de evaluación de diferentes programas de educación emocional y de motivación dirigidos a estudiantes de Secundaria y/o adolescentes a partir de los resultados de sus evaluaciones. Para ello se realiza una aproximación teórica a la evaluación y metaevaluación de programas y se lleva a cabo una revisión de la literatura académica de la que se seleccionan y compara evaluaciones de programas de educación emocional y motivación. Finalmente, se realiza una metaevaluación mediante la aplicación de estándares. Los resultados muestran que las evaluaciones se realizan, generalmente, de forma interna y sumativa, centrándose en los resultados de aplicación y no tanto en el proceso.

Palabras clave: metaevaluación; intervención; motivación; emoción; enseñanza secundaria.

Abstract

The present work aims to analyze the evaluation processes and results of different motivation and emotional education programs addressed to secondary students and/or adolescents from the results of their evaluation results. In order to achieve this goal, a theoretical approach to program evaluation and meta-evaluation is made, as well as a review of the academic literature from which evaluations of emotional education and motivation programs are selected and compared. Finally, a meta-evaluation is performed through the application of standards. The results show that evaluations are generally conducted internally and in a summative manner, focusing on the results of implementation rather than on the process.

Keywords: meta-evaluation; intervention; learning motivation; emotion; secondary education.

Introducción

Las emociones y la motivación tienen papeles esenciales a la hora de determinar el rendimiento académico de los estudiantes. Las emociones implican cambios en los sentimientos, cogniciones, tendencias de acción, comportamientos y aspectos fisiológicos a raíz de un efecto precedente (Lange et al., 2020) y poseen una influencia clave en el rendimiento ya que los procesos educativos siempre están acompañados de estados emocionales (Pulido y Herrera, 2017) que regulan las acciones de los estudiantes. Por su parte, la motivación es el impulso para actuar que sucede a raíz de la experimentación de dichas emociones (Anaya, 2014; Mirete et al., 2015) y ha centrado el interés de la comunidad científica no solo desde hace décadas (Järvela, 2001; Lopez, 2006; Rinaudo et al., 2003) sino, también, en los últimos tiempos (Barati, Joo y Aria, 2018; De Caso, 2014). Este constructo influye tanto en los procesos cognitivos y en las estrategias de aprendizaje de los estudiantes como en sus actitudes futuras hacia el aprendizaje (García, 2008).

Estos aspectos, si bien deben ser considerados en todas las etapas educativas, cobran especial relevancia en Educación Secundaria puesto que, en ella, los menores se encuentran experimentando cambios emocionales y motivacionales propios de la adolescencia (Burunat, 2004) la cual abarca desde los 10 años hasta los 19 (OMS, 2020) o según Sawyer et al (2016), hasta los 24. Por tanto, resulta clave promover medidas para su mejora.

Una de las herramientas que permiten mejorar estos aspectos a través de acciones coherentes y organizadas son los programas de intervención, definidos como unidades de acción que contribuyen a alcanzar de manera sistemática objetivos previamente planteados (Pérez, 2000, 2006) en función de las necesidades detectadas (Martínez, 2017). Para alcanzar dichos objetivos, es imprescindible que los programas estén diseñados e implementados de manera correcta, garantizando que con su aplicación se logren los resultados esperados. Por ello, uno de los elementos esenciales que debe considerarse a la hora de seleccionar un programa de intervención es si cuenta con un proceso de evaluación riguroso que incluya a la recogida de información válida y fiable y que permita determinar la calidad y los logros para tomar decisiones de mejora (Pérez, 2000) y la magnitud de los cambios producidos (Fryey, 2012).

La necesidad de conocer el impacto de los programas, así como de rendir cuentas a la sociedad, ha contribuido al aumento de la importancia de la práctica de la evaluación

(Méndez, 2018). Pese a ello, tal y como fue detectado por Hernández y Álvarez Rojo (1998) y comprobado posteriormente por Expósito et al. (2004), el desequilibrio entre la cantidad existente de programas de intervención evaluados y no evaluados es manifiesto. En esta línea, Pérez (2008) explica que la evaluación de programas en el ámbito educativo es limitada y contrasta con otros ámbitos como el de la agricultura o medicina. Para ello, pone como ejemplo la revisión llevada a cabo por Vallés y Vallés (2003), en la que únicamente uno de los catorce programas de educación emocional contemplados presentó estudios de validación. Además, Sánchez et al. (2016) realizan una revisión de programas de educación emocional en la que observan que estos no siempre cuentan con evaluaciones y, en caso de hacerlo, no poseen el nivel de profundidad adecuado.

La evaluación es una disciplina compleja. Según Barbier, el antecedente del tipo de evaluación que realizamos en este momento surgió en el siglo XIX, ante las necesidades de certificar, por una parte, los conocimientos que ofrecía el sistema educativo, y, por otra parte, y esto es importante, para establecer jerarquías que el mundo laboral pudiera leer. No es lo mismo tener primaria que tener secundaria, bachillerato o licenciatura, es decir, de esta manera se enviaba un mensaje al mundo laboral. En la evaluación convergen varias disciplinas, como la evaluación psicológica, la que proviene de las ciencias educativas y la visión de evaluación como disciplina, que proviene del pensamiento gerencial administrativo (Beltrán, 2014).

La metaevaluación, es decir, la necesidad de someter a la evaluación al mismo rigor que esta actividad supone; esto es, independientemente de cómo decidamos evaluar, en algún momento tendremos que, a revisar no solamente el diseño de la evaluación, sino su aplicación y las consecuencias que esta acción ha tenido en todos los participantes. (Rueda, 2014)

La motivación se refiere a la elección fines y medios, y, la satisfacción al nivel de agrado o bienestar experimentado en una situación o contexto. Las teorías actuales indican que tanto la motivación como la satisfacción dependen en gran parte de los valores y creencias a partir de lo que las personas establecen como estándares, interpretan motivos y evalúan las situaciones. Por ejemplo, la carga de trabajo, por si misma, no es un factor determinante en la motivación. (Pcheu, 2018)

Investigaciones previas muestran que las emociones están ligadas con el bienestar personal está conectado con un aprendizaje efectivo, existe poca investigación sobre cómo las emociones positivas se relacionan con las metas de logro y las medidas objetivas

de participación escolar. Sobre la base de la teoría de la expansión y construcción de emociones positivas, este estudio examina las asociaciones recíprocas entre las emociones positivas sobre metas de dominio y la participación académica estimada por el profesor (Datu, 2021)

La educación secundaria (también se denomina educación media, segunda enseñanza, enseñanza secundaria, enseñanza media, bachillerato o estudios medios) es la que tiene como objetivo capacitar al alumno para proseguir estudios superiores o bien para incorporarse al mundo laboral, además forma un espacio de enseñanza y aprendizaje con unas prácticas que se nutren de dos modelos epistemológicos y metodológicos tradicionales (Carrasco, 2021).

El presente trabajo tiene como objetivo analizar los procesos y resultados de evaluación de diferentes programas de educación emocional y de mejora de la motivación dirigidos a estudiantes de Secundaria y/o adolescentes.

Materiales y métodos

La muestra objeto de estudio se compone de un total de 19 documentos que describen la evaluación de programas de intervención en motivación y en educación emocional. La selección de la muestra partió de una búsqueda inicial en las bases de datos ERIC y Google Académico. Como criterio de inclusión, se seleccionaron aquellos documentos publicados con posterioridad al año 2010 en formato artículo, informe, tesis doctoral o actas de congresos, excluyendo los documentos en formato libro o manual. Las palabras claves utilizadas fueron: evaluación, programa, secundaria, motivación y emoción.

Tras el proceso de búsqueda inicial, se consideraron 374 resultados (fase 1), de los cuales se seleccionaron únicamente aquellos que cumplían los criterios relacionados tanto con las características de los programas como con su evaluación (Fase 2):

- Destinatarios: alumnado de Educación Secundaria y/o adolescentes (10-19 años).
- Contexto: instituciones y centros educativos de educación formal y no formal.
- Elementos considerados en la evaluación: aspectos emocionales y motivacionales.

Una vez aplicados estos criterios, el número de publicaciones se redujo a 153 documentos. Tras ello, se eliminaron aquellos documentos que abordaban intervenciones sobre poblaciones específicas (Alumnado con discapacidad, estudiantes con altas

capacidades...), excluyendo 56 artículos y configurando un total de 97 documentos (fase 3).

Finalmente, se realizó un último proceso en el que se seleccionaron aquellos artículos en los que se indicaba de forma identificable el nombre del programa evaluado. Además, se descartaron aquellos que mostraban evaluaciones de programas aplicados antes del año 2000 (fase 4). Estos criterios eliminaron una cifra elevada de artículos (78), estableciendo una muestra final de 19 documentos (figura 1).

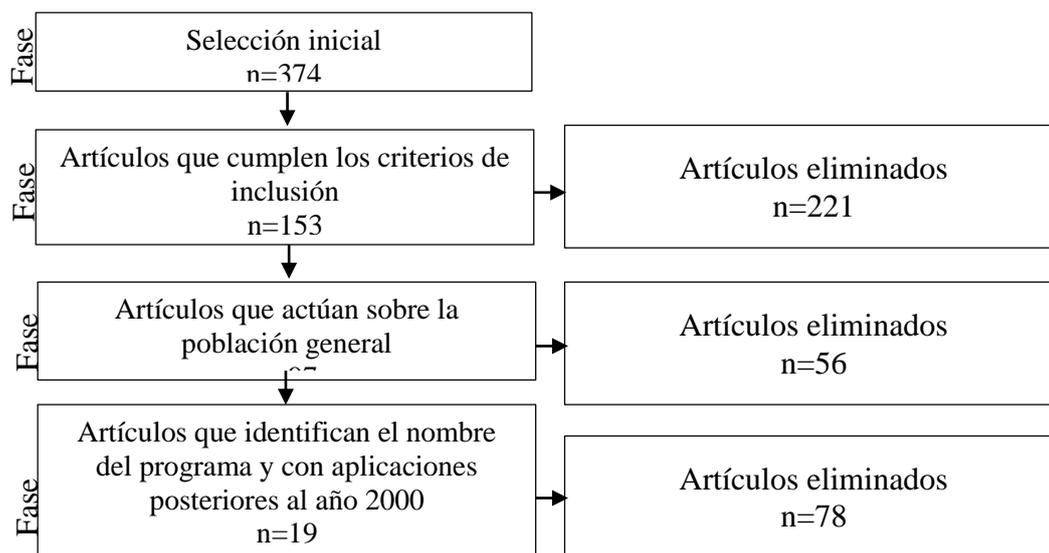


Figura 1. Proceso de identificación y selección

Una vez identificados y seleccionados los documentos, se procedió a analizarlos atendiendo a los siguientes criterios: año y lugar de realización de la evaluación, diseño, muestra, instrumentos, características del evaluador, función de la evaluación, perspectiva temporal y elementos evaluados (ver tabla 2)

Tabla 2. Criterios de clasificación y categorías de respuesta asociadas

Criterio	Categorías
Año de publicación de la evaluación	2000-2020
País de evaluación	Respuesta abierta
Diseño y metodología	Cualitativa, cuantitativa, mixta
Muestra	Respuesta abierta
Instrumentos	Respuesta abierta
Evaluador	Interno, externo
Función	Formativa, sumativa
Perspectiva temporal	Ex post, ex ante
Elementos evaluados	Necesidades, implementación, resultados

En el apéndice se muestran los resultados del análisis de cada una de las dimensiones para las evaluaciones contempladas.

De forma complementaria, se llevó a cabo una metaevaluación aplicando los estándares de la metaevaluación checklist (Scriven, 2013). Se trata, en todo caso, de una metaevaluación externa desarrollada tras la finalización del proceso de aplicación del programa y del proceso de evaluación a partir de la información disponible en los documentos, por lo que se seleccionan, únicamente, los estándares aplicables al presente trabajo (A3, B1, B2, B3, C1, C2, D1, D3 y D4) y se omiten aquellos que deben ser aplicados durante el proceso de evaluación o para los que no se dispone de información suficiente a través de las publicaciones seleccionadas (A1, A2, B4, B5, C3, C4, C5, D2 y D5).

Resultados

Análisis de la evaluación de los programas

A continuación, se presentan los resultados del estudio atendiendo a los criterios inicialmente señalados. Como consecuencia del criterio de actualidad aplicado, los documentos aparecen principalmente a partir del año 2012, y, más concretamente, se incrementa el número de publicaciones a partir de 2015 (Tabla 3)

Tabla 3. Distribución de porcentajes del año de publicación de los artículos de evaluación de los programas

Año	2008	2011	2012	2013	2015	2016	2018	2020
%	5.3	5.3	10.5	5.3	31.5	15.8	15.8	15.8

En lo que respecta al país de aplicación y realización de las evaluaciones, el 47,3% de los documentos identificados pertenecen a España y el 26,3% a Estados Unidos, mientras que el 5,3% de los programas se han evaluado en Portugal, Reino Unido, Macedonia, Letonia y Dinamarca, respectivamente.

Atendiendo al diseño de evaluación de resultados, la mayoría de las evaluaciones plantean un diseño pre test-post test, comprándose las medias tomadas antes de la aplicación del programa con aquellas resultantes de dicha aplicación (73,7%). En el 57,9% se concreta el diseño de dos grupos (control y experimental no equivalente en el 47,7% de los casos; control y experimental equivalente en el 10,5%), mientras que el 26,3% no cuenta con grupo de control. En la mayor parte de las evaluaciones se ha seguido una metodología cuantitativa (47,5%), mientras que el 36,8% han empleado una metodología mixta y tan solo el 5,3% una metodología cualitativa.

Entre los instrumentos de evaluación se han utilizado pruebas estandarizadas (26,3%) y no estandarizadas (21%), destacando la combinación de ambos tipos de instrumentos (42,1%), si bien el 10,5% no se explicita el tipo de instrumento.

Con respecto al evaluador, el 52,6% de los programas cuenta con un evaluador interno (generalmente, el diseñador y/o aplicador del programa), en el 21% la evaluación es externa y en el 26,4% restante la evaluación se ha llevado a cabo por evaluadores externos e internos de manera conjunta. Con respecto a la función de la evaluación, el 78,9% de los programas tiene una función sumativa, evaluándose, por tanto, sus resultados. Únicamente en el 5,3% la evaluación es formativa, mientras que en el 15,8% es tanto formativa como sumativa.

En los relativo a la perspectiva temporal, esto es, el momento en el que se ha llevado a cabo la evaluación, en el 78,9% de los programas la evaluación es realizada ex post p tras la aplicación del programa; en el 5,3% de los casos se lleva a cabo ex ante o previamente a la aplicación, y solo en el 15,8% se realiza en ambos momentos.

Por último, atendiendo a los elementos evaluados y vinculado con la perspectiva temporal, en el 57,9% de los programas se busca conocer los resultados de la aplicación, mientras que solo en el 5,3% se ha evaluado el diseño y en otro 5,3%, la implementación, si bien cabe destacar que en el 26,3% se ha evaluado tanto la implementación como los resultados y en el 5,3%, los tres elementos. La Tabla 4 ilustra los resultados en función del tipo de evaluador, la función, la perspectiva temporal y los elementos evaluados.

Tabla 4. Resultados del tipo de evaluador, función de la evaluación, perspectiva temporal y elementos evaluados.

Elemento	Tipología	%
Tipo de evaluador	Interno	52,5
	Externo	21
	Interno y externo	26,4
Función de la evaluación	Formativa	5,3
	Sumativa	78,9
	Formativa y sumativa	15,8
Perspectiva temporal	Ex ante	5,3
	Ex post	78,9
	Ex ante y ex post	15,8
Elementos evaluados	Diseño	5,3
	Implementación	5,3
	Resultados	57,9
	Implementación y resultados	26,3
	Diseño, implementación y resultados	5,3

Metaevaluación

Dada la importancia que posee la metaevaluación a la hora de garantizar la calidad y la honradez de las evaluaciones, se ha procedido a la aplicación de los estándares de la

metaevaluación Checklist (Scriven, 2013) en cada una de las 19 evaluaciones de programas, en la tabla 5 se recogen los estándares aplicados y el porcentaje de programas que cumplen cada uno de ellos.

Tabla 5. Estándares aplicados y porcentajes de cumplimiento.

Estándar	Se cumple
A3. Diseño y métodos. Determinación de la manera en la que se van a encontrar las respuestas a las diferentes preguntas planteadas en la evaluación	94,7%
B1. Contexto y antecedentes. Inclusión de factores contextuales que pueden determinar el éxito o el fracaso del programa.	52,7%
B2. Descripciones y definiciones. Explicitación de los elementos que forman parte del programa y van a ser evaluados	94,7%
B3. Clientes. Implicados en la evaluación.	100%
C1. Procesos. Valor de los medios que emplea la intervención. Se vincula con la evaluación de aquello que ocurre antes de que los resultados finales se produzcan	42,1%
C2. Efectos. Evaluación del impacto (efectividad del programa)	94,7%
D1. Síntesis. Resumen de la importancia relativa de cada dimensión	100%
D3. Responsabilidad y justificación. En caso de que la responsabilidad puede ser determinada y sea oportuno hacerlo	100%
D4. Informe y apoyos. Esclarecimiento de las conclusiones de manera apropiada, para los grupos, momento y lugar apropiados.	100%

De la metaevaluación realizada es posible concluir que, su bien las evaluaciones cumplen de manera completa los estándares seleccionados de forma general, existen excepciones. Como puede observarse, solamente en uno de los programas no consta que se cumpla totalmente el estándar A3 (Diseño y métodos), al no detallarse ni el método ni el diseño de la evaluación conviene recordar que la metaevaluación se basa en la información proporcionada por la literatura, por lo que la determinación de que los aspectos de un estándar no se cumplen puede no adecuarse a la realidad de la evaluación y deberse, en cambio, a la falta de información disponible.

En lo que respecta al estándar B1 (Contexto y antecedentes), en el 52,7% de las evaluaciones se ha considerado la influencia de los factores contextuales específicos en los que se desarrollan tanto los programas como los procesos de evaluación. El estándar B2 (Descripciones y definiciones) se cumple en el 94,7% de los casos, describiendo explícitamente las variables consideradas en el proceso, mientras que el 100% de ellas consideran las perspectivas de los interesados (B3, Clientes).

En lo relativo al estándar C1 (Procesos), tan solo el 42,1% de las evaluaciones lo cumplen, ya que en el resto se contemplan exclusivamente el diseño del programa o los resultados. En línea con lo anterior, el estándar C2 (efectos), relacionado con evaluación de la efectividad del programa, se contempla en el 94,7% de las evaluaciones. Finalmente, la totalidad de los procesos de evaluación cumplen los estándares D1 (Síntesis), D2

(Responsabilidad y justificación) y D4 (Informe y apoyos), puesto que en todos ellos se reflejan los resultados de las dimensiones evaluadas, se indican los responsables de las evaluaciones y se establecen de manera clara las conclusiones del proceso.

Discusión

Evaluación de programas

La literatura académica muestra diferentes perspectivas desde las que abordan la evaluación de programas. Por su relevancia, cabe destacar aquellos modelos que tradicionalmente han servido de base a la teoría evaluativa, como el modelo de cuatro niveles de Kirkpatrick (1959), el modelo Contexto, Input, Proceso y Producto (CIPP) de stufflebeam (1971, 1987) o el modelo de Stake (1975), centrado en los antecedentes, el proceso, las normas, los juicios y los resultados. Según Pérez Juste (1995, 2000, 2006), todo programa educativo ha de evaluarse en tres momentos, que constituyen la evaluación inicial, procesual o continua, y final o sumativa. Para ello, Pérez Juste (2006) desarrolla una pauta de evaluación que integra 74 indicadores agrupados en cuatro categorías relacionadas con estos tres momentos y la institucionalización de la evaluación.

Por su parte, Álvarez Rojo et al. (2002) determinan los supuestos metodológicos de la evaluación de programas, clasificando las evaluaciones en función del agente, la función de la evaluación, la perspectiva temporal y los elementos del programa que se estén evaluando (Tabla 1)

Tabla 1 Tipos de evaluación de programas

Criterio	Tipos	Características
Evaluador	Evaluación interna	Realizada por el mismo sujeto que diseña y/o aplica el programa
	Evaluación externa	Realizada por sujetos distintos de quien diseña y/o aplica el programa
Función de la evaluación	Evaluación formativa	Dirigida al desarrollo y mejora del programa
	Evaluación sumativa	Dirigida al control y valoración de resultados
Perspectiva temporal	Evaluación ex ante	Realizada antes de la aplicación para considerar la pertinencia y probabilidad de éxito
	Evaluación ex post	Realizada cuando el programa ha concluido
Elementos del programa	Evaluación de necesidades o del diseño	Realizada en la etapa de planificación
	Evaluación de la implementación la cobertura y el seguimiento	Realizada la etapa de aplicación

Evaluación de los Realizada en la etapa de finalización
resultados y el impacto

Fuente: Álvarez Rojo et al. (2002)

Metaevaluación de programas

El concepto de metaevaluación, acuñado por Scriven (1967), hace referencia a una investigación que, de manera sistemática, busca realizar un juicio de valor sobre la calidad y los méritos de una evaluación, teniendo como objetivo asegurar la calidad de los servicios evaluativos, evitar las prácticas ilegales e indicar el camino para perfeccionar la profesión evaluadora, así como favorecer su comprensión (Stufflebeam, 1987).

Con el fin de que la metaevaluación de programas se lleve a cabo con éxito, se han desarrollado una serie de estándares de evaluación que sirven como marco de trabajo, así como de guía sobre aquellos aspectos que permiten determinar su calidad. A la hora de realizar una metaevaluación, es necesario considerar que no todos los estándares han de ser aplicados en cada uno de los procesos de evaluación. Por el contrario, los profesionales encargados de su realización deben seleccionar y priorizar aquellos que se ajusten a las características de su evaluación (Samson y Solabarrieta, 2018) o que sean aplicables en el contexto en que esta se desarrolla (Martínez Mediano, 2017).

Cabe resaltar, debido a su actualidad, los Estándares del Joint Committee on Standards for Educational Evaluation (Yarbrough et al, 2011), compuestos por 30 ítems clasificados en cinco grupos: utilidad, viabilidad, adecuación, exactitud y responsabilidad; y la metaevaluación checklist (Scriven, 2013), que divide los criterios de metaevaluación en cuatro grupos: preliminares, fundamentos, subevaluaciones y conclusiones e implicaciones. El grupo de criterios preliminares se relaciona con la necesidad de explicitar los pasos y la metodología que sigue a un programa, así como la justificación de la evaluación; se compone de los estándares A1 (Resumen ejecutivo), A2 (Clarificaciones) y A3 (Diseño y métodos). El grupo fundamentos está formado por los estándares B1 (Antecedentes y contexto), B2 (Descripciones y definiciones), B3 (Consumidores), B4 (Recursos) y B5 (Valor), que se refieran a la especificación de los elementos del programa, tales como su contexto y su naturaleza. El tercer grupo de subevaluaciones se vincula con el momento de observación de las evidencias. Se contempla la evaluación tanto durante el proceso de aplicación como en relación con el impacto y la posibilidad de establecer comparaciones o aplicar el programa a otros contextos. Está formado por los estándares C1 (Procesos), C2 (Resultados), C3 (Costes), C4 (Comparaciones) y C5 (Generalizaciones).

En el último grupo de criterios, conclusiones e implicaciones, se relaciona con las consideraciones previas a la elaboración del informe de resultados y contempla también la realización de metaevaluaciones durante el proceso; se compone de los estándares D1 (Síntesis), D2 (Recomendaciones, explicaciones, predicciones y rediseños), D3 (Responsabilidad y justificación), D4 (Informe y apoyo) y D5 (Metaevaluación).

La evaluación de programas es un campo de estudio extenso y riguroso que requiere de un análisis y aplicación de criterios de evaluación bien seleccionados y definidos. La escasa muestra de trabajos analizados pone de manifiesto la necesidad de realizar evaluaciones rigurosas que contribuyan a constatar la calidad y eficacia de los programas.

El objetivo de este trabajo ha sido analizar los procesos y resultados de evaluación de diferentes programas de educación emocional y motivación dirigidos a estudiantes de Secundaria y/o adolescentes. Entre las evaluaciones identificadas se ha encontrado una gran diversidad de procedimientos de evaluación, así como variadas formas de abordar los resultados ya que, si bien existe algunos aspectos comunes a la mayoría de las evaluaciones, las diferencias en el modo de realizar estos procesos son notables, existiendo evaluaciones de tipo cualitativo como cuantitativo o mixto que emplean, por tanto, instrumentos de recogida de información muy variados. En ese sentido, aunque la disparidad de las características de las evaluaciones de los programas dificulta realizar generalizaciones, la mayoría de las evaluaciones son sumas activas y se centran en la determinación de resultados del programa, prestando menor atención al diseño y proceso de aplicación.

Este aspecto resta calidad a la aplicación de los programas ya que a través de la evaluación inicial y procesual se garantiza la calidad de su diseño, así como su ajuste a las necesidades iniciales y reajuste a aquellas surgidas durante su aplicación. Además, un elevado número de evaluaciones se han desarrollado de manera interna, de modo que las personas encargadas de su diseño eran, también, los propios evaluadores. Este hecho presenta la ventaja de que los evaluadores conocen profundamente las características del programa y sus destinatarios, pero cabría cuestionarse si una evaluación externa o mixta aumentaría la objetividad del proceso.

Además, es preciso hacer notar que la difusión de los resultados en medios científicos no se realiza hasta el año 2012, y más concretamente a partir del año 2015, mostrando esta escasa tradición de la literatura académica (Pérez, 2008).

Estos resultados están en línea con los hallados por Expósito et al. (2004), quienes analizan los patrones metodológicos de evaluación de 208 estudios publicados entre 1975 y 2000; la metodología empleada fue principalmente cuantitativa, utilizando un modelo evaluativo en la búsqueda de resultados de aplicación. Sin embargo, a diferencia del presente estudio, se emplearon principalmente instrumentos no estandarizados.

Otro de los elementos a considerar es la capacidad de generalización de los resultados de los programas, puesto que la aplicación a muestras reducidas o poco representativas no garantiza en absoluto su aplicabilidad; el tamaño reducido de las muestras y el uso generalizado de autoinformes puede no estar mostrando niveles reales de competencia (Torrijos et al., 2018), de modo que es preciso replicar los programas en otras poblaciones para poder asegurar sus resultados de evaluación. En este sentido, las evaluaciones identificadas presentan muestras de diversos tamaños que oscilan entre los 23 y los más de mil sujetos, teniendo estas últimas mayores garantías de calidad.

La metaevaluación permite extraer interesantes conclusiones. Así, en primer lugar, se ha podido determinar que la gran mayoría de evaluaciones cumplen el estándar A3, relativo al Diseño y métodos de la evaluación, y el B2 (Descripciones y definiciones) de la metaevaluación checklist (Scriven, 2013), determinado los aspectos metodológicos y la descripción de aquello que se evalúa de manera correcta y detallada. Además, todas las evaluaciones cumplen los estándares seleccionados del grupo conclusiones e implicaciones, contemplando el resultado de cada dimensión, las conclusiones de la evaluación los responsables de realizar la, así como el estándar B3 (Clientes) del grupo fundamentos. Se constata, asimismo que la mayoría de evaluaciones se sentarán en el determinar los estudios de los programas, olvidando tanto la evaluación del diseño como de los procesos.

Este aspecto merece una especial consideración ya que no realizar este tipo de evaluaciones implica, por una parte, aplicar programas de los que no se tiene certeza de que sean varios para alcanzar los objetivos planteados y por otra perder la oportunidad de realizar mejoras en el proceso de aplicación del programa. Finalmente es notable que el estándar Contexto y antecedentes (B1) se tenga en cuenta casi en la mitad de las ocasiones. Se trata de un elemento también considerado en los estándares del Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010) y con elevada importancia en la evaluación, puesto que contempla las características contextuales a la hora de evaluar y analizar los resultados de evaluación, y pasar por alto

este elemento puede suponer que las causas de los resultados se atribuyan a aspectos que no sean sus verdaderos causantes. Es posible que el cumplimiento casi total de determinados indicadores en todos los programas revisados (A3, B2, D1, D3 y D4) se deba, asimismo, a la exigencia de las publicaciones científicas que consideran necesario para su difusión contar con dicha información. Del mismo modo, podría incrementarse la exigencia en el cumplimiento de otros indicadores de evaluación si dichas publicaciones también así lo consideraran, de tal manera que se consiguiera asegurar que los programas aplicados han sido evaluados correctamente, independientemente del modelo de evaluación seleccionado y empleado para ello.

Este trabajo pretende ser un acercamiento a la evaluación de programas de educación emocional y motivación que, para confirmar estos resultados requerirá ampliarla muestra de evaluaciones seleccionadas a través de revisiones sistemáticas que permiten analizar las evaluaciones de los programas en profundidad y aplicar los estándares de meta evaluación de manera completa.

Dada la abundante publicación de programas de intervención, será necesario seguir desarrollando una cultura de investigación evaluativa entre los diferentes profesionales de la educación (Manzano, 2000), poniendo el énfasis en la formación correcta de futuros evaluadores (LaVelle, 2019), de tal manera que puedan conocer las bases teóricas de los modelos de evaluación que emplean para evaluar con confianza en función de las necesidades y de la complejidad de los programas (Frye y Hammer, 2012). En este sentido, sería necesario que no solo tuvieran conocimientos de los modelos y elementos de la evaluación sino también, de los diferentes procedimientos de metaevaluación existentes para garantizar su calidad y rigor. Para ello, se les debería proporcionar formación tanto teórica como práctica en contextos educativos reales. Además, sería de interés para los investigadores poner el foco en la evaluación de programas y no tanto en la creación o diseño de nuevos instrumentos de intervención que, en muchas ocasiones, se quedan sin posteriores aplicaciones.

Conclusiones

- 1. Contribuir a la comunidad científica no siempre supone la elaboración de un nuevo instrumento o programa; una adecuada contribución es alimentar la producción obteniendo y validando herramientas de valor para su aplicación real en el contexto educativo.***

2. *Desde un punto de vista más aplicado y práctico, y ante la importante tarea de selección y aplicación de programas de intervención que tienen por delante orientadores, pedagogos y educadores en general, dispones de estudios sólidos que acompañen los resultados de aplicación de los programas y que pongan su foco en la evaluación de los mismos, podrá ser una herramienta útil para la toma de decisiones.*

Referencias bibliográficas

1. Álvarez Rojo, V. G. (2002). *Diseño y evaluación de programas*. EOS.
2. Álvarez Rojo, V. y. (1998). El modelo de intervención por programas. *Revista de Investigación Educativa*, 79-123.
3. Barati, M. J. (2018). Studying the impact of metacognitive beliefs and self-regulatory strategies on students' educational motivation. *Iranian Journal of Social Sciences and Humanities Research*.
4. Beltrán, M. R. (2014). Introduction to special issue profiles 2013 educational evaluation in higher education. *Perfiles Educativos*, XXXVI(145). [http://dx.doi.org/10.1016/S0185-2698\(14\)70645-2](http://dx.doi.org/10.1016/S0185-2698(14)70645-2).
5. Burunat, E. (2004). El desarrollo del sustrato neurobiológico de la motivación y emoción en la adolescencia: ¿un nuevo período crítico? *Infancia y aprendizaje*, 87-104.
6. D., A. (2014). *Bases del aprendizaje y educación*. UNED.
7. Datu, A. J. (2021). La vida comprometida académicamente de los estudiantes orientados al dominio: Orden causal entre emociones positivas, metas de dominio y compromiso académico. *Revista de Psicodidáctica (English ed.)*, <https://doi.org/10.1016/j.psicoe.2021.02.001>.
8. De Caso, A. M. (2014). Pautas para el estudio de la motivación académica. *International Journal of Developmental and Educational Psychology. INFAD de Psicología*, 213-220.
9. Expósito, J. O. (2004). Patrones metodológicos en la investigación española sobre evaluación de programas educativos. *RELIEVE*, 185-209.
10. Frye, A. W. (2012). Program evaluation models and related theories: AMEE guide no. 67. *Medical teacher*, 288-299.
11. García, F. (2008). *Motivar para el aprendizaje desde la acción orientadora*. Madrid: Ministerio de Educación y Ciencia: Madrid. .
12. Järvelä, S. (2001). *Shifting research on motivation and cognition to an integrated approach on learning and motivation in context*.
13. Juste, P. (2000). La evaluación de programas educativos: conceptos básicos, planteamientos generales y problemática. *Revista de investigación educativa*, 261-287.
14. Kirkpatrick, D. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco.
15. Lange, J. D. (2020). Toward an Integrative Psychometric Model of Emotions. *Perspectives on Psychological Science*, 444-468.
16. LaVelle, J. (2019). Educating Evaluators 1976–2017: An Expanded Analysis of University-Based Evaluation Education Programs. *American Journal of Evaluation*, 1-16.
17. López, J. G. (2006). Aportaciones de la teoría de las atribuciones causales a la comprensión de la motivación para el rendimiento escolar. *Revista de la Facultad de Educación de Albacete*, 217–232.
18. M. Rueda, S. S.-B. (2014). La evaluación en la educación superior. *Perfiles Educativos*, XXXVI(145).
19. Manzano, N. (2000). Indicadores para evaluar programas de orientación educativa. *Revista de Orientación Educativa y Psicopedagógica*, 51-75.
20. Martínez Mediano, C. (2017). *Evaluación de programas*. UNED.
21. Méndez Pineda, J. M. (2018). *Evaluación externa del programa escuelas de tiempo completo*.
22. OMS. (2020). *OMS*. Obtenido de Adolescence.
23. Pcheu, A. (2018). ¿Existen diferencias en la motivación de distintas generaciones en Enfermería? *Revista Médica Clínica Las Condes*, 336 - 342.

24. Pérez-González, J. C. (2008). Propuesta para la evaluación de programas de educación socioemocional. *Electronic Journal of Research in Education Psychology*, 523-546.
25. Pulido Acosta, F. y. (2017). La influencia de las emociones sobre el rendimiento académico. *Ciencias Psicológicas*, 29-39.
26. Rinaudo, M. C. (2003). Motivación y uso de estrategias en alumnos universitarios. *Anales de psicología*, 107-119.
27. Samson, B. y. (2018). Improving evaluation of students: a meta-evaluation study on cce from the perspective of students evaluation standards. *MIER Journal of Educational Studies*.
28. Sánchez Calleja, L. G. (2016). Evaluación del diseño del programa AedEM de Educación Emocional para Educación Secundaria. *Revista Electrónica de Investigación y Evaluación Educativa*, 1-25.
29. Sánchez, J. J. (2015). El fracaso escolar y los enfoques de aprendizaje: medidas para la inclusión educativa. *Revista Electrónica Interuniversitaria de formación del profesorado*, 183-196.
30. Sawyer, S. M. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, 223-228.
31. Scriven, M. (197). *The methodology of evaluation: Perspectives of curriculum evaluation*. Rand McNally.
32. Stake, R. (1975). *Evaluating the Arts in Education: A Responsive Approach*. Columbus.
33. Stufflebeam, D. L. (1971). *The relevance of the CIPP evaluation model for educational accountability*. Annual meeting of the American Association of School Administrators.
34. Tejedor, F. J. (2000). El diseño y los diseños en la evaluación de programas. *Revista de investigación educativa*, 319-339.
35. Torrijos, P. T. (2018). Experimental evaluation of emotional development programmes for teachers in Secondary Education. *anales de psicología*, 68-76.
36. Yarbrough, D. S. (2010). *The Program Evaluation Standards: A guide for evaluators and evaluation users*. Thousand Oaks: Corwin Press.