

56

Fecha de presentación: marzo, 2023

Fecha de aceptación: mayo, 2023

Fecha de publicación: julio, 2023

EVALUACIÓN DE LA CALIDAD

DE LA FORMACIÓN DOCENTE UNIVERSITARIA: REVISIÓN Y ANÁLISIS DE MODELOS

EVALUATION OF THE QUALITY OF UNIVERSITY TEACHER TRAINING: REVIEW AND ANALYSIS OF MODELS

Milgjan Dixiana Martínez¹

E-mail: milgjan.martinez@unah.edu.hn

ORCID: <https://orcid.org/0000-0002-5715-0320>

¹Universidad Nacional Autónoma de Honduras, (UNAH). Honduras.

Cita sugerida (APA, séptima edición)

Dixiana Martínez, M. (2023). Evaluación de la calidad de la formación docente universitaria: revisión y análisis de modelos. *Universidad y Sociedad*, 15(4), 565-574.

RESUMEN

El objetivo del artículo es identificar las características principales, objetos de evaluación, roles, fases, dimensiones y categorías de análisis de 17 modelos de evaluación de programas de formación, clásicos y contemporáneos, con miras a aportar al análisis de la calidad de la formación del profesorado universitario. Se encontró que la mayor parte de los modelos fueron construidos para ámbitos distintos a la formación del profesorado, aunque permiten una adaptación pertinente en este campo; en ese sentido, predomina una evaluación de control, principalmente los modelos clásicos, y un enfoque orientado a la toma de decisiones, que muchas veces tiene implicaciones de mejora, aunque esto no siempre es así. Asimismo, se identificaron dimensiones similares en ambos grupos de modelos, como: los objetivos, contenidos, modalidades y estrategias, y mecanismos para el seguimiento y mejora de los procesos formativos en el marco de los programas. Se encontró que un modelo, el procesual para el desarrollo de estándares de calidad para programas de formación docente a través de aprendizaje abierto y a distancia, preparado por Marcelo en 2008, tiene un enfoque de calidad total, y comprende diferentes dimensiones, categorías y modalidades, buscando adaptarse a los distintos formatos en que se puede desarrollar la formación docente.

Palabras clave: calidad de la formación, evaluación de la calidad, formación docente, evaluación de programas, modelos de evaluación.

ABSTRACT

The objective of the article is to identify the main characteristics, objects of evaluation, roles, phases, dimensions and categories of analysis of 17 models of evaluation of training programs, classic and contemporary, with a view to contributing to the analysis of the quality of the training of the university teachers. It was found that most of the models were built for areas different than teacher training, although they allow pertinent adaptation in this field; in this sense, a control evaluation predominates, mainly the classic models, and a decision-oriented approach, which often has implications for improvement, although this is not always the case. Likewise, similar dimensions were identified in both groups of models, such as: the objectives, contents, modalities and strategies, and mechanisms for monitoring and improving the training processes within the framework of the programs. It was found that a model, the procedural model for the development of quality standards for teacher training programs through open and distance learning, prepared by Marcelo in 2008, has a total quality approach, and includes different dimensions, categories and modalities, seeking to adapt to the different formats in which teacher training can be developed.

Keywords: training quality, quality assessment, teacher training, program evaluation, evaluation models.

INTRODUCCIÓN

La evaluación es un proceso complejo, que implica la realización de juicios de valor sobre un aspecto determinado; en educación, esta complejidad resulta más evidente, dado que conlleva múltiples aristas que interactúan, de manera directa o indirecta, en el proceso de evaluación; por ello, al evaluar programas de formación, se tiene que delimitar el problema u objeto de la evaluación (Archibald, 2019), considerando que en el aprendizaje de los profesores pueden incidir muchos factores y dimensiones; en esa línea, pueden haber diferentes perspectivas: subjetivas, con mayor objetividad o fundamentalmente holísticas.

Según la literatura existe una multiplicidad de modelos de evaluación en respuesta a la generalizada tendencia a un limitado rango de experiencias que, quizás sean numerosas en cantidad, pero escasas en variedad; por ello resulta importante definir claramente el problema a investigar, así como establecer con precisión cada uno de sus componentes, con miras a generar información que aporte a la toma de decisiones en relación a la continuidad y necesidades de transformación de los programas de formación.

En ese orden de ideas, es importante acotar que, al evaluar los procesos de formación, se obtiene información de cómo se impacta en las prácticas que los profesores desarrollan en las aulas universitarias, qué tipo de actividades llevan a cabo, los materiales que emplean, la metodología que emplean, las inquietudes e intereses del alumnado, las relaciones, clima de clase y otros factores de suma importancia para la mejora de la calidad educativa.

Desde este contexto, dada la complejidad de la formación de las y los docentes, a través de la investigación de la cual surge el presente artículo, se abordó la problemática de la evaluación de los programas dirigidos a estos profesionales, entendiéndola también como un proceso farragoso y de mucho galimatías, por lo que se buscó dar respuesta a las siguientes interrogantes: ¿Qué modelos se han desarrollado durante las últimas décadas para evaluar programas de formación? ¿Cuáles son sus características? ¿En qué medida los mismos responden al ámbito de la educación superior? En esa línea, se realizó una investigación teórica, enfocada cualitativamente, a partir de la técnica del análisis documental bibliográfico, con el objetivo de caracterizar los modelos más empleados para evaluar programas de formación.

En línea con lo anterior, Pérez (2016) afirma desde su amplia experiencia en la evaluación de programas, que con frecuencia los modelos diseñados para ello, debido

a que intentan ser tan abarcadores, no resultan aplicables si no es consumiendo más tiempo, energía y recursos de los que prudentemente se les puede asignar en detrimento de las actividades al servicio del logro de lo que la evaluación pretende constatar; por tanto, la evaluación de programas debe tener en cuenta diversas técnicas, estrategias y criterios, ya que cada programa de formación cuentan con características únicas, por lo que es importante contar con modelos que consideren estas particularidades desde una perspectiva donde prime un pensamiento evaluativo que identifique pertinentemente los elementos que representa el problema u objeto de la evaluación, como plantea Archibald (2019).

Dado que la evaluación de programas ha sido mayormente estudiada para evaluar la efectividad y el impacto de la formación en otros campos distintos a la educación, como la empresa, la salud, el desarrollo comunitario, la innovación social (Hargraves & Denning, 2018).

Un estudio realizado por Parra & Ruiz (2020) en el que se analizan los aspectos fundamentales, modelos y perspectivas actuales relacionadas con la evaluación de impacto de los programas formativos, encuentran que actualmente el énfasis se da en la medición a nivel básico, de la satisfacción de los participantes, sin profundizar en los niveles superiores como la transferencia o el impacto en el desarrollo de competencias o en la práctica profesional, por lo que la información que se genera a partir de este tipo de evaluaciones no recoge evidencias sustanciales respecto a la calidad de los aprendizajes de los participantes, limitando así la toma de decisiones sobre factores como, por ejemplo, la inversión pública y privada que reciben los programas de formación, o la política formativa de las unidades y entidades responsables.

En el campo de la salud, destaca el estudio de Hamza & Oandasan (2020), quienes evaluaron, desde un modelo similar al CIPP de Daniel Stufflebeam, la educación médica basada en competencias (CBME, por sus siglas en inglés), de la cual, si bien se ha tenido un gran auge en los últimos años, es poca la literatura existente en relación al impacto que la misma produce en los estudiantes. A partir de los resultados de la evaluación realizada por los autores, se evidenció que emparejar deliberadamente la evaluación con el cambio es una actividad importante y, cuando se logra, produce información valiosa de las experiencias de quienes implementan y participan en un programa.

Desde la mirada de Pérez (2016), estudioso de la evaluación de programas, la población participante en los procesos de evaluación, entendidos estos como una forma de investigación evaluativa, puede estar constituida por

educadores, profesores y formadores, en diferentes contextos, por tanto, es necesario estudiar no solamente los cambios que se dan en las prácticas luego de que los profesores participan en programas de formación, sino también los niveles en que suceden todos estos cambios, en vista de que esta información permitirá valorar el impacto de estos programas.

DESARROLLO

Para el desarrollo de la investigación se realizó una revisión de literatura científica sustentada en una estrategia basada en palabras clave (modelos de evaluación, modelos de evaluación de programas, modelos clásicos de evaluación, modelos contemporáneos de evaluación), así como en criterios de selección, inclusión y exclusión (se consideraron modelos tanto anglosajones, como hispanos e incluso francófonos, con relevancia a nivel teórico e importantes aportes en el campo de la educación).

Para la revisión de los modelos, se empleó la técnica del análisis documental bibliográfico, la cual según Casasempere & Vercher (2020), propicia la obtención de un máximo rendimiento al realizar revisiones de literatura en investigaciones cualitativas. Desde este contexto, cabe señalar que, en línea con los objetivos de la investigación desde la cual surge el presente artículo, el estudio no se enfocó desde métodos cuantitativos ni estadísticos, sino meramente cualitativos.

Partiendo de lo anterior, los 17 modelos seleccionados para la realizar el análisis, comprendieron un periodo de desarrollo que abarcó 73 años, desde 1942 hasta 2015; estos modelos fueron: ver tabla 1

Tabla 1.- Modelos seleccionados para realizar el análisis documental bibliográfico

| Modelos clásicos | Autor | Modelos contemporáneos | Autor |
|---|--------------------|--|---|
| Modelo de verificación de logros | Ralph Tyler | Modelo para valorar el impacto de la formación | Pamela Wade |
| Modelo de los cuatro niveles para la evaluación de la formación | Donald Kirkpatrick | Modelo de evaluación | Serge Barzucchetti y Jean-Francois Claude |
| Modelo de evaluación responsable | Robert Stake | Modelo con cinco niveles críticos de evaluación | Thomas Guskey |
| Modelo de evaluación por estándares | | Modelo de evaluación de programas | José Tejada |
| Modelo CIPP | Daniel Stufflebeam | Modelo de evaluación | Ramón Pérez Juste |
| Modelo de evaluación retributiva | Michael Scriven | Modelo procesual para el desarrollo de estándares de calidad para programas de formación docente a través de aprendizaje abierto y a distancia | Carlos Marcelo |
| Modelo de evaluación por inducción | | Modelo Kirkpatrick del Nuevo Mundo | Wendy y Jim Kirkpatrick |
| Modelo sin referencia a metas | | Modelo de evaluación evolutiva | Urban, Hargraves y Trochim |
| Modelo orientado a la planeación | Lee Cronbach | | |

Fuente: Elaboración propia

Para el análisis se tomaron en cuenta las principales características de cada modelo, sus objetos de evaluación, la metodología que emplean, los participantes involucrados en el proceso de evaluación y las fases, dimensiones y vínculos con distintas tipologías y enfoques de evaluación y teorías del aprendizaje, entendiendo que el aprendizaje en los participantes es uno de los propósitos de cualquier programa.

Al indagar sobre los primeros teóricos que propusieron modelos para evaluar programas de formación, salta a la palestra el nombre del experto norteamericano Ralph Tyler, quien en 1942 propuso un modelo de verificación de logros. Para este autor, la evaluación de la formación se limita a recopilar datos al respecto, y comparar los mismos con los objetivos previamente establecidos, con miras a comprobar qué participantes alcanzan los objetivos; en ese

sentido, los beneficiarios de la información que se genera son los planificadores y directivos del programa (Gómez & Valdéz, 2019).

Otro de los clásicos que han desarrollado modelos para evaluar programas de formación es el también norteamericano, Donald Kirkpatrick quien, en el año de 1959 planteó cuatro niveles de evaluación de la formación. El primer nivel que propone Kirkpatrick, como señala Cahapay (2021), es la evaluación de la reacción de los participantes ante las acciones formativas. En segundo lugar, el aprendizaje de los participantes, demostrada por el cambio de actitudes, ampliación de conocimientos y/o mejora de capacidades como consecuencia de asistir a las acciones formativas. El tercer nivel, estrechamente relacionado con el segundo, tiene que ver con la conducta, en tanto la medida en que se han dado los cambios en el comportamiento del participante evidenciados en el segundo nivel.

Como cuarto nivel, se encuentra la evaluación de los resultados finales que ocurren debido a que los participantes han asistido a las acciones formativas. Estos resultados pueden ser: aumento de la producción, mejora de la calidad, menores costos, reducción de la frecuencia y/o de la gravedad de los accidentes, incremento de las ventas, reducción de la rotación de la plantilla y mayores beneficios, entre otros (figura 1).



Figura 1.- El Modelo de Kirkpatrick para la Evaluación de la Formación. Fuente: Elaboración propia

En 1963, el psicólogo educativo norteamericano Robert Stake, planteó un modelo de evaluación de programas al que denominó “respondiente”. Este modelo fue bastante particular en su tiempo, ya que el autor lo formuló de una manera eminentemente cualitativa, contemplando una valoración de los resultados de los programas de formación en tres momentos: en una etapa inicial se consideran los antecedentes del programa, así como la descripción de este; en un segundo momento, los resultados

son tomados en cuenta para realizar un juicio y establecer nuevos antecedentes para desarrollar las “transacciones”, para luego realizar una última evaluación de los resultados. En esa línea, la evaluación contempla la historia del programa, sus antecedentes, operaciones, efectos secundarios, logros diferidos y resultados, de los cuales se benefician, además de los participantes, los expertos (Bundi & Pattyn, 2022).

Por su parte, y casi al mismo tiempo que Stake, el también norteamericano Daniel Stufflebeam formulaba en 1965 su modelo de toma de decisiones, también conocido como “modelo CIPP” (contexto, input, proceso y producto). El modelo está conformado por cuatro elementos clave: el análisis de las tareas, el plan de obtención de la información, el plan para el informe sobre los resultados, y el plan para la administración del estudio. En ese orden de ideas, el énfasis del modelo está en proporcionar conocimientos suficientes y una base valorativa para tomar y justificar las decisiones. Para ello se indaga sobre cómo planificar, ejecutar y reutilizar la información para promover el crecimiento y el desarrollo a un coste razonable (Nouraey et al., 2020).

Contemporáneo de Stufflebeam, el matemático y filósofo norteamericano Michael Scriven propuso en 1966 tres modelos para el abordaje de la evaluación de programas de formación, como indican Kitivo & Kavulya (2021): el modelo de evaluación retributiva, el de evaluación por inducción, y el modelo sin referencia a metas. El primero considera un análisis de las metas educativas y los efectos que provoca el programa de formación, a partir de los cuales se considera un perfeccionamiento de las metas para evitar el abandono del programa de formación por los participantes.

El modelo de evaluación por inducción se enfoca en la metaevaluación de los procesos formativos, considerando que los programas de formación deben contar con una evaluación no solo sumativa, que permita la certificación de los participantes, sino también formativa, que permita a los participantes recibir la retroalimentación necesaria para que los aprendizajes sean más significativos y que se logren los objetivos del programa; en ese sentido, el autor plantea como uno de los elementos de este modelo, la valoración de los resultados, los cuales se deben plasmar en informes que permitan considerar, en función con otras alternativas, si el programa formativo puesto en marcha tiene posibilidades de generalización.

El tercer modelo propuesto por Scriven, parte de una valoración de necesidades formativas, a partir de la cual se realiza la evaluación de las metas del programa, comparando las características de este con otras alternativas,

con miras a determinar los costos y la efectividad, y se culmina con la evaluación del personal y del propio programa. A través de este modelo, Scriven pretendía juzgar los méritos relativos de bienes y servicios alternativos, y su relación con los costes, las necesidades de los participantes y los valores de la sociedad en general, por lo que los beneficiarios directos de esta evaluación son los participantes.

El último de los modelos clásicos revisados respecto a la evaluación de programas de formación fue el orientado a la planeación, formulado por Lee Cronbach en 1973, el cual plantea dos fases: una divergente y otra convergente. En ese sentido, se basa en tres factores principales: las Unidades (U), que hacen referencia a los individuos o grupos objeto de la evaluación; los Tratamientos (T), a los que se expondrán las unidades; y, las Operaciones (O), que se refieren a las observaciones que se realizan antes, durante y después de exponer las unidades a los tratamientos específicos (Shaw, 2017).

Luego de revisar las características de cada uno de los modelos de evaluación objeto de análisis, y en línea con lo que plantea la literatura acerca de la tipología y enfoques de evaluación, y de las teorías de aprendizaje, como aspecto determinante en relación al logro de los objetivos de los programas de evaluación, se encontró que algunos de los modelos anteriores, como por ejemplo el de Donald Kirkpatrick y Wendy y Jim Kirkpatrick, abarcan los 3 tipos de evaluación: orientada al control, al proceso y al impacto. Asimismo, se observó que, si bien existe en cierta medida una evaluación de control, algunos modelos se orientan a dos tipos de evaluación: de proceso y de impacto, ejemplo de ello son el Modelo de evaluación por estándares, de Robert Stake y el Modelo de evaluación sin referencia a metas, de Michael Scriven.

También se encontraron algunos modelos con una orientación predominantemente hacia el control, como ser el Modelo de Tyler, el CIPP y el de Lee Cronbach, los cuales se centran en medir la consecución de los logros para considerar si un programa de formación debe continuar o no. Modelos como el respondiente, planteado por Stake y el de evaluación retributiva, de Scriven, tienen una tendencia hacia la evaluación del impacto.

En cuanto a los enfoques de evaluación, los modelos de Tyler, el de Donald Kirkpatrick, el modelo CIPP, de Daniel Stufflebeam, el de evaluación retributiva de Michael Scriven, y el de Lee Cronbach, se orientan a la toma de decisiones, ya que postulan que los resultados de la evaluación benefician a los gestores de los programas, quienes se encargan de hacer juicios de valor para determinar

el curso a seguir o, en su defecto, la necesidad de cancelar los programas.

El modelo de Cronbach percibe a los participantes como *objeto de evaluación* o insumos/fuentes de información, que permite obtener referentes para la toma de decisiones. Al respecto, el modelo CIPP también considera al sujeto como un *insumo*, una materia prima que debe ser transformada en un *producto* a través de un *proceso*, tomando en cuenta su *contexto*, siendo beneficiarios de la evaluación los tomadores de decisiones. Por su parte, el modelo de evaluación retributiva se centra en las metas y resultados, es decir, en la *eficiencia terminal* del programa, mientras que el modelo de evaluación evolutiva busca conocer la eficacia de los programas, para la toma de decisiones. En contraste, los modelos de evaluación por inducción y sin referencia a metas, de Michael Scriven, y los modelos respondientes y por estándares, de Robert Stake (Bundi & Pattyn, 2022), se enfocan en la mejora institucional.

El Modelo de evaluación por inducción, de Scriven, contempla que los participantes reciban la retroalimentación necesaria para que los aprendizajes sean más significativos, además del logro de objetivos; en esa línea, se centra en reforzar las prácticas educativas de los participantes como parte de la mejora institucional, mientras que el Modelo sin referencia a metas, también propuesto por el mismo autor, valora las necesidades formativas, por lo que los beneficiarios directos del proceso de evaluación son los participantes.

En cuanto a los modelos de Robert Stake, el de evaluación por estándares, que tiene una visión cualitativa, se relaciona con el enfoque de mejora institucional ya que considera las necesidades de los participantes, mientras que el modelo respondiente, también cualitativo, se centra en los resultados y logros diferidos del programa, considerando los antecedentes y la historia del programa.

En relación a las teorías del aprendizaje, en el entendido de que el aprendizaje de los participantes es uno de los principales aspectos que se evalúa de los programas, se percibe que los modelos de Tyler, Kirkpatrick, el modelo respondiente, el CIPP, el de evaluación retributiva y el modelo de Lee Cronbach se alinean más al conductismo, dado que se enfocan en el logro de los resultados, sin considerar de manera auténtica a los participantes.

En contraste, los modelos de inducción y sin referencia a metas, de Michael Scriven, agregan un componente de retroalimentación a los participantes, para mejorar sus aprendizajes, por lo que responderían mejor al cognitivismo.

Luego de casi una década sin plantear ninguna novedad en cuanto a la evaluación de programas de formación, Pamela Wade propuso en 1994 un modelo para valorar el impacto de la formación, enfocado en calcular y medir los resultados de los programas. En esa línea, la autora concibe la evaluación como la medición del valor que la formación aporta a la organización. El modelo contempla cuatro niveles: respuesta, acción, resultados e impacto. El primer nivel valora las reacciones de los participantes ante la formación y el aprendizaje; el nivel de acción se refiere a la transferencia de aprendizajes al puesto de trabajo; el tercer nivel se enfoca en los efectos de la formación en la institución, medidos mediante indicadores cuantitativos o duros y cualitativos o blandos; y, el cuarto nivel valora el impacto de la formación en la organización, a través del análisis del coste-beneficio (Parra & Ruiz, 2020).

Al mismo tiempo, los autores franceses, Serge Barzucchetti y Jean-Francois Claude proponían en 1995 un modelo con tres niveles: condiciones de explotación, situación profesional y situación de formación. El primer nivel evalúa las condiciones de funcionamiento y explotación de la organización; el nivel relacionado con la situación profesional analiza el grado de creación o evolución de las competencias individuales o colectivas necesarias para el trabajo, mientras que el tercer nivel, vinculado a la situación de formación, valora las transformaciones del participante logradas en la interfaz pedagógica, respecto a una posible nueva representación de la realidad y nuevos conocimientos que se pueden posteriormente transferir a su desempeño laboral (Barzucchetti y Claude, 1995, como se citó en UNICEF, 2019).

Ya a inicios del siglo XXI, Thomas Guskey proponía un modelo con cinco niveles críticos de evaluación del desarrollo profesional, como refieren McChesney & Aldridge (2018): en el primer nivel se valoran las reacciones de los participantes (también denominada coeficiente de felicidad), en cuanto a la experiencia de formación en general; más que valorar “el entretenimiento”, en este nivel se pretende conocer la satisfacción inicial de los participantes, lo cual puede contribuir a mejorar el diseño y entrega de los programas o actividades de forma válida. La información es recopilada generalmente a través de cuestionarios que se aplican al finalizar cada sesión o actividad, aunque también suelen utilizarse grupos focales, entrevistas, registros de aprendizajes y otros.

El segundo nivel del modelo de Guskey se enfoca en valorar el conocimiento, habilidades y actitudes que los participantes hayan adquirido, en función de las metas del programa o actividad formativa; esto requiere definir criterios e indicadores específicos previos al comienzo

de la experiencia de desarrollo profesional. Esto permite identificar, además del logro de los aprendizajes como parte del diseño, la obtención de aprendizajes no intencionados, positivos o negativos, así como las preocupaciones de los participantes. Los resultados recopilados (ya sea al final o durante las acciones) permiten mejorar el contenido, el formato y la organización del programa.

El apoyo y cambio en la institución, que corresponde al tercer nivel del modelo propuesto por Guskey, depende de cada organización, confluyendo diversas variables, como el soporte, la disponibilidad hacia el cambio, las políticas institucionales, que pueden determinar el éxito o fracaso de cualquier esfuerzo de desarrollo profesional. En ese sentido, pueden encontrarse escasos resultados positivos en este nivel, mismos que no se deben a una pobre formación, sino mayormente a la incompatibilidad en cuanto a la implementación de esfuerzos. Para evitar que esto sea contraproducente respecto al resto de niveles, resulta importante considerar toda la información lo cual, si bien es complejo, puede lograrse aplicando procedimientos diversos, como análisis de registros, cuestionarios, grupos focales y entrevistas estructuradas, entre otras.

En el cuarto nivel, concerniente a la aplicación de los nuevos conocimientos y habilidades por los participantes en el trabajo, la clave es la clara especificación de los indicadores que revelen tanto el grado como la calidad de la ejecución. Además de las técnicas aplicadas en los niveles anteriores para recopilar la información, en este nivel Guskey plantea la utilización de la observación directa, de la manera más discreta posible; dicha recolección no puede realizarse en una sola sesión, sino que debe realizarse luego de un cierto tiempo (e idóneamente en varios momentos) después del cual los participantes puedan adaptar y aplicar a sus prácticas y a su entorno los nuevos aprendizajes, de manera gradual, dependiendo del interés en el uso continuado o permanente, para reestructurar los programas y actividades futuras, para facilitar la aplicación de una mejor y más coherente forma.

Como puede observarse hasta este punto, los teóricos de la evaluación de programas, principalmente los angloparlantes, habían planteado modelos orientados a determinar, fundamentalmente, si la inversión que realizaban las organizaciones, alcanzaba los resultados necesarios para mantener en funcionamiento los programas de formación, lo cual, con el auge del enfoque de formación por competencias, se entendía como el desarrollo o fortalecimiento de las competencias necesarias en los empleados para desempeñarse de manera efectiva en sus puestos laborales.

Sin embargo, algunos autores venían analizando los elementos de estos modelos para contextualizarlos al ámbito de la educación. Uno de esos autores es el psicólogo y pedagogo José Tejada, quien en 1997 planteaba una serie de criterios e indicadores para evaluar programas, tomando en cuenta la pertinencia, actualización, objetividad, aplicabilidad, suficiencia, eficacia, eficiencia, comprensividad, relevancia y coherencia de los programas. Desde este contexto, el autor esbozó un modelo que considera tres momentos en la evaluación del programa: inicial, procesual y de resultados; para cada uno valora tres dimensiones: la conceptual (el qué y el para qué de la evaluación), la metodológica (el cómo de la evaluación), y la operativa (el con qué realizar la evaluación). Particularmente respecto al tercer momento, en el que según Tejada se evalúan los resultados del programa, se propone una valoración con tres finalidades: verificar la consecución (logros) de los objetivos, valorar los cambios previos y no previstos (impacto), y verificar la valía (costes) del programa de cara a la satisfacción de las necesidades (Tejada & Navío, 2019).

Bajo esta misma óptica, Ramón Pérez Juste, quien luego de muchos años de reflexionar sobre la docencia y la investigación planteaba en el año 2000 desde su modelo de evaluación, que, para evaluar programas, es fundamental contar con planteamientos metodológicos simples y de fácil acceso, para que los que los evaluadores no desperdicien el potencial de mejora de la evaluación (Pérez Juste, 2006, como se citó en Pérez, 2016). Por ello, el autor aboga por una complementariedad metodológica, en tanto a la apertura para consultar diversas fuentes de información, y el uso de diferentes técnicas de recogida y análisis de datos, en función de su naturaleza y de los objetivos de cada caso. En ese orden de ideas, el autor propone un modelo de evaluación de programas de formación con tres momentos: el momento previo, en el que se valoran los objetivos del programa; el momento denominado como “núcleo de la evaluación”, en el que se realiza la recogida y valoración de la información, a partir de la cual se toman decisiones orientadas a la mejora; y el momento posterior, en el que se brinda seguimiento de los planes de mejora.

Respecto a los estándares que, según Pérez Juste y Tejada deben considerarse al evaluar programas, Robert Stake, quien desde los años 60's tenía una visión cualitativa de la evaluación, planteó en 2006 un modelo de evaluación por estándares, que según Bundi & Pattyn (2022), considera no solo los objetivos del programa, sino también las necesidades de los receptores, sus rendimientos y actuaciones y, por supuesto, los estándares

de evaluación con sus respectivas ponderaciones, y los costes del programa.

En esta misma línea, el español Carlos Marcelo preparó para la oficina regional de la UNESCO en América Latina y el Caribe, un modelo procesual para el desarrollo de estándares de calidad para programas de formación docente a través de aprendizaje abierto y a distancia, el cual busca un buen grado de flexibilidad que permita adaptarse a los distintos formatos de formación docente (Marcelo, 2008, como se citó en Delgado, 2019). Las dimensiones que contempla el modelo son: el contexto, el diseño, la producción, la puesta en marcha, la implementación y el seguimiento, cada una de las cuales parte de un estándar general, que luego se subdividen en estándares más específicos, de acuerdo con la amplitud y complejidad del estándar.

El último de los modelos revisados fue el de evaluación evolutiva de los programas, propuesto por Jennifer Urban, Mónica Hargraves y William Trochim en 2014, a través del cual de acuerdo con Szijarto & Cousins (2019), se busca analizar y juzgar a profundidad el dinamismo y estatismo de los programas, según diversos objetivos y variadas metodologías, a partir de rigurosos criterios y estándares científicos, y para distintos propósitos como la planificación, la mejora, la certificación, la acreditación, el diagnóstico, la reforma, la incentivación, entre otros.

Para culminar con la revisión de los modelos de evaluación de programas de formación nos pareció importante traer a colación la reformulación de los planteamientos que Donald Kirkpatrick hiciera en los años 50's. En el año 2015, Jim Kirkpatrick, hijo de Donald Kirkpatrick y su esposa Wendy Kirkpatrick, rediseñaron el modelo clásico, denominándolo “Modelo Kirkpatrick del nuevo mundo”, el cual según refiere Gandomkar (2018), mantiene como parte de su base los cuatro niveles postulados a mediados del siglo pasado por Donald Kirkpatrick, y agrega nuevos elementos con el propósito de aportar a una operacionalización efectiva. ver figura 2

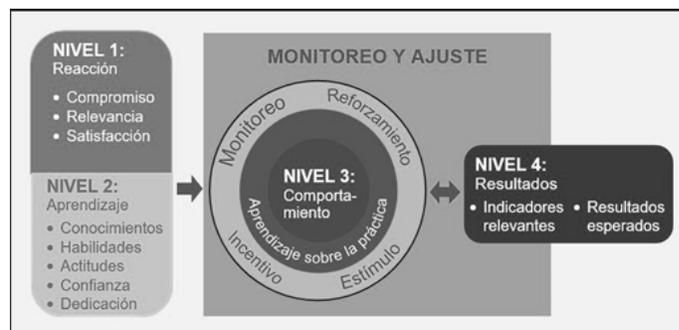


Figura 2.- El Modelo Kirkpatrick para el Nuevo Mundo

Fuente: Traducido del Modelo de Kirkpatrick para el Nuevo Mundo, referido por Gandomkar (2018)

Para cada uno de los cuatro niveles planteados en el modelo básico de Kirkpatrick, el nuevo modelo agrega algunos indicadores que, según los autores, son relevantes para valorar la efectividad de los procesos formativos. En ese sentido, el modelo de Kirkpatrick para el Nuevo Mundo, así como su precursor, pretende valorar en qué medida se producen los resultados específicos como resultado de los eventos de aprendizaje y el posterior reforzamiento; asimismo, busca indagar en qué medida los participantes adquieren los conocimientos, habilidades y actitudes previstos, y hasta qué punto los aplican al regresar a sus puestos de trabajo; también identifican la reacción de los participantes en relación a los eventos formativos.

De acuerdo con la revisión realizada, modelos como el de Barzucchetti y Claude, además de otros más clásicos, como el de evaluación por inducción, propuesto por Scriven, tienen una tendencia a evaluar el proceso, dado que consideran un componente de retroalimentación a los participantes, con miras a que sus aprendizajes mejoren, mientras que el modelo de Pamela Wade y el de Urban, Hargraves y Trochim, tiene se orientan a evaluar el impacto, ya que buscan valorar los resultados que alcanzan los programas, principalmente desde una relación costo-beneficio, con miras a considerar algunos aspectos para la mejora.

En esa línea, al revisar los enfoques de evaluación, se observó que el Modelo de los Cinco Niveles Críticos de la Evaluación del Desarrollo Profesional, de Thomas Guskey, que considera las preocupaciones de los participantes, está vinculado tanto al enfoque de toma de decisiones como al de mejora institucional, ya que busca, a partir de los resultados del proceso evaluativo, plantear posibilidades de mejora en los aprendizajes de los participantes, con miras a fortalecer el desempeño y las prácticas educativas. El Modelo Kirkpatrick del Nuevo Mundo, de Wendy y Jim Kirkpatrick, y el de evaluación evolutiva de Urban, Hargraves y Trochim, también se plantean desde un enfoque orientado hacia la mejora institucional, ya que continúa buscando cambios o aprendizajes en los participantes; por tanto, los evaluadores buscan recoger información para tomar decisiones al respecto.

Otros modelos con enfoque de mejora institucional son el de José Tejada, Ramón Pérez Juste, y el de Barzucchetti y Claude. El modelo de evaluación de programas de José Tejada se enfoca en los aprendizajes alcanzados en los participantes, con tres finalidades: verificar logros, valorar impacto y verificar costes; para ello, analiza el

coste-beneficio y hace una valoración de aspectos curriculares del programa. De manera similar, el modelo de Pérez Juste valora la complementariedad metodológica en función de criterios, enfocándose en valorar la capacidad de mejora de los programas, para lo cual plantea el seguimiento a los planes de mejora. Por su parte, el modelo planteado por Serge Barzucchetti y Jea-Francois Claude busca valorar el desarrollo o evolución en las competencias de los participantes, por tanto, se orienta hacia una mejora institucional a través del fortalecimiento de la práctica educativa.

Finalmente, de los modelos revisados, se identificó que el planteado por Carlos Marcelo se enfoca en la calidad total, desarrollando un proceso de evaluación por estándares dirigido a diferentes modalidades de formación. En cuanto a las teorías del aprendizaje, se reconoce que el modelo planteado por Pamela Wade tiene una orientación conductista, mientras que el modelo de Barzucchetti & Claude se ubica en el constructivismo, ya que valora las transformaciones de los participantes logrados en la interfaz pedagógica, respecto a una posible nueva representación de la realidad y nuevos conocimientos que posteriormente pueden ser transferidos a su desempeño laboral. De manera similar, los modelos de Thomas Guskey, José Tejada, Ramón Pérez Juste y Carlos Marcelo, evalúan no solo el conocimiento, sino también las habilidades y actitudes adquiridas por los participantes, su potencial de cambio, innovación, mejora y necesidades.

CONCLUSIONES

Luego de revisarse los 17 modelos para la elaboración de este artículo, se llega a la conclusión que la mayoría de ellos fueron elaborados para un ámbito ajeno a la formación del profesorado universitario; sin embargo, los modelos desarrollados para este propósito, rescatan múltiples elementos planteados por los autores clásicos, como ser los momentos o etapas de la evaluación: al inicio, durante y al finalizar la formación. Asimismo, se observa que los modelos clásicos dan particular interés a los resultados de los programas de formación en relación a los costos asignados para el desarrollo de los mismos.

Se identificó también que los modelos contemporáneos contemplan dimensiones similares a las de los modelos clásicos como, por ejemplo, las metas u objetivos de los programas de formación, los contenidos de estos, las modalidades y estrategias de estos, y los mecanismos de seguimiento para la mejora de los programas. De la misma forma, estos modelos consideran nuevos enfoques del aprendizaje, como ser la enseñanza y el aprendizaje por competencias, y el aprendizaje a lo largo de la vida. Cabe señalar que algunos modelos, como el de Robert

Stake y los de Pérez Juste, José Tejada y Carlos Marcelo, plantean una evaluación por estándares, considerando una lista bastante detallada y específica de los aspectos que se deben tomar en cuenta para evaluar cada dimensión los programas de formación.

Modelos como el de los cuatro niveles de Kirkpatrick para evaluar la formación, ha sido uno de los más populares a lo largo de las siete décadas de desarrollo teórico sobre la evaluación de programas, por lo que ha servido de base para otros teóricos de este campo; sin embargo, dado que actualmente tiene más de 60 años, la forma en que las personas aprenden y en que las organizaciones trabajan ha sufrido cambios, los cuales en algunas ocasiones, son bastante significativos, ejemplo de ello es el cambio en el término "formación" por "desarrollo profesional", principalmente en el ámbito de la educación.

A partir de la revisión puede identificarse también que, en función de las condiciones que deba cumplir la evaluación de programas, cuando las instituciones son las que promueven la evaluación del impacto de los programas, se perciben mayores aportes para la mejora de los procesos de formación, siempre y cuando dicha evaluación se realice de manera honesta y ética.

Es importante destacar que todos los modelos de evaluación de programas de formación revisados, tanto los clásicos como los contemporáneos, coinciden en que la evaluación debe analizar el grado de alcance de las metas u objetivos planteados en su diseño, y que debe brindar información que permita analizar aspectos para la mejora de los diferentes aspectos de los programas.

REFERENCIAS BIBLIOGRÁFICAS

- Archibald, T. (2019). What's the Problem Represented to Be? Problem Definition Critiques as a Tool for Evaluative Thinking. *American Journal of Evaluation*, 41(1):6-19. <https://doi.org/10.1177%2F1098214018824043>
- Bundi, P. & Pattyn, V. (2022). Citizens and Evaluation: A Review of Evaluation Models. *American Journal of Evaluation*, 0(0) OnlineFirst. <http://dx.doi.org/10.1177/10982140211047219>
- Cahapay, M. (2021). Kirkpatrick Model: Its Limitations as Used in Higher Education Evaluation. *International Journal of Assessment Tools in Education*, 8(1):135-144. <https://doi.org/10.21449/ijate.856143>
- Casasempere, A. & Vercher, M. (2020). Análisis Documental Bibliográfico. Obteniendo el Máximo Rendimiento a la Revisión de la Literatura en Investigaciones Cualitativas. *New Trends in Qualitative Research*, 4:247-257. <https://doi.org/10.36367/ntqr.4.2020.247-257>
- Delgado, S. (2019). Perspectivas en torno a la Formación Docente y la Posibilidad de una Capacitación y Actualización Constante: Una Mirada desde los Actores en una Universidad Mexicana. *Panorama*, 13(24). <https://journal.poligran.edu.co/index.php/panorama/article/view/1204/1022>
- Gandomkar, R. (2018). Comparing Kirkpatrick's original and new model with CIPP evaluation model. *Journal of Advances in Medical Education & Professionalism*, 2(2),94-95. <https://dx.doi.org/10.30476/jamp.2018.41014>
- Gómez, L. & Valdés, M. (2019). The Evaluation of Teacher Performance in Higher Education. *Propósitos y Representaciones*, 7(2):479-515. <https://doi.org/10.20511/pyr2019.v7n2.255>
- Hamza, D.; Ross, S. & Oandasan, I. (2020). Process and Outcome Evaluation of a CBME Intervention Guided by Program Theory. *Journal of Evaluation in Clinical Practice*, 26(4):1096-1104. <https://doi.org/10.1111/jep.13344>
- Hargraves, M. & Denning, C. (2018). Visualizing Expertise: Collaborative Pathway Modeling as a Methodology for Conveying Community-driven Strategies for Change. *Journal of Agriculture, Food Systems, and Community Development*, 8(Suppl. 1):101-115. <https://doi.org/10.5304/jafscd.2018.08A.005>
- Kitivo, E. & Kavulya, J. (2021). Evaluation of Training Programmes. *International Journal of Educational Theory and Practice*, 4(1-4):1-16. <https://www.finessejournals.com/index.php/IJETP/article/view/28/26>
- McChesney, K. & Aldridge, J. (2018). A review of practitioner-led evaluation of teacher professional development. *Professional Development in Education*, 45(2). <https://doi.org/10.1080/19415257.2018.1452782>
- Nouraey, P.; Al-Badi, A.; Riasati, M. & Maata, R. (2020). Educational Program and Curriculum Evaluation Models: A Mini Systematic Review of the Recent Trends. *Universal Journal of Educational Research*, 8(9): 4048-4055. <http://dx.doi.org/10.13189/ujer.2020.080930>
- Parra, R. & Ruiz, C. (2020). Evaluación de Impacto de los Programas Formativos: Aspectos Fundamentales, Modelos y Perspectivas Actuales. *Revista Educación*, 44(2). <https://doi.org/10.15517/revedu.v44i2.40281>
- Pérez Juste, R. (2016). ¿Quo vadis, evaluación? Reflexiones pedagógicas en torno a un tema tan manido como relevante. *Revista de Investigación Educativa*, 34(1):13-30. <http://dx.doi.org/10.6018/rie.34.1.239381>

- Shaw, I. (2017). Lee Cronbach (1916–2001) – A critical appreciation. *Qualitative Social Work*, 17(1): 152-163. <https://doi.org/10.1177%2F1473325017744496>
- Szijarto, B. & Cousins, B. (2019). Mapping the Practice of Developmental Evaluation: Insights from a Concept Mapping Study. *Evaluation and Program Planning*, 76. <https://doi.org/10.1016/j.evalprogplan.2019.05.001>
- Tejada, J. & Navío, A. (2019). Valoración de la Adquisición de Competencias Profesionales en el Prácticum a través del Contrato de Aprendizaje por Parte de los Alumnos: Caso del Grado de Pedagogía. *Revista Iberoamericana de Evaluación Educativa*, 12(2):67-87. <https://doi.org/10.15366/riee2019.12.2.004>
- UNICEF (2019). *Cote D'Ivoire: Plan de Renforcement des Capacites des Points Focaux Membres du Conseil National pour la Nutrition dans le Cadre de la Mise en Oeuvre de la Plateforme Nationale Multisectorielle D'Information sur la Nutrition en Cote D'Ivoire*. Fondo de las Naciones Unidas para la Infancia. http://www.pnmin.gouv.ci/fichier/doc/PLAN%20DE%20FORMATION%20DES%20POINTS%20FOCAUX%20MEMBRE%20DU%20PNMIN_12062019_1.pdf