

## Evolución de la Evaluación. Revisión bibliográfica

**Vicente Jaime Hernández y León**  
Universidad Autónoma de Zacatecas  
[vicentejhl@hotmail.com](mailto:vicentejhl@hotmail.com)

**Martín Jorge Montes Nava**  
Universidad Autónoma de Zacatecas  
[montesnavamartin@hotmail.com](mailto:montesnavamartin@hotmail.com)

**Georgina del Pilar Delijorge González**  
Universidad Autónoma de Zacatecas  
[georgina.delijorge@uaz.edu.mx](mailto:georgina.delijorge@uaz.edu.mx)

### Resumen

Realizar una revisión bibliográfica de tipo descriptivo, es decir, proporcionar al lector una puesta al día sobre los conceptos, concepciones, teorías, objetivos, finalidades y procedimientos que se han esgrimido en la constante evolución de la evaluación del aprendizaje en concreto.

La estrategia de búsqueda bibliográfica y criterios de selección de los artículos, así como el análisis de la variabilidad, fiabilidad y validez de los mismos se fundamentó en un criterio principalmente cronológico que permitió destacar a los principales teórico del tema.

Para la localización de los documentos bibliográficos se utilizaron varias fuentes documentales y en el Internet. Una vez seleccionada las bases de datos, se eligieron los descriptores o palabras clave.

Desde Tyler. (Padre de la evaluación educativa). Da una visión metódica de la misma, superando desde el conductismo la mera evaluación psicológica.

Expone su idea de “currículum”: ¿Qué objetivos se desean conseguir? ¿Con qué actividades se pueden alcanzar? ¿Cómo pueden organizarse eficazmente estas experiencias? ¿Cómo se puede comprobar si se alcanzan los objetivos?

En esta época aparecen instituciones y servicios educativos de todo tipo, se avanza en la tecnología de la medición y en los principios estadísticos del diseño experimental; aparecen las famosas taxonomías de los objetivos educativos de Bloom.

Múltiples perspectivas, múltiples medidas de resultados, y métodos tanto cuantitativos como cualitativos para recoger y analizar la información se han dado hasta la fecha con la evaluación por competencias.

**Palabras clave** Evaluación, Evaluación sumativa, Calificación.

---

## Introducción

Considero que una aproximación histórica puede dar claridad al desarrollo de la conceptualización y para ello me permito seguir a Miriam González Pérez, investigadora cubana que hace un recorrido puntual al respecto (González Pérez, 2000) que asevera que

la evaluación es inherente a toda actividad humana ya sea desde el punto de vista social o individual y aplicada a la enseñanza como *“una actividad intencional y organizada a los efectos de producir determinados aprendizajes relevantes para el individuo”* tiene particularidades la problemática propia de la educación, de las instituciones, el proceso enseñanza aprendizaje.

### **Orígenes**

La autora refiere que al revisar los orígenes históricos de la evaluación muchas de las funciones que hoy día se le atribuyen no están asociadas con su inicio, ya que el mismo está más bien relacionado con el examen y que esta práctica ha ocupado el espacio de la evaluación.

Se cita en el informe mencionado que autores como (Durkheim, 1938; V.A. Judges, 1971; Díaz Barriga, 1993) señalan que el examen surge de *la burocracia china*, por la dinastía Han, que en su afán de seleccionar a los funcionarios de su vasto territorio, *aplicaron instrumentos para escoger a sus mandarines para que administraran los recursos para bien de las Dinastías, en el año 2375 antes de nuestra era. Estos exámenes de competencia permanecieron por un largo tiempo, hasta su desaparición en 1905. Lo que constituyó una innovación frente a las formas de organización de las sociedades arcaicas, dónde el poder y el privilegio eran prerrogativas hereditarias. Estudio de Ebel Robert (como se cita en Díaz Barriga, 1993).*

En la antigüedad los Griegos establecieron competencias en el arte de la persuasión y la elegancia en el estilo era muy apreciada, aunque no se llegaron a establecer listas de aprobados ni clasificación alguna, según Judges, V.A. citado en la compilación de *“El Examen”* (Díaz Barriga, 1993) esta situación fue transferida a la Roma imperial. Igualmente se tienen algunos antecedentes aunque no muy precisos de la Universidad de Constantinopla y en las grandes escuelas de Alejandría, Nicea y Tesalónica.

### **Precedentes: Antes de los tests y de la medición**

De la antigüedad se da un gran salto hasta la Edad Media en la información sobre la evaluación dentro de la enseñanza, probablemente debido a que la educación estaba financiada regularmente por las autoridades imperiales.

Los exámenes como tales al igual que los grados aparecieron en la Edad Media, en la universidad medieval como producto de la organización de cuerpos académicos en instituciones que debían seleccionar a sus miembros en el proceso de admisión. El mismo *Durkheim* (citado por González Pérez, 2000) *“subraya que el examen, en el escenario educativo de la universidad medieval, se realiza con la finalidad de mostrar la competencia adquirida por el alumno y no como instrumento de certificación o de promoción; sólo accedían al examen aquellos alumnos que estuviesen en condiciones de exhibir el aprendizaje logrado de modo satisfactorio.”*

Es decir desde ese tiempo el sistema de enseñanza se preocupa por cumplir con su función tradicional de conservar y transmitir la cultura heredada del pasado y legitimar una cultura dominante que beneficia a los sectores de mayor poder en la sociedad. (Bourdieu y Passeron, 1977 citados por González Pérez).

Extendiendo la reseña de Miriam González la investigadora cubana y con Judges citado por Díaz Barriga, que nos apuntan:

“Hacia el siglo XII se identifican lo que se plantean como las primeras prácticas universitarias que tienen cierto parecido con las condiciones de los exámenes modernos; concomitantes con el resurgimiento del trivium<sup>1</sup> y el

---

<sup>1</sup> Gramática, retórica y dialéctica (o lógica). En las universidades medievales, el trivium comprendía los tres aspectos básicos del conocimiento: La gramática como mecánica del lenguaje; La lógica (o dialéctica — lógica y dialéctica llegaron a ser sinónimos en el tiempo) que es la "mecánica" del pensamiento y el análisis; y la retórica que es el uso del lenguaje para instruir y persuadir. En la

quadrivium (las siete ciencias liberales). Se trataba de ejercicios profesionales y de erudición para discriminar entre sus estudiantes en cuanto a grados de competencia para la promoción jerárquica y el otorgamiento de títulos y grados. Consistían en los estudios preliminares de textos, resúmenes o extractos de pasajes literarios, seguida por una controversia con los compañeros y maestros, empleando la lógica y argumentaciones, a lo que seguía el estudio de los materiales; la combinación de ambos complementaba el fundamento cultural, para proseguir estudios más elevados y aplicados, que más tarde condujeron a los niveles de master y doctorados.

En occidente se establecieron las primeras prácticas para proporcionar grados de competencia en Bolonia, en París y más tarde en Oxford, mismas que tenían gran similitud con las prácticas chinas (basadas en los escritos y citas de Confucio) ya que se debía poseer el conocimiento de textos, extractos de colecciones, de pasajes literarios, etc. Esto se vio reforzado por el aumento en la demanda en el acceso a la educación y de comprobar los meritos individuales y de las instituciones, lo que originó la introducción de los exámenes escritos.

En este acercamiento a la historia de la evaluación no se puede dejar de mencionar a la escuela jesuita que establecía que “si las materias son susceptibles de ser enseñadas también lo son de ser examinadas, primero por los tests de controversia y tests orales y después por medio de ejercicios escritos” por lo que se inició con la implementación de los exámenes escritos, apartándose de lo usual en esa época y ha establecer “el más alto grado de efectividad a la que se hubiera llegado hasta entonces o desde entonces”(Brubacher, citado por Judges, 1971), incluso publicaron un texto(1586) sobre educación dónde incluyeron un conjunto de reglas para la organización de los exámenes.

---

Escolástica medieval se consideran campos preparatorios para el quadrivium, el cual comprende la aritmética, la geometría, la música y la astronomía

De ello se pueden distinguir dos cosas en la iniciación de los métodos modernos de exámenes: Primero, el curso según el Ratio Studiorum (Plan de enseñanza), era amplio dado que abarcaba a la filosofía escolástica y un aprendizaje teológico superior, lo cual daba origen a lo segundo, para poder obtener el grado se debía recitar y memorizar el trabajo realizado previamente es decir se puso énfasis en la medición del rendimiento. En estas escuelas dice Judges (Díaz Barriga, 1993) los jesuitas fueron admirados y temidos en toda Europa y la “clasificación, esencia del método, se transformó en un refinado arte.

Afirma el historiador inglés que se puede constatar que aun en este tiempo se tienen reminiscencias de esas prácticas al mantener los grados de “bachelor” y “master” en el otorgamiento de grados.

Más tarde en los siglos XIII al XIX con el desarrollo de los planes de estudio y la delimitación de la enseñanza secundaria y la universitaria se requirió de filtrar la entrada a la universidad mediante la realización de exámenes finales de aquel nivel de enseñanza y a la entrada a la universidad.

De tal suerte el examen desempeña un doble papel, el de servir de filtro para el acceso a un nivel superior, y el de certificar un nivel general alcanzado. Surge en Alemania durante el periodo de Federico el Grande, el abitur, examen final de la escuela clásica, de elevadas normas y prestigio para avalar la entrada a la universidad y el nivel logrado en la escuela. Similar práctica de examen se desarrolla en Gran Bretaña y más tarde en Francia. Para dar seguimiento esquemático al desarrollo de la evaluación ver lámina no. 1 sobre la Evolución histórica de la evaluación.

En Estados Unidos un funcionario de educación Horace Mann en 1845, observó la manera en que se practican los exámenes a los niños y propuso que se cambiaran a formato escrito ya que los exámenes orales tendían a degradarse en ceremonias rutinarias y en cambio los escritos podían tener algunas ventajas como: lograr más elementos de juicio

sobre el rendimiento de cada alumno; se dispone de una constancia por escrito; se da igual trato a todos los alumnos, al preguntarse lo mismo y hay menos posibilidades de favoritismo. Estudio de Ebel, Robert (citado por Díaz Barriga, 1993)

A finales del s.XIX el interés por la medición científica de las conductas humanas se acentúa debido a diversos factores que condicionan la actividad evaluativa, como la aparición de las teorías evolucionistas, el desarrollo de los métodos estadísticos y en general del desarrollo de la sociedad industrial.

A continuación me referiré a un extenso recorrido que se hace desde la antigüedad hasta nuestros días presentado en la revista RELIEVE (Escudero Escorza, 2003).

Desde la antigüedad se han venido creando y usando procedimientos instructivos en los que los profesores utilizaban referentes implícitos, sin una teoría explícita de evaluación, para valorar y, sobre todo, diferenciar y seleccionar a estudiantes. Dubois (1970) y Coffman (1971) citan los procedimientos que se empleaban en la China imperial, hace más de tres mil años, para seleccionar a los altos funcionarios. Otros autores como Sundbery (1977) hablan de pasajes evaluadores en la Biblia, mientras Blanco (1994) se refiere a los exámenes de los profesores griegos y romanos. Pero según McReynold (1975), el tratado más importante de evaluación de la antigüedad es el Tetrabiblos, que se atribuye a Ptolomeo. También Cicerón y San Agustín introducen en sus escritos conceptos y planteamientos evaluadores.

En la Edad Media se introducen los exámenes en los medios universitarios con carácter más formal. Hay que recordar los famosos exámenes orales públicos en presencia de un tribunal, aunque sólo llegaban a los mismos los que contaban con el visto bueno de sus profesores, con lo que la posibilidad de fracaso era prácticamente inexistente. En el Renacimiento se siguen utilizando procedimientos selectivos y Huarte de San Juan, en su

Examen de ingenios para las ciencias, defiende la observación como procedimiento básico de la evaluación (Rodríguez y otros, 1995).

En el siglo XVIII, a medida que aumenta la demanda y el acceso a la educación, se acentúa la necesidad de comprobación de los méritos individuales y las instituciones educativas van elaborando e introduciendo normas sobre la utilización de exámenes escritos (Gil, 1992).

Entrado el siglo XIX se establecen los sistemas nacionales de educación y aparecen los diplomas de graduación, tras la superación de exámenes (exámenes del Estado). Según Max Weber (Barbier, 1993), surge un sistema de exámenes de comprobación de una preparación específica, para satisfacer las necesidades de una nueva sociedad jerárquica y burocratizada. En los Estados Unidos, en 1845, Horace Mann comienza a utilizar las primeras técnicas evaluativas del tipo «tests» escritos, que se extienden a las escuelas de Boston, y que inician el camino hacia referentes más objetivos y explícitos con relación a determinadas destrezas lecto-escritoras. Sin embargo, no se trata todavía de una evaluación sustentada en un enfoque teórico, sino más bien, algo que responde a prácticas en buena medida rutinaria y con frecuencia basada en instrumentos poco fiables.

Al final del siglo XIX, en 1897, aparece un trabajo de J. M. Rice, que se suele señalar como la primera investigación evaluativa en educación (Mateo y otros, 1993). Se trataba de un análisis comparativo en escuelas americanas sobre el valor de la instrucción en el estudio de la ortografía, utilizando como criterio las puntuaciones obtenidas en los tests.

Con el auge de las teorías positivistas en la Psicología a finales del s. XIX, la utilización de tests se intensificó en la medición del aprovechamiento de los alumnos, con énfasis en los resultados finales del proceso y en el dominio de los contenidos.

González Pérez (González Pérez, 2000) extiende la narración y señala:

Las ideas psicológicas conductistas enriquecieron y fortalecieron la evaluación, tanto teórica como instrumentalmente, pero sin representar variaciones significativas en su esencia. A la luz de dichas ideas surgió y se enraizó la denominada Pedagogía por objetivos, de fuerte repercusión en la práctica educativa y cuyo impacto se extiende hasta nuestros días.

Bajo esta concepción, los objetivos y la evaluación se constituyeron en componentes priorizados del proceso de enseñanza aprendizaje, relegando en cierto sentido a los demás elementos, incluso al contenido que fuese aspecto de consideración central hasta entonces. Valorar el cumplimiento de los fines, del modo más preciso, detallado y objetivo posible, se erigió en un principio ineludible y en motor impulsor para la aplicación de vías y procedimientos “técnicos” que posibilitaron tales exigencias.

En su ayuda aparecen los enfoques importados de la práctica productiva relativos a la planificación y control de los procesos industriales y otros sectores laborales, que dio lugar al planeamiento detallado del proceso de enseñanza, siempre en función de asegurar el logro de los objetivos previamente concebidos y especificados. La evaluación continuó en función de medir los resultados, parciales o finales, y cuantificar el aprendizaje, aunque con un avance no despreciable: la valoración referida a criterios. Estos criterios los aporta el modelo de los objetivos, es decir, el patrón de referencia que antes se sustentaba en la distribución de resultados según la curva normal, deja de ser lo esperado, lo deseable.

En los comienzos del s. XX (1903) se publicó el primer libro de texto sobre medición educacional por Edgard L. Thorndike, quién creía en el valor del experimento y la medición y apuntaba que *“Todo lo que existe, existe en cierta cantidad y conocerlo exhaustivamente implica conocer su cantidad lo mismo que su calidad”* (citado por Ebel en Díaz Barriga,

1993) quién expresa que ningún otro merece la designación de “padre de las modernas mediciones educacionales”

Igualmente importante es mencionar a Alfred Binet quien en 1904 a pedido del Ministerio de Educación francés comenzó a *“desarrollar una medida válida, objetiva y práctica de educabilidad, apta para identificar entre los que no aprendían, a aquellos cuyas capacidades mentales eran demasiado limitadas para permitir el progreso normal del aprendizaje, esta escala ha sido utilizada y perfeccionada en posteriores mediciones, incluyendo el concepto de edad mental.”*

En el periodo de expansión y optimismo que Stufflebeam y Shinkfield (1987) no han dudado en calificar de “irresponsabilidad social”, por el gran despilfarro consumista tras una época de recesión. Se trata de la etapa conocida como la de la inocencia (Madaus y otros, 1991). Se extienden mucho las instituciones y servicios educativos de todo tipo, se producen cantidad de tests estandarizados, se avanza en la tecnología de la medición y en los principios estadísticos del diseño experimental (Gulliksen, 1950; Lindquist, 1953; Walberg y Haertel, 1990) y aparecen las famosas taxonomías de los objetivos educativos (Bloom y otros, 1956; Krath-wohl y otros, 1964).

a) *Diferentes conceptos de evaluación.* Por una parte existe la clásica definición dada por Tyler: *“la evaluación como el proceso de determinar el grado de congruencia entre las realizaciones y los objetivos previamente establecidos, a la que corresponden los modelos orientados hacia la consecución de metas”.* Contrasta esta definición con aquella más amplia que se propugna desde los modelos orientados a la toma de decisiones: *“la evaluación como el proceso de determinar, obtener y proporcionar*

*información relevante para juzgar decisiones alternativas, defendida por Alkin (1969), Stufflebeam y otros (1971), MacDonald (1976) y Cronbach (1982)”.*

Además, el concepto de evaluación de Scriven (1967), *como el proceso de estimar el valor o el mérito de algo*, es retomado por Cronbach (1982), Guba y Lincoln (1982), y House (1989), con objeto de señalar las diferencias que comportarían los juicios valorativos en caso de estimar el mérito o el valor.

b) *Diferentes criterios.* De las definiciones apuntadas anteriormente se desprende que el criterio a utilizar para la valoración de la información también cambia. Desde la óptica de la consecución de metas, una buena y operativa definición de los objetivos constituye el criterio fundamental. Desde la perspectiva de las decisiones y situados dentro de un contexto político, Stufflebeam y colaboradores, Alkin y MacDonald llegan a sugerir incluso la no valoración de la información por parte del evaluador, siendo el que toma las decisiones el responsable de su valoración. (ver Tabla XI).

Las definiciones de evaluación que acentúan la determinación del «mérito» como objetivo de la evaluación, utilizan criterios estándares sobre los que los expertos o profesionales están de acuerdo. Se trata de modelos relacionados con la acreditación y el enjuiciamiento profesional (Popham, 1980).

Los autores (Stake, 1975; Parlett y Hamilton, 1977; Guba y Lincoln, 1982; House, 1983) que acentúan *el proceso de evaluación al servicio de determinar el –valor- más que el –mérito- de la entidad u objeto evaluado, abogan por que el criterio de valoración fundamental sean las necesidades contextuales en las que ésta se inserta.* Así, Guba y Lincoln (1982) refieren los términos de la comparación valorativa; de un lado, las características del objeto evaluado y, de otro, las necesidades, expectativas y valores del grupo a los que les afecta o con los que se relaciona el objeto evaluado.

c) Pluralidad de procesos evaluativos dependiendo de la percepción teórica que sobre la evaluación se mantenga. Los modelos de evaluación citados y otros más que pueden encontrarse en la bibliografía, representan diferentes propuestas para conducir una evaluación.

d) Pluralidad de objetos de evaluación. Como dice Nevo (1983 y 1989), existen dos conclusiones importantes que se obtienen de la revisión de la bibliografía sobre la evaluación. Por un lado, *cualquier cosa puede ser objeto de evaluación y ésta no debería limitarse a estudiantes y profesores y, por otro, una clara identificación del objeto de evaluación es una importante parte en cualquier diseño de evaluación.*

e) Apertura, reconocida en general por todos los autores, de la información necesaria en un proceso evaluativo para dar cabida no sólo a los resultados pretendidos, sino a los efectos posibles de un programa educativo, sea pretendido o no. Incluso Scriven (1973 y 1974) propone una evaluación en la que no se tenga en cuenta los objetivos pretendidos, sino valorar todos los efectos posibles. Apertura también respecto a la recogida de información no sólo del producto final, sino también sobre el proceso educativo. Y apertura en la consideración de diferentes resultados de corto y largo alcance. Por último, apertura también en considerar no sólo resultados de tipo cognitivo, sino también afectivos (Anderson y Ball, 1983).

f) Pluralidad también reconocida de las funciones de la evaluación en el ámbito educativo, recogiendo la propuesta de Scriven entre evaluación formativa y sumativa, y añadiéndose otras de tipo socio-político y administrativas (Nevo, 1983).

g) Diferencias en el papel jugado por el evaluador, lo que ha venido a llamarse evaluación interna vs. evaluación externa. No obstante, una relación directa entre el

evaluador y las diferentes audiencias de la evaluación es reconocida por la mayoría de los autores (Nevo, 1983; Weiss, 1983; Rutman, 1984).

h) Pluralidad de audiencia de la evaluación y, por consiguiente, pluralidad en los informes de evaluación. Desde informes narrativos, informales, hasta informes muy estructurados (Anderson y Ball, 1983).

i) Pluralidad metodológica. Las cuestiones metodológicas surgen desde la dimensión de la evaluación como investigación evaluativa, que viene definida en gran medida por la diversidad metodológica.

### **El nuevo impulso alrededor de Stufflebeam**

Para terminar este recorrido analítico-histórico desde los primeros intentos de medición educativa hasta la actual investigación evaluativa en educación, queremos recoger las recomendaciones que más recientemente nos viene ofreciendo una de las figuras señeras de este campo en la segunda mitad del siglo XX.

Me refiero a Daniel L. Stufflebeam, proponente del modelo CIPP (el más utilizado) a finales de los sesenta, desde 1975 a 1988 presidente del «Joint Committee on Standards for Educational Evaluation» y actual director del «Evaluation Center» de la Western Michigan University (sede del Joint Committee) y del CREATE (Center for Research on Educational Accountability and Teacher Evaluation), centro auspiciado y financiado por el Departamento de Educación del gobierno americano.

Recogiendo estas recomendaciones (Stufflebeam, 1994, 1998, 1999, 2000 y 2001), en las que se han ido integrando ideas de diversos evaluadores también notables, no sólo ofrecemos una de las últimas aportaciones a la actual concepción de la investigación

evaluativa en educación, sino que completamos en buena medida la visión del panorama actual, rico y plural, tras analizar la cuarta generación de Guba y Lincoln.

Stufflebeam *“invoca a la responsabilidad del evaluador, que debe actuar de acuerdo a principios aceptados por la sociedad y a criterios de profesionalidad, emitir juicios sobre la calidad y el valor educativo del objeto evaluado y debe asistir a los implicados en la interpretación y utilización de su información y sus juicios.”* Sin embargo, es también su deber, y su derecho, estar al margen de la lucha y la responsabilidad política por la toma de decisiones y por las decisiones tomadas.

Michael Scriven (Gran Bretaña, 1928) es un académico estadounidense, titulado en matemática y doctorado en filosofía. Realizó contribuciones significativas en los campos de la filosofía, la psicología, el pensamiento crítico, y, sobre todo, la evaluación educativa. Ha producido más de 400 publicaciones eruditas y trabajado como revisor editorial para 42 revistas especializadas. Scriven es ex-presidente de la Educativa Americana de Investigación y la Asociación Americana de Evaluación. También es redactor y cofundador del *Journal of Multidisciplinary Evaluation* (Revista de Evaluación Multidisciplinaria).

Scriven, diferenciará entre evaluación formativa (orientada a la mejora) y sumativa (centrada en el impacto y los resultados del programa). Por otra parte, hay tres formas de establecer la relación entre el desarrollo del programa y la evaluación: Causal: siguiendo el modelo de objetivos. Fortuita: la relación se atribuye al azar o a variables externas. Independencia: Scriven propone que los evaluadores desconozcan los objetivos del programa para centrarse en los efectos reales.

La elección de un modelo evaluativo, supone elegir una teoría epistemológica y de concepción de la realidad. Los modelos sirven para conceptualizar el campo de la evaluación y proporcionar un lenguaje común que permita debatir puntos y temas conflictivos.

## Conclusiones

### **1. Modelo de objetivos y comportamiento (Enfoque de Tyler/Bloom)**

La evaluación se encarga de definir los resultados y los objetivos del programa, de manera que la tarea del evaluadora, consiste en ver hasta qué punto esos objetivos fueron alcanzados.

Este modelo de evaluación utiliza objetivos medibles como parámetro comparativo para lo cual utiliza diseños experimentales y cuasiexperimentales.

Debilidad: la definición de los objetivos a veces es arbitraria. El conocimiento previo de los objetivos puede sesgar el resultado de la evaluación (Scriven).

### **2. Análisis de Sistemas (Rossi y Freeman)**

Este modelo se basa en la microeconomía aplicada al área de servicios humanos. Se basa en el análisis de costo y beneficio.

La evaluación debe ser objetiva de manera tal que cualquier otro grupo de investigadores, arribe a las mismas conclusiones. En este modelo, son sinónimos la objetividad y la fiabilidad.

Los datos son reducidos a unos pocos indicadores para realizar el cálculo costo-beneficio y determinar la eficacia del programa.

Debilidad: La complejidad del programa queda muy reducida e ignora los intereses de los que participan en el programa.

### **3. Toma de decisiones (Stufflebeam/Patton)**

Este modelo, también denominado CIPP (porque considera las variables Contexto, Input, Proceso, Producto), se caracteriza por estar orientado a la toma de decisiones. Así, la evaluación se estructura en función de la decisiones que deben tomarse.

Debilidad: aumenta el riesgo en el uso de resultados.

#### 4. Enfoque libre de objetivos (Scriven)

El evaluador no debe conocer los objetivos para evitar un determinado perfil en el estudio de los resultados. El evaluador debe ser un experto capaz de evitar un perfil determinado debido a sí mismo, como a factores externos para determinar causalidades en condiciones no experimentales.

Debilidad: el evaluador puede sustituir los objetivos del programa por los propios. Scriven lo resuelve centrando el resultado de la evaluación en el producto que el programa ha generado, que es más objetivable, desvinculando la subjetividad del evaluador.

#### Bibliografía

Alonso Sánchez, M., D. Gil Pérez, et al. (1996) *La evaluación y la calificación en una enseñanza constructivista de las ciencias*. Departamento de Didáctica de las Ciencias. Universitat de València. Investigación en la Escuela, 30, 15-26,

Ander-Egg, E. (1995). *Técnicas de Investigación Social*. Buenos Aires, Argentina.

Barberá Gregori, E., Ed. (1999). *Evaluación de la Enseñanza, evaluación del aprendizaje*. Barcelona, España, edebé.

Barraza Barraza, L. (2007). *El Ser y el deber Ser de la participación social*. Durango, Dgo. Sría de Educación del Estado de Durango.

Blanco Gutiérrez, O. (2003). *La práctica de la evaluación en la carrera de Educación Básica integral de la ULA, Táchira*. Departamento de Evaluación y Estadística. Táchira, Universidad de Los Andes-Táchira. Venezuela. ULA-Táchira.: 17-31.

Bolseguí, M. (2004). *Propuesta de un Plan Estratégico de Evaluación*. Colegio Universitario de los Teques "Cecilio Acosta". UPEL-Maracay, Venezuela.

Bustíos Rivera, P. (2002). *Niveles de aprendizaje cognitivo programados y evaluados por los docentes de las Escuelas Académico-Profesionales de Obstetricia de las Universidades del Perú*. Facultad de Educación .Unidad de Postgrado. Lima, Perú., Universidad Nacional Mayor de San Marcos

Casanova, M. A. (1998). *La Evaluación Educativa*. Escuela Básica. Editado por la S.E.P.

Casanova María, A. (1992). *La evaluación, garantía de calidad para el centro educativo*. Zaragoza, España.

Cox, R. (1989). "Educación superior: evaluación de los estudiantes." En Enciclopedia Internacional de la Educación (Trad. del inglés) Vol. 4, pp. 1970-1978. Barcelona

Chadwick, C. B. & Rivera, N. I. (1997). *Concepto e importancia de la evaluación y la evaluación formativa*. México, SEP.

Choppin, B. H. (1989). "Evaluación, valoración y medida.En " Enciclopedia Internacional de la Educación (Trad. del inglés) Vol. 5, pp. 2586-2587. Barcelona

Díaz Barriga, Á., et al. (1993). *El examen: Texto para su historia y debate*. México, UNAM.

Escudero Escorza, T. (2003) Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *Revista ELección de Investigación y EValuación Educativa* v. 9, 11-43.

Española, Real Academia de la Lengua. (2005). *Diccionario de la Real Academia de la Lengua Española*, 22<sup>a</sup> ed. Madrid: Autor.

García García, M. Á.; Fernández Pariente I., (2005). *Una experiencia de evaluación alternativa en la asignatura de "Fundamentos de Ciencia de los Materiales" en la EPS de Ingeniería de Gijón. Oviedo, Gijón, España., Universidad de Oviedo: 105-122.*

González Pérez, M. (2000). "Evaluación del Aprendizaje en la Enseñanza Universitaria." *Revista Cubana de Educación Superior.* Universidad de la Habana. 20(1), 47-67.

Guzmán Arredondo, A.; Alvarado Cabral, J. J. (2008). *Fases y operaciones metodológicas en la investigación educativa.* Centro de Actualización del Magisterio. Durango, Dgo. México.

Hambleton, R. K. (1989). "*Medidas referidas al criterio.*" Ministerio de Educación y Ciencia / Vicens-Vives.: Vol. 6, *Enciclopedia Internacional de la Educación* (Trad. del inglés): pp. 3839-3844. Barcelona.

Hernández, J. ; Pérez, J. (2001). "*Evaluación de la docencia a nivel universitario.*" Memorias del VI Congreso Nacional de Investigación Educativa. Colima, COMIE.

Hernández Nodarse, M. (2004). "*La evaluación del aprendizaje de la Química; una experiencia desarrollada en grupos de alto rendimiento del IPVCE "Lenin".*" Escuela Internacional de Educación Física y Deportes, La Habana, Cuba.

Hernández Sampieri, R., Fernández Collado, C. et al. (2006). *Metodología de la Investigación.* México, Mc. Graw Hill.

La Enciclopedia libre and B. Virtual. (2000). *La Enciclopedia Libre.* La enciclopedia Libre

Larroyo, F. (1982). Diccionario Porrúa de Pedagogía y Ciencias de la Educación. *Diccionario Porrúa de Pedagogía y Ciencias de la Educación*. México, D.F., Editorial Porrúa, S.A: 601.

López Frías, B. S., Hinojosa Kleen, E. M. (2001). *Evaluación del aprendizaje: Alternativas y nuevos desarrollos*. México. ITESM.

Mainer Baqué, J. (2002). "Pensar históricamente el examen, para problematizar su presente. Notas para una sociogénesis del examen." *Revista Jerónimo de Uztáriz de la Universidad Pública de Navarra*, nº 17-18(1).

Moreno Olivos, T. (2007). "La evaluación del aprendizaje en educación superior. El caso de la carrera de Derecho." *Red de revistas científicas de América Latina y el Caribe, España y Portugal*. REDALYC 048: 61-67.

Nevo, David (1983). "La conceptualización de la Evaluación Educativa." *Revista de Investigación en Educación*, 53, 1: 117-128.

Nevo, David (1989). "Papel del evaluador." *En Enciclopedia Internacional de la Educación*. (Trad. del inglés). Ministerio de Educación y Ciencia / Vicens-Vives.Barcelona.: Vol. 5pp. 2587-2589.

Nodarse Hernández, M. (2006). *La evaluación del aprendizaje: ¿estímulo o amenaza?* . La Habana, Cuba, Escuela Internacional de Educación Física y Deportes. *Revista Iberoamericana de Educación*

Pastrana, N. E. (1981). "Clasificación de los modelos de evaluación." UNAM.México.

Pimienta Prieto, J. H. (2007). *Metodología Constructivista*. México, Pearson Educación de México, S.A. de C.V.

Rivera, D., et al. (2006). "Evaluar: tener, poder y valer." *Revista Iberoamericana de Educación*. n.º 40/4(10 de noviembre de 2006).

Román Reyes, A. (2002). *Diccionario Crítico de Ciencias Sociales*. U. Complutense. Madrid, España., Disponible en: <http://www.ucm.es/info/eurotheo/diccionario>.

Rueda Beltrán, M. (2003). "Problemas y retos de la docencia desde el ámbito de las políticas públicas." *Revista de la Educación Superior* .Sitio de ANUIES, Vol. XXXII(3), No. 127.

Saavedra R, Manuel, S. (2001). *Evaluación del Aprendizaje. Conceptos y Técnicas*. México, D.F., ed. PAX.

Sandoval, S. C. (2001). "Hacia una evaluación de la modernización educativa." *Revista Mexicana de Investigación Educativa* 6(13): 401-423.

Santillana (1996). *Currículo y aprendizaje*. Colombia.

Stufflebeam, D. L., Shinkfield, A. J. (1985). *Evaluación Sistemática: Guía teórica y práctica*. España, Ministerio de Educación y Ciencia.

Tyler Ralph, W. (1950). *Principios Básicos del Currículo "Basic Principles of Curriculum and instruction"*. Argentina, Ediciones Troquel