

## **PRINCIPALES CORRIENTES DE EVALUACIÓN COMO POLÍTICAS EDUCATIVAS EN EL SIGLO XX: sus repercusiones en las IEAS de México en el siglo XXI**

**Liberio Victorino Ramírez \***

Para citar este artículo puede utilizar el siguiente formato:

Liberio Victorino Ramírez (2018): "Principales corrientes de evaluación como políticas educativas en el siglo XX: sus repercusiones en las IEAS de México en el siglo XXI", Revista TECSISTECATL (diciembre 2018).  
En línea: <https://www.eumed.net/rev/tecsistecatln23/politicas-educativas.html>  
Handle: <http://hdl.handle.net/20.500.11763/tecsistecatln23politicas-educativas>

### **Resumen**

En el presente artículo, la temática de investigación es el desarrollo de las principales corrientes de evaluación como políticas educativas, con una revisión de su evolución histórica que ha tenido durante el siglo XX; este documento se presenta como resultado de una investigación concluida que fungió como requisito para apoyar a los estudiantes del Doctorado en Ciencias de Educación Agrícola Superior de la Universidad Autónoma Chapingo (UACH) y el Instituto Iberoamericano para el Apoyo de la Agricultura (IICA).

Como objetivo central, este estudio busca dar respuesta a la siguiente interrogante: ¿Cómo se ubican las Instituciones de Educación Agrícola Superior (IEAS) con respecto a los modelos de evaluación institucional en las Universidades e Instituciones de Educación Agrícola Superior?

Palabras Clave: corrientes de evaluación, políticas educativas, educación agrícola superior.

### **Introducción**

En las ciencias de la educación, la aproximación histórica es una vía fundamental para comprender la concepción, estructura, funciones, ámbito, entre otros; este hecho es especialmente evidente en el caso de la evaluación como práctica profesional, ya que ha sufrido profundas transformaciones conceptuales y funcionales a lo largo de la historia y, sobre todo a lo largo del siglo XX en el que principalmente se ubica el desarrollo del presente trabajo.

El estudio se describe mediante la base de tres planteamientos en recientes referencias sobre el tema (Stufflebeam y Shinkfield, 1987; Madaus y otros, 1991). Cabrera (1986); Salvador (1992) y (Guba y Lincoln, 1989), y que se usó indistintamente a través de un proceso matizado, por sucesivas aproximaciones y focalización progresiva del objeto de estudio; complementados por la investigación de Victorino (2006) y la

---

\* Profesor Investigador de la Universidad Autónoma Chapingo. Investigador Nacional por el Sistema Nacional de Investigadores (SIN), Conacyt de México. CE: [victorinoramrezliberio@yahoo.com.mx](mailto:victorinoramrezliberio@yahoo.com.mx)

influencia de éstas corrientes en los cuadernillos de Comités Interinstitucionales de Evaluación en Educación Superior (CIEES) sobre evaluación educativa. Esto es con la finalidad de caracterizar la evaluación interinstitucional aplicada a la educación superior, y acercarse a dibujar el escenario tendencial que se prefigura en las Instituciones de Educación Agrícola Superior (IEAS).

El primer planteamiento, quizás el más utilizado es el contexto, que ofrecen Stufflebeam y Shinkfield, 1987; Madaus y otros, 1991), que en sus trabajos suelen establecer seis épocas, empezando su análisis desde el siglo XIX. Donde se estructuran en: a) época de la reforma (1800-1900), b) época de la eficiencia y del “testing” (1900-1930), c) época de Tyler (1930-1945), d) época de la inocencia (1946-1956), e) época de la expansión (1957-1972) y f) época de la profesionalización (desde 1973) que enlaza con la situación actual.

El segundo planteamiento es de los autores Cabrera (1986) y Salvador (1992) que citan tres grandes épocas, tomando como punto de referencia central la figura de Tyler en el segundo cuarto del Siglo XX. A la época de Tyler se le denomina de nacimiento, a las anteriores de precedentes o antecedentes y a la posterior de desarrollo.

El tercer planteamiento es el de (Guba y Lincoln, 1989), el cual destacan cuatro generaciones, que según ellos se apoya la primera generación en la medición, que llega hasta el primer tercio de este siglo, la segunda es la descripción, la tercera es la del juicio o valoración y la cuarta con enfoque paradigmático constructivista como base para determinar la información que se necesita; la cual consideran los autores que es en la que estamos en este momento.

Por su parte Victorino (2000), nos precisa la última época, en donde se destaca el caso de México y América Latina con su coincidencia en las IEAS de pasar de la política de evaluación centrada, en política indicativa desde los 50's a los 70's a la transición como política prioritaria y exigible como rendición de cuentas en los 80's y 90's.

## METODOLOGÍA

Centrado en técnicas de revisión documental, especialmente obras publicadas en donde se trata de caracterizar los principales rasgos de estas corrientes y autores que tienen que ver con los procesos de evaluación.

En el análisis histórico de estos tres planteamientos e investigaciones, se resume los enfoques evaluativos más relevantes de los distintos modelos que, con mayor o menor fuerza, intentan acotar lo que hoy es la evaluación en la educación, así como mostrar las tendencias que expresan el escenario en las IEAS identificando regularidades y heterogeneidades, con la pertinencia de llevar a cabo procesos de evaluación a través del Análisis de Redes Sociales (ARS) con el cual me identifiqué como enfoque a seguir; y centrar su profundización en las brechas entre instituciones y la cada vez más acentuada desigualdad en su interior.

Como objetivo central, este estudio busca dar respuesta a la siguiente interrogante: ¿Cómo se ubican las IEAS con respecto a los modelos de evaluación institucional?

## DESARROLLO DE LAS ÉPOCAS Y EL TIPO DE EVALUACIÓN PREDOMINANTE

Época Precedente: de la Reforma (1800-1900), la eficiencia y del “testing” (1900-1930),

1ª Generación de la medida (técnica)

Desde la antigüedad se han venido creando y usando procedimientos e instructivos en los que los profesores utilizaban referentes implícitos, la evaluación no era desconocida, pero tampoco era un

movimiento reconocible ya que no tenía una teoría explícita para valorar y, sobre todo, diferenciar y seleccionar a estudiantes.

Los procedimientos que se empleaban en la China imperial hace más de tres mil años, para seleccionar a los altos funcionarios. Los pasajes evaluadores en la Biblia, que se refiere a exámenes de profesores griegos y romanos; como Sócrates que utilizaron cuestionarios evaluativos como parte de su metodología didáctica. También Cicerón y San Agustín introducen en sus escritos conceptos y planteamientos evaluadores. El tratado más importante de evaluación de la antigüedad es el Tetrabiblos, que se atribuye a Ptolomeo.

En la Edad Media en Inglaterra, se utilizaron comisiones reales para evaluar los servicios públicos y se introducen los exámenes en los medios universitarios con carácter más formal. Hay que recordar los exámenes orales públicos en presencia del 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.

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.

Entrado el siglo XIX se establecen los sistemas nacionales de educación y aparecen los diplomas de graduación, tras la superación de exámenes del Estado. Según Max Weber, 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 dirigió una evaluación, basada en tests de rendimiento, donde se comienza a utilizar las primeras técnicas evaluativas del tipo “tests” escritos, que se extienden a las escuelas de Boston, para conocer bien a sus estudiantes e iniciar 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 suele señalarse como la primera investigación evaluativa en educación (Mateo y otros, 1993). Realizando el estudio de conocimientos en ortografía a estudiantes de un amplio sector escolar y sacó la conclusión de que la enseñanza de la ortografía, que entonces estaba en boga, no había producido avances sustanciales en el aprendizaje. Este estudio está generalmente reconocido como la primera evaluación formal sobre un programa educativo realizado en América. 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.

A finales del siglo XIX, se despierta un gran interés por la medición científica de las conductas humanas. Esto es algo que se enmarca en el movimiento renovador de la metodología de las ciencias humanas, al asumir el positivismo de las ciencias físico-naturales. En este sentido, la evaluación recibe las mismas influencias que otras disciplinas pedagógicas relacionadas con procesos de medición, como la pedagogía experimental y la diferencial (Cabrera, 1986).

La actividad evaluativa se vio condicionada de forma decisiva por diversos factores, tales como:

- El florecimiento de las corrientes filosóficas positivistas y empíricas, que apoyaban a la observación, la experimentación, los datos y los hechos como fuentes del conocimiento verdadero. Aparece la exigencia del rigor científico y de la objetividad en la medida de la conducta humana y se potencian las pruebas escritas como medio para combatir la subjetividad de los exámenes orales.
- La influencia de las teorías evolucionistas y los trabajos de Darwin, Galton y Cattell, apoyando la

- medición de las características de los individuos y las diferencias entre ellos.
- El desarrollo de los métodos estadísticos que favorecía decisivamente la orientación métrica de la época.
- El desarrollo de la sociedad industrial que potenciaba la necesidad de encontrar unos mecanismos de acreditación y selección de alumnos, según sus conocimientos.

Consecuentemente con este estado de cosas, en este periodo entre finales del siglo XIX y principios del XX, se desarrolla una actividad evaluativa intensa conocida como “testing”, que se define por caracterizarse en:

- Medición y evaluación resultaban términos intercambiables. En la práctica sólo se hablaba de medición.
- El objetivo era detectar y establecer diferencias individuales, dentro del modelo del rasgo y atributo, para determinar la posición relativa del sujeto dentro de la norma grupal.
- Los tests de rendimiento, sinónimo de evaluación educativa, se elaboraban para establecer discriminaciones individuales, olvidándose en gran medida la representatividad y congruencia con los objetivos educativos. La evaluación y la medida tenían poca relación con los programas escolares. Los tests informaban algo sobre los alumnos pero no de los programas con los que se les había formado.

En el campo educativo destacan algunos instrumentos de aquella época, como las escalas de escritura, redacción, ortografía, cálculo, lectura y de aritmética.

Sin embargo, fue en los tests psicológicos donde los esfuerzos tuvieron mayor impacto en los comienzos del siglo XX. En Francia destacan trabajos sobre tests de capacidades cognitivas.

Años más tarde, con las necesidades de reclutamiento en la Primera Guerra Mundial, se construyen tests colectivos de inteligencia general e inventarios de personalidad.

Tras la contienda, los tests psicológicos se ponen al servicio de fines sociales. La década entre 1920 y 1930 marca el punto más alto del “testing”, pues se idean multitud de tests estandarizados para medir toda clase de destrezas escolares con referentes objetivos externos y explícitos, basados en procedimientos de medida de la inteligencia, para utilizar con grandes colectivos de estudiantes.

Estas aplicaciones estandarizadas se acogen muy bien en los ámbitos educativos en que los profesores construyan sus propias pruebas objetivas, para no tener que confiar exclusivamente en las propuestas por especialistas externos.

Este movimiento estuvo vigente en paralelo al proceso de perfeccionamiento de los tests psicológicos con el desarrollo de la estadística y del análisis factorial. El fervor por el “testing” decreció a partir de los años cuarenta e, incluso, empezaron a surgir algunos movimientos hipercríticos con estas prácticas.

Guba y Lincoln (1989) se refieren a esta evaluación como a la primera generación, denominada como la generación de la medida. El papel del evaluador era técnico, como proveedor de instrumentos de medición. Según estos autores, esta primera generación permanece todavía viva, pues existen textos y publicaciones que utilizan de manera indisoluble evaluación y medida.

Antes de Ralph W. Tyler, en Francia se inicia en los años veinte una corriente independiente conocida como docimología, que supone un primer acercamiento a la verdadera evaluación educativa. Se criticaba, sobre

todo, el divorcio entre lo enseñado y las metas de la instrucción. La evaluación se dejaba, en último término, en manos de una interpretación totalmente personal del profesor. Como solución se proponía: a) elaboración de taxonomías para formular objetivos, b) diversificación de fuentes de información, exámenes, expedientes académicos, técnicas de repesca y tests, c) unificación de criterios de corrección a partir del acuerdo entre los correctores de las pruebas y d) revisión de los juicios de valoración mediante procedimientos tales como la doble corrección, o la media de distintos correctores. Como puede verse, se trata de criterios en buena medida vigentes actualmente y, en algún caso, incluso avanzados.

Época de Nacimiento: de R. Tyler (1930-1945); de la Inocencia (1946-1956)

## 2ª Generación de la evaluación (analista)

En los primeros años de la década de los treinta, Ralph Tyler acuñó el término evaluación educacional y publicó una amplia y renovadora visión del currículo y la evaluación; desarrolló sus intuiciones hasta que conformaron un método que suponía una clara alternativa a otras perspectivas.

La principal característica se centraba en objetivos claramente fijados. De hecho, definía la evaluación como algo que determina si han sido alcanzados ciertos objetivos. Como consecuencia de esta definición, los evaluadores debían ayudar a quienes elaboraran currículos a clarificar el rendimiento de los estudiantes, que había sido precisamente determinado a través de la realización de un currículo. Los objetivos resultantes, referidos al rendimiento, proporcionaban entonces la base para elaborar un currículo y un test de desarrollo. La elaboración del currículo quedaba así influida tanto por los contenidos que debían aprenderse como por la evolución del rendimiento de los estudiantes. La tecnología de los tests de desarrollo iba a servir de base para la elaboración de tests relacionados con los objetivos, y también a los relacionados con las diferencias individuales y con las normas nacionales o del estado.

Durante la década de los treinta, se cayó en las profundidades de la Gran Depresión. Las escuelas, como otras instituciones públicas, se estancaron en una ausencia total de recursos y optimismo, se intentó contribuir a que la educación se convirtiera en un sistema dinámico e innovador, que a la vez se fuera renovando a sí mismo. Con el nombre de Educación Progresiva, este movimiento reflejó la filosofía del pragmatismo y empleó los instrumentos de la psicología conductista.

Tyler (Joint Committee, 1981), fue ser el primero en dar una visión metódica de la evaluación educativa, superando desde el conductismo, la mera evaluación psicológica. Entre 1932 y 1940, dirigió la investigación Eight-Year Study of Secondary Education para la Progressive Education Association, y plantea la necesidad de una evaluación científica que sirva para perfeccionar la calidad de la educación, pensó en examinar la efectividad de ciertos currículos renovadores y estrategias didácticas que se estaban empleando en treinta escuelas a lo largo de toda América. El estudio resulta notable porque ayudó a difundir, ensayar y demostrar su concepción de la evaluación educacional de una manera inmediata. Desde el momento en que la evaluación tyleriana incluía comparaciones internas entre los resultados y los objetivos, ya no necesitaba recurrir a costosas y problemáticas comparaciones entre grupos de control y grupos experimentales.

El método se centra en determinar el grado de éxito, por lo que se opone a métodos indirectos que determinan factores como la calidad de la enseñanza, el número de libros que hay en la biblioteca, los materiales y la participación de la comunidad. Y además, las evaluaciones tylerianas no se ven demasiado afectadas por la fiabilidad de las diferencias entre las puntuaciones individuales de los alumnos ya que generalmente cubren una franja más amplia de variables exógenas que las cubiertas por los tests tipificados.

La obra de síntesis la publica unos años después (Tyler, 1950), exponiendo de manera clara su idea de "currículum", e integrando en él su método sistemático de evaluación educativa, como el proceso surgido

para determinar en qué medida han sido alcanzados los objetivos previamente establecidos.

El “currículum” viene delimitado por las cuatro cuestiones siguientes: ¿Qué objetivos se desean conseguir? ¿Con qué actividades se pueden alcanzar? ¿Cómo pueden organizarse eficazmente estas experiencias? Y ¿Cómo se puede comprobar si se alcanzan los objetivos?

La buena evaluación precisa condiciones, como la propuesta clara de objetivos; Determinación de situaciones en que se deben manifestar las conductas esperadas; Elección de instrumentos apropiados de evaluación; Interpretación de resultados de las pruebas; Determinación de la fiabilidad y objetividad de las medidas.

Esta evaluación ya no es una simple medición, porque supone un juicio de valor sobre la información recogida. Se alude, aunque sin desarrollar a la toma de decisiones sobre los aciertos o fracasos de la programación, en función de los resultados de los alumnos.

Para Tyler, la referencia central en la evaluación son los objetivos preestablecidos, que deben ser cuidadosamente definidos en términos de conducta, teniendo en cuenta que deben marcar el desarrollo individual del alumno, pero dentro de un proceso socializador.

El objeto del proceso evaluativo es determinar el cambio ocurrido en los alumnos, pero su función es más amplia que el hacer explícito este cambio a los propios alumnos, padres y profesores; es también un medio para informar sobre la eficacia del programa educacional y también de educación continua del profesor. Se trata, según Guba y Lincoln (1989), de la segunda generación de la evaluación.

A pesar de lo anterior y de que las reformas tylerianas no siempre se aplicaron de inmediato, las ideas de Tyler fueron muy bien acogidas por los especialistas en desarrollo curricular y por los profesores. Su esquema era racional y se apoyaba en una tecnología clara, fácil de entender y aplicar (Guba y Lincoln, 1982; House, 1989) y encajaba perfectamente en la racionalidad del análisis de la tarea que comenzaba a usarse con éxito en ámbitos educativos militares (Gagné, 1971).

Tras la Segunda Guerra Mundial se produce un periodo de expansión y optimismo que Stufflebeam y Shinkfield (1987) lo califican de “irresponsabilidad social”, por el gran despilfarro consumista tras una época de recesión. Se trata de la etapa conocida como 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 y aparecen las famosas taxonomías de los objetivos educativos. Sin embargo, en esta época, la aportación de la evaluación a la mejora de la enseñanza es escasa debido a la carencia de planes coherentes de acción. Se escribe mucho de evaluación, pero con escasa influencia en el perfeccionamiento de la labor instruccional.

Además de ser una época de plenitud, este período fue un tiempo de pobreza y desesperación en las ciudades del interior y áreas rurales. Fue un período de profundos prejuicios raciales y de segregación; de exorbitado consumo y despilfarro de los recursos naturales, y de un abrumador desarrollo de la industria y capacidad militar, con muy pocas precauciones contra los perjuicios que iban a sufrir el paisaje y las futuras generaciones.

Más que de la evaluación educacional, se trató de una expansión de ofertas educacionales, de personal y de facilidades. Se construyeron nuevos edificios, aparecieron nuevos tipos de instituciones educativas, como los colegios comunitarios. Los distritos escolares se unieron a otros con el fin de proporcionar un amplio abanico de servicios educativos, incluyendo servicios para la salud mental y física, consejos escolares, comedores, enseñanzas musicales, programas deportivos, educación comercial, técnica y educación social. Se incrementaron desmesuradamente las inscripciones en programas educativos para maestros y, en



general, las inscripciones en los colegios aumentaron espectacularmente.

Este panorama general de la sociedad y la educación se reflejó también en la evaluación educacional. Mientras se producía esta gran expansión de la educación, la sociedad no parecía tener un gran interés en formar profesores competentes, en localizar y solucionar las necesidades de los menos privilegiados o en identificar y solucionar los problemas del sistema educativo.

Los educadores escribían sobre la evaluación y sobre la recopilación de los datos más importantes, pero, al parecer, no traducían estos esfuerzos en intentos de mejorar los servicios educativos. Esta falta de objetivos provocó también el retraso de los aspectos técnicos de la evaluación. Hubo un considerable desarrollo de los instrumentos y estrategias aplicables a los distintos métodos evaluativos, como los tests, la “experimentación comparativa” y la “coincidencia entre resultados y objetivos”. Como consecuencia, los educadores dispusieron de nuevos servicios de tests y de nuevas maneras de calificarlos, de algoritmos para designar objetivos de comportamiento, de taxonomías estadísticas para analizar los datos educativos. Pero estas contribuciones no se derivaban de un análisis de la información necesaria para valorar y perfeccionar la educación, ni representaban un desarrollo de la experiencia escolar.

El pensamiento de Tyler se resume en: La necesidad de analizar cuidadosamente los propósitos de la evaluación, antes de ponerse a evaluar; El propósito en la evaluación de los alumnos es guiar su aprendizaje, esto es, ayudarles a que aprendan; El “portafolio” es un instrumento valioso de evaluación, pero depende de su contenido; La evaluación debe ser idiosincrásica, adecuada a las peculiaridades del alumno y el centro. En rigor, la comparación de centros no es posible; Los profesores deben rendir cuentas de su acción educativa ante los padres de los alumnos. Para ello, es necesario interaccionar con ellos de manera más frecuente y más informal.

Medio siglo después de que Tyler revolucionara el mundo de la evaluación educativa, se observa la fortaleza, coherencia y vigencia de su pensamiento; sus ideas básicas, convenientemente actualizadas, se entroncan fácilmente en las corrientes más actuales de la evaluación educativa.

Durante este período, las evaluaciones educacionales dependían, sobre todo, de los límites de los distritos escolares locales. Las escuelas, dependían del interés y la habilidad local para poder hacer o no una evaluación. Las agencias federales y estatales aun no estaban demasiado introducidas en el campo de la evaluación de programas. Los fondos destinados a las evaluaciones provenían de las arcas locales, de fundaciones o de organizaciones profesionales. Esta ausencia de ayudas exteriores y de apoyos para la realización de evaluaciones en todos los niveles de la educación iba a concluir con la llegada del próximo período de la historia de la evaluación.

Época del Desarrollo: de Expansión (1957-1972); de la Profesionalización (1973)

3ª Generación de valoración (juez)

La época de la inocencia en la evaluación tuvo brusco final con el auge en la década de los años 50 y principios de los 60, de las evaluaciones de proyectos de currículos a gran escala, financiadas por estamentos federales. Los educadores se encontraron, durante este período, con que ya no podían hacer o no hacer evaluaciones a su antojo, y que los futuros progresos de las metodologías evaluativas deberían estar relacionadas con conceptos como utilidad y relevancia.

Su violento despertar, en este período, marcó el final de una era de complacencia y ayudó a emprender profundos cambios, guiados por el interés público y dependiendo del dinero de los contribuyentes para su financiamiento, lo cual condujo a que la evaluación se convirtiera en una industria y en una profesión.

Como consecuencia del avance tecnológico se proporcionó nuevos programas educativos en matemáticas, ciencias e idiomas, y creó servicios asesores y programas de encuestas en los distritos escolares. Aparecieron, así, un cierto número de proyectos nacionales para el desarrollo de currículos, en áreas de ciencias y matemáticas. Finalmente, se asignaron fondos para evaluar estos programas.

Los 4 métodos evaluativos expuestos en las evaluaciones que se realizaron en este período fueron: Primero, el método Tyler, usado para definir objetivos de los nuevos currículos y valorar el grado en que los objetivos quedaban alcanzados. Segundo, los nuevos tests de los contenidos de los nuevos currículos. Tercero, el método del criterio profesional usado para valorar las propuestas y verificar periódicamente los resultados. Finalmente, muchos evaluadores se comprometieron a evaluar los resultados del desarrollo del currículo a través de pruebas concretas.

La comunidad evaluativa educacional participó para evaluar los nuevos currículos que estaban adecuadamente financiados y aplicaban cuidadosamente la tecnología que había ido desarrollándose durante la pasada década o incluso antes. A pesar de todo esto, empezaron a darse cuenta de que sus esfuerzos no tenían éxito.

Esta valoración negativa quedó perfectamente reflejada en un histórico artículo de Cronbach (1963). Examinando los esfuerzos evaluativos del pasado reciente, criticó duramente los conceptos en que se habían basado las evaluaciones por su falta de relevancia y utilidad, y aconsejó a los evaluadores, que abandonaran sus inclinaciones de entonces y se dedicaran a hacer evaluaciones basadas en comparaciones de resultados obtenidos de tests tipificados, realizados por grupos experimentales y de control. Cronbach arguía que el análisis y relación de resultados de un test podrían ser, probablemente, de más utilidad para los profesores que informarles sobre el promedio del total de las calificaciones.

Como resultado de la creciente inquietud producida por los esfuerzos realizados en las evaluaciones con resultados siempre negativos, el Phi Delta Kappa creó el National Study Committee on Evaluation (Stufflebeam y otros, 1971). Tras examinar el panorama, este comité concluyó que la evaluación educacional “era víctima de una grave enfermedad” y recomendó el desarrollo de nuevas teorías y métodos de evaluación, así como nuevos programas para preparar a los propios evaluadores. Al mismo tiempo, empezaron a aparecer nuevas conceptualizaciones sobre la evaluación Provus (1971), Hammond (1967), Eisner (1967) y Metfessel y Michael (1967) propusieron una reforma del modelo Tyler. Glaser (1963), Tyler (1967) y Popham (1971) recomendaron los tests basados en criterios como una alternativa a los tests basados en normas. Cook (1966) pidió la utilización del método de análisis sistemático para la evaluación de programas. Scriven (1967), Stufflebeam (1967, 1971) y Stake (1967) crearon nuevos modelos de evaluación que se alejaban radicalmente de los métodos anteriores.

Estas conceptualizaciones, reconocían la necesidad de evaluar las metas, examinar las inversiones, y analizar el perfeccionamiento y la prestación de los servicios, así como determinar los resultados que se desean (o no se desean) obtener del programa. También insistieron en la necesidad de enjuiciar el mérito o el valor del objeto de la evaluación.

En esa época aparece un cierto desencanto con la escuela pública y crece la presión por la rendición de cuentas (MacDonald, 1976; Stenhouse, 1984). En 1958 se promulga la ley de defensa educativa que proporciona muchos programas y medios para evaluarlos. En 1964 se establece el Acta de educación primaria y secundaria (ESEA) y se crea el National Study Committee on Evaluation, creándose una nueva evaluación no sólo de alumnos, sino orientada a incidir en los programas y en la práctica educativa global (Mateo y otros, 1993; Rodríguez y otros, 1995).

Para mejorar la situación y retomar la hegemonía científica y educativa, se destinaron millones de dólares de los fondos públicos a subvencionar nuevos programas educativos e iniciativas del personal de escuelas públicas americanas encaminadas a mejorar la calidad de la enseñanza. (Popham, 1983;



Rutman y Mowbray, 1983; Weiss, 1983). Este movimiento se vio también potenciado por el desarrollo de medios tecnológicos (audiovisuales, ordenadores...) y el de la enseñanza programada, cuyas posibilidades educativas despertaron el interés entre los profesionales de la educación (Rosenthal, 1976).

De la misma forma que la proliferación de programas sociales en la década anterior había impulsado la evaluación de programas en el área social, los años sesenta serán fructíferos en demanda de evaluación en el ámbito de la educación. Esta nueva dinámica en la que entra la evaluación, hace que, aunque ésta se centraba en los alumnos como sujeto que aprende, y el objeto de valoración era el rendimiento de los mismos, sus funciones, su enfoque y su última interpretación variará según el tipo de decisión buscada.

La recesión económica que caracteriza los finales años sesenta, y, sobre todo, la década de los setenta; hizo que la población civil, como contribuyentes, y los propios legisladores se preocupasen por la eficacia y el rendimiento del dinero que se empleaba en la mejora del sistema escolar. A finales de los años sesenta, y como consecuencia de lo anterior, entra en escena un nuevo movimiento, la era de la «Accountability», de la rendición de cuentas (Popham, 1980 y 1983; Rutman y Mowbray, 1983), que se asocia fundamentalmente a la responsabilidad del personal docente en el logro de objetivos educativos establecidos.

De hecho, en 1973, la legislación americana instituyó la obligación de controlar el logro de los objetivos educativos y la adopción de medidas correctivas en caso negativo (MacDonald, 1976; Wilson y otros, 1978). Es comprensible que, planteado así, este movimiento de rendición de cuentas, de responsabilidad escolar, diera lugar a una oleada de protestas por parte del personal docente.

Otra dimensión de la responsabilidad escolar nos la ofrece Popham (1980), cuando se refiere al movimiento de descentralización escolar. Los grandes distritos escolares se dividieron en áreas geográficas más pequeñas, y, por consiguiente, con un control ciudadano más directo sobre lo que ocurría en las escuelas.

Como consecuencia se amplió considerablemente el fenómeno de la evaluación educativa. El sujeto directo de la evaluación siguió siendo el alumno, pero también todos aquellos factores que confluyen en el proceso educativo (el programa educativo en un sentido amplio, profesor, medios, contenidos, experiencias de aprendizaje, organización, etc.), así como el propio producto educativo.

Como resultado de estas nuevas necesidades de la evaluación, se inicia durante esta época un periodo de reflexión y de ensayos teóricos con ánimo de clarificar la multidimensionalidad del proceso evaluativo. Estas reflexiones teóricas enriquecerán decisivamente el ámbito conceptual y metodológico de la evaluación. dará lugar al nacimiento de esa nueva modalidad de investigación aplicada que hoy denominamos como investigación evaluativa.

Como hitos de la época hay que destacar dos ensayos por su decisiva influencia: el artículo de Cronbach (1963), Course improvement through evaluation, y el de Scriven (1967), The methodology of evaluation. Las ideas evaluativas expuestas en estos trabajos, obligan a que se refieran a ellas:

Del análisis que Cronbach del concepto, funciones y metodología de la evaluación:

- a) Asociar el concepto de evaluación a la toma de decisiones b) La evaluación para mejorar un programa mientras éste se está aplicando.
- c) Los estudios evaluativos sean de tipo comparativo con criterios de tipo absoluto, y defender la valoración con relación a objetivos bien definidos y no la comparación con otros grupos.
- d) Se ponen en cuestión estudios a gran escala, puesto que las diferencias entre tratamientos

pueden impedir discernir con claridad las causas de los resultados. Con estudios más analíticos, bien controlados, que pueden usarse para comparar versiones alternativas de un programa.

- e) Metodológicamente la evaluación debe incluir: 1) estudios de proceso; 2) medidas de rendimiento y actitudes y 3) estudios de seguimientos.
- f) Las técnicas de evaluación no se limitan a los tests de rendimiento. Los cuestionarios, las entrevistas, la observación sistemática y no sistemática, las pruebas de ensayo.

Reflexiones del ensayo de Scriven (1967). Sus fecundas distinciones terminológicas ampliaron el campo semántico de la evaluación, a la vez que clarificaron el quehacer evaluativo:

- a) El objetivo de la evaluación es invariante, supone en definitiva el proceso por el cual estimamos el valor de algo que se evalúa, mientras que las funciones de la evaluación pueden ser enormemente variadas. Estas funciones se relacionan con el uso que se hace de la información recogida.
- b) Funciones que puede adoptar la evaluación: la formativa y la sumativa. Propone el término de evaluación formativa para calificar aquel proceso de evaluación al servicio de un programa en desarrollo, con objeto de mejorarlo, y el término de evaluación sumativa para aquel proceso orientado a comprobar la eficacia del programa y tomar decisiones sobre su continuidad.
- c) Énfasis en la consecución de objetivos previamente establecidos, porque si los objetivos carecen de valor, no tiene ningún interés saber hasta qué punto se han conseguido. Resalta la necesidad de que la evaluación debe incluir tanto la evaluación de los propios objetivos como el determinar el grado en que éstos han sido alcanzados.
- d) Formas diferentes de valorar un elemento de la enseñanza. La evaluación intrínseca valora el elemento por sí mismo, mientras que la evaluación extrínseca valora el elemento por los efectos que produce en los alumnos.
- e) Scriven adopta una posición contraria a Cronbach, defendiendo el carácter comparativo que deben presentar los estudios de evaluación. Admite con Cronbach los problemas técnicos que los estudios comparativos entrañan y la dificultad de explicar las diferencias entre programas, pero Scriven considera que la evaluación como opuesta a la mera descripción implica emitir un juicio sobre la superioridad o inferioridad de lo que se evalúa con respecto a sus competidores o alternativas.

Estas dos aportaciones comentadas influyeron decisivamente en la comunidad de evaluadores, incidiendo no sólo en estudios en la línea de la investigación evaluativa, a la que se referían preferentemente, sino también en la evaluación orientada al sujeto, en la línea de evaluación como «assessment» (Mateo, 1986). Estamos ante la tercera generación de la evaluación que, según Guba y Lincoln (1989), se caracteriza por introducir la valoración, el juicio, como un contenido intrínseco en la evaluación. Ahora el evaluador no sólo analiza y describe la realidad, además, la valora, la juzga con relación a distintos criterios.

A pesar de las voces críticas con la operativización de objetivos (Eisner, 1967 y 1969; Atkin, 1968), no sólo por la estructura de valor que en ello subyace, sino también por centrar la valoración del aprendizaje en los productos más fácilmente mensurables, a veces los más bajos en las taxonomías del dominio cognoscitivo, y de que se prestaba escasa atención a los objetivos del dominio afectivo, que presentan mayor dificultad de tratamiento operativo, el modelo evaluativo de Tyler se enriquecería mucho en estos años, con trabajos

sobre los objetivos educativos que continuarían y perfeccionarían el camino emprendido en 1956 por Bloom y colaboradores, 1971.

Stake (1967) propuso su modelo de evaluación, The countenance model, que sigue la línea de Tyler, pero es más completo al considerar las discrepancias entre lo observado y lo esperado en los «antecedentes» y «transacciones», y posibilitar algunas bases para elaborar hipótesis acerca de las causas y los fallos en los resultados finales.

Metfessell y Michael (1967) presentaron también un modelo de evaluación de la efectividad de un programa educativo en el cual, aún siguiendo el modelo básico de Tyler, proponían la utilización de una lista comprensiva de criterios diversos que los evaluadores podrían tener en cuenta en el momento de la valoración y, por consiguiente, no centrarse meramente en los conocimientos intelectuales alcanzados por los alumnos.

Suchman (1967) profundiza en la convicción de que la evaluación debe basarse en datos objetivos que sean analizados con metodología científica, matizando que la investigación científica es preferentemente teórica y, en cambio, la investigación evaluativa es siempre aplicada. Su principal propósito es descubrir la efectividad, éxito o fracaso de un programa al compararlo con los objetivos propuestos y, así, trazar las líneas de su posible redefinición. Esta investigación evaluativa para Suchman debe tener en cuenta: a) la naturaleza del destinatario del objetivo y la del propio objetivo, b) el tiempo necesario para que se realice el cambio propuesto, c) el conocimiento de si los resultados esperados son dispersos o concentrados y d) los métodos que han de emplearse para alcanzar los objetivos. Además, defiende la utilización de evaluadores externos para evitar todo tipo de tergiversación de los profesores muy implicados en los procesos instruccionales.

El énfasis en los objetivos y su medida traerá también la necesidad de una nueva orientación a la evaluación, la denominada evaluación de referencia criterial. La distinción introducida por Glaser (1963) entre mediciones referidas a normas y criterios tendrá eco al final de la década de los sesenta, precisamente como resultado de las nuevas exigencias que a la evaluación educativa se le planteaban. Así, por ejemplo, cuando Hambleton (1985) estudia las diferencias entre tests referidos al criterio y tests referidos a la norma, señala para los primeros, además de los conocidos objetivos de describir la ejecución del sujeto y tomar decisiones sobre si domina o no domina un contenido, otro objetivo como es el de valorar la eficacia de un programa.

Desde finales de los sesenta los especialistas se pronunciarán decisivamente a favor de la evaluación criterial, en cuanto que es el tipo de evaluación que suministra una información real y descriptiva del estatus del sujeto o sujetos respecto a los objetivos de enseñanza previstos, así como la valoración de ese estatus por comparación con un estándar o criterio de realizaciones deseables, siendo irrelevantes, al efecto de contraste, los resultados obtenidos por otros sujetos o grupo de sujetos.

En las prácticas evaluativas de esta década de los sesenta se observan dos niveles de actuación. Uno de evaluación orientada hacia los individuos, fundamentalmente alumnos y profesores. El otro nivel, de evaluación orientada a la toma de decisiones sobre el «instrumento» o «tratamiento» o «programa» educativo. Este último nivel, impulsado por la evaluación de programas en el ámbito social, será la base para la consolidación educativa en evaluación de programas y de la investigación evaluativa.

En 1973, el campo de la evaluación empezó a cristalizar y a emerger como una profesión diferenciada de las demás, relacionada a sus antecedentes de investigación y control, aunque es instructivo considerar su desarrollo en el contexto del período anterior. En aquellos momentos, los evaluadores enfrentaban una crisis de identidad. No sabían cuál era su papel, si debían ser investigadores, administradores de tests, profesores, organizadores ó filósofos. Cualquier calificación especial que se les pudiera dar no resultaría clara. No existen organizaciones profesionales dedicadas a la evaluación entendida como especialidad, ni

tampoco publicaciones especializadas a través de los cuales pudieran intercambiar información acerca de su trabajo. Esencialmente, no existía literatura acerca de la evaluación educacional, excepto trabajos inéditos que circulaban únicamente entre círculos de especialistas. Había pocas oportunidades de prepararse para ejercer un servicio educativo. Las normas claras para una buena práctica se veían confinadas a las encuestas educacionales y psicológicas.

El campo de la evaluación era amorfo y fragmentado. Muchas evaluaciones estaban dirigidas por personal sin preparación, otras por investigadores metodológicos que intentaban sin éxito adaptar sus métodos a las evaluaciones educativas. (Guba, 1966).

Los estudios evaluativos estaban cargados de confusión, ansiedad y hostilidad. La evaluación educacional como especialidad tenía poca talla y ninguna dimensión política.

Con este telón de fondo, el progreso realizado por los evaluadores educacionales para profesionalizar su especialidad durante los años 70's es totalmente detestable. Algunas publicaciones, aparecieron en aquella época, y luego han demostrado ser excelentes vehículos para recoger y difundir información acerca de las diversas facetas de la evaluación educacional. A diferencia de lo que ocurría 15 años atrás, numerosos libros y monografías tratan ahora exclusivamente de la evaluación. De hecho, el problema actual no es encontrar literatura sobre la evaluación, sino seguir el ritmo de lo que se publica.

Muchas universidades han empezado a ofrecer por lo menos un curso de metodología evaluativa. Unas pocas han elaborado programas para graduarse en evaluación.

De un modo creciente, la especialidad está considerando la metaevaluación (Scriven, 1975; Stufflebeam, 1978) como un medio para asegurar y comprobar la calidad de las evaluaciones. Este considerable desarrollo profesional de la evaluación educacional ha producido diversos resultados; aunque la preparación y titulación de evaluadores ha mejorado con el fin de que las instituciones puedan disponer de servicios realizados por personas cualificadas, algunos lamentan que este desarrollo pueda reducirse a una especie de club cerrado y exclusivista (Stake, 1981).

Aunque se ha incrementado la comunicación entre los partidarios de los métodos positivistas/cuantitativos y los que proponen métodos fenomenológicos/cualitativos, existe, actualmente, el peligro de que se produzca una polarización entre estos dos campos.

A pesar de la creciente búsqueda de métodos adecuados, del incremento de la comunicación y el entendimiento entre los principales teóricos, y del desarrollo de nuevas técnicas, la práctica actual de la evaluación ha cambiado muy poco en la mayoría de los casos.

Según Stufflebeam y Shinkfield (1987), a pesar de la gran proliferación de modelos sobre evaluación educativa, no todos gozan de igual fortaleza y aceptación. La razón radica en que asuman la naturaleza misma de la evaluación, la cual consiste en determinar mediante un juicio, el valor y mérito de un objeto. Desde esta perspectiva, Stufflebeam clasifica los métodos en pseudoevaluativos, cuasievaluativos y verdaderamente evaluativos.

- Los métodos pseudoevaluativos, aún al recoger la información no revelan las verdaderas conclusiones, las falsifican o las transmiten de modo selectivo. Hay dos tipos: de investigaciones encubiertas y fundamentadas en las relaciones públicas.
- Las cuasievaluaciones son estudios cuyo objetivo es resolver un problema concreto mediante una metodología sin juzgar o enjuiciar el valor y mérito de lo analizado; como los estudios basados en objetivos y fundamentados en la investigación educativa.
- Las verdaderas evaluaciones evitan dar resultados tendenciosos, pues no se enfocan sobre un

seguimiento de la realidad, sino que realizan investigaciones de conjunto que llevan a determinar el valor y el mérito de algo; como la planificación evaluativa (Cronbach, 1982); la evaluación orientada hacia el perfeccionamiento (Stufflebeam, 1976); el método evaluativo de Stake, R. (1967); el método de Owens y Wolf (1973); la evaluación iluminativa o método holístico (Hamilton, D., Parlett, M., Mac Donald, B. y Stake, R. 1977), y el método de Scriven (1982).

Si con algo se podría caracterizar las aportaciones teóricas que nos ofrecen los especialistas durante los años setenta es con la proliferación de toda clase de modelos de evaluación que expresan la propia óptica del autor que los propone sobre qué es y cómo debe conducirse un proceso evaluativo. Se trata, por tanto, de una época caracterizada por la pluralidad conceptual y metodológica. Guba y Lincoln (1982) nos hablan de más de cuarenta modelos propuestos en estos años, y Mateo (1986) se refiere a la eclosión de modelos. Estos enriquecerán considerablemente el vocabulario evaluativo.

House (1989), tienden a clasificar estos modelos en dos grandes grupos, cuantitativos y cualitativos. Esas dos tendencias se observan hoy en las propuestas evaluativas, pero los diferentes modelos se diferencian más por destacar o enfatizar alguno o algunos de los componentes del proceso evaluativo y por la particular interpretación que a este proceso le dan. Los enfoques modélicos, más que modelos, terminan construyéndose en cada investigación evaluativa, en función del tipo de trabajo y las circunstancias (Escudero, 1993).

En este movimiento de propuestas de modelos de evaluación cabe distinguir dos épocas con marcadas diferencias conceptuales y metodológicas. En una, las propuestas por Tyler en su planteamiento, que ha venido a llamarse de "Consecución de Metas". En esta época destacan la propuesta de Hammond (1983) y el Modelo de Discrepancia de Provus (1971). Para estos autores los objetivos propuestos siguen siendo el criterio fundamental de valoración, pero enfatizan la necesidad de aportar datos sobre la congruencia o discrepancia entre las pautas de instrucción diseñadas y la ejecución de las mismas en la realidad del aula.

Otros modelos consideran el proceso de evaluación al servicio de las instancias que deben tomar decisiones. El CIPP (contexto, input, proceso y producto), propuesto por Stufflebeam y colaboradores (1971) y el CES (toma sus siglas del Centro de la Universidad de California para el Estudio de la Evaluación) dirigido por Alkin (1969). Estos autores van más allá de la evaluación centrada en resultados finales, puesto que en sus propuestas suponen diferentes tipos de evaluación, según las necesidades de las decisiones a las que sirven.

Una segunda en la proliferación de modelos, representada por modelos alternativos, que con diferentes concepciones de la evaluación y de la metodología a seguir comienzan a aparecer en la segunda mitad de esta década de los setenta. La Evaluación Responsable de Stake (1975 y 1976), la Evaluación Democrática de MacDonald (1976), la Evaluación Iluminativa de Parlett y Hamilton (1977) y la Evaluación como crítica artística de Eisner (1985). Este segundo grupo de modelos evaluativos enfatiza el papel de la audiencia de la evaluación y de la relación del evaluador con ella, que debe ser «transaccional y fenomenológica». Con una evaluación de tipo etnográfica, que consideran adecuada a la antropología social (Parlett y Hamilton, 1977; Guba y Lincoln, 1982; Pérez 1983).

Este resumen de modelos de la época de eclosión se aproxima al amplio abanico conceptual teórico y metodológico que hoy se relaciona con la evaluación. Nevo (1983 y 1989) pretende realizar una conceptualización de la evaluación, a partir de la revisión de la literatura especializada, atendiendo a los tópicos ¿Qué es la evaluación? ¿Qué funciones tiene? ¿Cuál es el objeto de evaluación? No encuentra una única respuesta a estas cuestiones. Pero también es cierto que bajo esta diversidad subyacen diferentes concepciones teóricas y metodológicas sobre la evaluación. Diferentes concepciones que han dado lugar a una apertura y pluralidad conceptual en el ámbito de la evaluación en varios sentidos (Cabrera, 1986). Puntos sobresalientes de esta pluralidad

conceptual:

- a) Diferentes conceptos de evaluación.
- b) Diferentes criterios a utilizar para la valoración de la información
- c) Pluralidad de procesos evaluativos
- d) Pluralidad de objetos de evaluación.
- e) Apertura, 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.
- f) Pluralidad de las funciones en el ámbito educativo, entre evaluación formativa y sumativa, y añadiéndose otras de tipo socio-político y administrativas (Nevo, 1983).
- g) Diferencias por el evaluador, en evaluación interna vs. evaluación externa. (Nevo, 1983; Weiss, 1983; Rutman, 1984).
- h) Pluralidad de audiencia de la evaluación y, pluralidad en los informes de evaluación. i) Pluralidad metodológica.

A finales de los ochenta, tras todo este desarrollo antes descrito, Guba y Lincoln (1989) ofrecen una alternativa evaluadora, que denominan cuarta generación, pretendiendo superar deficiencias de las tres generaciones anteriores, tales como una visión gestora de la evaluación, una escasa atención al pluralismo de valores y un excesivo apego al paradigma positivista. La alternativa la denominan respondente y constructivista, integrando de alguna manera el enfoque respondente propuesto por Stake (1975), y la epistemología postmoderna del constructivismo (Russell y Willinsky, 1997).

Las demandas, las preocupaciones y los asuntos de los implicados o responsables (stakeholders) sirven como foco organizativo de la evaluación, que se lleva a cabo dentro de los planteamientos metodológicos del paradigma constructivista.

La propuesta de Guba y Lincoln (1989) se extiende en la explicación de la naturaleza y características del paradigma constructivista.

Para juzgar la calidad de la evaluación, se nos ofrecen tres enfoques que se denominan paralelo, el ligado al proceso hermenéutico y el de autenticidad.

Los criterios paralelos, de confianza, se denominan así porque intentan ser paralelos a los criterios de rigor utilizados muchos años dentro del paradigma convencional. Estos criterios han sido validez interna y externa, fiabilidad y objetividad. En el caso de la cuarta generación los criterios que se ofrecen son los de credibilidad, transferencia, dependencia y confirmación.

El análisis dentro del propio proceso, encaja con el paradigma hermenéutico, a través de un proceso dialéctico. Los criterios que denominan de autenticidad, incluyen los siguientes: a) imparcialidad, justicia, ontológica, educativa, catalítica y táctica.

Este análisis de la cuarta generación de podemos terminarlo con los rasgos en: La evaluación es un



proceso sociopolítico; de colaboración; de enseñanza/aprendizaje; continuo, recursivo y altamente divergente; emergente; con resultados impredecibles; que crea realidad.

En esta evaluación, se retienen las características del evaluador fruto de las tres primeras generaciones, esto es, la de técnico, la de analista y la de juez, pero estas deben ampliarse con destrezas para recoger e interpretar datos cualitativos (Patton, 1980), con la de historiador e iluminador, con la de mediador de juicios, así como un papel más activo como evaluador en un contexto socio-político concreto.

Para terminar este recorrido analítico-histórico desde los primeros intentos de medición educativa hasta la actual investigación evaluativa en educación, el modelo más utilizado a finales de los sesenta, desde 1975 a 1988 fue el de Daniel L. Stufflebeam, proponente del modelo CIPP. Y que se ve como una «transdisciplina», pues es aplicable a muchas disciplinas diferentes y a muchos objetos diversos (Scriven, 1994).

Tanto en el contexto internacional como en el nacional la educación superior ha adquirido un papel determinante en el desarrollo social, económico, político y cultural.

La sociedad se está volviendo más participativa y exige cambios en los diversos sectores del sistema educativo, por lo que las IES deben actuar con mayor responsabilidad social y responder a tales exigencias.

Fenómenos como la expansión desordenada del sistema de educación superior en América Latina, el surgimiento de IES privadas de baja calidad y el deterioro de la educación pública han suscitado preocupación por el mejoramiento de la calidad de la educación superior. En la década de los setenta algunas IES mexicanas se vieron en la necesidad de definir las pautas de su desarrollo y surgieron los primeros acercamientos a la evaluación institucional conocidos como autoestudios, cuyo propósito principal era la elaboración de sus planes de desarrollo y la capacitación de las IES para regular su futuro (Carrión y Fernández, 1989). A partir de 1990, en el contexto de una tendencia internacional muy marcada, han sido cada vez más notorios los esfuerzos por implantar mecanismos de evaluación; sin embargo, estos intentos no se han consolidado plenamente y suelen estar débilmente articulados a la toma de decisiones. Los métodos de evaluación de la calidad de la IES han sido diversos, desde la evaluación interna, que se gesta en el seno de las propias instituciones, hasta la evaluación externa, que involucra a organismos ajenos a las IES.

Durante este proceso algunas IES avanzaron y se hicieron mejores, al punto de adquirir el calificativo de “buenas” o de “buena calidad”, mientras que otras, quizá con una visión más estrecha, no tuvieron avances tan significativos e incluso fueron calificadas como de “mala calidad”. Varias IES sólo vieron en la evaluación un mecanismo de descripción de sus funciones y quehaceres para satisfacer una demanda.

Ante un panorama de resultados tan contrastados, las principales interrogantes que motivan este proyecto de investigación son las siguientes: ¿Qué significa ser una “buena” institución? ¿Es posible medir o captar de manera precisa las características que hacen “buena” a una IES? ¿Qué es lo que hace que una institución sea mejor que otra? ¿Frente a una IES en qué se distingue otra que es “mejor”? ¿Cómo debe ser entendido el término “mejor”? ¿Qué requisitos es necesario cubrir para que una institución pueda ser considerada con un nivel aceptable de calidad? ¿Instituciones con menos recursos y experiencia que otras, pueden tener niveles aceptables de calidad? ¿Cuáles son los referentes de calidad que se deben considerar para tener un nivel aceptable? Y, finalmente, ¿Cómo puede mejorar una institución?

En un contexto nacional e internacional de transformación social continua la educación superior debe ser capaz de ofrecer alternativas y propuestas adecuadas para poder enfrentar los problemas que genera el desarrollo acelerado, de ahí que los siguientes elementos constituyan el eje de esta investigación:

Primero, las exigencias de la sociedad de que las IES respondan a sus necesidades y enfrenten los

nuevos retos del mundo globalizado.

Segundo, la necesidad de dar cauce a la preocupación por la calidad de la educación superior con bases más sólidas y con mecanismos de evaluación más eficaces.

Tercero, la superación del proceso de evaluación concebido solamente como una descripción sistematizada y el arribo a una valoración efectiva del desarrollo institucional, mediante la utilización de sistemas de evaluación congruentes con los rasgos y metas de las IES.

Finalmente, dado que la educación superior opera en un nuevo escenario de competencia mundial, se vuelve necesario desarrollar sistemas de indicadores para contextos más amplios cuyos parámetros sean apropiados para la redefinición de las funciones y objetivos que persiguen las IES. Es importante desarrollar modelos que tomen en cuenta las características de las IES, pues adoptar modelos que no correspondan a ciertas particularidades puede resultar más perjudicial que benéfico.

Principales orientaciones e impacto de las políticas en la educación agrícola superior. La instrumentación de la evaluación a las IES se dio en 1990 con la aprobación del documento "Propuestas de Lineamientos para la Evaluación de la Educación Superior (ANUIES, 1990) con tres mecanismo elementales a llevar a cabo:

- 1) Procesos de evaluación institucional, a cargo de las propias IES.
- 2) Procesos de evaluación interinstitucional, mediante evaluación externa, incluyendo la evaluación de pares académicos.
- 3) Evaluación realizada por especialistas.

Es un proceso de revisión continua, sistemática y organizada, a partir de objetivos y metas propuestas; la cual es realizada por los mismos participantes de los programas académicos. Se toma como base la información directa, en hechos, situaciones, opiniones, recomendaciones vertidas por sus integrantes, para lograr el mejoramiento y desarrollo académico deseable.

En la evaluación interinstitucional, se ha institucionalizado por los Comités Interinstitucionales de Evaluación de la educación Superior (CIEES) denominado evaluación diagnóstica. Se basa en un enfoque constructivista, cuya responsabilidad consiste en detectar logros y deficiencias en un proyecto, programa educativo y área académico-administrativa.

Este tipo de evaluación interinstitucional en el contexto de las políticas públicas universitaria, a decir de Victorino (2000), parte del supuesto de que políticas tales como el financiamiento, contención salarial y deshomologación, así como asuntos de descentralización, fusión y cierre de instancias e instituciones gubernamentales y educativas, el gobierno mantuvo un impacto fuerte. En tanto que las acciones de reforma académica, planeación y evaluación a las instituciones de educación agrícola superior fue menos contingente sobre todo en la idea de intervenir desde un inicio en algunas universidades a las que les reconocía su papel autónomo.

Esquema. Transición de la política a las políticas de educación superior (EAS)		
Orientaciones hasta 1989		Desde 1989 a 1994
Financiamiento, salarios y contratación de personal académico	Política Única (1985): Incrementos anuales basados en la matrícula y negociación política. No competitivo e ind. De resultados. (desde 1985): Subsidio ordinario por la SESIC. Salarios (contenidos). Pactos: inflacionarios (1985) Sin contratación de PA y administrativos.	Continuidad de criterios en Subsidios y salarios. Diversificación de financiamiento FOMES. Carrera PA (1990) Ingresos propios, colegiaturas, rentabilidad de servicios.
Descentralización (Desconcentración, planeación)	SINAPES (1978) Medidas indicativas. Exigencias a gobiernos estatales.	PND-Programas sectoriales. PDI Diagnósticos y prospecciones.
Evaluación	Seguimiento superficial de las funciones universitarias. Prescripción sin juicios ni condicionamientos.	Seguimiento END-EXOG, Aut. Institucional. ANUIES-IES-CONAEVA Evaluación Externa de Programas (CIEES) Evaluación. Desempeño Académico por actores (IEAS) Evaluación del SES (CONAEVA)

En la evaluación de 1988 hay un seguimiento superficial de las funciones universitarias. Las IES cubren ciertas prescripciones sin juicios ni condicionamientos. Desde 1989 se pasa de la planeación a la evaluación como política dominante por el control de las IES (Victorino, 2000). En este contexto se vive una euforia por la evaluación, de forma tal que distintas instancias oficiales solicitan evaluaciones externas e internacionales. (Coombs, 1991) nacionales, así como autoevaluaciones a los establecimientos educativos (CONAEVA, 1991), considerando una mínima participación de los profesores mediante los pares académicos.

En el periodo de 1989-1994 se habla insistentemente de la evaluación a las Universidades Públicas e incluso se da continuidad hasta el año 2000. El desenvolvimiento de las políticas evaluativas hacia la educación superior se da mediante lo siguiente:

- De 1990 a 1992 se avanza en la autoevaluación de las IES
- Después, de la evaluación interna o institucional, con intercambio de información entre instituciones afines a nivel internacional.
- Para 1994 y a mediados de 1995 se concretan las acciones de la CONAEVA, se conoce también como evaluación externa.
- En 1996 la evaluación se centra más en los pares académicos, y a publicar los resultados de evaluaciones por funciones universitarias.

El SINAPES también se fortalece con las medidas de desconcentración y descentralización, a la vez que sus intencionalidades incluyendo una serie de cambios curriculares buscan fundamentalmente readecuar la educación a las nuevas condiciones internacionales.

Las evaluaciones de distinto alcance y tipo internacional y nacional, coincidían en que había en general una baja calidad; por lo que la estrategia central del PROMEP, 1989-1994, así como sus documentos derivados por el nivel educativo y funciones propusieron diversas estrategias para elevar la calidad, de las fundamentales destacan las siguientes: revisar contenidos, renovar métodos, privilegiar la formación de maestros, articular los diversos niveles educativos y vincularlos procesos pedagógicos con los avances

científicos y tecnológicos. Estas estrategias también incluyen al posgrado nacional.

En 1988 se daba cierta apertura e iniciativa de cambios propios a las IEAS, su órgano corporativo era la Asociación Mexicana de Educación Agrícola Superior (AMEAS). Pero a partir de 1989 dicha organización es absorbida por las políticas nacionales propuestas por la ANUIES, incluyendo toda su estrategia evaluativa, así como la ampliación de la cobertura de esta última organización.

No obstante intentos de AMEAS para la creación del Sistema Nacional de Educación Agrícola en 1984 y la fallida Ley de Educación Agrícola propuesta por el Ejecutivo Federal, todavía en 1989 se publicó el Plan de Desarrollo de Educación Agrícola Superior (PDEAS) con 16 programas y 23 proyectos, muy parecido a PROIDES de la ANUIES en 1986.

Aún con estos criterios, procedimientos y modalidades de evaluación fueron orientados a la Comisión Nacional de Evaluación (CONAEVA) junto con la ANUIES.

En el contexto de las distintas facetas de evaluación nacional la educación agrícola superior, pasó de ser un área más en las distintas esferas del conocimiento de la educación superior y de su propio sistema.

Victorino aborda en (La evaluación y universidades. Las IEAS de la zona centro del país ante la evaluación de los CIEES, 2004) sobre la contextualización de las políticas de evaluación emanadas del programa para la Modernización Educativa (SEP, 1989) y el documento de Estrategias para la evaluación de la Educación Superior (ANUIES, 1990), al mismo que se le da continuidad con el sexenio del Dr. Ernesto Zedillo Ponce de León (1995-2000). En el nivel de educación superior se analiza el contexto mexicano en el ambiente internacional; se ubican los programas de evaluación y sus distintos mecanismos, etapas y modalidades: autoevaluación, evaluación institucional y evaluación externa. Así los mecanismos para lograr la acreditación y certificación de los programas académicos y de los profesionales egresados de una carrera de agronomía o cualquier profesión, respectivamente.

Un cuestionamiento central del trabajo debe consistir en documentar por qué las políticas evaluativas no tuvieron buena recepción por parte de las Universidades Públicas Autónomas, lo cual provocó varios casos de rechazo y resistencia a esas políticas durante los años 90's.

Ni fue lo mismo a finales de los 90's cuando varias IEAS accedieron a la evaluación diagnóstica por parte de organismos oficiales como los CIEES.

Si en los años 90 las políticas evaluativas hacia las IEAS estuvo más centrada en rasgos y criterios de corte cuantitativo, para la primera década del siglo XXI (2001-2010), dichos criterios transitaron más hacia la evaluación con criterios cualitativos y de mejora en las condiciones de las propias, trayendo como consecuencia de los cinco años antes habían logrado su acreditación y en 2011 debería buscar su re acreditación por un lustro más.

## CONCLUSIONES

Después de este apretado recorrido sobre la evaluación a las IEAS, es importante que los evaluadores, al planificar sus estudios, tengan en cuenta que sus resultados pueden ser utilizadas en el contexto de una investigación que proponga una comprensión más amplia de los fenómenos examinados en evaluación, que pueden servir no sólo como guía de programas operativos, o como resumen de sus contribuciones, sino también para ilustrar investigaciones concretas que incluyan cuestiones teóricas y metodológicas del ámbito que se evalúa.

Así pues, parece claro que la evaluación es un concepto altamente complicado. A la hora de planificar o

dirigir una evaluación, deben tenerse en cuenta una serie de consideraciones:

1. Tras el análisis del desarrollo de la evaluación a lo largo del Siglo XX, es importante resaltar la importancia de los principales modelos, planteamientos metodológicos, diseños, perspectivas y visiones de la evaluación en la actualidad. Su análisis, es un complemento necesario para una visión como la histórica que, por su linealidad, tiene el riesgo de ofrecer una imagen disciplinar artificialmente fraccionada.
2. La eclosión de propuestas evaluativas que tradicionalmente han sido denominadas como modelos y en algunos casos como diseños de investigación evaluativa. Se sabe que existieron varias decenas de estas propuestas muy concentradas en los años ochenta, que parece un tema cerrado desde hace cuatro lustros, ya que no han surgido nuevos modelos o propuestas, salvo alguna excepción según la diversidad de carreras profesionales que se someten a estos procesos de evaluación.
3. Hoy en día (2011) se sigue hablando de modelos, métodos y diseños en la literatura especializada, sobre todo buscando su clasificación de acuerdo con diversos criterios, origen paradigmático, propósito, metodología, entre otros atributos. Con gran diversidad, lo que prueba es que, además de dinamismo académico en la investigación evaluativa, todavía existe cierta debilidad teórica en los procesos de evaluación para la acreditación de carreras en ámbito agropecuario.
4. Muchos de los acercamientos a la conceptualización de la evaluación se les ha denominado indebidamente como modelos a pesar de que ninguno de ellos tenga el grado de complejidad y de globalidad que debería acarrear el citado concepto. Lo que un texto clásico en evaluación denomina como «modelos contemporáneos de evaluación» el propio Stake (1981) dice que sería mejor denominarlo como «persuaciones» mientras que House (1983) se refiere a «metáforas».
5. Norris (1993) apunta que el concepto de modelo se utiliza con cierta ligereza al referirse a concepción, enfoque o incluso método de evaluación. De Miguel (1989), por su parte, piensa que muchos de los llamados modelos solamente son descripciones de procesos o aproximaciones a programas de evaluación. Darling-Hammond y otros (1989) utilizan el término «modelo» por costumbre, pero indican que no lo hacen en el sentido preciso que tiene el término en las ciencias sociales, esto es, apoyándose en una estructura de supuestos interrelacionales fundamentada teóricamente. El propio autor del modelo CIPP, solamente utiliza esta denominación de manera sistemática para referirse a su propio modelo (Stufflebeam y Shinkfield, 1987), utilizando los términos de enfoque, método entre otros.
6. Quizás en nuestra percepción el término enfoque evaluativo sea el más apropiado, aunque aceptemos seguir hablando de modelos y diseños por simple tradición académica.
7. Nuestra idea es que a la hora de plantearnos una investigación evaluativa, no contamos todavía con unos pocos modelos bien fundamentados, definidos, estructurados y completos para elegir uno de ellos, pero sí tenemos distintos enfoques modélicos y un amplio soporte teórico y empírico, que permiten ir respondiendo de manera adecuada a las distintas cuestiones que le va planteando el proceso de investigación, ayudándole a configurar un plan global, un organigrama coherente, un «modelo» científicamente robusto para llevar a cabo una evaluación con la finalidad de caracterizar la evaluación interinstitucional aplicada a la educación superior, y acercarse a dibujar el escenario tendencial que se prefigura en las Instituciones de Educación Agrícola Superior (IEAS).

Como puede observarse a lo largo del artículo, los procesos de evaluación a las IEAS llegaron para continuar evaluando. No obstante varias instituciones educativas no han accedido a este proceso bajo la política educativa de la calidad académica a través de la acreditación, debido a que no cuentan con las condiciones sobre todo de infraestructura (laboratorio, centro de computo, edificios entre otros), además del personal docente calificado, de tiempo completo y con estudios de posgrados.

Por lo que el escenario tendencial del enfoque de la evaluación interinstitucional será el poder lograr cambios en los indicadores de calidad (incremento de profesores de tiempo completo con nivel de doctorado pertenecientes al Sistema Nacional de Investigadores (SIN) y habilitados en un cuerpo académico en proceso de consolidación (PROMEP, 2004) guiados por la búsqueda de recursos adicionales para el

desarrollo de programas de fortalecimiento en docencia, investigación, y extensión para incrementar o bien mantener la matrícula mediante el aseguramiento de la calidad.

En la actualidad (2011) se han instrumentado estrategias para el fortalecimiento institucional con programas estratégicos como Programa de Fortalecimiento Institucional (PIFI) y la ampliación del Padrón Nacional de Posgrado de Calidad (PNPC) hacia los programas de posgrado por parte del Consejo Nacional de Ciencia y Tecnología (CONACYT), los cuales solicitan continuar con modelos de evaluación para mejorar la calidad de los estudios posgraduales.

En un análisis retrospectivo se puede considerar que las corrientes de evaluación implementadas a los largos del siglo XX, algunas de ellas presentan una continuación en su aplicabilidad a los procesos de evaluación en la IEAS durante la primera década del siglo XXI. Las cuales se han promovido y reforzado con la integración de los CIEES en una perspectiva de mediano y largo plazos.

#### Literatura citada

- Alkin, M. (1969). Evaluation theory development, *Evaluation Comment*, 2, 1, 2-7.
- ANUIES (1990) "estrategias para la evaluación de la educación superior" Universidad futura, 2(5), pp 61-68.
- Atkin, J. M. (1968). Behavioral Objectives in curriculum design: A cautionary note, *The Science Teacher*, 35, 27-30.
- Bloom, B. S., Hastings, J. TH. Y Madaus, F. (1971). *Handbook on formative and summative evaluation of student learning*. New York : McGraw-Hill.
- Bloom, B.S., Hastings, T. y Madaus G. (1975). Evaluación del aprendizaje. Buenos Aires: Troquel.
- Cabrera, F. (1986). Proyecto docente sobre técnicas de medición y evaluación educativas. Barcelona: Universidad de Barcelona.
- Carrión, C y Fernández, A (1989) "Experiencias en la autoevaluación institucional en México" Revista de la Educación Superior, No 69, Pp. 75-103.
- Consejo Nacional de Evaluación,(CONAEVA) Grupo Técnico (1991), Propuesta metodológica para la autoevaluación institucional anual, Manuscrito.
- Coombs, P.H. (1991), El futuro de la educación no formal en un mundo cambiante. Fundación Santillana. Madrid, España.
- Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers College Record*, 64, 672-683.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. Chicago : Jossey-Bass.
- Darling-Hammond, L., Wise, AE, & Pease, SR (1989). Teacher evaluation in the organizational context: A review of the literature. En E. R. House (Ed.) *New directions in educational evaluation* (pp.203-253). London: The Falmer Press.
- De Miguel, M. (1989). Modelos de investigación sobre organizaciones educativas. *Revista de Investigación Educativa*, 7, 13, 21-56.
- Eisner, E. W. (1967). Educational objectives: Help or hindrance? *The School Review*, 75, 250-260.
- Eisner, E. W. (1969, Instructional and expressive educational objectives: their formulation and use in curriculum. En J. Popham (Ed.), *Instructional objectives* (pp. 1-18). Chicago : AERA.
- Eisner, W. E. (1985). *The art of educational evaluation*. London: The Falmer Press.
- Escudero, T. (1993). Enfoques modélicos en la evaluación de la enseñanza universitaria, *Actas de las III Jornadas Nacionales de Didáctica Universitaria «Evaluación y Desarrollo Profesional»* (pp. 5-59). Las Palmas: Servicio de Publicaciones, Universidad de Las Palmas.
- Fernández de Castro, J. (1973). La enseñanza programada. Madrid: CSIC.



- Gagne, R. M. (1971). Las condiciones del aprendizaje. Madrid: Aguilar.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: some questions. *American Psychologists*, 18, 519-521.
- Guba, E. G. y Lincoln, Y. S. (1989). Fourth Generation Evaluation. Newbury Park, Ca. Sage Publications
- Guba, E.G. (1966). Study of Title III activities: Report on evaluation. Bloomington: Indiana University and the National Institute for the Study of Educational Change.
- Guba, G. E. y Lincoln, Y. S. (1982). Effective evaluation. San Francisco: Jossey Bass Publishers.
- Hambleton, R. K. (1985). Criterion-referenced measurement. En *Encyclopedia of Educational Research*. New York : McMillan.
- Hamilton, Mac Donald, Parlett y Stakes, 1977
- Hammond, R. L. (1967) "Evaluation at the local level." Address to the Miller Committee for the National Study of ESEA Title III. HOLMES, E.G.A.
- Hammond, R. L. (1983). Evaluation at the local level. En B. R. Worthen y J. R. Sanders, Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Company
- House, E. R. (1983). How we think about evaluation. En House, E. R., *Philosophy of evaluation*. San Francisco , Ca.: Jossey-Bass.
- House, E. R. (1989). Evaluating with validity. Newbury Park, Ca. Sage.
- J. M. Rice citado por Mateo, 1993 Joint Committee on Standards for Educational Evaluation (1981, Standards for evaluations of educational programs, projects, and materials. New York.: McGraw-Hill.
- La evaluación iluminativa o método holístico citado por Stufflebeam, D. L. y Shinkfield, A. J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Barcelona: Paidós/MEC.
- Macdonald, B. (1971). The evaluation of the Humanities Curriculum Project: a holistic approach. *Theory into Practice*, 10, 3, 163-169.
- Madaus, G. F. y otros (1991). Evaluation Models. Viewpoints on Educational and Human Services Evaluation. Hingham , Mass.: Kluwer-Nijhoff Publishing.
- Mateo, J. (1986). *Proyecto docente e investigador*. Barcelona: Universidad de Barcelona.
- Mateo, J. y otros (1993). La evaluación en el aula universitaria. Zaragoza: ICE-Universidad de Zaragoza.
- Mateo, J. y otros (1993). *La evaluación en el aula universitaria*. Zaragoza: ICE-Universidad de Zaragoza.
- Metfessel, N. S. y Michael, W. B. (1967). A paradigm involving multiple criterion measures for the evaluation of effectiveness of school programs. *Educational and Psychological Measurement*, 27, 931-943.
- Nevo, D. (1989). The conceptualization of educational evaluation: An analytical review of the literature. En E. R. House (Ed.), *New directions in educational evaluation*. London : The Falmer Press, 15-29.
- Norris, N. (1993). *Understanding educational evaluation*. London : Kogan Page/CARE, School of Education , University of East Anglia.
- Owens, T. R. & Wolf, R. L (1973). Educational evaluation by adversary proceedings. En E.R. House (Comp.), *School Evaluation: The Politics and Process*. Berkeley : McCutchan.
- Parlett, M. y Hamilton , D. (1977). Evaluation as illumination: A new approach to the study of innovative programmes, en Hamilton D. y otros (Eds.), *Beyond the numbers game*. London : MacMillan.
- Patton, M. Q. (1980). *Qualitative Evaluation methods*. Beverly Hills , Ca.: Sage.
- Pérez, A. (1983). Modelos contemporáneos de evaluación. En J. Gimeno y A. Pérez, *La enseñanza: su teoría y su práctica* (pp. 426-449). Madrid.
- Popham, W. J. (1971). *Problemas y técnicas de la evaluación educativa*. Madrid: Anaya.
- Popham, W. J. (1983). Evaluación basada en criterios. Madrid: Magisterio Español, S. A..

- Provus, M. (1971). Discrepancy evaluation. For educational program improvement and assessment. Berkeley, Ca.: McCutchan Publishing Co.
- Rodríguez, T. y otros (1995). Evaluación de los aprendizajes. Aula Abierta Monografías 25. Oviedo: ICE-Universidad de Oviedo.
- Russell, N. y Willinsky, J. (1997). Fourth generation educational evaluation: The impact of a post-modern paradigm on school based evaluation. *Studies in Educational Evaluation*, 23, 3, 187-199.
- Rutman, L. y Mowbray, G. (1983). Understanding program evaluation. Beverly Hills, Ca.: Sage.
- Rutman, L., (Ed.) (1984). *Evaluation research methods: A base guide*. Beverly Hills , Ca.: Sage.
- Salvador, L. (1992). Proyecto docente., Universidad de Cantabria.
- Scriven, M. (1967). The methodology of evaluation. En Perspectives of Curriculum Evaluation, (pp. 39-83). AERA Monograph 1. Chicago : Rand McNally and Company.
- Scriven, M. (1975). Evaluation in science education. Washington DC: National Institute of Education.
- Scriven, M. (1982). Professorial ethics. *Journal of Higher Education*, 53(3), pp. 307-317.
- Scriven, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, 20, 1, 147-166.
- Stake, R. E. (1967). The countenance of educational evaluation. Teacher College Record, 68, 523-540.
- Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach*. Ohio: Merrill.
- Stake, R. E. (1981). Setting standars for educational evaluators. *Evaluation News*, 22, 148-152.
- Stenhouse, L. (1984). Investigación y desarrollo del curriculum. Madrid: Morata.
- Stufflebeam, D. L. (1971). ¿ *The relevance ofthe CIPP evaluation model for educational accountability?* Joun1al of Research and Policy Analysis, 2 (3), 5-20.
- Stufflebeam, D. L. (1976). *Meta-evaluation*. Kalamazoo, MI: The Evaluation Center, Western Michigan University.
- Stufflebeam, D. L. y Shinkfield, A. J. (1987). Evaluación sistemática. Guía teórica y práctica. Barcelona: Paidós/MEC.
- Stufflebeam, D. L. y Shinkfield, A. J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Barcelona: Paidós/MEC.
- Stufflebeam,D.L.(1978). Meta-evaluation: An overview. *Evaluation and the Health Professions*, 1,17-43.
- Suchman, E. A. (1967). *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. New York : Russell Sage Foundation.
- Taba, H. (1962). Curriculum development. Theory and practice. New York: Harcourt Brace .
- Tyler, R. W. (1950). Basic principles of curriculum and instruction. Chicago: University of Chicago Press .
- Tyler, R. W. (1967). Changing concepts of educational evaluation. En R. E. Stack (Comp.), Perspectives of curriculum evaluation. AERA Monograph Series Curriculum Evaluation, 1. Chicago: Rand McNally,.
- Tyler, R. W., (Ed.) (1969). Educational evaluation: New roles, new means. Chicago: University of Chicago Press .
- Victorino Ramírez, L. 2006. Procesos de evaluación en la Universidad hispanoamericana, Castellanos Editores-UACH, México
- Weiss, C. H. (1983). Investigación evaluativa. Métodos para determinar la eficiencia de los programas de acción. México: Trillas.
- Wilson, A. R. J. (1978). La evaluación de los objetivos. En J. A. R. Wilson (Ed.), *Fundamentos psicológicos del aprendizaje y la enseñanza* (pp. 549-578). Madrid: Anaya.