

**PERSPECTIVA DA AVALIAÇÃO EDUCACIONAL AO LONGO DA HISTÓRIA –  
DOS TESTES AO MODELO CONSTRUTIVISTA**

Luciana Kellen de Souza Gomes | lucianakellen2004@hotmail.com

Maria Isabel Filgueira Lima Ciasca | Isabelfil@yahoo.com.br

Lourdes Rafaella Santos Florencio | rafaellaflorencio@gmail.com

## **INTRODUÇÃO**

A avaliação é uma área do conhecimento que sofreu profundas mudanças conceituais e funcionais ao longo do século XX, mudanças essas que para efeito de análise das diversas abordagens da avaliação no contexto histórico, Escudero divide em seis períodos históricos as diferentes perspectivas que a avaliação assumiu ao longo da história: a) período da reforma (1800-1900), b) a eficiência de tempo e "testar" (1900-1930), c) era de Tyler (1930-1945), d) idade da inocência (1946-1956), e) era de expansão (1957-1972) e f) da era profissional (desde 1973), que refere-se a situação atual.

Gil (1992), afirma que no século XVIII, o aumento da demanda e do acesso à educação, sublinha a necessidade de verificar os méritos individuais e das instituições educacionais, quando são desenvolvidas e introduzidas regras sobre a utilização de provas escritas.

No final do século XIX, surge grande interesse na medição científica do comportamento humano. Isso é algo que faz parte do movimento renovador da metodologia das ciências humanas, ao assumir o positivismo das ciências físicas e naturais. Neste sentido, a avaliação recebe as mesmas influências como outras disciplinas pedagógicas relacionadas com os processos de medição, como pedagogia experimental e diferencial.

Nesse período, influenciaram a avaliação: a) Pensamento positivista e empirista, apoiado na observação, experimentação, dados e fatos como fontes de conhecimento verdadeiro, b) A influência das teorias evolucionistas e os trabalhos de Darwin, Galton e Cattell, apoiando a medição das características dos indivíduos e as diferenças entre eles. c) O desenvolvimento de métodos estatísticos que favoreceram decisivamente a orientação métrica do tempo, d) O desenvolvimento da sociedade e a necessidade de seleção de estudantes, de acordo com o seu conhecimento.

Nesse período, surgiu o “teste”, caracterizado pela similaridade dos termos avaliação e

medição, busca pelo estabelecimento de diferenças individuais em função de objetivos educacionais, enfim, avaliava os alunos e não os programas educacionais.

No entanto, foi nos testes psicológicos onde os esforços tiveram maior impacto, provavelmente o trabalho de Thorndike, o mais influente no início do século XX, seguido dos testes de habilidade cognitiva de Binet .

Após a guerra, os testes psicológicos estão a serviço de fins sociais. A década entre 1920 e 1930 marca o ponto mais alto da 'testing', já que haviam muitos testes padronizados para medir todos os tipos de habilidades escolares com relativa objetividade e de aplicação a grande grupos de alunos.

O fervor do uso dos testes diminuiu a partir dos anos quarenta até começaram a surgir alguns movimentos hiper-críticos com essas práticas. Guba e Lincoln (1989) referem-se a essa avaliação como a primeira geração que pode legitimamente ser chamado a geração da medida.

## **OS PRIMÓRDIOS ESTUDOS DA AVALIAÇÃO**

Tyler é considerado o pai da avaliação. Tudo que se fala sobre ela é considerado antes e depois de Tyler que criticou especialmente a distância entre o que é ensinado e os objetivos de instrução.

São propostas de Tyler: desenvolvimento de taxonomias para formular objetivos, a diversificação das fontes, instrumentos e testes, a unificação de critérios de correção das provas, a revisão das avaliações por meio de métodos como por diferentes corretores.

Tyler foi o primeiro a dar uma visão metódica, indo além do behaviorismo, muito em voga na época, a mera avaliação psicológica. Ele levanta a necessidade de uma avaliação científica que serve para melhorar a qualidade da educação, indicando claramente sua ideia de "currículo" e integrando-o método sistemático de avaliação educacional. O "currículo" é delimitado pelas quatro questões seguintes: Que objetivos são desejados? Que habilidades podem ser desenvolvidas? Como você pode organizar de forma eficaz essas experiências? Como você pode verificar se os objetivos são alcançados?

Para Tyler a avaliação exige algumas condições, são elas: Clareza de objetivos, Identificação de situações que devem mostrar os comportamentos esperados, escolha de instrumentos de avaliação adequados, interpretação dos resultados do teste, Determinação da confiabilidade e objetividade das medidas. Contudo, essa avaliação não é uma medida

simples, porque é um juízo de valor sobre as informações coletadas. Para Tyler, a referência central na avaliação são as metas pré-determinadas, que devem ser cuidadosamente definidas, em termos de comportamento.

O propósito do processo de avaliação é determinar a aprendizagem nos alunos, mas também um meio para informar sobre a eficácia do programa educacional e a competência do professor. Este é, de acordo com Guba e Lincoln (1989), a segunda geração da avaliação.

Após a Segunda Guerra Mundial, um período de expansão e otimismo que Stufflebeam e Shinkfield (1987) não hesitaram em descrever como "irresponsabilidade social", é conhecida como a fase de inocência, neste momento, a contribuição da avaliação para melhorar a educação é escassa devido à falta de planos educacionais coerentes. Muito é escrito de avaliação, mas com pouca influência sobre a melhoria do trabalho educacional.

Ralph W. Tyler morreu em 18 de fevereiro de 1994, mais de 90 anos, depois de sete décadas de contribuições e serviços frutíferos para a avaliação, dentre tantas contribuições destacam-se a necessidade de analisar cuidadosamente o propósito da avaliação antes de começar a avaliar. O fator mais importante na avaliação dos alunos é orientar a sua aprendizagem, ou seja, ajudá-los a aprender; Para ele, o "portfolio" é uma ferramenta valiosa para a avaliação, mas depende de seu conteúdo. Em qualquer caso, devemos ser cautelosos com a preponderância de um processo de avaliação, incluindo o "portfolio" por sua incapacidade de cobrir todo o espectro de aspectos avaliáveis; A avaliação verdadeira deve ser idiossincrática, adequada às características do aluno e da escola. Os professores devem ser responsáveis pela interação frequente com os pais dos alunos.

Meio século depois de Tyler que revolucionou o mundo da avaliação educacional, suas ideias até hoje são fonte de inspiração nas discussões sobre a avaliação.

Da mesma forma que a proliferação de programas sociais na década anterior tinha solicitado a avaliação dos programas na área social, os anos sessenta seriam fecundos na procura de avaliação no domínio da educação. Esta nova dinâmica na avaliação centrada nos alunos valorizando o desempenho deles, prestando contas com o que era gasto no setor público.

A avaliação continua a verificar o desempenho do aluno, mas também todos aqueles fatores que convergem no processo educativo (o programa educacional em um sentido amplo, professor, mídia, conteúdo, experiências de aprendizagem, organização, etc.) e o próprio produto educacional. Surge a perspectiva da multidimensionalidade da avaliação, essa reflexão teórica, impulsiona estudos agora chamados de pesquisa em avaliação.

Como marcos da época destacamos dois ensaios para a sua influência decisiva: o artigo de Cronbach (1963). Na obra de Cronbach, encontramos as seguintes sugestões:

- a) A avaliação para Tomada de Decisão: decisão para melhoria do programa e instrução, decisão para reter ou promover os alunos e por fim, decidir sobre o regulamento administrativo, sobre a qualidade do sistema, professores, organização, etc.
- b) A avaliação é usada para melhorar um programa enquanto este está a ser aplicado, contribui mais para o desenvolvimento da formação de avaliação utilizado para estimar o valor de saída de um programa completo.
- c) Afirma a necessidade de uma avaliação com referência aos objetivos bem definidos e não a comparação com outros grupos.
- d) Não concorda com avaliação em larga escala, uma vez que as diferenças entre os processos podem ser muito grandes e impedem discernir claramente os resultados. Defende os estudos analíticos, que podem ser utilizados para comparar diferentes programas.
- f) A partir desta perspectiva, técnicas de avaliação não podem ser limitadas a testes de desempenho. São sugeridos questionários, entrevistas, observação sistemática e não sistemática, testes, e de acordo com o autor, desempenham um papel importante na avaliação. (p.673).

Se essas reflexões por Cronbach foram marcantes, não menos importante, o ensaio de Scriven (1967):

- a) A avaliação é uma atividade metodológica, independente do que estamos avaliando. O objetivo da avaliação é invariante, em última análise, é o processo pelo qual se estima o valor de algo que é avaliado, enquanto que as funções da avaliação podem ser muito variadas. Estas funções relacionadas com a utilização das informações coletadas.
- b) Aponta duas funções diferentes que a avaliação pode adotar: a formativa e somativa. Propõe-se o prazo de avaliação formativa para qualificar esse processo de avaliação a serviço de um programa de desenvolvimento, a fim de melhorar, e o prazo de avaliação somativa é o processo orientado a verificar a eficácia do programa e tomar decisões sobre a sua continuidade.
- c) Outra contribuição importante de Scriven é a crítica à ênfase dada à avaliação da verificação dos objetivos pré-determinados. Destaca a necessidade de a avaliação incluir tanto a avaliação dos objetivos a si mesmos como determinar a medida em que estes tenham sido alcançados. (Scriven, 1973 e 1974).
- d) Clareadora é também a distinção de Scriven entre a avaliação intrínseca e extrínseca, como duas formas diferentes de valorizar um elemento de educação. Em uma avaliação intrínseca do elemento é avaliado para si mesmo, enquanto a avaliação extrínseca o elemento é avaliado para os efeitos que tem sobre os alunos. (p.83)

Scriven adota uma posição contrária à Cronbach, defendendo o caráter comparativo a ser apresentado por estudos de avaliação. Ele admite problemas técnicos que envolvem estudos comparativos e a dificuldade de explicar a diferença entre os programas, mas Scriven considera que a avaliação em oposição à mera descrição implica um julgamento sobre a

superioridade ou inferioridade do que é avaliado com relação a concorrentes ou alternativas.

Estamos na terceira geração da avaliação, de acordo com Guba e Lincoln (1989), que é caracterizado pela introdução de avaliação, julgamento, como um conteúdo intrínseco na avaliação. Agora, o avaliador não se limita a analisar e descrever a realidade, ele também avalia e julga-lo em relação a diferentes critérios.

Durante os anos sessenta muitas outras contribuições moldaram uma nova concepção avaliativa, percebe-se que o núcleo conceitual da avaliação era a avaliação da aprendizagem do aluno como resultado de uma situação educacional sistemática. Além disso, ela começa a prestar atenção não só para os resultados pretendidos, mas também os efeitos colaterais, e até mesmo para resultados ou efeitos a longo prazo.

Stake (1967) propôs o seu modelo de avaliação, que segue a linha de Tyler, mas é mais completo quando se considera as discrepâncias entre o observado e esperado nos "contexto" e "transações" e permite desenvolver hipóteses sobre as causas e as falhas nos resultados finais.

Nas práticas de avaliação dos anos sessenta são observados dois níveis de atividade. Um nível pode ser classificado como a avaliação orientada para os indivíduos, fundamentalmente alunos e professores. O outro nível é o da tomada de decisão orientado sobre o "instrumento" ou "tratamento" ou "programa" avaliação educacional. Este último nível, também impulsionado pela avaliação dos programas na área social, será a base para consolidação no campo educacional da avaliação de programas e pesquisa de avaliação.

Se algo poderia caracterizar as contribuições durante os anos setenta é a proliferação de todos os tipos de modelos de avaliação que inundam o mercado bibliográfico, modelos de avaliação que expressam variadas formas de conduzir um processo avaliativo.

Neste movimento de propostas modelo de avaliação, podemos distinguir duas fases com diferenças conceituais e metodológicas: Em um primeiro momento as propostas foram expostas por Tyler em sua abordagem, o que veio a ser chamado de "Conquista de Metas", outros modelos consideram o processo de avaliação para atender instâncias que devem tomar decisões. Os exemplos notáveis destes são provavelmente o mais famoso e amplamente utilizado de todos o CIPP (de contexto, insumo, processos e produto), proposto por Stufflebeam e colaboradores.

Na segunda metade dos anos setenta surgem a avaliação responsável de Stake, que Guba e Lincoln (1982) concordam porque em geral, este segundo grupo de modelo avaliativo enfatiza o papel da audiência da avaliação. A avaliação do público-alvo é prioritária nestes

modelos orientados para tomada de decisão, e responsável pelo currículo ou os objetivos de desenvolvimento. A relação entre o avaliador e o público nas palavras de Guba e Lincoln (1989) deve ser "transacional e fenomenológica." Estes são modelos que defendem uma avaliação etnográfica.

Este resumo de modelos do período de incubação é suficiente para abordar o vasto leque conceitual teórico e metodológico que hoje está relacionada com a avaliação. São eles: a clássica definição dada por Tyler, avaliação como o processo de determinar o grau de congruência entre as formas de realização e os objetivos previamente estabelecidos. Esta definição contrasta com o modelo orientado para tomada de decisão: avaliação como o processo de identificação, obtenção e fornecimento de informações relevantes para julgar decisões alternativas, por Stufflebeam (1971) e Cronbach (1982). Além disso, o conceito de avaliação Scriven (1967) e o processo de estimar o valor ou o mérito de algo, que é retomado por Cronbach (1982) e Guba e Lincoln (1982). A partir das definições acima salientado que os critérios utilizados para avaliar as informações também mudam, e do ponto de vista da realização dos objetivos, uma boa definição de trabalho dos objetivos constitui o critério fundamental.

Enfim, a pluralidade de objetos de avaliação. Como Nevo (1983) diz, existem duas importantes conclusões extraídas da revisão da literatura sobre a avaliação. Por um lado, qualquer coisa pode ser avaliada e isso não deve ser limitada a alunos e professores e, por outro, uma identificação clara do objeto da avaliação é uma parte importante de qualquer projeto de avaliação. A pluralidade também reconheceu as funções de avaliação na educação, recolhendo a proposta de Scriven entre a avaliação formativa e somativa, e adicionando outro tipo sócio-política e administrativa (Nevo, 1983).

No final dos anos oitenta, depois de todo esse desenvolvimento descrito acima, Guba e Lincoln (1989) oferecem uma alternativa de avaliação que chamou de quarta geração, tentando superar o que segundo esses autores são deficiências das três gerações anteriores, eles criticam a gestão da avaliação, valorizam mais o pluralismo de valores e criticam bastante o positivismo. A alternativa de Guba e Lincoln é amparada no construtivismo, que se integra de alguma forma a proposta de Stake (1975), onde as demandas, preocupações e questões das partes interessadas servem como foco organizacional de avaliação (como uma base para determinar que informação é necessária), que é realizado dentro das posições metodológicas do paradigma construtivista.

É necessário o uso de reivindicações, preocupações e questões envolvidas, de acordo com Guba e Lincoln, porque eles são grupos de risco para a avaliação onde seus problemas devem ser devidamente observados, para que eles se sintam protegidos contra tal risco.

Essa mudança de paradigma é justificada por esses autores, considerando que a metodologia padrão não contempla a necessidade de identificar as demandas, preocupações e questões envolvidas, deseja-se uma descoberta, em vez de verificação, típica do positivismo; Considera fatores contextuais onde são realizados métodos diferentes para cada caso e a suposta neutralidade da metodologia convencional é de utilidade duvidosa quando juízos de valor são olhados a partir de um objeto social.

Nesta base, o avaliador é responsável por determinadas tarefas, que sequencialmente ou em paralelo, há construção de um processo de trabalho ordenada e sistemática. As responsabilidades básicas do avaliador da quarta geração são a de identificar todos os envolvidos com a avaliação de risco, atenção para cada grupo envolvido, suas concepções sobre o que foi avaliado e suas demandas e preocupações, além de proporcionar um contexto e uma metodologia hermenêutica para compreender e criticar as diferentes exigências e preocupações e gerar o maior acordo de possíveis exigências em tais preocupações. Organizar uma agenda para a negociação sobre as questões não acordadas, coletar e fornecer as informações necessárias para a negociação, formar e mediar um "fórum" de negociação com os envolvidos, desenvolver e elaborar relatórios para cada grupo envolvido em diversos acordos e resoluções e refazer a avaliação sempre que houver pendências.

A proposta de Guba e Lincoln (1989) se estende bem para explicar a natureza e as características do paradigma construtivista, em oposição à positivista, ao falar das etapas ou fases de avaliação nesta quarta geração, seus proponentes mencionam doze etapas ou fases, com diferentes subfases em cada uma delas. E, para avaliar a qualidade da avaliação, são definidos os critérios de credibilidade, agência de transferência e de confirmação (Lincoln e Guba, 1986). Outra forma de avaliar a qualidade da avaliação é a análise do próprio processo, que se encaixa com o paradigma hermenêutico, através de um processo dialético.

Esta análise da quarta geração, podemos sintetizar com as características com que Guba e Lincoln definem avaliação: a avaliação é um processo sócio-político, o processo de avaliação é um conjunto de colaboração, a avaliação é um processo de ensino / aprendizagem, o processo de avaliação é contínuo, recursivo e muito divergente, a avaliação é um processo emergente, a avaliação é um processo com resultados imprevisíveis, a avaliação é um

processo que cria a realidade.

Russell e Willinsky (1997) defendem as potencialidades da abordagem da quarta geração para desenvolver formulações alternativas de avaliação prática entre os envolvidos, aumentando a probabilidade de que a avaliação servirá para melhorar a educação escolar. Isto requer dos professores o reconhecimento de outras posições, além de sua própria, a implicação de todos, desde o início do processo e, além disso, o desenvolvimento de abordagens mais pragmáticas para a conceituação de Guba e Lincoln, adaptadas aos diferentes realidades escolares.

Para completar essa jornada analítico-histórica das primeiras tentativas de avaliação educacional para pesquisa avaliativa atual na educação, queremos reunir as recomendações que vêm, mais recentemente, oferecendo uma das figuras imponentes deste campo na segunda metade do século XX. Estamos nos referindo a Daniel L. Stufflebeam, proponente do modelo CIPP (o mais utilizado) até o final dos anos sessenta. O autor invoca a responsabilidade do avaliador, que deve agir de acordo com princípios aceitos pela sociedade e critérios profissionais, fazendo julgamentos sobre a qualidade e o valor educativo do objeto avaliado e deve ajudar as pessoas envolvidas na interpretação e uso de suas informações e julgamentos.

Para avaliar a educação em uma sociedade moderna, Stufflebeam (1994) nos diz que eles devem tomar algumas referências básicas, tais como: as necessidades educacionais. É necessário perguntar se o ensino ministrado atende às necessidades dos alunos e suas famílias em todas as áreas, à luz dos direitos fundamentais.

Mesmo no campo formativo (Scriven, 1991), os estudos de avaliação devem ser capazes de avaliar em que medida os professores e instituições de ensino são responsáveis e responsável pelo cumprimento de seus deveres e responsabilidades profissionais, estudos de avaliação deverão fornecer instruções para a melhoria, não é suficiente fazer um juízo sobre o mérito ou valor de alguma coisa coletando os fatos anteriores, todo o estudo avaliativo deve ter uma componente de avaliação formativa e uma somativa.

A avaliação do contexto (necessidades, oportunidades, problemas em uma área...) deve ser utilizado de forma prospectiva para localizar as metas e objetivos e definir bem as prioridades. Sendo assim, o processo de avaliação deve ser utilizado de modo prospectivo para melhorar o plano de trabalho, mas também de forma retrospectiva para avaliar em que medida a qualidade do processo de determinada porque os resultados são de um nível ou outro (Stufflebeam e Shinkfield de 1987).

A avaliação do produto é o meio para identificar os resultados intencionais e não intencionais no participante ou afetados pelo objeto avaliado, estudos de avaliação devem apoiar a comunicação e a inclusão substantiva e funcional daqueles envolvidos (stakeholders) com as principais questões, critérios, conclusões e implicações da avaliação, bem como na promoção da aceitação e utilização de resultados.

Por fim, estudos de avaliação devem empregar as várias perspectivas múltiplas medidas de resultados, ambos os métodos quantitativos e qualitativos para coletar e analisar informações. A pluralidade e complexidade do fenômeno educativo sugere usar abordagens múltiplas e multidimensionais em estudos de avaliação e, estudos de avaliação devem ser avaliados, incluindo meta-avaliações para melhorar a sua qualidade e seu uso e meta-avaliações somativas para auxiliar os usuários na interpretação de seus resultados e fornecer sugestões para melhorar futuras avaliações .

## **CONSIDERAÇÕES FINAIS**

Após esta análise do desenvolvimento da avaliação ao longo do século XX, afigura-se adequado, em jeito de resumo e conclusão, recolher e destacar aqueles que são considerados os principais modelos, abordagens metodológicas, projetos, perspectivas e visões de avaliação no hoje.

Stake, na sua primeira fase, é tyleriano relativo aos objetivos centrados, propondo o método de avaliação, como um quadro total da avaliação. Em meados dos anos setenta, Stake se afasta da tradição tyleriana de preocupação com os objetivos e revisa seu método de avaliação em direção a uma posição que ele qualifica como "sensível", assumindo que os objetivos do programa podem modificados, a fim de fornecer uma visão completa e holística do programa e responder aos problemas reais e as questões levantadas pelos envolvidos. De acordo com Stufflebeam , este modelo feito Stake é o líder de uma nova escola de avaliação que requer uma abordagem pluralista, flexível, interativo, holística, subjetiva e orientada a serviços. Este modelo sugere "serviço ao cliente", proposto por Scriven (1973), avaliando suas necessidades e expectativas.

Quanto aos conteúdos da avaliação, Cronbach (1983) propôs que são planejados e controlados os seguintes elementos: Unidades (U) que estão sujeitas a avaliação, os indivíduos ou grupos envolvidos, Tratamento (T) da avaliação, Operações (O) que realizou a

avaliação para a recolha e análise de dados, e para tirar conclusões e Contexto em que o programa e a sua avaliação ocorre. Assim, em uma pesquisa de avaliação específico pode ser dado várias unidades, vários tratamentos e várias operações, em última análise, vários (UTO), dentro de um universo de situações UTO aceitável.

A visão construtivista da quarta geração é o último destas visões sendo mantido por Guba e Lincoln (1989) e seguido por muitos avaliadores americanos e britânicos. Nós já vimos que esta visão rejeita a avaliação da qualidade de pesquisa orientada, mérito, valor, etc., e promove a ideia de que este é o resultado de uma construção por indivíduos e grupos de negociação.

Scriven (1994) analisa criticamente as visões e se aproxima do modelo Stufflebeam CIPP e suas abordagens, ele diz que é o mais próximo de toda visão do significado comum, pois para ele este modelo que chamou transdisciplinar e que ele qualifica como significativamente, é radicalmente diferente dos outros.

Na perspectiva transdisciplinar a pesquisa de avaliação tem dois componentes: o conjunto de campos de aplicação da avaliação e do conteúdo da própria disciplina. Algo semelhante ao que acontece com disciplinas como estatísticas e medição. Em suma, a pesquisa de avaliação é uma disciplina que inclui o seu próprio conteúdo e muitas outras disciplinas; análise e melhoria preocupação se estende a muitas disciplinas, tornando-se transdisciplinar.

Outra característica distintiva da visão transdisciplinar de avaliação é que é uma visão técnica. Avaliação não só requer suporte técnico a partir de muitas outras disciplinas, mas também tem sua própria metodologia. Esta perspectiva transdisciplinar de pesquisa de avaliação Scriven (1994), em grande parte em linha com a abordagem que apenas têm defendido em outros momentos (Escudero, 1996). Não temos um conflito com outros pontos de vista na mesma medida que as posições Scriven e, de fato, considerar a partir de uma posição pragmática que todas as visões têm pontos fortes e que em qualquer caso, fornecer algo de útil para a compreensão conceitual e desenvolvimento de pesquisas de avaliação. No entanto, nós pensamos que esta visão moderna de Scriven é sólida e consistente e amplamente aceita hoje.

## REFERENCIAS BIBLIOGRÁFICAS

- Escudero E. T. (2003). **Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de inteso desarrollo de la evaluación en educación.** RELIEVE:, v. 9, n. 1, p. 11-43. Disponível em [http://www.uv.es/RELIEVE/v9n1/RELIEVEv9n1\\_1.htm](http://www.uv.es/RELIEVE/v9n1/RELIEVEv9n1_1.htm)
- Cronbach, L. J. (1963). **Course improvement through evaluation.** Teachers College Record, 64, 672-683.
- Gil, E. (1992). **El sistema educativo de la Compañía de Jesús.** La «Ratio Studiorum». Madrid: UPCO.
- Guba, G. E. y Lincoln, Y. S. (1982). **Effective evaluation.** San Francisco: Jossey Bass Publishers.
- Guba, E. G. y Lincoln, Y. S. (1989). **Fourth Generation Evaluation.** Newbury Park, Ca.: Sage
- Nevo, D. (1983). **The conceptualization of educational evaluation:** An analytical review of the
- 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. (1991, **Evaluation Thesaurus.** Newbury Park , Ca.: Sage
- Stufflebeam, D. L. y Shinkfield, A. J. (1987). **Evaluación sistemática.** Guía teórica y práctica. Barcelona: Paidós/MEC.
- Tyler, R. W. (1950). Basic principles of curriculum and instruction. Chicago: University of Chicago Press.**