

EVOLUÇÃO CONCEITUAL DA AVALIAÇÃO EDUCACIONAL: UM PERCURSO INTERNACIONAL

WIRLA RISANY LIMA CARVALHO

FACED/UFC/Bolsista CAPES/DS. E-mail: wirlar@gmail.com

ALINE MARIA GOMES LIMA

FACED/UFC/Bolsista CAPES/OBEDUC E-mail: alinegomeslima@yahoo.com.br

JOSÉ AIRTON DE FREITAS PONTES JÚNIOR

FACED/UFC/BolsistaCAPES/PROPAG. E-mail: joseairton.junior@yahoo.com.br

Introdução

A Avaliação Educacional tem sua proto-história enfatizada sob o conceito de medida como fundamento da avaliação. Nesse transcurso o avaliador visto como sujeito principal foi responsabilizado pela medição do conhecimento e por um julgamento de valor (CARVALHO, 2012; ESCUDERO, 2003; VIANNA, 2000).

O advento da teoria de Ralph Tyler (1902-1994) inaugurou os termos “avaliação educacional” e “avaliação da aprendizagem”, da mesma forma que fez com o exame, a seleção e sobre a própria avaliação, mantendo-a com a mesma denominação. Tecendo naquele momento uma visão equivocada, no entanto, primordial, em vista do que se tornou mais adiante na história da avaliação, o que chamamos de “cultura de avaliação” (CARVALHO, 2012; LUCKESI, 2011; VIANNA, 2000).

Carvalho (2012, p.23) aborda essa questão trazendo que:

Essa constatação não ocorre apenas no cenário brasileiro, mas também mundialmente, uma vez que a compreensão teórica da avaliação tem crescido e as inovações pedagógicas vêm sendo aplicadas desde o século XVI até o momento presente. Entretanto, a prática pedagógica da avaliação ainda não foi efetivamente modificada, pelos educadores, no cotidiano escolar.

Entender, portanto, a “[...] construção do conceito de avaliação, para que fique bem definido cada passo da avaliação da apren-

dizagem, com o objetivo de se concretizarem práticas avaliativas mais concatenadas com o desenvolvimento pleno do aprendente.” é importante nesse momento (CARVALHO, 2012, p.23).

Delinear esse percurso histórico dos fundamentos conceituais da Avaliação Educacional é fundamental auxílio para o entendimento da dimensão e das reflexões atuais sobre a Avaliação do ensino-aprendizagem “[...] suas práticas ainda rudimentares até os conceitos mais complexos, que são hoje abordados no contexto mundial e nacional.” (CARVALHO, 2012, p.23).

Intenciona-se, portanto, como objetivo deste trabalho, discutir sobre o caminho conceitual da avaliação e os novos rumos no decorrer da História da Educação em favor de uma prática de desenvolvimento pleno do educando e do próprio educador.

Para tanto, no cenário internacional, foram identificadas formas de avaliação no decurso da História: da Antiguidade à Idade Média, durante o Renascimento, assim como o nascimento formal da teoria da avaliação por volta do final do século XIX e início do século XX, legados como a Psicometria, com suas escalas contribuindo para as percepções e estudos sobre o desempenho escolar dos estudantes, passando pela Docimologia, até a criação mais específica e firmada pelos teóricos da avaliação de algo que se delineou com o tempo uma cultura (CARVALHO, 2012; VIANNA, 2000).

Inicia-se, pois, essa história da avaliação pelo contexto mundial em que ela apresentou suas primeiras formulações conceituais. A avaliação, assim como o conhecimento do homem, está constantemente em modificação, pois deve acompanhar para exercer um juízo de valor, todas as etapas da humanidade. No entanto, aqui nesse trabalho, parte-se de sua formalização como área de conhecimento.

O Início das Concepções Avaliativas no Contexto Mundial

A sistematização da avaliação foi formalizada em meados de 1930, nos Estados Unidos com Ralph Tyler (1902-1994), embora a

história da avaliação date desde à Idade Antiga, quando não haviam sistemas teóricos que pudessem embasar uma prática pedagógica. A China imperial foi destaque nesse contexto, utilizando-se da avaliação exclusivamente para seleção de funcionários de alto escalão nos anos 3.000 a.C. A Grécia e Roma atuou mais de forma pedagógica dando ao professor essa tarefa de avaliar. E na Idade Média, em pleno Renascimento, já se defendia uma ideia de avaliação que se utilizava de exames orais nas universidades e que trazia a observação como procedimento importante para fins de seleção.

Esses seriam os fatos históricos antigos mais citados como pressupostos para uma teorização da avaliação nascida por volta do final do século XIX e início do século XX, quando houve um aumento da demanda educacional e um maior acesso às instituições de ensino, fatores provenientes da Revolução Industrial (CARVALHO, 2012; ESCUDERO, 2003).

A classificação e a aferição de valores nesse início conceitual, superficial à primeira vista é, no entanto, claramente importante para o contexto histórico da avaliação. Tornou-se, desta feita, uma preocupação iminente, pois não havia uma abordagem reflexiva até então sobre

[...] um processo avaliativo que privilegiasse o envolvimento de todos os agentes educativos. A avaliação era um procedimento realizado para se obter uma classificação. Desse modo, não era importante pensar sobre o aluno, o professor ou qualquer agente educacional envolvido na seleção de pessoas, seja na escola, seja em âmbitos governamentais (CARVALHO, 2012, p. 25).

A Psicometria Rumo à Avaliação Educacional atual

Sobre esses primeiros passos, Alfred Binet (1857-1911), psicólogo francês, no início do século XX, a partir de 1905, influenciou os Estados Unidos e a Inglaterra criando o primeiro teste psi-

cométrico de inteligência, a escala Binet-Simon. Pois, mostrou-se descontente com os testes inteligência centrados em estímulos sensoriais e habilidades simples até então utilizados (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

A *Psicometria* vem nessa época como campo de estudos sobre a medida de inteligência e foi inicialmente idealizada em relação às medidas educacionais, gerando repercussões na avaliação da aprendizagem dos alunos e em seu rendimento escolar (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

Em seus trabalhos, Binet (1857-1911) tinha como objetivo indicar uma medida ou um nível de inteligência após identificação e estabelecimento de relações e diferenças individuais entre alunos, comparando-as ao desempenho de crianças classificadas como normais em relação à sua faixa etária. Ainda com relação aos seus estudos sobre educação de crianças com atraso intelectual, Binet criou, juntamente com seu colega Theodore Simon (1873-1961), três testes de inteligência, denominados depois de escalas Binet-Simon, datadas em 1905, 1908 e 1911 (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

Sobre essas três escalas, Carvalho (2012, p.25-26) expõe que:

A primeira escala, em 1905, foi considerada um instrumento preliminar e experimental, que consistia em trinta problemas (testes) organizados em ordem crescente de dificuldade e tinha como finalidade avaliar algumas funções complexas do intelecto, com ênfase nas áreas de julgamento, compreensão e raciocínio. Predominavam os testes de conteúdo verbal, todavia também havia testes sensoriais e perceptuais em proporções menores. Essa escala inicial não possuía nenhum método objetivo preciso de obtenção de um escore total.

A escala de 1908 trouxe como diferencial um número maior de testes, a eliminação de alguns outros testes considerados insatisfatórios da escala inicial e o agrupamento das questões (problemas) em níveis de idade. Esse último procedi-

mento possibilitou que o escore da criança testada fosse expresso como um *nível mental*, correspondente à idade das crianças normais, obtidas como desempenho padrão para comparação.

Enfim, a terceira versão da escala, em 1911, não apresentou modificações essenciais, apenas foram revisados os testes e acrescentados outros específicos, além de aumentar o número de testes para cada nível de idade. A única particularidade dessa escala é o fato dela ter sido ampliada para o nível adulto. Como vemos, Binet avançou com seus estudos se os compararmos à visão de avaliação utilizada nos primórdios para classificação e seleção de pessoal.

Gardner (2001, p.23) afirma sobre os testes de inteligência de Binet, que este “começou com perguntas de base sobretudo sensorial, mas logo descobriu a superioridade do poder de previsão de outras questões, mais ‘acadêmicas’”. E que “Desde a época de Binet, os testes de inteligência são elaborados principalmente para aferir a memória verbal, o raciocínio verbal, o raciocínio numérico, a apreciação de sequências lógicas e a capacidade de dizer como resolver problemas do cotidiano” (Op. cit., p. 23).

Tendo como base os estudos de Binet, em 1912, sob a proposta do psicólogo alemão Wilhelm Stern (1871-1938) nasceu o Quociente de Inteligência (QI), que consistia em dividir a idade mental pela idade cronológica. Contudo, só foram concretizadas instruções específicas e criadas normas que padronizaram os testes para aplicação a um grande número de pessoas simultaneamente com o psicólogo americano Lewis Terman (1877-1956) e o professor de Harvard, Robert Yerkes (1876-1956) (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

Essa padronização facilitava que os resultados fossem aferidos através de máquinas específicas para esse propósito psicometricamente mais refinada, a Escala Binet-Simon foi superada pela Stanford-Binet, proposta por Terman e utilizada ainda hoje. Os resultados desta última escala eram apresentados na forma de

QI, onde a razão entre a idade mental e a cronológica passou a ser multiplicada por 100, no intuito de evitar que os valores fossem números fracionados como acontecia com o QI de Stern (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

Carvalho (2012) apresenta que a Psicometria acreditava que a inteligência era transmitida geneticamente, portanto, o teste de QI acabou propagando a ideia de invariabilidade da própria inteligência. Gerando negativamente uma concepção de inteligência de natureza inata e estática, favorecendo preconceitos e classificações, pois os indivíduos que se encontravam abaixo da média, eram tidos como deficientes e incapazes de melhorar sua aprendizagem, fadados ao insucesso escolar e social. Da mesma forma, os outros “acima da média” eram valorizados e tinham todas as oportunidades de crescimento intelectual e social aumentadas. Essa visão sobre o QI foi muito contestada ao longo das décadas após sua criação.

Carvalho (2012, p. 27) afirma que os alunos eram “medidos” na escola, até então, por testes psicométricos, que avaliavam seus desempenhos em habilidades próprias e convencionadas na época, sendo classificados abaixo, iguais ou acima do QI tido como ideal:

Segundo uma distribuição normal, quando a idade mental equivale à idade cronológica, o resultado da divisão é igual a 1 e o QI igual a 100, indicando um intelecto mediano. A normalidade ou desempenho conforme a média varia dentro de um intervalo de vinte e nove pontos em relação a esse valor-para mais ou para menos: o indivíduo que apresentar QI entre 71 e 129 é considerado de inteligência mediana ou normal. Já os resultados situados abaixo e acima desse intervalo demarcariam pontos de corte para identificar intelectos situados fora da média: QI igual ou menor a 70 indica deficiência intelectual e QI igual ou superior a 130, altas habilidades/superdotação (ANASTASI; URBINA, 2000; GARDNER, 2001).

Dentro do todo exposto, vimos que a ideia de fracasso escolar era uma responsabilidade transferida inteiramente para o aluno e não para o sistema educacional, nem para a escola e muito menos

para o professor (CARVALHO, 2012; ESCUDERO, 2003; VIANNA, 2000).

Contradizendo essa visão na atualidade, Anastasi (1977, p. 251), afirma que “QI é uma expressão do nível de habilidade de um indivíduo, num determinado momento com relação a suas normas de idade [...] o QI não é fixo e invariável; e é suscetível a modificações por intervenções ambientais”.

Carvalho (2012, p.27) dessa forma assevera que “isso quer dizer que a visão original de QI, ao longo dos tempos, perdeu força; hoje os testes de QI são utilizados para medição de habilidades acadêmicas, e não representam mais um conceito ou medida da capacidade total de inteligência de um indivíduo.” Sendo hoje entendida como uma capacidade que pode se modificar de acordo com a idade e o universo biopsicossocial do indivíduo, apresentando diferentes níveis durante a vida de uma pessoa (ANASTASI; URBINA, 2000; CARVALHO, 2012; GARDNER, 2001).

Analisando a evolução conceitual da avaliação, o período referido à Psicometria demonstra que os testes avaliativos se voltam ao aluno, porém responsabilizando-o plenamente por seus sucessos e fracassos na aprendizagem, numa perspectiva apenas genética, em que não havia qualquer consideração do professor ou da escola como agentes colaboradores do desenvolvimento cognitivo e do rendimento escolar. Comparada à visão de avaliação dos tempos remotos, houve certo progresso com a Psicometria, embora o olhar fadado à medição e a uma inteligência de caráter inato ainda estivesse fortemente presente.

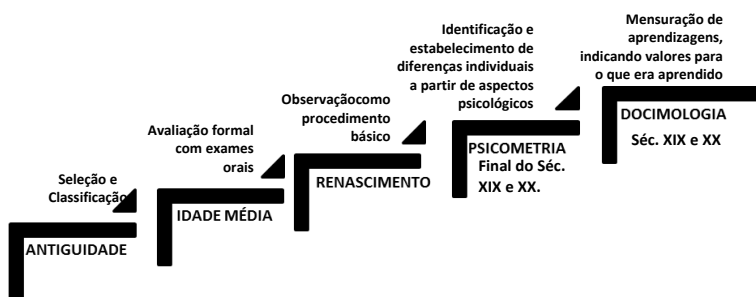
A Docimologia como Ciência dos Exames

ADocimologia, em 1922, nasceu na França, como uma corrente que trouxe uma visão mais metódica da avaliação. Superou os conceitos trabalhados pelos testes psicométricos, que até então não dispunham de recursos para avaliar os sistemas educacionais,

nem tampouco as aprendizagens dos discentes de uma forma mais adequada, correlacionando todos os agentes envolvidos no processo educativo (CARVALHO, 2012).

Esse estudo tornou-se a ciência dos exames e, com isso, foi considerada a primeira corrente a se aproximar de um modelo de Avaliação Educacional, no momento em que se utilizou de métodos e instrumentos estruturados para a mensuração de aprendizagens, que indicassem o valor do que era aprendido pelos alunos (ver figura 1) (CARVALHO, 2012; ESCUDERO, 2003; HADJI, 2001; MIRANDA, 1980).

Figura 1 – Evolução conceitual segundo os períodos históricos



Fonte: CARVALHO, 2012, p.30.

Dessa forma, Carvalho (2012, p.27) lembra que “a Docimologia trouxe uma visão diferenciada entre o que era ensinado e o que era meta educacional, assim como fomentou novos questionamentos acerca do papel do professor como detentor absoluto da avaliação”.

Dentro dessa visão, novos questionamentos acerca do processo avaliativo no contexto escolar instigavam os estudiosos da época. Tyler (1930), no entanto, lançou-se à exploração dessa área logrando grande êxito.

Os Pais da Avaliação Educacional: de Tyler a Stufflebeam

Por sua grande contribuição, ainda está presente na cultura pedagógica norte-americana e em países como o Brasil e, portanto, como inovador para a época, Tyler é considerado o “pai da avaliação educacional” (TYLER 1942 apud VIANNA, 2000).

Tyler, a partir de 1930, propôs um olhar diferenciado para avaliar a instituição como um todo, sob a visão de que os objetivos de um programa de Avaliação Educacional deviam ser estabelecidos em função de um todo integrado. Em sua visão, a avaliação deveria ser retirada de um patamar secundário que a limitava à atribuição de notas aos estudantes e à elaboração de relatórios aos pais (CARVALHO, 2012; VIANNA, 2000).

Para Tyler, seus fundamentos teóricos deveriam orientar melhor o processo educativo, através de uma avaliação baseada na idealização e realização dos objetivos educacionais, ou seja, utilizando-se desta para uma aplicação de um currículo bem estruturado, em verdade, reavaliando o currículo e seus objetivos. Outro intuito da avaliação deste estudioso seria promover uma segurança a todos os envolvidos no processo educacional, quanto à sua eficiência em promover a aprendizagem. Dessa forma, sua concepção extrapolou as atividades relacionadas aos testes, medidas e avaliação, gerando grande influência na Educação em muitos países (CARVALHO, 2012; VIANNA, 2000).

Nasua concepção de avaliação, Tyler (1942 apud VIANNA, 2000) enfatizou a importância da relação professor-aluno no processo avaliativo. Suas proposições consideravam que o professor devia esclarecer aos alunos o que pretendia deles em relação às práticas curriculares em sala de aula. Em vista disso, os alunos seriam motivados a promover comportamentos mais adequados ao processo de Avaliação Educacional, em conformidade com os objetivos estabelecidos. “Para Ralph Tyler, a Educação seria, portanto, um processo capaz de criar ou modificar padrões de conduta nos indivíduos” (CARVALHO, 2012, p. 28).

Diante de todo o exposto, infere-se que “[...] esta deve apresentar uma diversidade instrumental para avaliar múltiplos comportamentos – como conhecimentos, habilidades, modos de pensar, atitudes e interesses – não se concentrando apenas em elementos isolados.”(CARVALHO, 2012, p. 28). Nessa visão, a avaliação não deveria trabalhar apenas com os alunos, mas, de forma solidária, também com todos os agentes envolvidos no processo educativo, como os professores, pais e, inclusive, os administradores (TYLER, 1942 apud VIANNA, 2000).

Viu-se, portanto, que em 1930, a avaliação ganhou mais complexidade conceitual e aproximou-se das bases utilizadas atualmente no cenário mundial e nacional. Essa visão de envolvimento apresentada por Tyler foi primordial nos conceitos educacionais hoje, não só na área de avaliação, mas em todas as instâncias da instrução escolar e das instituições que a pensam e executam (CARVALHO, 2012; SAUL, 1995).

Dando continuidade à evolução conceitual da avaliação, Cronbach (1963 apud VIANNA, 2000) trouxe, como conceito de avaliação, uma atividade diversificada, que exigiria uma série de tomadas de decisão, utilizando-se de um grande número de diferentes informações. Em 1960, formatou reflexões acerca da avaliação com vistas ao aprimoramento do currículo, de cursos, do sistema de ensino e dos professores, também apresentando grandes contribuições à temática histórica da Avaliação Educacional.

No intuito de possibilitar maior entendimento aos alunos, Cronbach (1963 apud VIANNA, 2000, p. 68-69) assinala que a avaliação pode proporcionar as seguintes decisões: i) a determinação dos métodos de ensino e o material instrucionaisão realmente eficientes; ii) a identificação das necessidades dos alunos, para possibilitar o planejamento da instrução; julgamento do mérito dos estudantes, para fins de seleção e agrupamento; fazer com que os estudantes conheçam seu progresso e suas deficiências; e iii) julgamento da eficiência do sistema de ensino e dos professores, entre outras coisas.

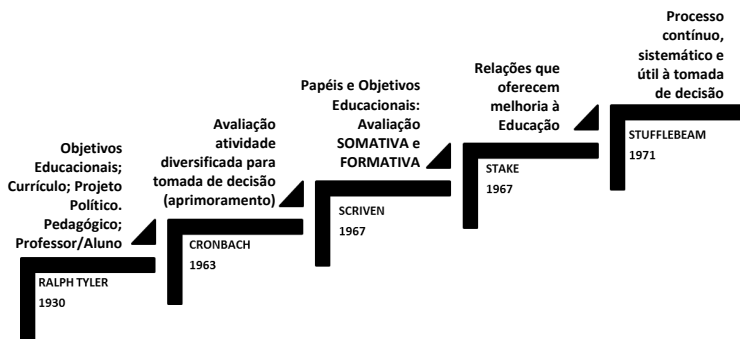
Importante ressaltar, de acordo com Carvalho (2012) que o acima exposto remete ao que é importante hoje para a avaliação do ensino-aprendizagem, enfatizando algumas características a serem observadas no desenvolvimento de uma boa ação pedagógica em sala de aula. Credo que a avaliação é utilizada em favor de uma aprendizagem mais significativa e da própria eficiência do sistema de ensino e dos professores.

Em 1967, Scriven fundamentou que a avaliação desempenhava muitos papéis (*roles*), mas que possuía apenas um objetivo, o de determinar o valor ou o mérito do que estava sendo avaliado. Sua maior contribuição foi de fato diferenciar papéis (*roles*) e objetivos (*goals*) no contexto educacional, diferenciando, dessa forma, dois tipos de avaliação: a *somativa*, realizada ao final do processo educativo, e a *formativa*, executada durante esse processo (VIANNA, 2000).

Conferindo maior importância à avaliação formativa, Scriven mostrou que esta devia ocorrer ao longo do desenvolvimento do programa, projetos e produtos educacionais, gerando informações úteis para o aprimoramento do que estava sendo implantado e, tendo como objetivo maior, realizar modificações necessárias no decorrer do processo educativo (SCRIVEN, 1967 apud VIANNA, 2000).

Em 1967, Stake formatou que a avaliação seria uma procura das relações que permitiriam uma melhoria da Educação. Assim, apresentou diversas críticas às práticas convencionais da época e, da mesma forma, Stufflebeam (1971) ressaltou que a avaliação é um processo contínuo, sistemático e que serve para a tomada de decisões, bem como esclareceu que a avaliação não queria provar, mas trazer melhorias à Educação (STAKE, 1967; STUFFLEBEAM, 1971 apud VIANNA, 2000).

Figura 2 – Evolução conceitual da teoria da avaliação no mundo



Fonte: CARVALHO, 2012, p. 30.

Conforme a gravura acima, Carvalho (2012) constatou que o modelo de Tyler foi incrementado de diversas formas, pois consistia basicamente numa análise de dados. Nas visões subsequentes de Cronbach, Scriven, Stake e Stufflebeam que foram além, até mesmo questionando a própria Docimologia, pode-se notar que elevaram a avaliação a um elemento de aprimoramento educacional e considerando-a imprescindível à tomada de decisões (VIANNA, 2000).

Considerações Finais

Indubitavelmente, todo o exposto nesse artigo, foi um grande passo para o desenvolvimento dos conceitos atuais da avaliação de ensino-aprendizagem. Os autores abordados acima são sem contradição grandes colaboradores da temática de Avaliação Educacional sob uma perspectiva de maior colaboração da avaliação para a aprendizagem em sala de aula e num contexto mais amplo (CARVALHO, 2012).

A avaliação formativa, nessa visão, com caráter contínuo e processual, foi um legado desse percurso histórico conceitual e tem sido apontada como uma prática a ser seguida, na atualida-

de, para obtenção de qualidade efetiva no processo pedagógico. Devendo-se evitar, contudo, na visão de Carvalho (2012, p.31) “[...] que se torne um instrumento de controle dos alunos, pois essa presença constante, gradual e contínua pode transformar-se em um disciplinamento exagerado, tornando-se, com isso, tão aversiva quanto a avaliação tradicional.” (AFONSO, 2000; HADJI, 2001; PERRENOUD, 1999).

A avaliação apesar de ainda existir como um instrumento de mensuração e notas é considerada como elemento componente do ato pedagógico, colaborando para a realização de objetivos educacionais e da melhoria na qualidade da aprendizagem dos alunos no decorrer do cotidiano escolar. Conclui-se, pois, que a avaliação cresceu, indubitavelmente, na própria formulação de sua teoria ao longo de seu processo de construção histórica, ainda estando esse caminho sendo trilhado de forma dinâmica e não finalizado (CARVALHO, 2012; LUCKESI, 2011).

Referências Bibliográficas

- AFONSO, A. J. *Avaliação educacional: regulação e emancipação*. São Paulo: Cortez, 2000.
- ANASTASI, A.; URBINA, S. *Testagem psicológica*. Porto Alegre: Artmed, 2000.
- CARVALHO, W.R.L. *Da alvorada ao ocaso: estudo de caso sobre a avaliação da aprendizagem em cursos de pedagogia diurno e noturno em uma instituição de Ensino Superior da rede pública*. 2012. 148f. Dissertação (Mestrado em Educação Brasileira) – Universidade Federal do Ceará, Ceará, 2012.
- ESCUADERO, E. T. Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación em educación. *Relieve*, Zaragoza, v. 9, n. 1, p. 11-43, 2003.
- GARDNER, H. *Inteligência: um conceito reformulado*. Rio de Janeiro: Objetiva, 2001.

- HADJI, C. *Avaliação desmistificada*. Porto Alegre: Artmed, 2001.
- LUCKESI, C. C. *Avaliação da aprendizagem*: componente do ato pedagógico. São Paulo: Cortez, 2011.
- MIRANDA, M. J. C. *A docimologia em perspectiva*. 1980. 71f. Tese (Doutorado em Psicologia Diferencial)– Universidade de Lisboa, Lisboa, 1980.
- PERRENOUD, P. *Avaliação*: da excelência à regulação das aprendizagens – entre duas lógicas. Porto Alegre: Artmed, 1999.
- SAUL, A. M. *Avaliação emancipatória*: desafio à teoria e à prática de avaliação e reformulação de currículo. São Paulo: Cortez, 1995.
- VIANNA, H. M. *Avaliação educacional*: teoria, planejamento, modelos. São Paulo: IBRASA, 2000.