

AVALIAÇÃO EDUCACIONAL: ALGUMAS IDÉIAS PRECURSORAS

Heraldo Marelím Vianna *

1.0. Introdução: – Tyler e Cronbach

O conceito de avaliação associado à verificação dos objetivos de um programa educacional estruturou-se, basicamente, a partir dos trabalhos de Ralph W. Tyler, na década de 40. A sua obra teve grande impacto sobre a teoria e a prática da avaliação educacional, influenciando alguns modelos, como, por exemplo, o de Hammond (s.d.) e o de Metfessel e Michael (1967), e grandes projetos de avaliação, como o *National Assessment of Educational Progress*, nos anos 70. O pensamento nuclear de Tyler acha-se no artigo – *General Statement on Evaluation* –, que se tornou um clássico no gênero e continua a ser de validade no presente, apesar de publicado em 1942.

Lee J. Cronbach em seu artigo – *Course improvement through evaluation* (1963) – ainda que não apresente um modelo de avaliação, oferece um conjunto de idéias altamente provocadoras que tiveram grande impacto na década de 60, influenciando trabalhos como o de Scriven (1967) e o de Stake (1967), que, por sua vez, repercutiram, profundamente, na prática da avaliação educacional. O ensaio de Cronbach, publicado em 1963, é de grande valor e discute, sobretudo, quatro pontos: – a associação entre avaliação e o processo de tomada de decisão; os diferentes papéis da avaliação educacional; o desempenho do estudante como critério de avaliação de cursos; e, finalmente, analisa algumas técnicas de medida à disposição do avaliador educacional.

As idéias de Tyler (1942) e Cronbach (1963) são apresentadas de forma resumida no presente trabalho, que se baseia, amplamente, nos artigos desses dois educadores norte-americanos, procurando ser, na medida do possível, fiel ao pensamento dos mesmos.

*Do Departamento de Seleção de Recursos Humanos da Fundação Carlos Chagas.

2.0. Avaliação educacional – a abordagem de Tyler

As múltiplas atividades dos educadores, no dia-a-dia de suas várias e complexas atividades, nem sempre permitem que considerem, adequadamente, os objetivos de um programa de avaliação educacional, que devem ser estabelecidos em função da educação como um todo integrado, com vistas a abarcar todo o sistema educacional. A avaliação, entretanto, na maioria das vezes, fica limitada a aspectos secundários do processo, ainda que importantes na sua individualidade. Assim é que atribuir notas a estudante, elaborar relatórios aos pais sobre seu progresso, entre outras, constituem importantes aspectos do trabalho de avaliação, sem dúvida, mas um programa de avaliação deve visar a alcançar objetivos bem mais amplos.

Tyler (1942), há vários decênios, estabeleceu uma linha de abordagem, hoje clássica, que merece ser analisada e debatida amplamente, numa perspectiva crítica. Assim, a avaliação objetiva, em primeiro lugar, verificar, periodicamente, a eficiência de uma instituição como agente promotora da educação. A avaliação, portanto, propicia elementos para uma crítica institucional, uma discussão sobre a sua atuação, um debate sobre os seus próprios fundamentos teóricos, que orientam o processo educativo. A coleta de elementos através da avaliação permitiria, pois, aprimorar o programa institucional, reformulando-o naquilo em que fosse deficiente ou inoperante, na concepção de Tyler.

A elaboração de um currículo, e sua posterior adoção por uma instituição educacional, parte de uma seqüência de hipóteses. A sua estruturação baseia-se em crenças e valores que, supostamente, a comunidade acadêmica acredita sejam um reflexo de uma comunidade maior em cujo contexto se insere a instituição educacional. A validação dessas hipóteses seria, por conseguinte, o segundo objetivo de um programa de avaliação. A comunidade acadêmica supõe que o currículo estruturado seja um reflexo das aspirações, crenças e valores que predominam na cultura de uma sociedade e imagina, igualmente, que a sociedade julgue perfeitamente satisfatório o conjunto curricular. Age, pois, a comunidade acadêmica, conforme Tyler (1942), a partir de suposições, de hipóteses que precisariam ser verificadas e validadas. A avaliação permitiria estabelecer a congruência entre as hipóteses que fundamentaram o currículo e a realidade do meio social que ele procura refletir.

As instituições educacionais têm como um de seus objetivos maiores, pelo menos supostamente, a orientação do estudante. Ora, isso somente é possível se, por intermédio da avaliação, forem levantadas informações sobre o seu desempenho, o seu crescimento e desenvolvimento. A avaliação para esse fim não se pode limitar a algumas dimensões do indivíduo, mas deve ser suficientemente abrangente de modo a caracterizar todos os aspectos realmente significantes do seu desempenho, a fim de que, com precisão, possa ser orientado quanto ao seu progresso e às dificuldades que por acaso venha enfrentando.

Outro objetivo da avaliação, na discussão de Tyler (1942), refere-se à segurança de todos os envolvidos no processo educacional quanto à eficiência do mesmo em promover a aprendizagem. Alunos, professores, pais e administradores necessitam de informações que traduzam, inequivocamente, que a instituição está alcançando os seus objetivos educacionais. A avaliação sistemática permite eliminar qualquer tipo de insegurança que se apossa da comunidade, tornando-a mais confiante no êxito de seu currículo e de suas atividades programáticas.

A escola, e, em geral, toda e qualquer instituição educacional, existe no contexto de uma comunidade social. Uma não pode ignorar a outra: – a escola, porque se tornaria alienada ao ignorar a sociedade em que se situa; a comunidade, porque se veria impedida de garantir a continuidade de seus valores, tão importantes quanto a transmissão do patrimônio cultural na sobrevivência da cultura social. Assim, a avaliação objetivaria, igualmente, acentuar o nexo que existe entre a escola e a sociedade, exercendo uma função de relações públicas, por intermédio da divulgação de seus resultados, que informariam à comunidade sobre a excelência da escola e de seu programa.

A avaliação objetivaria, ainda, oferecer elementos para que alunos e professores possam esclarecer de modo insofismável os seus reais propósitos, o que pretendem efetivamente conseguir. O programa de avaliação visaria a permitir que professores e estudantes definissem mais claramente os seus objetivos: – o professor poderia demonstrar o que espera conseguir dos alunos; estes, por sua vez, ficariam habilitados a apresentar suas exigências em relação à escola. Um programa de avalia-

ção deve, por conseguinte, resultar de um esforço cooperativo para que possa atingir plenamente as suas finalidades.

A estruturação de qualquer programa de avaliação, diante do que foi brevemente discutido até agora, segundo a abordagem de Tyler (1942), deve ser precedida de um amplo debate sobre os seus objetivos, pois somente após a sua caracterização é que outras decisões serão tomadas a fim de estabelecer o que vai ser avaliado, qual a tecnologia a empregar no processo de avaliação e, finalmente, como sintetizar e interpretar os resultados de todo o processo de avaliação.

Um programa de avaliação deve, necessariamente, basear-se em alguns pressupostos, que orientarão todo o seu desdobramento. Ainda dentro da linha geral de Tyler (1942), podemos resumir da forma seguinte:

1º) A educação é um processo que visa a modificar os padrões de conduta dos indivíduos. É meridianamente claro que há uma diferença entre os que sofreram o processo educacional e os que não tiveram essa possibilidade, sendo essa diferença, presumivelmente, uma decorrência da experiência educacional. As mudanças decorrentes da educação são modificações na maneira como o indivíduo reage, ou seja, são mudanças na sua maneira de comportar-se. Somente na medida em que teriam ocorrido modificações, o processo educacional demonstraria o seu êxito.

2º) A instituição educacional pretende conseguir diferentes mudanças no padrão de comportamento, as quais são, na verdade, seus objetivos educacionais. O programa educacional não se pode limitar apenas a conteúdos curriculares, a metodologias de ensino ou, então, a procedimentos de ensino, pois são apenas meios para a concretização de objetivos específicos, sendo estes últimos as mudanças desejadas pelo processo educacional; assim, os programas educacionais devem apresentar as modificações a serem alcançadas expressas em termos de comportamentos, a fim de que possam ser adequadamente avaliadas.

3º) Um programa educacional é avaliado pela determinação do êxito em concretizar os objetivos previamente estabelecidos. Se considerarmos, como foi anteriormente apresentado, segundo Tyler (1942), que os programas educacionais visam a promover mudanças nos comportamentos dos alunos e que estes comportamentos são, fundamentalmente, os objetivos educacionais, depreende-se, por conseguinte, que a avaliação de um programa educacional visa a estabelecer quais as mudanças que efetivamente ocorreram no estudante.

4º) Um outro pressuposto importante, em qualquer programa de avaliação, diz respeito à maneira como o estudante organiza os padrões de comportamento. A tendência mais comum consiste em avaliar partes isoladas de comportamentos, sem estabelecer relações entre os mesmos. Em princípio, os programas procuram verificar as mudanças que o estudante sofreu nas suas informações, nas suas habilidades, nos diferentes modos de pensar, nas suas atitudes e interesses, mas o comum é verificar cada uma dessas dimensões isoladamente, sem tentar relacioná-las, tendo em vista que o ser humano reage de modo mais ou menos unificado.

5º) Os métodos de avaliação geralmente se fixam em exames escritos, quando, na verdade, é válido, para fins de avaliação, todo e qualquer instrumento ou maneira que permita verificar a situação do estudante em relação à concretização dos objetivos educacionais. Um dos pressupostos da avaliação baseia-se no fato de que há necessidade de uma diversidade de técnicas para medir e avaliar diferentes tipos de comportamentos.

6º) A avaliação não é um processo isolado, destaca Tyler (1942), concentrado exclusivamente no estudante. Exige um esforço cooperativo de professores, estudantes e pais, a fim de que se possa extrair um máximo de proveito do programa. Isso porque todos participam da fixação dos objetivos, procuram evidências que demonstrem o progresso do estudante e, finalmente, têm interesse na interpretação dos resultados.

A prática de um programa de medidas exige, naturalmente, diversos procedimentos, que devem ser estruturados de maneira adequada para que válidos e precisos sejam os seus resultados. Tyler (1942) esquematizou o problema da seguinte forma:

1º) formulação e classificação dos objetivos, segundo seus níveis de generalidade e especificidade;

2º) definição de cada objetivo em termos de comportamento;

- 3º) identificação de situações que demonstrem os comportamentos apresentados nos objetivos;
- 4º) seleção e experimentação de métodos adequados para coleta de informações sobre cada tipo de objetivo (ou construção de novos instrumentos, caso não existam instrumentos já prontos e pré-testados);
- 5º) elaboração de critérios para interpretação e uso dos resultados dos vários instrumentos.

3.0. Avaliação educacional segundo a perspectiva de Cronbach

O especialista em currículo frequentemente solicita a colaboração de elementos na área da avaliação; entretanto, o entendimento entre ambos nem sempre é perfeito. Além disso, muitas vezes aquele especialista não possui uma idéia perfeita do tipo de colaboração que o avaliador lhe poderá proporcionar. Por outro lado, conforme Cronbach (1963), muitas técnicas e concepções teóricas da avaliação não se ajustam perfeitamente aos problemas de currículo.

A avaliação, no seu sentido mais amplo, pode ser definido como um processo que visa à coleta e ao uso de informações que permitam tomar decisões sobre um programa educacional. A avaliação, portanto, segundo Cronbach (1963), deve ser entendida como uma atividade diversificada, que exige a tomada de diversos tipos de decisões e o uso de uma grande variedade de informações. A avaliação com vistas ao aprimoramento de currículos não deve ser confundida, como muitos o fazem, com a construção de instrumentos de medida e a obtenção de escores fidedignos, processos que, eventualmente, podem entrar no contexto da avaliação, mas que não são indispensáveis para que ela possa atingir os seus objetivos.

Cronbach (1963) mostra que a avaliação é usada com o objetivo de tomar três tipos de decisões:

1. determinar se os métodos de ensino e o material instrucional utilizado no desenvolvimento de um programa são realmente eficientes;
2. identificar as necessidades dos alunos para possibilitar o planejamento da instrução; julgar o mérito dos estudantes para fins de seleção e agrupamento; fazer com que os estudantes conheçam seu progresso e suas deficiências; e
3. julgar a eficiência do sistema de ensino e dos professores etc.

Assim, de acordo com Cronbach (1963), no primeiro caso, a avaliação permitiria decisões que levariam ao aperfeiçoamento do currículo; no segundo, referir-se-ia aos alunos submetidos a determinado programa; e, finalmente, no terceiro e último caso, as decisões seriam de natureza administrativa.

Cronbach (1963), em seu artigo, discute, particularmente, alguns pontos fundamentais que não podem ser ignorados pelo avaliador educacional. Inicialmente, enfatiza o seguinte: — *quando a avaliação visa ao aprimoramento de cursos, seu principal objetivo é verificar quais os efeitos do curso, ou seja, quais as mudanças que produz no estudante*. O problema, segundo a perspectiva de Cronbach, não está em determinar se um curso é eficiente ou ineficiente apenas. É preciso lembrar que os resultados da instrução são multidimensionais e, desse modo, a avaliação deve promover o mapeamento de todos os efeitos do curso em cada uma de suas dimensões. Um erro freqüente está na concentração em um único escore de diversos desempenhos esperados após a realização de um curso. Isso pode ser enganador, pois um insucesso numa dimensão pode ser compensado pelo sucesso em outra. Os escores compostos englobam, e muitas vezes ocultam, julgamentos sobre a importância de vários resultados; desse modo, para fins de avaliação, é importante que os resultados sejam apresentados separadamente, a fim de que se tenha uma idéia real das mudanças que estariam ocorrendo no estudante como decorrência da influência exercida pelo currículo.

A avaliação presta um grande serviço quando identifica os aspectos do curso que necessitam de revisão, no dizer de Cronbach (1963). É evidente que qualquer especialista em currículo gostaria de apresentar evidências sobre a eficiência do seu produto, entretanto, conforme observa Cronbach, costumam relutar na aceitação de uma avaliação externa. O procedimento habitual consiste

em submeter o produto à avaliação somente depois de terminado, com vistas a uma confirmação do que foi previamente estabelecido. Esse comportamento não proporcionará bons resultados, além de traduzir um menosprezo pelo importante papel que o avaliador pode efetivamente desempenhar. A fim de apresentar um papel influente, a avaliação deve acompanhar o desenvolvimento do currículo, quando o mesmo ainda se acha em estado fluido, para usar palavras de Cronbach (1963). A avaliação, assim entendida, possibilita a criação de conhecimentos sobre a natureza das capacidades que constituem os objetivos educacionais do projeto.

A comparação de cursos não deve ser o objetivo dominante da avaliação, na concepção de Cronbach (1963). A comparação de resultados de avaliações de cursos deve ser cautelosa para evitar que decisões errôneas sejam tomadas. As diferenças entre os escores médios de diferentes cursos, geralmente, costumam ser pequenas, em virtude da grande diferença que há entre e intra-grupos submetidos ao mesmo curso. A impossibilidade de equiparar (*equate*) diferentes grupos prejudica a interpretação dos resultados e representa um problema difícil de superar nos estudos comparativos de cursos. Além do mais, ainda conforme Cronbach (1963), em experimentos educacionais é difícil ocultar dos estudantes o fato de que integram um grupo experimental, sendo igualmente complexo o controle do viés dos professores numa situação experimental. Ocorrem mudanças comportamentais e, assim, nem sempre se pode afirmar com convicção que um determinado resultado decorre, efetivamente, da própria inovação ou se decorre do fato de alunos e professores terem sido colocados diante de uma situação nova e experimental.

A comparação entre grupos pode oferecer resultados equívocos; desse modo, Cronbach (1963) propõe que estudos formais sejam planejados sobretudo para determinar o desempenho após o curso por um grupo perfeitamente conhecido, a fim de verificar objetivos importantes e a ocorrência de efeitos colaterais. Cronbach (1963) chama a atenção para o fato de que um experimento em que os tratamentos comparados diferem em inúmeros aspectos, a ocorrência de uma pequena diferença numérica em favor da situação nova não significa grande coisa e não contribui para o aumento de nossos conhecimentos. Estudos analíticos, em escala menor, mas em condições controladas, de versões alternativas de um mesmo curso, oferecem melhores resultados do que pesquisas de campo que aplicam tratamentos dissimilares a grupos diferentes.

Um problema posto em destaque por Cronbach (1963) refere-se aos procedimentos de medida, que variam em função do tipo de decisão a ser tomada. O instrumento usado para uma decisão administrativa deverá ser compreensivo e possuir todas as características psicométricas desejáveis, o mesmo ocorrendo para decisões individuais; contudo, para a avaliação de um curso, o emprego da técnica de amostragem de itens (*item sampling*) oferece elementos suficientes para um julgamento preciso.

A avaliação de um curso é freqüentemente identificada com a aplicação de um teste; no entanto, conforme acentua Cronbach (1963), existem diferentes maneiras de analisar o desempenho de um estudante, e este desempenho, por sua vez, não constitui o critério único para a avaliação de um curso. A avaliação pode ser realizada por intermédio de levantamentos de opiniões, relatórios de professores usuários do curso (observação de comportamento), observações sistemáticas, medidas ou estudos do processo, medidas de proficiência, medidas de atitudes e estudos de seguimento (*follow-up*), destaca Cronbach em seu artigo.

O levantamento de opiniões de especialistas sobre um determinado curso constitui um procedimento perfeitamente legítimo, ainda que implique diferentes problemas, sobretudo se forem influenciadas por preconceitos, o que, naturalmente, dará margem a controvérsias. O posicionamento em relação a muitas metodologias de ensino pode ser polêmico, sendo necessária a sua experimentação para conhecimento de seus efeitos educativos. A opinião de diferentes especialistas pode servir, segundo Cronbach (1963), para avaliar esses efeitos.

As opiniões manifestas por professores, durante o desenvolvimento de um currículo, são importantes, pois graças aos seus relatórios, que traduzem observações de comportamentos, positavam-se os pontos críticos, as deficiências e todos os demais elementos que estão a exigir maiores cuidados, talvez substituição ou mesmo supressão total da parte comprometida. A observação do professor, ainda que não seja de natureza sistemática, reveste-se de grande importância para a ava-

liação educacional (Cronbach, 1963).

A observação sistemática é mais imparcial e, freqüentemente, consegue apresentar diferentes aspectos com maior profundidade; contudo, apresenta, também, alguns problemas, como o de custos e o referente à demora que geralmente introduz entre o momento do ensino e a análise dos resultados. Este tipo de observação é particularmente útil durante a fase intermediária de desenvolvimento do currículo, quando as deficiências mais evidentes já foram eliminadas.

O estudo do processo refere-se aos eventos que ocorrem em sala, sendo, pois, de especial valor para o aprimoramento do curso. Nesse estudo, o registro de eventos não se limita àqueles relativos ao estudante durante a fase de instrução, procura, também, enfocar o procedimento do professor em sala de aula. Este conjunto de atividades dos alunos e dos professores possibilita ao construtor de currículos diagnosticar deficiências e propor modificações.

A medida da proficiência, ressalta Cronbach (1963), utiliza de várias técnicas. O uso da amostragem de itens — itens diferentes para alunos diferentes — oferece mais vantagens do que o emprego dos mesmos itens para todos os estudantes, na avaliação de cursos. Este tipo de medida pode utilizar-se, ainda, de questões de redação, de questões de resposta aberta curta e, se for o caso, de observação de indivíduos ou de grupos na abordagem de problemas de pesquisa ou na tentativa da solução de um problema complexo.

Cronbach (1963) destaca o fato de considerar mais importante os dados relativos a um item do que os escores do teste. O escore global pode dar confiança ou não em relação a um determinado currículo, mas pouco informa como aprimorá-lo. Outro aspecto importante em avaliação de currículo, conforme mostra Cronbach (1963), refere-se à ajustagem do instrumento de medida ao currículo, que não deve ser motivo de preocupação para o avaliador. Ainda que isso possa parecer surpreendente, porque em outras situações isso não deve ocorrer, na avaliação de currículo, contudo, e em termos ideais, todos os tipos desejáveis de proficiência devem ser medidos e não apenas os de alguns objetivos selecionados pelo construtor do currículo. Se há um interesse em saber se o currículo alcança os seus objetivos, impõe-se o ajustamento do teste ao currículo, mas se a intenção é a de verificar até que ponto o currículo atende a interesses mais amplos, o ajustamento não é necessário, tendo em vista o desejo de que sejam verificados todos os objetivos possíveis.

Outro ponto enfocado por Cronbach (1963) diz respeito à distinção entre testes factuais e testes para verificar processos mentais complexos, segundo a terminologia de Bloom. A classificação dos itens de acordo com as categorias da Taxonomia é difícil e às vezes impossível. Se uma resposta traduz conhecimento ou raciocínio depende de como o aluno foi ensinado e não apenas da questão apresentada. Aplicar um teste apenas para verificar se o aluno “sabe” ou “não sabe” um certo assunto não é inteiramente relevante para fins de avaliação de um curso, conforme Cronbach (1963), importando, isto sim, medir o conhecimento em termos de profundidade, relacionamento a capacidade de aplicá-lo a novas situações.

A questão da especificidade das questões é considerada por Cronbach (1963), que discute, detalhadamente, o problema. A especificidade, na maior parte das vezes, concentra-se no uso de uma terminologia própria do curso, que somente é compreendida por aqueles que tiveram a oportunidade de assisti-lo. Ainda que o conhecimento dessa terminologia seja importante, para fins de avaliação, entretanto, é mais importante a medida de compreensão, de relação e de outras variáveis do curso, que, em princípio, poderiam ser verificadas, também, em quem não foi diretamente submetido ao curso em questão.

A transferência de aprendizagem constitui um dos principais objetivos de qualquer currículo novo, ou seja, o desenvolvimento da capacidade de aprender novos elementos numa área específica. Ainda que não seja fácil, como reconhece Cronbach (1963), construir instrumentos que meçam a capacidade de “aprender a aprender”, esse é um aspecto que não pode ser ignorado pelo avaliador educacional.

Voltando, novamente, aos diferentes modos de avaliar um curso, Cronbach (1963) discute, ainda, as medidas de atitude e os estudos do tipo *follow-up*.

A formação de atitudes parece ser uma preocupação constante de todos os responsáveis por desenvolvimento de currículos, pois a expectativa é a de que cada atividade relacionada à aprendi-

zagem contribua para o desenvolvimento de atitudes além da área curricular específica. A mensuração de atitudes pode ser feita de diferentes modos: entrevistas, questionários e outras técnicas semelhantes, sem dúvida valiosas, mas cujos resultados, como adverte Cronbach (1963), devem ser aceitos com as devidas cautelas, porque sujeitos a distorções.

Os estudos de *follow-up* possuem pontos de contacto com as técnicas de observação, no entanto, conforme Cronbach (1963), muitas vezes se mostram de pouca valia para fins de aprimorar um curso ou explicar seus efeitos, em virtude do distanciamento temporal da fase de instrução inicial. Outro problema no *follow-up* é que este avalia os efeitos de um curso como um todo, não sendo possível detectar, por exemplo, desvios da situação ideal, diferenças na eficiência das várias partes do curso, ou, ainda, diferenças de um item a outro item do programa, no dizer de Cronbach (1963). A necessidade, às vezes impossível, de igualar os grupos experimentais e os de controle, constitui outro problema nos estudos de *follow-up*, que, apesar de apresentarem esses pontos críticos, são importantes, especialmente para fins de pesquisas sobre a construção de novos currículos.

4.0 Conclusões

Tyler apresenta uma abordagem centrada na cuidadosa formulação de objetivos, que devem ser definidos a partir de uma cuidadosa consideração de três variáveis: – o estudante, a sociedade e a área de conteúdo, de acordo com os princípios da psicologia da aprendizagem e uma filosofia da educação bem definida.

Os objetivos educacionais devem ser apresentados de forma a que possam ser quantificados; assim, sua apresentação deve refletir comportamentos terminais, que serão verificados ao término de um programa a fim de constatar a congruência entre os objetivos idealizados e os efetivamente alcançados, o que refletirá o sucesso de um programa educacional.

O trabalho de Cronbach permite várias conclusões gerais, destacando-se as seguintes:

1. a avaliação educacional requer a descrição de resultados; dessa forma, certas preocupações das medidas educacionais para a produção de escores precisos visando a comparar indivíduos, ou a comparação de escores médios de diferentes cursos, pouco contribui para a avaliação educacional, cuja descrição dos resultados deve ser a mais ampla possível, ainda que às custas do sacrifício de uma suposta justiça e precisão;
2. a avaliação educacional deve estabelecer quais as mudanças realmente ocorridas no estudante por influência do curso, e identificar os aspectos deste mesmo curso que precisam ser revistos;
3. a análise do desempenho em itens isolados ou em certos tipos de problemas fornece mais informações do que a análise de escores compósitos;
4. o emprego do mesmo instrumento para todos os estudantes, no caso da avaliação educacional, não constitui uma necessidade, podendo-se, para esse fim, usar a amostragem de itens, em que diferentes estudantes responderão a diferentes itens;
5. o objetivo da avaliação educacional não consiste em simplesmente aquilatar o valor de cursos, rejeitando-os ou aceitando-os, mas, sim, em ser uma parte fundamental no processo de desenvolvimento de currículos, através da coleta e do uso de dados que possibilitem uma compreensão mais profunda do processo educacional.

5.0 REFERÊNCIAS BIBLIOGRÁFICAS

- CRONBACH, L.J. (1963) – Course improvement through evaluation. *Teachers College Records*, 64. Columbia University.
- HAMMOND, D.L. (s.d.) – Evaluation at the local level. Tuckson, Arizona. EPIC Evaluation Center.

- METFESSEL, N.S. e MICHAEL, W.B. (1967)** – A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs. *Educational and Psychological Measurement*, 27.
- SCRIVEN, M. (1967)** – The methodology of evaluation, in Stake, R.E. (Ed.) – *Curriculum evaluation*. AERA monograph series on evaluation nº 1. Chicago, Rand McNally.
- STAKE, R.E. (1967)** – The countenance of educational evaluation. *Teachers College Record*, 68.
- TYLER, R.W. (1942)** – General statement on evaluation. *Journal of Educational Research*, 35.