

# NOTAS SOBRE A AVALIAÇÃO DE PESQUISAS SÓCIO-EDUCACIONAIS (\*)

(Subsídios para a discussão sobre avaliação de pesquisas  
educacionais, em reunião promovida pelo CNPq).

JACQUES ROCHA VELLOSO

Universidade de Brasília, Associação Nacional  
de Pós-Graduação em Educação (ANPED).

## ADVERTÊNCIA E INTRODUÇÃO

Esse texto foi originalmente preparado para trabalhos preliminares de um projeto de avaliação de pesquisas sócio-educacionais, a ser desenvolvido pelo INEP-MEC. Seu conteúdo está voltado para um projeto específico, e difere do documento original somente em algumas modificações de forma e uma ou outra de conteúdo. Mas pareceu-nos que estas notas poderiam trazer algum subsídio para uma reunião em que se debaterá a avaliação de pesquisas educacionais no Brasil.

O texto inicialmente trata das dificuldades da avaliação de pesquisas, de dificuldades na definição de critérios para

---

\* Agradecemos os comentários de Oscar Serafini a uma versão anterior deste trabalho. As imperfeições que permaneçam são, contudo, de nossa responsabilidade.

essa avaliação e discute um esquema de avaliação. Num segundo momento, aborda problemas específicos de um roteiro ("pauta") de avaliação de pesquisas, elaborado por uma instituição financiada pela OEA e adaptada pelo INEP. Como alguns problemas específicos desse roteiro são ilustrativos de problemas mais gerais na avaliação de pesquisas, manteve-se no texto esta segunda parte do trabalho; sugere-se ao leitor um rápido exame do roteiro (em anexo).

Ao pensarmos na avaliação de projetos ou de trabalhos de pesquisa sócio-educacionais, uma das primeiras perguntas que surge é "avaliação para que?". Ao sugerirmos uma resposta para esta pergunta, deixaremos de lado as pesquisas custeadas por recursos privados. Ao pesquisador, como indivíduo, cabe decidir sobre o custo de oportunidade de seu próprio tempo. A este indivíduo e às entidades privadas cabe julgar da rentabilidade profissional e/ou monetária do investimento de seus recursos. Voltamo-nos então para as pesquisas financiadas com recursos públicos. A resposta à pergunta "avaliação para que?" baseia-se na suposição de que a avaliação de projetos de pesquisa, propostos ou em desenvolvimento, é operacional e viável. Aceita esta suposição, a resposta em si depende em parte do tipo de avaliação considerada.

### *Finalidades da Avaliação de Pesquisas*

Ao considerarmos uma proposta de pesquisa, desejamos saber, através da avaliação, se ela é: a) viável e b) se atende a determinados critérios preestabelecidos. Quando consideramos uma pesquisa em desenvolvimento, desejamos saber se a) ela está sendo desenvolvida de modo eficaz — se promete atingir seus objetivos — e b) se está sendo conduzida de modo eficiente — se as atividades realizadas permitem atingir os objetivos com o menor dispêndio de recursos<sup>1</sup>

<sup>1</sup> Entende-se recursos no sentido econômico e amplo da palavra: quantidade/tempo de recursos humanos e quantidade de recursos materiais, os quais traduzem-se em custos diretos e indiretos.

possível. Se pensarmos numa pesquisa já realizada, podemos desejar saber se ela merece ser publicada: foi "bem feita?", constitui contribuição para a ciência?, etc. Caso a pesquisa já tenha sido publicada, as duas perguntas acima também constituem-se num dos objetivos de avaliação. Podemos estar pensando em usar seus resultados para fins práticos como, por exemplo, para a política educacional. Em qualquer dos casos acima citados, temos em mente uma decisão a ser tomada quanto à pesquisa proposta, em desenvolvimento ou realizada.

Em suma, a avaliação serviria para tomar uma decisão quanto à pesquisa. Esta decisão tem geralmente por finalidade satisfazer a alguma necessidade pessoal ou institucional. Analisemos a necessidade institucional.

A decisão a ser tomada pode referir-se à concessão ou negação de recursos financeiros para a pesquisa, para continuar o financiamento já concedido ou para custear sua publicação. Estas decisões têm que pautar-se por um ou mais critérios, como se pode depreender dos exemplos enumerados acima. Além disso, tratando-se de instituições, espera-se que estes critérios sejam objetivos, isto é, que possam ser utilizados por indivíduos diversos numa ou mais instituições e, quando aplicados a uma mesma pesquisa, produzam resultados semelhantes senão iguais.<sup>2</sup>

### *Critérios para a Avaliação de Pesquisas*

A definição objetiva deste critérios constitui talvez o calcanhar de Aquiles da avaliação de pesquisas. Na verdade,

<sup>2</sup> Uma definição mais rigorosa de objetividade dos critérios a serem aplicados é aquela sugerida por Kerlinger para um método de observação objetivo. Tal método satisfaz à condição de que seja mínima a variância dos observadores (ou avaliadores, na nossa análise). Esta condição é preenchida quando é nula a variância devido às diferenças na atribuição de numerais a objetos, por parte de observadores diferentes. Estendendo a definição de Kerlinger à nossa análise, um critério seria objetivo quando os resultados de sua aplicação a um mesmo caso, por diferentes observadores, tivesse a variância igual a zero. Veja-se Fred Kerlinger, *Foundations of Behavioral Research*, 2.ª ed. N. York, Wiley, 1973, pág. 491.

esta definição objetiva provavelmente é a questão fundamental em qualquer processo de avaliação e mormente na avaliação de pesquisas, como ficará claro a seguir.

A avaliação de pesquisas é uma exigência natural do processo decisório, como o é a avaliação do desenvolvimento e dos resultados de qualquer atividade de uma instituição. Este processo decisório é, por sua vez, o cerne das atividades de administração de qualquer instituição. A boa administração recomenda que as decisões sejam tomadas tendo em vista os objetivos da instituição.<sup>3</sup> No caso de uma empresa privada, seus objetivos podem ser os de maximizar os lucros (ou de aumentar sua renda bruta). No caso de uma instituição pública, esperar-se-ia que seus objetivos fossem derivados de alguma função de bem-estar social. Como a definição não subjetiva desta função não é viável, conforme nos ensina a teoria econômica, contentemo-nos, para discutir a avaliação da pesquisa educacional, com uma função cujo conteúdo fosse definido subjetivamente. Faltaria ainda, entretanto, definir a forma e a ponderação dos elementos que fariam parte desta função. Sejamos, no momento, modestos, e suponhamos que alguém já definiu esta função. Estão assim definidos os objetivos gerais e específicos que devem cumprir as instituições públicas.

Estes objetivos indicam a direção que devem tomar os gastos públicos, a fim de que eles sejam socialmente rentáveis. Trata-se, portanto, de escolher entre gastos (ou investimentos) alternativos que se apresentam ao decisor. Uma escolha racional seria aquela que se baseasse no critério de maior rentabilidade social. Supondo que fosse possível medir esta rentabilidade social, a tarefa do decisor seria então simples quando lhe apresentassem gastos alternativos. A dificuldade na verdade está, como em qualquer análise custo-benefício, na mensuração dos custos e sobretudo dos benefícios.

<sup>3</sup> Daí a necessidade de que sejam avaliados projetos de pesquisa. O decisor necessita, entre outras informações, saber se o financiamento de determinada pesquisa atende aos objetivos de sua instituição.

No caso de gastos com pesquisas, é extremamente difícil — senão impossível — traduzir os objetivos em unidades monetárias, de modo que os custos de um projeto de pesquisa possam ser comparados a seus benefícios. Portanto, na avaliação da pesquisa educacional, seria necessário definir padrões através dos quais pudéssemos comparar os custos de diferentes projetos de pesquisa com os benefícios que estes projetos trariam. Para tanto seria necessário, antes de mais nada, definir o que consideraríamos como benefícios.

Poderíamos considerar como benefícios as contribuições que os resultados do projeto trariam para o desenvolvimento da ciência na área. Faltaria então definir objetivamente como medir os diversos graus de importância destas contribuições. Ou poderíamos considerar como benefícios as contribuições que estes resultados trariam para a política educacional em geral. Os objetivos desta política, por sua vez, teriam que ser definidos *a priori*, através de nossa função de bem-estar social: seria uma política voltada para a eficiência interna do sistema educacional (produzir mais graduados com menos recursos) ou para a eficiência externa do sistema (graduados “mais competentes profissionalmente”, cidadãos “mais integrados na sociedade”, homens que “se realizassem mais plenamente como pessoas humanas” — e como avaliar isto?). Seria uma política voltada para objetivos de eqüidade (“mais igualdade de oportunidades educacionais” — como definí-la?) ou para objetivos de eficiência e também de eqüidade, e com que ponderação relativa para cada um destes grupos de objetivos? Supondo que os objetivos da política educacional estivessem definidos operacionalmente e que nos restasse apenas avaliar em que medida determinada pesquisa facilitaria a consecução destes objetivos, ainda assim teríamos à frente uma tarefa de difícil realização.<sup>4</sup>

<sup>4</sup> Embora a tarefa possa ser realizável para determinadas pesquisas, em relação a determinados objetivos da política educacional, é pouco provável que isto possa ser obtido para toda e qualquer pesquisa e para todos os objetivos de uma política educacional. O raciocínio do texto e desta nota também é aplicável à relação pesquisa/contribuição para a ciência.

Mesmo que esta tarefa fosse realizável, faltaria ainda definir objetivamente como *relacionar* estes benefícios com os custos de cada projeto. Por outras palavras, seria necessário que pudéssemos, para cada projeto, comparar seu custo com o número de unidades de benefícios associadas a cada projeto, supondo que cada unidade tivesse igual valor. Ou então que pudéssemos ponderar cada unidade de benefícios, em cada projeto, segundo determinados critérios que posteriormente nos permitissem dividir a soma dos benefícios pelo custo do projeto. É claro que tal tarefa é irrealizável.

### *Um Esquema de Avaliação*

Talvez pudéssemos ser mais modestos — ou mais realistas — e esquecer-nos do desejo de imprimir objetividade total ao trabalho de elaboração e de posterior aplicação de uma metodologia abrangente de avaliação de trabalhos de pesquisa. Nesta ordem de idéias, poderíamos dividir o processo de avaliação em três níveis.<sup>5</sup> No primeiro nível do processo talvez predominasse a avaliação subjetiva do especialista, tendo por base sua experiência e os objetivos da instituição financiadora. Procuraríamos então imprimir maior objetividade ao segundo nível deste processo. No terceiro nível predominaria, mais uma vez, a perspectiva do especialista.

No primeiro nível seriam definidas algumas linhas gerais (critérios) para que o árbitro ou avaliador pudesse julgar da conveniência da pesquisa. Um critério básico poderia ser o grau de congruência das finalidades da pesquisa com os objetivos da instituição no tocante à pesquisa, independentemente dos méritos metodológicos do projeto apresentado. Este poderia ser chamado de nível de decisão social.

Neste nível, a primeira tarefa seria identificar os objetivos gerais da instituição, em termos de pesquisa. No caso de pesquisa sócio-educacional, poder-se-ia discutir, por exem-

<sup>5</sup> Trata-se aqui de pesquisas candidatas à obtenção de financiamento para sua execução e também de pesquisas já realizadas, para que se avalie a conveniência de custear sua publicação. Não se considera, neste momento, a avaliação de pesquisas em desenvolvimento.

plo, em que medida são objetivos da instituição: desenvolver, através de pesquisas, as "ciências da educação"; obter resultados que permitam decisões para aumentar a eficiência interna e externa bem como a eqüidade do processo educacional.

A segunda tarefa seria definir operacionalmente os objetivos da instituição quanto às pesquisas. Ocorre, entretanto, que não é possível definir de modo claro e não ambíguo todo e qualquer objetivo do tipo enunciado acima. Sendo inviável definir operacionalmente todo e qualquer objetivo desse tipo, parece mais racional excluir essa preocupação do trabalho de avaliação. Assim, na avaliação a este nível, haveria uma elevada dose de subjetividade no julgamento acerca da congruência entre as finalidades da pesquisa e as da instituição, no tocante à pesquisa.

No segundo nível, que poderíamos chamar de decisão acadêmica ou metodológica, definir-se-iam critérios que permitissem avaliar se o projeto de pesquisa ou a pesquisa realizada é acadêmica ou metodologicamente satisfatória.

Aí seriam definidos critérios para tópicos tais como: formulação do problema; esquema de referência conceitual; fontes de dados; amostra e população; etc. A "pauta de avaliação" de um projeto da OEA em Buenos Aires, adaptada pelo INEP (em anexo), sugere critérios para cada um desses tópicos. Exemplos:

- O problema é definido de forma sucinta, clara e correta, do ponto-de-vista lógico e gramatical?
- As premissas ou postulados em que se baseia o estudo são indicados explicitamente?
- Indica-se a população objeto de estudo à qual se devam referir as generalizações produzidas pela pesquisa?

No terceiro nível dever-se-iam definir os critérios pelos quais seriam combinadas as decisões do primeiro e do segundo nível. Poder-se-ia pensar em ponderações para as decisões tomadas nos níveis anteriores, a fim de alcançar-se uma ava-

liação global. Trata-se, por exemplo, da avaliação final de uma pesquisa (ou um projeto) altamente congruente com os objetivos da instituição mas com algumas falhas metodológicas. Trata-se, ainda, de comparar esta pesquisa com outra não distante desses objetivos mas metodologicamente sólida. Embora se possa pensar em estabelecer pesos para cada um desses aspectos, seu cômputo final evidentemente dependerá muito da perspectiva pessoal de cada avaliador.

### Problemas da "Pauta de Avaliação"

Em princípio seria conveniente, num processo de avaliação de pesquisa, valer-se de tópicos e itens como aqueles contidos na pauta de avaliação. Temos várias sugestões a fazer quanto a itens específicos da pauta, mas o fundamental é chamar a atenção para alguns problemas metodológicos mais amplos desse documento e da tentativa do uso de critérios objetivos na avaliação de pesquisas.

A pauta é composta de vários itens. Os itens da pauta são critérios para avaliar determinados aspectos de um trabalho de pesquisa, sobretudo da pesquisa empírica.<sup>6</sup> Atribuem-se pesos aos itens de acordo com o grau de congruência entre o critério nele estipulado e o que apresenta a pesquisa com relação ao item (os pesos variam de zero a três). A avaliação final é feita computando-se a média aritmética dos pesos em cada item.<sup>7</sup>

São vários os problemas metodológicos aqui encontrados. Em primeiro lugar, é preciso reconhecer que existe uma elevada dose de subjetividade na decisão de atribuir um determinado ponto ("nota") a cada aspecto da pesquisa (item). É natural que as preferências ou preconceitos de cada avaliador se façam sentir quando este procurar atribuir uma nota a um aspecto de pesquisa, por mais objetivo que o avaliador procure ser. Poder-se-ia diminuir a subjetividade através da "soma de subjetividades". Uma vez definida a pauta de avaliação, vários juízes participariam do processo de atri-

<sup>6</sup> A pesquisa dita teórica está implicitamente excluída.

<sup>7</sup> Excluem-se os itens não aplicáveis a cada tipo de pesquisa.

buição de pontos aos vários aspectos de cada pesquisa, cada um trabalhando independentemente do outro. Este método<sup>8</sup> tem a vantagem de permitir que, quando da análise de um aspecto da pesquisa, uma nota muito elevada devida, por exemplo, a algum "preconceito positivo", seja compensada por uma nota muito baixa, devida a um "preconceito negativo". Estamos esperando, com este procedimento, que a lei do acaso funcione a nosso favor, ou seja, que os extremos tendenciosos sejam mutuamente compensados. O processo é tanto mais seguro quanto maior for o número de juízes. A desvantagem deste procedimento é o seu elevado custo operacional para avaliar-se um grande número de pesquisas.<sup>9</sup>

Outra dificuldade desta pauta está na construção da escala para cada item. Determinados itens podem sugerir a um juiz uma avaliação numa escala de 1 a 3 (1, 2, 3) enquanto que outros itens lhe pareceriam adequados se possuíssem uma escala de 1 a 5 (1, 2, 3, 4, 5). O problema talvez pudesse ser resolvido com pré-testes da pauta, como habitualmente se faz com escalas de atitude. Entretanto, deve-se notar que com uma escala de atitudes estamos procurando medir algo subjetivo (atitude) de modo bastante objetivo (escala). Aqui, ao contrário, estamos procurando medir algo supostamente objetivo (aspecto da pesquisa) de modo bastante subjetivo (escala preferida pelo juiz, no pré-teste). A dificuldade talvez esteja em encontrar um denominador objetivo e comum para esta última subjetividade. Estas dificuldades enfatizam a necessidade de um pré-teste cuidadoso da pauta, caso ela venha a ser utilizada.

Por outro lado, estas mesmas dificuldades sugerem a possibilidade de que o modo de utilização da pauta ficasse a critério de cada avaliador. Assim, alguns poderiam preferir valer-se da média aritmética dos pontos atribuídos a cada

<sup>8</sup> O método é essencialmente o mesmo freqüentemente empregado, por exemplo, para cada avaliação de conteúdo de itens de um teste. Veja-se Fred Kerlinger, *op. cit.* pág. 458.

<sup>9</sup> Sem considerarmos que pode haver discordância, entre os juízes, quanto aos casos de "não se aplica", talvez invalidando a média das médias, a não ser que o acaso funcione a nosso favor com um pequeno número desses casos.

item para decidir se a pesquisa satisfaz ao seu (do avaliador) critério de "aprovação". Outro juiz poderia preferir usar a pauta apenas como orientação geral para chegar a uma decisão final quanto à pesquisa. A desvantagem deste método é que ficaria a cargo de *cada* avaliador definir *seu* critério de aprovação. A vantagem é evitar-se uma avaliação final pretensamente objetiva, onde de fato predominaria a subjetividade na construção da escala de cada item e na atribuição de "notas" (pontos) a cada aspecto da pesquisa.

Um outro problema da pauta proposta refere-se à insuficiente definição de critérios para que um item (aspecto de pesquisa) possa ser julgado e portanto avaliado de acordo com a escala proposta (ou outra qualquer). Não nos referimos aqui a uma insuficiência que possa ser sanada através de análise profunda e exaustiva de cada item. Isto porque esta análise provavelmente indicaria tão grande quantidade de subcritérios, e de combinações destes, que tornariam praticamente inexecuível sua aplicação. Exemplificamos com o item (b) do tópico "Análise dos dados" (pág. A-11).

"As generalizações *vão além* do que permitem realmente os dados" (grifo nosso).<sup>10</sup>

A primeira pergunta que nos vem à mente é: uma resposta positiva a esta pergunta implica em pontos positivos ou em pontos negativos? A pauta propõe os seguintes pesos ou pontos para todos seus itens:

- aspecto plenamente atendido = 3 pontos
- razoavelmente atendido = 2 pontos
- sofrivelmente atendido = 1 ponto
- insuficientemente atendido = 0 pontos
- não se aplica = (item não é computado na média aritmética).

É certo que nossa pergunta poderia ser facilmente res-

10 É possível que este item seja atípico, mas uma inspeção superficial dos itens não fornece indicações neste sentido.

pondida através de uma reformulação do item. Este não é um problema de difícil solução mas sua existência serve para ilustrar a necessidade de cuidadosa revisão da pauta preliminar. O problema realmente sério está na dificuldade do estabelecimento de critérios simples e objetivos, que permitam atribuições de pesos de forma não ambígua.

Assim, numa escala de 0 a 3, o juiz teria que decidir, em primeiro lugar, o que seria "muito além", "bastante além" "pouco além" e "nada além". Em segundo lugar, supondo que o pesquisador tivesse ido "algo além do permitido" *mais de uma vez*, o juiz teria que decidir se ir "algo além", cometendo este erro várias vezes, é pior ou melhor do que ir "muito além" uma única vez. Em terceiro lugar, ao examinar mais de uma pesquisa, se o avaliador encontrasse numa delas apenas uma generalização, como ele julgaria a incidência de erros relativos e absolutos?

Tomemos vários itens para ilustrar o raciocínio. Suponhamos que em duas pesquisas diferentes, na definição das hipóteses, na elaboração das perguntas do questionário, na operacionalização das variáveis, no pré-teste do instrumento e nas generalizações dos resultados encontramos um percentual constante (digamos 33%) de erros. Mas na pesquisa A existem 6 hipóteses, 18 operacionalizações de variáveis, 30 perguntas no questionário, um pré-teste de 3 instrumentos e 3 generalizações. Na pesquisa B estes números são, respectivamente, 3, 6, 10, 1 e 1. Como um avaliador consideraria esta situação? Suponhamos que a pesquisa A era mais difícil de ser realizada por incluir um maior número de hipóteses, de variáveis, de instrumentos. Deste modo, o percentual de 33% de erros nessa pesquisa equivale a um percentual maior na pesquisa B?<sup>11</sup> Ou o avaliador consideraria simplesmente o número absoluto dos erros? Como ele poderia tomar uma decisão que pretende ser objetiva, sem conhecer as outras pesquisas que ele examinaria no futuro, e como a pesquisa

11 Deve-se notar que cada uma destas alternativas combina-se com cada uma das alternativas mencionadas ao parágrafo anterior.

presentemente examinada seria comparada com as demais (futuras), nestes aspectos? <sup>12</sup>

Esta dificuldade nos leva a outra, igualmente seria. Trata-se do problema da atribuição de pesos relativos, porém agora não mais *dentro* de cada item, mas sim a *cada* item no cômputo global. O problema deriva da possibilidade de que diferentes itens da pauta de avaliação devam ser ponderados diferencialmente para o cômputo da média global da pauta. Exemplo: o não atendimento ao item *X* do tópico *Y* é considerado erro de maior gravidade do que não atendimento ao item *Z* do mesmo tópico. Se há graduações de gravidade nos erros cometidos — e não cremos que se possa considerar todos os erros como de igual gravidade — então quais são os pesos relativos que devemos atribuir a cada item e/ou cada tópico?

Na prática, segundo a experiência nos indica, o julgamento acerca da seriedade do erro depende do contexto, da natureza da pesquisa. Um erro *X* na pesquisa *A* pode não ser tão importante quanto o mesmo erro *X* na pesquisa *B*. Um dos critérios — talvez o principal — para definir a gravidade do erro seria o grau de invalidade que teriam os resultados (e/ou conclusões) em virtude do erro cometido. Este processo de avaliação é feito através do confronto da percepção subjetiva que tem o avaliador dos diversos aspectos de uma pesquisa, com o “critério” acima mencionado. É em função de sua experiência como profissional que um pesquisador define o grau de invalidade dos resultados e conclusões e, a partir deste, a seriedade do erro ou inadequação encontrada no relatório da pesquisa estudada.

Assim, existe potencialmente um grande número de combinações possíveis de erro/grau de invalidade dos resultados. Além disto, estas possíveis combinações dificultam qualquer sistematização *a priori* simplesmente porque cada caso é um caso. Poder-se-ia pensar, é claro, numa tarefa her-

12 O avaliador, como indivíduo, poderia usar sua experiência profissional para comparar a pesquisa sendo avaliada com outras que já leu anteriormente, tomando-as como padrão. Mas neste caso o critério de avaliação seria essencialmente pessoal.

cúlea de sistematização de todas estas possibilidades, inicialmente através de consultas a pesquisadores e, num segundo momento, através da observação da experiência, numa atividade em que dezenas de profissionais trabalhassem por alguns anos com vistas a esta sistematização. Não cremos que seja possível lograr uma sistematização completa, mas de qualquer modo tal projeto está fora de cogitação, dadas as limitações de recursos disponíveis. Em suma, não se considera que seja possível sistematizar a seriedade relativa de erros, inadequações e omissões, e mesmo que isto pudesse ser feito a atribuição de pesos relativos a esta seriedade provavelmente continuaria a ser um processo subjetivo.

#### *Um Resumo e uma Proposta*

Apresentamos alguns problemas de avaliação de pesquisas sócio-educacionais, bem como alguns problemas específicos de uma “pauta de avaliação” de pesquisas. Estes últimos, na sua maioria, são decorrentes dos primeiros. São dificuldades que naturalmente surgiriam quando da tentativa de elaboração de qualquer roteiro de avaliação de pesquisas que buscasse ter a precisão procurada pela pauta sugerida.

Os problemas apresentados são de natureza conceitual e metodológica. Não se tentou uma separação rigorosa destes dois aspectos devido à natureza preliminar destas notas. Vale lembrar, entretanto, que tal separação nem sempre será possível em outra oportunidade, já que muitas vezes aqueles dois aspectos estão intimamente interligados.

Passando em breve revista as questões abordadas, verifica-se que a questão fundamental é, sem dúvida, a contradição entre objetividade e subjetividade. Tem-se, de um lado, o desejo de alcançar objetividade de critérios e de julgamento em qualquer processo de avaliação, inclusive no de pesquisas. Do outro, tem-se a dificuldade — intransponível em muitos dos casos — de alcançar objetividade na definição dos critérios de avaliação de pesquisas e sobretudo nos julgamentos destas pesquisas.

O primeiro problema na definição de critérios objetivos é a dificuldade de obter critérios que nos permitissem decidir

se uma pesquisa deveria ser realizada (ou, no caso de pesquisas já realizadas, se deveria ser publicada), independentemente de sua viabilidade e qualidade metodológica, aspectos que seriam analisados num segundo momento. Sugermos acima que tal decisão ficasse a cargo de especialistas que, com sua experiência, julgassem do mérito da pesquisa tendo em vista a congruência entre os objetivos da pesquisa e os da instituição financiadora.

O segundo problema refere-se ao elenco de dificuldades metodológicas na construção e aplicação de uma pauta de avaliação para pesquisas sócio-educacionais. Procuramos mostrar que há uma elevada dose de subjetividade na atribuição de "notas" (pontos) aos diferentes aspectos de uma pesquisa, mesmo que forneçamos, ao avaliador, critérios objetivos para seu trabalho. Em seguida verificamos que, em boa parte dos casos, tal objetividade não parece ser atingível. Por outras palavras, não parece seja possível definir — com suficiente amplitude, profundidade e ausência de ambiguidade — critérios para avaliação de diversos aspectos de pesquisas sócio-educacionais.

Verificamos também que, caso fossem superadas as duas dificuldades iniciais — e não nos parece que possam ser superadas — restaria ainda o problema de atribuir pesos relativos a cada um dos itens de uma pauta de avaliação, considerando-se a seriedade relativa dos erros, das inadequações e das omissões porventura encontradas nas pesquisas. Nossa discussão anterior sugere que a atribuição destes pesos é um processo basicamente subjetivo. Além disso, esta atribuição de pesos seria baseada numa enorme combinação de possibilidades, e parece inviável obter-se uma sistematização destas contribuições.

Em resumo, os problemas brevemente analisados sugerem a inviabilidade de uma pauta de avaliação de pesquisas sócio-educacionais que seja objetivamente elaborada e objetivamente aplicada. Nossa sugestão, portanto, é que se pense a) num esquema de avaliação onde predominasse a experiência do profissional que realizasse a avaliação e b) num programa de treinamento para "avaliadores juniors", desti-

nado a profissionais com pelo menos alguma experiência pessoal em pesquisa. Nesse programa, procurar-se-ia dar aos treinandos condições de avaliarem pesquisas básicas, a partir de um roteiro bastante amplo porém igualmente flexível.

No esquema de avaliação proposto temos três níveis. No primeiro nível predominaria, como já foi sugerido, a experiência do avaliador. Propõe-se que no segundo e terceiro níveis — embora com menor intensidade no segundo — também predomine a experiência do avaliador. Este se valerá de um roteiro indicativo para o processo de avaliação, ficando entretanto a seu critério definir as minúcias deste processo. O grau de detalhamento deste roteiro é tema para discussões que não cabem nestas notas. Nossa impressão é que o roteiro poderia alcançar o detalhamento contido na pauta de avaliação, acrescido de indicações para o primeiro e terceiro níveis de avaliação.

Do processo de avaliação participariam "decisores senior" e "decisores junior". Os últimos deveriam inicialmente receber, para avaliação, pesquisas de menor complexidade, e deveriam recorrer aos "decisores senior" para discussão de temas julgados pertinentes. Supõe-se que a equipe de avaliadores será diversificada quanto à sua área de experiência, já que também são diversificadas as áreas abrangidas pelas pesquisas financiadas pelo INEP. Não obstante, seria conveniente contar com consultores de diversas áreas (antropologia, ciência política, economia, estatística, psicologia, sociologia etc.), para discussão de tópicos controversos ou complexos.

O ponto básico da nossa questão é que devemos alocar mais tempo para o treinamento de avaliadores do que para a elaboração do roteiro. Julgamos que a utilização de um roteiro preliminar dará origem a um roteiro final, e que este será o produto da experiência dos avaliadores (senior e junior) e de seu trabalho com os consultores. Parece-nos que um roteiro indicativo e preliminar não demandaria tanto tempo para sua elaboração quanto exigiria o treinamento de "avaliadores junior", a ser realizado mediante aulas, seminários e trabalhos supervisionados.

## A N E X O

### LISTA DE CRITÉRIOS GERAIS PARA AVALIAÇÃO DE PESQUISA OU ESTUDO

A lista foi ordenada de modo que cada um dos critérios expressos constitua uma "pauta de avaliação" para julgar ou valorar a forma em que determinado requisito teria sido atendido no trabalho. Leva-se também em conta o caso em que a pauta indicada, por qualquer circunstância, não seja aplicável ao estudo em questão. Dividindo-se o total de pautas entre o número de pautas aplicáveis, pode-se obter uma média para valorar ou qualificar o trabalho que se está analisando. A lista pode ser reduzida ou ampliada segundo se deseje apreciar o trabalho em seus aspectos mais gerais ou mais específicos e pode servir como simples guia nos próprios trabalhos.

#### Escala proposta:

- este aspecto é coberto de maneira explícita e adequada (3);
- implícita, um tanto vaga (2);
- muito vagamente (1);
- o aspecto é ignorado (0);
- dadas as características do estudo, não se aplica (N).

#### Escala adotada:

- aspecto plenamente atendido (3);
- razoavelmente atendido (2);
- sofrivelmente atendido (1);
- insuficientemente atendido (0);
- não se aplica (N).

Fonte: OEA. *Investigaciones Educativas en America. Reseñas analíticas I-II-III*. Buenos Aires, Proyecto Multinacional de Investigación Educativa e Ministério de Cultura y Educación, 1976, p. 121-128. Adaptação do INEP-MEC.

### PAUTAS DE AVALIAÇÃO

#### 1. *Formulação do problema: antecedentes*

- a) Fez-se revisão completa de toda a documentação educacional disponível e das publicações que têm relação com os fatores do problema e sintetizou-se o que já se conhece sobre o assunto?
- b) Foram analisadas, com o devido cuidado, as relações existentes entre os diversos fatores que intervêm no problema?
- c) Apresentou-se um resumo de estudos anteriores em forma integrada que sugira um marco conceitual, melhor do que em forma cronológica?

#### 2. *Formulação do problema: justificativa*

- a) A apresentação dos antecedentes revela a existência de um problema que não foi solucionado ou que demanda solução?
- b) Demonstra-se a significação do problema relacionando-o com as necessidades da instituição, do sistema, da região, do país, ou de um ramo determinado do conhecimento? De que forma a solução dos problemas contribuiria para satisfazer tais necessidades?

#### 1. FORMULAÇÃO DO PROBLEMA: ANTECEDENTES

	3	2	1	0	N
a)					
b)					
c)					

#### 2. FORMULAÇÃO DO PROBLEMA: JUSTIFICATIVA

	3	2	1	0	N
a)					
b)					

3. *Formulação do problema: definição*

- a) A formulação do problema abarca todos os elementos e relações previamente descritos?
- b) O problema é definido de forma sucinta, clara e correta, do ponto de vista lógico e gramatical?

4. *Formulação do problema: alcance e limites*

- a) Especifica-se claramente quais são os aspectos do problema estudados e quais ficam fora do âmbito do estudo, seja por limitações relacionadas com o caráter da população ou pelo tipo de relações ou enfoques escolhidos?

5. *Esquema de referência conceitual: premissas*

- a) As premissas ou postulados que baseiam o estudo são indicados explicitamente?
- b) Existe encadeamento lógico entre tais premissas e postulados e o restante da explicação proposta?

3. FORMULAÇÃO DO PROBLEMA: DEFINIÇÃO

	3	2	1	0	N
a)					
b)					

4. FORMULAÇÃO DO PROBLEMA: ALCANCE E LIMITES

	3	2	1	0	N
a)					
b)					

5. ESQUEMA DE REFERÊNCIA CONCEITUAL: PREMISSAS

	3	2	1	0	N
a)					
b)					

6. *Esquema de referência conceitual: hipóteses*

- a) As hipóteses propostas estão bem fundamentadas, tanto em postulados como em alguma teoria sólida que se relaciona com a matéria?
- b) Se já existem outras hipóteses sobre o problema, a nova hipótese proposta é mais ampla quanto à explicação de um maior número de fenômenos?
- c) A explicação proposta é verificável mediante fatos empíricos?
- d) Contribuiria a hipótese para prognosticar novos fatos e relações?

7. *Esquema de referência conceitual: definição de termos*

- a) Há uma definição clara de cada uma das variáveis (fatores) mais importantes do estudo, e de termos cujo sentido pode estar sujeito a interpretação?
- b) As definições dessas variáveis aparecem expressas em forma operacional, isto é, ligada a fatos e relações empíricas comprovadas?

6. ESQUEMA DE REFERÊNCIA CONCEITUAL: HIPÓTESES

	3	2	1	0	N
a)					
b)					
c)					
d)					

7. ESQUEMA DE REFERÊNCIA CONCEITUAL: DEFINIÇÃO DE TERMOS

	3	2	1	0	N
a)					
b)					

- c) Todos os termos são utilizados e empregados com o mesmo sentido ao longo do estudo, sem a introdução sub-reptícia de ligeiras variações que possam invalidar a argumentação?
- d) A parte de definições é apresentada de forma separada e visível desde o início do estudo?

8. *Fontes e Métodos. Procedimento geral*

- a) Indica-se a Metodologia Geral com suficiente detalhe, de modo a permitir seu exame ou verificação e são apresentados argumentos que justifiquem sua escolha?
- b) Esses métodos levam na devida conta as condições locais: características ou relações especiais que o meio apresenta?
- c) Os métodos escolhidos oferecem uma forma econômica de coletar a informação ou existe outra forma mais simples e igualmente eficiente?
- d) Os métodos são suficientemente refinados e precisos para o tipo de inferências e generalizações que o estudo pretende alcançar?
- e) Tais métodos se harmonizam com as premissas e postulados que originaram o estudo?

3      2      1      0      N

c)				
d)				

8. *FONTES E MÉTODOS. PROCEDIMENTO GERAL*

3      2      1      0      N

a)				
b)				
c)				
d)				
e)				

- f) Foram eliminadas deficiências cometidas em estudos anteriores e se indicam algumas possíveis deficiências ou limitações possivelmente existentes na metodologia adotada?
- g) A época e o local onde se coletou a informação são especificados?
- h) Especificam-se as características dos sujeitos e dos materiais utilizados na pesquisa, bem como as razões pelas quais alguns deles não intervieram em alguma parte do estudo?
- i) Caso se tenha realizado estudo preliminar ou piloto, indicam-se as variações adotadas quanto à metodologia e suas razões?
- j) Há justificativa sobre os procedimentos estatísticos ou matemáticos empregados na elaboração dos dados em vista do grau de precisão que estes apresentam?

9. *Amostragem*

- a) Indica-se a população objeto do estudo e à qual se devam referir as generalizações produzidas pela pesquisa?

3      2      1      0      N

6. *ESQUEMA DE REFERÊNCIA CONCEITUAL: HIPÓTESES*

3      2      1      0      N

f)				
g)				
h)				
i)				
j)				

9. *AMOSTRAGEM*

3      2      1      0      N

a)				
----	--	--	--	--

- b) A amostra escolhida da mencionada população é realmente representativa?
- c) São também representativos o grupo testado e os grupos experimentais da referida população?
- d) Especifica-se o método utilizado na seleção da amostra e dos demais grupos? Coincidem estes com o método de seleção?
- e) O critério de seleção garantirá uma amostra representativa?
- f) Há indícios de que alguma prova preliminar que se aplicou aos sujeitos ou que o conhecimento que esses teriam da experimentação contribuíram para alterar seu comportamento a ponto de comprometer a amostra e invalidar as generalizações aplicadas a toda a população?

#### 10. Provas

- a) Foram verificadas a validade e a fidedignidade estatística das provas empregadas, assim como as normas básicas aplicáveis à população em estudo?

	3	2	1	0	N
b)					
c)					
d)					
e)					
f)					

#### 10. PROVAS

	3	2	1	0	N
a)					

- b) Correspondem as provas ao tempo disponível para respondê-las, às condições físicas do meio e ao conhecimento que os juízes têm de seu uso?
- c) Caso exista um júri qualificador, conhecem todos os juízes os antecedentes do estudo? Estão claras as bases de qualificação? Há razões para suspeitar de parcialidade ou preferência?
- d) Existem circunstâncias que limitam a amplitude ou a liberdade de respostas?

#### 11. Questionários: conteúdo das provas

- a) É necessário fazer todas e cada uma das perguntas?
- b) As perguntas são suficientemente precisas a fim de motivar as respostas desejadas?
- c) Cobrem as perguntas, em conjunto, aspectos essenciais das informações necessárias?
- d) Há necessidade de perguntas mais específicas ou concretas para que se possa obter uma informação mais precisa da situação? Ou, pelo contrário, devem ser mais gerais para que configurem um quadro global da situação?

	3	2	1	0	N
b)					
c)					
d)					

#### 11. QUESTINÁRIOS: CONTEÚDO DAS PROVAS

	3	2	1	0	N
a)					
b)					
c)					
d)					

- e) As perguntas são oportunas e estão isentas de uma orientação parcial que já sugira a resposta?
- f) Há suficientes alternativas de respostas, a fim de contribuir para a plenitude e a exatidão de expressão?

#### 12. Questionários: formas das perguntas

- a) Cada pergunta está redigida de forma breve e clara, isenta de tecnicismos e ambigüidades?
- b) A pergunta é ambígua porque deixa de fora alternativas importantes, ou porque o marco de referência é insuficiente?
- c) Empregam-se frases estereotipadas, emotivas ou superlativas que contribuem para comprometer a resposta?
- d) Empregam-se frases impróprias ou rudes que, por provocarem sentimentos hostis, podem contribuir para invalidar as respostas?
- e) Existe alguma forma mais impessoal de se fazer a pergunta de modo a ampliar as possibilidades de obter uma informação mais precisa?

3      2      1      0      N

e)				
f)				

#### 12. QUESTIONÁRIOS: FORMAS DAS PERGUNTAS

3      2      1      0      N

a)				
b)				
c)				
d)				
e)				

#### 13. Questionários: ordem das perguntas

- a) As perguntas iniciais preparam o terreno para as seguintes?
- b) Estão ordenadas de modo a facilitar a atenção e vencer resistências?

#### 14. Questionários: Respostas

- a) Convém que as respostas sejam dadas pela simples marcação no questionário? Mediante duas ou três palavras, ou resposta livre?
- b) Se a pergunta deve ser marcada, deve-se empregar dicotomias, múltipla-escolha ou escalas?
- c) Levam-se em conta as diferenças de precisão nas perguntas segundo se utilizem escalas de ordem de intervalo ou de razão?
- d) O questionário foi elaborado com vistas a facilitar a tabulação dos dados?
- e) Houve o cuidado de se dar instruções concisas e claras para cada operação e mediante o uso de exemplos quando for o caso?

#### 13. QUESTIONÁRIOS: ORDEM DAS PERGUNTAS

3      2      1      0      N

a)				
b)				

#### 14. QUESTIONÁRIOS: RESPOSTAS

3      2      1      0      N

a)				
b)				
c)				
d)				
e)				

- f) Quando são apresentados itens de múltipla escolha, as diferentes alternativas são colocadas ao acaso, a fim de reduzir as possibilidades de erros sistemáticos?
- g) Escolheu-se uma amostra aleatória dos sujeitos que não responderam, a fim de obter suas respostas e compará-las com as respostas dadas?

15. *Questionários: Pré-teste*

- a) O questionário foi previamente testado com grupo reduzido de sujeitos?
- b) Os objetivos e as perguntas foram plenamente elucidados durante esse teste?
- c) Depois das modificações introduzidas no documento, se verificou a confiabilidade das respostas?

16. *Classificação dos dados*

- a) Desde o início da pesquisa, pensou-se na forma de ordenar e classificar os dados?
- b) Os dados não aparecem, ao mesmo tempo, em mais de uma categoria?

	3	2	1	0	N
f)					
g)					

15. *QUESTIONÁRIOS: PRÉ-TESTE*

	3	2	1	0	N
a)					
b)					
c)					

16. *CLASSIFICAÇÃO DOS DADOS*

	3	2	1	0	N
a)					
b)					

- c) A informação disponível é insuficiente? Ou pelo contrário se apresentam dados em demasia para o estudo?

17. *Compilação dos dados*

- a) Houve parcialidade do observador em seus julgamentos por ter conhecido previamente a atuação dos sujeitos em estudo?
- b) Todos os aspectos da coleta foram verificados e comprovados, tais como anotações, cômputos e procedimentos?
- c) Analisou-se criticamente a autenticidade e a fidelidade das fontes? Tais fontes foram mencionadas de modo a facilitar a consulta se necessário?
- d) Utiliza o informe meios gráficos ou ilustrativos quando necessário? Sua elaboração é correta?

18. *Análise dos dados*

- a) A análise é formulada de maneira lógica e objetiva?
- b) As generalizações vão além do que permitem realmente os dados?

17. *CAMPILAÇÃO DOS DADOS*

	3	2	1	0	N
c)					

18. *ANÁLISE DOS DADOS*

	3	2	1	0	N
a)					
b)					
c)					
d)					

	3	2	1	0	N
a)					
b)					

- c) Apresenta a análise contradições, incongruências, noções vagas ou exageros?
- d) Faz-se confusão entre fatos empíricos, inferências e opiniões?
- e) Omite o pesquisador informação que possa estar em desacordo com sua hipótese?

19. *Conclusões*

- a) As conclusões são claras e precisas, indicando o contexto a que se aplicam?
- b) Tais conclusões decorrem de fato da análise dos dados, sem a introdução indevida de novos dados?
- c) São sugeridos novos aspectos relativos ao problema que devem ser pesquisados?

20. *Algumas anotações sobre estudos de tipo histórico*

- a) O estudo tem por base fontes originais?

3      2      1      0      N

c)				
d)				
e)				

19. *CONCLUSÕES*

3      2      1      0      N

a)				
b)				
c)				

20. *ALGUMAS ANOTAÇÕES SOBRE ESTUDOS DE TIPO HISTÓRICO*

3      2      1      0      N

a)				
----	--	--	--	--

- b) Fez-se estudo paralelo com o fim de comprovar, verificar ou avaliar a seriedade, a competência, os motivos, etc. dos testemunhos, quando e como foram registradas suas observações?
- c) A autenticidade e a fidedignidade das fontes foram examinadas de forma crítica?
- d) Fez-se uma interpretação lingüística correta dos documentos, de acordo com a concepção e a ideologia da época em que foram produzidos?
- e) Em casos necessários, recorreu-se à opinião de peritos em disciplinas auxiliares?
- f) Os problemas de autenticidade, data e local de publicação das fontes originais foram devidamente considerados?

21. *Considerações sobre estudos de tipo descritivo*

- a) Utiliza-se um esquema preciso de pesquisa que permita coletar a informação necessária para examinar a hipótese, em vez de uma coleção indiscriminada de dados?
- b) Houve de alguma forma sistematização da compilação de observações com vistas a evitar erros de percepção, memória ou apreciação?

3      2      1      0      N

b)				
c)				
d)				
e)				
f)				

21. *CONSIDERAÇÕES SOBRE ESTUDOS DE TIPO DESCRIPTIVO*

3      2      1      0      N

a)				
b)				

- c) Identifica-se com precisão a condição, a relação e o fato que o observador procura indicar?
- d) Emprega-se uma taxonomia precisa, de modo a evitar ambigüidades que permitam classificar o mesmo dado, simultaneamente, em mais de uma categoria?
- e) A taxonomia é apropriada para ressaltar, caso se queira, as semelhanças, as diferenças ou as relações em estudo?
- f) As categorias utilizadas são apropriadas, segundo o caráter dos dados, para sua qualificação?
- g) O estudo se limita aos aspectos mais superficiais das condições, deixando por acaso de considerar as relações internas dos fenômenos ou relações causais?

\*\*\*

## 22. Considerações sobre estudos de tipo experimental

- a) Apresenta o estudo um esquema de pesquisa claramente formulado que corresponda às hipóteses do estudo?
- b) Contém o esquema os elementos de controle indispensáveis para assegurar a validade dos resultados do estudo?

	3	2	1	0	N
c)					
d)					
e)					
f)					
g)					

## 22. CONSIDERAÇÕES SOBRE ESTUDOS DE TIPO EXPERIMENTAL

	3	2	1	0	N
a)					
b)					

- c) Analisam-se cuidadosamente as possíveis deficiências do estudo que limitam sua validade interna e externa?
- d) O pesquisador tem condições de manipular a variável independente ou se encontra em situação *post facto*, na qual só lhe resta limitar-se à análise comparativa?
- e) Caso os sujeitos participantes do estudo sejam voluntários, verificam-se os efeitos dessa circunstância quanto à equivalência dos grupos ou quanto à validade externa do estudo?
- f) Conta-se com os necessários subgrupos experimentais para responder a diferentes perguntas do estudo?
- g) Comprovaram-se as características dos sujeitos que abandonaram a experimentação?
- h) Foram violados alguns dos princípios que baseiam o uso de técnicas estatísticas nesse tipo de estudos?
- i) A hipótese experimental é definida com precisão e decorre logicamente das hipóteses do problema?
- j) Especificou-se de antemão o grau de significação estatística aceitável para rejeitar a hipótese experimental?

	3	2	1	0	N
c)					
d)					
e)					
f)					
g)					
h)					
i)					
j)					

23. Considerações sobre organização e forma \*

- a) O título identifica o problema com exatidão e concisão?
- b) Foram observadas todas as normas da ABNT de um modo geral?
- c) \_\_\_\_\_ do emprego de sinopses?
- d) \_\_\_\_\_ do emprego de sumários das normas de citação?
- e) \_\_\_\_\_ das normas de referenciamento bibliográfica?
- f) Cada seção contém as partes que a integram e é encabeçada pelo correspondente subtítulo?
- g) A linguagem usada é correta?
- h) A linguagem usada é objetiva?
- i) A linguagem usada é clara?
- j) A linguagem usada é direta?
- l) Relega-se toda a documentação excedente a um apêndice?

23. CONSIDERAÇÕES SOBRE ORGANIZAÇÃO E FORMA

3      2      1      0      N

a)				
b)				
c)				
d)				
e)				
f)				
g)				
h)				
i)				
j)				
l)				

- m) O informe encontra-se expurgado de frases inúteis, tabelas e documentos desnecessários como elemento de apoio?
- n) As instruções são adequadas?
- o) Deu-se a devida extensão aos temas de maior importância e não houve exagero quanto à extensão daqueles de menor importância?

3      2      1      0      N

m)				
n)				
o)				