

Algumas Definições de Avaliação

DILVO I. RISTOFF*

* Universidade Federal de Santa Catarina

Recebimento: 01/04/03

Aprovação: 22/05/2003

Resumo: O propósito deste texto é discutir algumas das definições de avaliação educacional, buscando explorar o seu significado e importância no contexto da educação superior brasileira.

Palavras-chave: avaliação; educação superior; universidade.

Abstract: The purpose of this paper is to discuss some definitions of educational evaluation, trying to explore their meaning and significance within the context of higher education in Brazil.

Key words: evaluation; higher education; university.

A definição de “avaliação” é, na verdade, uma questão mais complexa do que normalmente se imagina. David Nevo, após rápida discussão da bibliografia sobre o assunto, referenda a definição do Joint Committee, de 1981, qual seja, a de que a avaliação deve ser definida como sendo “a investigação sistemática do valor e do mérito de um objeto” (16), alertando, no entanto que há pelo menos uma grande discordância em torno desta definição que é a do Stanford Evaluation Consortium que, liderado por Cronbach, se recusa a aceitar o papel de juiz implícito na definição. Para Cronbach e seu grupo o avaliador é antes de tudo um educador cujo sucesso se mede pelo que os outros aprendem. O avaliador argumentam, não é um juiz de futebol contratado para decidir quem está certo ou errado. Para Cronbach, portanto, avaliar significa simplesmente conduzir um estudo sistemático do que ocorre com um programa ou instituição e em consequência deles. Seu objetivo é melhorar o programa ou instituição e instrumentalizar

outros programas ou instituições com objetivos similares. Feitas as contas, no entanto, Nevo entende que o que predomina na literatura é mesmo a idéia de que é, sim, preciso emitir juízos, que a neutralidade é uma fantasia e que expressar os juízos é uma obrigação de quem conhece melhor o objeto avaliado, embora, admita, esta definição nem sempre seja a mais simpática.

Com relação a esta definição é ainda necessário acrescentar que ela, além de se basear na idéia de que avaliar é uma atividade de pesquisa sistemática e não uma mera expressão de opiniões e palpites de iluminados (Os Deuses e Chefes expressam opiniões, os homens e subalternos fazem avaliação), é uma tarefa orientada para a identificação de valor e mérito. Aqui cabe perceber que o Joint Committee trabalha com as palavras “worth” e “merit”, algo que poderíamos traduzir por valor de mercado e valor intrínseco. Esta distinção é extremamente importante, especialmente no meio universitário ou educacional onde o valor não pode sujeitar-se unicamente ao mercado, devendo buscar o mérito intrínseco, as qualidades intrínsecas do objeto. O que seria da grande maioria das áreas humanas e das ciências básicas se o seu valor fosse medido exclusivamente pelo valor de mercado?

Ainda com relação a esta questão de mérito e valor, convém lembrar que em 1999 a Academia Real da Suécia concedeu o prêmio Nobel de Física aos Professores Gerardus't Hooft, da Universidade de Utrecht, Holanda, e ao Professor Emérito Marinus J.G. Veltman, de Bilthoven, também na Holanda. Após ter recebido o prêmio, o professor Gerardus't Hooft foi entrevistado por um jornalista que lhe perguntou para que serviam as suas descobertas. A sua resposta foi categórica e deste teor: “não serve para nada, que eu saiba!” Se analisarmos bem, para o jornalista interessado em alimentar o senso comum das massas, com notícias do imediato, do utilitário, do visível e do palpável, esta talvez tenha sido a resposta mais honesta. Não serve para nada mesmo! Este “Não serve para nada”, que nos remete à *República* de Platão, com os devidos adendos, bastaria hoje para expulsar do campus bem mais do que poetas. Teríamos que acrescentar aí também os filósofos, sociólogos, lingüistas, críticos de arte, historiadores e, como podemos notar pela resposta de Hooft, os físicos e talvez matemáticos, químicos, biólogos, etc. etc. Se, no entanto, a pesquisa é mesmo inútil, por que diabos os pesquisadores receberam o mais alto prêmio que pode ser concedido a um cientista? Segundo a Academia, os dois pesquisadores receberam o prêmio por terem colocado a teoria da física de partículas sobre uma base matemática mais sólida. Mostraram, entre outras coisas, como a teoria pode ser usada para desenvolver cálculos precisos das quantidades físicas. Várias experiências com aceleradores de laboratórios recentemente teriam confirmado os resultados calculados. Nas palavras da própria Academia, o prêmio foi concedido “pela elucidação da estrutura quântica das interações eletrofracas na física!”.

E, por falar em prêmio Nobel, o prêmio para Medicina de 1999 foi para Günter Blobel, um professor da Rockefeller University, em Nova Iorque, por descobrir que “proteínas possuem sinais intrínsecos que governam o seu deslocamento e localização na célula”. O interessante é perceber que Blobel estava trabalhando nesta direção há trinta anos e que já em 1970 ele havia descoberto o tal sinal intrínseco das proteínas. O que lhe rendeu o Nobel, no entanto, foi o fato de que nos últimos vinte anos ele detalhou os mecanismos moleculares que subjazem a estes processos, descobrindo inclusive que os princípios que orientam as proteínas são universais, e que operam de forma semelhante em células animais e vegetais.

O que estes dois casos têm a nos ensinar sobre a avaliação das universidades? No primeiro caso, que a universidade operacional de que falava Marilena Chauí, jamais terá espaço para pesquisadores como Hooft e seu colega, laureados com o maior prêmio da Academia, pois está presa ao utilitarismo, ao produtivismo, ao eficientismo, à relação custo-benefício. Esta universidade será incapaz de ver o útil no belo, na verdade revelada, na descoberta desinteressada.

As humanidades estão acostumadas a serem rotuladas como inúteis e de apelar para os sentimentos inferiores do ser humano (aos instintos e às paixões) desde os tempos platônicos e os seus defensores já aprenderam a viver com o peso desta acusação. Em geral respondem como Vítor Manuel de Aguiar e Silva e Theophile Gautier. Silva há alguns anos ressaltava que “se a origem e medida determinante dos valores fossem as necessidades práticas e imediatas do homem, o poeta situar-se-ia em grau inferior ao do sapateiro; mas o homem alberga exigências de outra ordem, inexplicáveis para uma mentalidade utilitarista... Para que a beleza das mulheres, se qualquer mulher, fisiologicamente normal poderia satisfazer as necessidades dos economistas? Para que a beleza tão frágil e tão pura das rosas, os coloridos profundos e aveludados das tulipas, se as couves e as batatas é que saciam a fome? Para que Mozart e a graciosidade cristalina da sua música? E, no entanto só almas bárbaras recusariam sob o pretexto da sua inutilidade, a beleza das mulheres e das rosas e da música de Mozart”. Gautier é ainda mais radical: “Não há verdadeiramente belo senão o que não pode servir a nada; tudo o que é útil é feio, porque é expressão de qualquer necessidade, e as necessidades do homem são ignóbeis. O local mais útil de uma casa são as latrinas”, arremata.

... o valor não pode sujeitar-se unicamente ao mercado, devendo buscar o mérito intrínseco, as qualidades intrínsecas do objeto. O que seria da grande maioria das áreas humanas e das ciências básicas se o seu valor fosse medido exclusivamente pelo valor de mercado?

Raramente chegaríamos a este nível de radicalismo para definir o belo, colocando-o como necessariamente antagônico ao útil, mas a contraposição é eficaz para mostrar que no outro extremo, com o mesmo radicalismo, se manifesta o utilitarismo. Que um prêmio Nobel continue sendo concedido a algo “inútil” como a descoberta de Hooft é sem dúvida um motivo para crer que nem tudo está perdido.

Com relação ao segundo exemplo, cabe apenas uma pergunta: Quantas empresas estariam dispostas a investir num sujeito com tanta falta de pressa, como este Blobel, que levou trinta anos desenvolvendo as suas pesquisas, isto é, levou quase o tempo de vida da totalidade das grandes universidades brasileiras? A resposta é óbvia: nenhuma! Novamente, cabe a observação: o fato de que gente sem pressa como os Blobels conquistem os mais altos prêmios da academia mostra que nem tudo está perdido e que professores universitários podem continuar a sonhar com pesquisas desatreladas das duas grandes pragas que infernizam a vida das universidades de verdade: o utilitarismo e o imediatismo.

A título de curiosidade, vale a pena discutir um pouco mais algumas das várias definições que têm historicamente sido expressas pelos que pensam a avaliação.

1. “Avaliação é um processo para determinar até que ponto os objetivos educacionais foram realmente alcançados” (Ralph Tyler, 1950).

Como podemos verificar, aqui a avaliação, na visão de Tyler, está fortemente centrada nos objetivos. Em consequência, um programa, currículo ou instituição é bem avaliado ou avaliada se os objetivos expressos foram atendidos, não cabendo ao avaliador qualquer questionamento sobre os mesmos.

2. “[Avaliação é a coleta de] informações com vistas à tomada de decisões” (Cronbach, 1963).

Nesta definição Cronbach já estabelece uma função instrumental importante ao processo, qual seja a de coletar informações úteis que possam municiar os administradores para tomada de decisões. Na implícita utilidade das informações a serem coletadas para fins gerenciais e de tomada de decisões, desmorona a resistência de Cronbach à expressão de juízos. Ou seja, ele não precisa julgar os objetos, pois o juízo já está implícito no tipo de indicadores que selecionará.

3. “A avaliação é a descoberta da natureza e do valor de alguma coisa. Os propósitos da avaliação podem ser muitos, mas a avaliação sempre tenta descrever algo e indicar os seus méritos e deméritos... A avaliação não é uma busca de causas e efeitos, um inventário do status presente, ou uma previsão de sucesso futuro. Ela é tudo isto, mas somente na medida que contribuir para a identificação da substância, função e valor” (Stake, 1969).

O interessante na definição de Stake é que ele recoloca a idéia de que avaliar é descobrir o valor e a natureza de um objeto. Avaliar implica em descrever o objeto e ao fazê-lo implica em identificar méritos e deméritos do objeto. Stake insiste na idéia de que identificar causas e efeitos e tirar conclusões sobre o estado presente e fazer previsões para o futuro são mera consequência e que só fazem sentido se precedidos da identificação do que denomina de substância do objeto, de sua função e de seu valor. O que fica evidente em Stake é que os méritos, deméritos e valores são sempre encontrados nos objetos e não são nunca uma imposição do observador – um objetivismo frequentemente questionado nos tempos pós-modernos.

4. “A investigação sistemática do valor e do mérito de algum objeto” (Joint Committee on Standards for Evaluation, 1981).

Esta definição continua a predominar na literatura, talvez por ter sido resultado de um esforço conjunto de várias universidades americanas no início da década de 80 e também por ser uma definição que contempla alguns dos elementos essenciais do efetivo trabalho avaliativo: o caráter investigativo, a natureza sistemática, e o reconhecimento de que todos os objetos têm valor intrínseco e valor de mercado.

5. “Avaliação educacional é o estudo concebido e conduzido para ajudar o público a julgar e a aperfeiçoar o valor de algum objeto educacional” (Daniel Stufflebeam, 1983— *Evaluation Models*).

A definição de Stufflebeam está centrada na idéia do propósito de julgar o valor, ou seja, de *atribuir* valor e de uma vez executada esta tarefa, *aperfeiçoar* o valor. Valor, portanto, para Stufflebeam, não é uma qualidade imutável e de intrinsecidade exclusiva. Ela não pertence exclusivamente ao objeto; pertence também a quem o estuda e valoriza. Esta definição é sensivelmente diferente da de Stake, cuja lógica ainda era a de *identificação* e não de *construção* de valores.

Esta definição de Stufflebeam se completa com outra do mesmo ano, qual seja a de que a avaliação deve ser entendida como um “estudo orientado para a tomada de decisões. Seu uso deve ser proativo (para aperfeiçoar um programa) e também retroativo (para julgar o seu valor)” e ainda por outra, também do mesmo ano, qual seja, a de que avaliar significa realizar “estudos para determinar se uma

... a conquista dos mais altos prêmios da academia mostra que nem tudo está perdido e que professores universitários podem continuar a sonhar com pesquisas desatreladas das duas grandes pragas que infernizam a vida das universidades de verdade: o utilitarismo e o imediatismo.

dada instituição e seus atores estão aptos a desempenhar as funções sociais para as quais foram designados” (Stufflebeam, 1983). Fica evidente que Stufflebeam tenta acrescentar especialmente dois aspectos relativos à utilidade da avaliação, isto é, a sua natureza formativa e somativa, a sua capacidade de interferir nos processos para determinar resultados diferenciados e também a idéia da avaliação como prestação de contas e como responsabilização dos atores envolvidos—noções que hoje estão totalmente incorporadas às definições em uso.

6. Um pouco diferente é a definição do australiano Stephen Kemmis. Para Kemmis, “Avaliação é o processo de organização de informações e argumentos que permitam aos indivíduos ou grupos participarem do debate crítico sobre programas específicos” (1986). Embora retome a idéia de avaliação como processo, estabelecida por Tyler, Kemmis sustenta que o processo é de organização de informações. Ou seja, Kemmis não apenas raciocina em termos de dados brutos coletados, mas de dados já transformados em informação e em argumentos. Os seus dados são, pois já contaminados por inferências e juízos e têm por objetivo não estabelecer verdades, mas abrir canais para participação de indivíduos no debate sobre programas específicos. Este lado participativo e de construção coletiva, parece indicar uma alteração significativa no conceito de legitimação política do processo junto aos grupos ou comunidades onde se realiza.

7. Esta definição avança em 1989, com Guba e Lincoln, no seu livro *Fourth Generation Evaluation*. Para estes autores, “não existe uma maneira ‘correta’ de definir avaliação, uma maneira que, se encontrada, poria fim à argumentação sobre como ela deve proceder e sobre quais são os seus propósitos. Entendemos que as definições de avaliação são construtos mentais humanos, cuja correspondência com alguma ‘realidade’ não é importante e não *deve* ser importante”. Ou seja, para Guba e Lincoln, definir significa impor para nós mesmos uma compreensão específica da realidade, buscando de certa forma congelar os seus procedimentos e os seus objetivos, que na visão dos autores devem permanecer indefinidos em nome da criatividade e da negociação ou transação que deve ser inerente ao processo.

8. Alguns anos mais tarde, em 1991, Michael Scriven, escreve que “Avaliação é o processo através do qual se determina o mérito, o valor de mercado e o valor das coisas”. Scriven usa as palavras “merit”, “worth” e “value”, onde “value” está associado à percepção da importância do objeto avaliado em contextos específicos. O que distingue a definição de Scriven da definição do Joint Committee é mais do que a palavra “value”. Para ele, o processo avaliativo *determina* valor ao invés de estudá-lo, como um container que determina o conteúdo, como um copo que dá forma à água que contém. Em

outras palavras, Scriven deixa explícita uma visão de avaliação cuja valoração ocorre não a partir do que é *encontrado*, mas a partir do que se *constrói* sobre o encontrado, com a participação decisiva de uma carga valorativa já *a priori* incorporada às estruturas conceituais do avaliador, como se o processo avaliativo fosse, repito, um container que moldasse o conteúdo. Daí a sua insistência em distinguir avaliação de mensuração (em geral de natureza puramente descritiva) e de sua preferência por avaliações centradas no usuário, como “o valor de algo para alguém ou para alguma instituição”. Ou seja, embora a “avaliação [seja] o que distingue comida de lixo, mentiras de verdades, e ciência de superstição” (140), é ela também que diz que a erva daninha só é daninha se vista no contexto do jardim no qual a sua presença é indesejada.

9. No Brasil, estas definições adquirem nuances próprias. Em 1993, o PAIUB definiu avaliação como “um processo contínuo de aperfeiçoamento acadêmico; uma ferramenta para o planejamento da gestão universitária; um processo sistemático de prestação de contas à sociedade; um processo de atribuição de valor... a partir de parâmetros derivados dos objetivos; um processo criativo de autocrítica”. Como é possível verificar, temos aí a preocupação com a natureza formativa da avaliação, a preocupação com o caráter sistemático do processo, a idéia da prestação de contas à sociedade, a idéia da avaliação como instrumento útil à administração educacional e a idéia da avaliação centrada nos objetivos. Talvez a única novidade do PAIUB, no plano específico da definição, em comparação com as definições anteriormente citadas, tenha sido a idéia da continuidade. O que torna a definição do PAIUB bem mais complexa é a sua preocupação em estabelecer princípios norteadores, objetivos e características que ajudam a melhor definir o Programa como um todo. Assim princípios como a continuidade, a adesão voluntária, o respeito à identidade institucional, a não-punição e premiação, a legitimidade, globalidade e comparabilidade, colocados lado a lado, tornaram-se elementos definidores de uma compreensão de avaliação institucional até então inexistente. Neste sentido, o PAIUB, embora visivelmente influenciado pela compreensão de avaliação estabelecida pela bibliografia internacional, passa a ser um programa com grau significativo de originalidade na sua concepção.

10. José Dias Sobrinho (1995) define avaliação institucional como “um empreendimento sistemático que busca a compreensão global da universidade, pelo reconhecimento e pela integração de suas diversas dimensões”. A definição de Sobrinho apresenta três aspectos importantes: (1) a idéia de que a avaliação, pela sua dimensão, não é apenas mais um estudo sistemático, mas um empreendimento, ou seja, algo laborioso e difícil, algo que necessita de

decisão e vontade política para ser executado; (2) a resistência a avaliações centradas em compreensões parciais, baseadas em alguns poucos indicadores que, ato contínuo, tornam-se representativos do todo para decretar o estado de uma instituição. Como a definição de Sobrinho tem em mente especificamente a universidade, depreende-se que ele imagina a compreensão global como a única forma de fazer justiça a esta instituição de múltiplas funções e sabidamente complexa; e (3) a compreensão de que não basta apenas reconhecer as diferentes dimensões de uma instituição; é também necessário ver até que ponto estas funções estão integradas ou organicamente articuladas entre si. Assim como não é possível conceber uma administração universitária sem ver as suas implicações sobre o ensino, a pesquisa e a extensão, não é possível ver as atividades-fim da universidade sem as suas articulações com toda a dimensão administrativa, de infra-estrutura, desde a qualidade das bibliotecas, das livrarias, até as oportunidades concretas de produção e disseminação do saber e a dimensão política. As contribuições de Sobrinho para a Avaliação Institucional da educação superior brasileira são tantas que seria necessário acompanhar as suas reflexões sobre todo o processo avaliativo para se ter uma compreensão mais completa dos pensamentos que subjazem a esta definição. Isto extrapolaria os limites a que estamos aqui sujeitos.

11. Isaura Belloni, outro nome que deixou a sua marca na história da avaliação institucional em nosso país, em uma de suas muitas discussões sobre o assunto, define Avaliação assim: “A avaliação institucional... é um empreendimento que busca a promoção da tomada de consciência sobre a instituição. Seu objetivo é melhorar a universidade. A autoconsciência institucional constitui importante subsídio para o processo da tomada de decisão, tanto em nível individual quanto em nível coletivo, da instituição como um todo, com vistas ao seu aperfeiçoamento, e tem como ponto de fundamental importância a intensa participação de seus membros tanto na forma de encaminhar a avaliação na identificação de critérios e procedimentos, como na utilização dos resultados” (1995). Quatro aspectos chamam a nossa atenção nesta definição: (1) a preocupação com a tomada de consciência sobre a instituição; (2) o caráter instrumental da avaliação, isto é, avaliação como processo voltado à instrumentalização dos tomadores de decisão; (3) o caráter formativo e de aperfeiçoamento individual e institucional e (4) a participação coletiva em todo o processo avaliativo. Tendo participado do Comitê Assessor do PAIUB, Isaura Belloni subscreve a toda uma concepção de universidade lá expressa e a definição aqui escolhida, como no caso de Dias Sobrinho, é insuficiente para explicitar todo o seu pensamento. Mesmo assim, o que pa-