

Avaliação quantitativa, avaliação qualitativa: interações e ênfases

José Dias Sobrinho

*Nada se edifica sobre a pedra, tudo sobre a areia, mas
nosso dever é edificar como se fora pedra a areia*
Borges, *Evangelho Apócrifo*, nº 41

Parece-nos que isso é bastante adequado para o trabalho do educador. Construir permanentemente, muitas vezes em situações difíceis e condições precárias, mas sempre com o sentido positivo, otimista e integrador da construção. Não do jeito de quem empilha tijolos, mas sim de quem edifica uma casa. De propósito, o texto começa e termina com metáforas de construção, mais propriamente de construção de significações, de perguntas e respostas.

Queremos passar a idéia de que a avaliação é parte intrínseca e constituinte do complexo fenômeno de construção humana. Falaremos aqui da avaliação, tratando mais propriamente da avaliação institucional e até mais especificamente da universidade. Dada a ênfase que atribuímos à função formativa e à natureza social e pública das instituições educativas, estamos acentuando a importância da avaliação qualitativa, sem diminuir o valor da avaliação quantitativa como parte integrante desse processo. Em outros textos, tratamos mais amplamente dessa temática. Remetemos o eventual leitor interessado no assunto à Bibliografia no final deste texto.

Como bem observa Isabel Torroba Arroyo, tudo leva a crer que nossos antepassados paleolíticos teriam muita coisa a ensinar sobre avaliação qualitativa, especialmente quanto à observação participante. Suas experiências de sobrevivência, as atividades de seu cotidiano, as metas atingidas, as técnicas aplicadas em cada situação singular, muitas registradas em suas pinturas rupestres, certamente não são reações instintivas e descontroladas, mas ações dirigidas por um cérebro apto a desenvolver avaliações, transferir experiências e participar de observações sobre a realidade. "Suas pinturas não são só a expressão de um espírito criador (...) porém nelas se representa a expressão gráfica de uma cultura, de

uma forma de entender o mundo e de desenvolver atividades habituais nesse mundo e nessa cultura.” (Torroba Arroyo, 1993, p. 295).

As pinturas rupestres, que demonstram o grande domínio anatômico da natureza dos animais, para além de sua dimensão descritiva, são o esforço de atribuição de significados, ou seja, de qualificação do mundo. A caça não é ação guiada só por instinto. Não se extingue em si mesma, uma vez realizada. O homem primitivo avalia sua força, suas habilidades, seu senso de orientação espacial e sua capacidade técnica e os registra enquanto cultura, modos de ver, sentir e fazer o mundo. A natureza torna-se cultura pela atribuição e organização dos sentidos.

Avaliação implica um fundamentado conhecimento daquilo sobre o que interrogamos e atribuição de significados aos fatos, dados e informações que colhemos. Para além dos fatos e a partir deles, a produção dos juízos de valor. Avaliar é uma ação que não admite neutralidade. Ultrapassa as descrições objetivas e as análises de coerência interna da realidade tomada por objeto. É um processo de forte conteúdo ético, pois indaga sobre valores e significados sociais. Atribuir significações e emitir juízos de valor, ou seja, avaliar, é reconhecer o mundo da produção humana e as diferenças, é responder às perguntas que fazemos a respeito de seus valores ou de suas qualidades. “De uma cidade, não aproveitamos as suas sete ou setenta e sete maravilhas, mas a resposta que dá às nossas perguntas.” (Calvino, 1990, p. 94). Como no símbolo concebido por Calvino, as cidades simbolizando a complexa existência humana, buscamos na avaliação os significados que precisamos compreender para, no limite, melhor compreendermos a nós mesmos.

Para além da contabilidade e da mensuração dos produtos e de todos os índices quantificáveis e capturáveis em tabelas e estatísticas, avaliar uma instituição é compreender as suas finalidades, os projetos, a missão, o clima, as pessoas, as relações sociais, a dinâmica dos trabalhos, a disposição geral, os grupos dominantes e as minorias, os anseios, os conflitos, os valores, as crenças, os princípios, a cultura. Então, é um empreendimento ético e político. Cada instituição tem a sua própria casuística e cada avaliação tem as suas próprias indagações. Não há, portanto, um modelo de avaliação pronto para uso geral e indiscriminado, como se houvesse uma forma única de instituição e como se o que se pretende avaliar fosse sempre e em todo lugar uma só e a mesma coisa.

A avaliação deve, assim, compreender e respeitar a identidade institucional em seu permanente dinamismo. Os pontos de referência de um processo avaliativo são a instituição em particular, tomada como objeto, mas através de seus

agentes constituindo-se também como sujeito de avaliação, contrastando-se o seu ser e o seu dever ser, a realização e o projeto. Obviamente, uma universidade participa dos princípios universais das instituições dessa natureza. Há, portanto, critérios gerais construindo os significados da universidade. Entretanto, não há um modelo único e universal de universidade. Sabe-se que as instituições de ensino superior experimentam uma crescente diversificação. As diferenciações vão das estruturas institucionais aos processos de ensino e aprendizagem, dos conteúdos, métodos e formação profissional às formas de inserção na sociedade e aos modos de compreender as suas funções sociais, podendo ser caracterizadas, quanto aos tipos, de centro de educação superior universitário e não universitário (pesquisa e amplitude dos cursos), quanto à magnitude (tamanho, localização, formas de organização), quanto ao perfil acadêmico (áreas mais fortes, níveis de estudos, etc.), quanto aos estudantes (diferentes perfis), quanto às fontes de financiamento e natureza jurídica (instituições públicas, privadas, mistas, religiosas). Há instituições com caráter internacional, outras cumprem missões mais concernidas às necessidades regionais. Como bem postula a Unesco, seria errôneo apresentar qualquer dos tipos de centro, de programa ou de instituição mencionados como se fossem respostas ou modelos ótimos. A diversificação é desejável. “Mas todas as decisões destinadas a fomentar a diversificação devem inspirar-se na preocupação pela *qualidade* das instituições e dos programas, a *equidade*, quanto ao acesso e à preservação da *missão* e à função da educação superior, com pleno respeito da *liberdade acadêmica* e da *autonomia institucional*.” (Unesco, 1995, 20).

Cada tipo de instituição se realiza de acordo com determinações históricas e tem suas justificativas ideológicas. As questões que emergem em cada instituição demandam respostas compatíveis com a identidade que cada uma delas constrói. É levando-se em conta essas identidades próprias que devem ser consideradas a qualidade, a função social, a pertinência. Não se pode adotar os mesmos critérios, instrumentos e procedimentos para a avaliação de uma grande universidade de pesquisa, de caráter universal, possuidora de laboratórios e de bibliotecas adequados e de docentes de grande capacitação, e para a avaliação de uma instituição regional, carente de infra-estrutura para o desenvolvimento da pesquisa, embora possa realizar o seu projeto com forte enraizamento em seu contexto social, portanto com relativa qualidade e pertinência. A meta da avaliação é julgar o valor. Mas, não há um valor absoluto e definitivo. O valor de uma grande universidade de pesquisa deve ser julgado de acordo com parâmetros adequados a esse perfil. Uma instituição centrada no ensino e na missão regional deve ser avaliada conforme critérios que dêem conta dessa realidade particular. Importante é que “a avaliação há de dar voz aos que participam

e, inclusive, há de representar a voz daqueles que não têm a possibilidade de aceder a essas atividades.” (Santos Guerra, 1993, p. 62).

A avaliação que propomos não deve servir como organizadora social das universidades, a separar aquelas que supostamente estão acima de um hipotético nível de qualidade daquelas outras que estariam abaixo e, portanto, condenadas ao desaparecimento, pelas leis de um darwinismo social que dualiza a sociedade, restringe e circunscreve a população que pode merecer os benefícios e os privilégios. Como organizador social, esse tipo de avaliação que estabelece a divisão de águas através do conceito escorregadio e vagamente instituído de “nível”, apresentado como se fosse uma realidade natural e preexistente às próprias universidades e às demais instituições de educação formal, cumpre uma função ideológica de grande poder instrumental. Baudelot e Establet (1994, p. 143) elaboram bem essa questão:

Com as suas dimensões contraditórias, a noção de “nível” nada tem a ver com um conceito científico, adequado para definir protocolos de medida indiscutível. Mas não é por isso um desses termos ideológicos, como “criação”, “arte” ou “liberdade”, que adquirem para refletir experiências difusas e cronologicamente estranhas, essa pluralidade de sentidos que lhes proporciona simultaneamente valor e consistência. A metáfora do nível possui, logo de início, a evidência de um fim pertinazmente perseguido: trata-se de um organizador social no qual se exprime, com força e simplicidade, o plano da sociedade escolar que se quer construir.

O nível não é simplesmente altitude ou um indicador métrico, embora assim apareça, pois, tradicionalmente se tem vinculado a avaliação à mensuração, à medida, à comparação e à classificação. Sempre se operou uma relação muito forte entre avaliação e número. A quantidade, além de ser a dimensão mais visível e fácil de se operar, sempre se mostrou como a mais confiável e capaz de emprestar um caráter científico à avaliação. Mas, diferente das aparências, o nível não é simplesmente uma medida ou um indicador numérico, é, sobretudo, um instrumento de controle e de seleção sociais. Trata-se de fenômeno que ultrapassa o campo do escolar e efetua uma normalização da sociedade, a qual permite “avaliar a maioria das realidades humanas e sociais ou morais em termos de uma quantificação, real ou imaginária, por referência a uma linha de nível médio” (Baudelot e Establet, 1994, p. 131).

O nível, indicador arbitrário de qualidade, é a linha de corte, acima da qual estão a salvação e os privilégios, abaixo da qual estão o desprestígio no imaginário social e o abandono dos poderes públicos. A grande carga de significação do nível de qualidade está, pois, nos efeitos que produz: as hierarquias

sociais. Os efeitos desqualificam a ilusão de cientificidade: os critérios de nível de qualidade amoldam-se conforme as necessidades e flexibilizam-se ao sabor das conseqüências para melhor cumprirem sua função de ordenamento e hierarquização.

Por mais que o nível se ornamente das aparências mais exigentes de um instrumento de medida objetivo e universal, a originalidade desta medida padrão reside no fato de nada medir. O nível louva-se ou deplora-se, decreta-se ou recusa-se, eleva-se ou baixa-se, alcança-se ou recupera-se: nunca se mede. (Baudelot e Establet, 1994, p. 23)

Com essa função de ornamento e hierarquização social notadamente vinculada à escolarização, a noção de nível surge no século XIX, como explicam Baudelot e Establet. Quem lhe dá consistência institucional é o *baccalauréat*, primeiro diploma na burocracia francesa moderna baseado em exame geral e nacional ao término da escolaridade média;

um sistema de exigências impessoais, equitativamente propostas a todos (...) o nível adquire no cenário do exame a existência autônoma da norma: assim poder-se-á confrontar-lhe seres concretos tanto individuais como coletivos: “nível” do candidato ou da classe, “nível” dos estabelecimentos e do ensino que proporciona. (p. 134)

A discussão do nível é interessante, não só porque nos permite entender melhor a função desses indicadores de medida, de classificação e de seleção e como eles têm uma longa história associada às práticas de avaliação no âmbito escolar, mas também porque tem força de atualidade e presença. Com efeito, a sistemática que o MEC está instituindo para a avaliação das instituições de ensino superior brasileiras consiste, sobretudo, num exame nacional que supostamente mediria o nível do estudante concludente e presumivelmente do respectivo curso em relação ao mercado. Desde logo, registramos um de seus efeitos, recuperando um comentário crítico que já tem mais de cem anos a respeito de uma sistemática antiga e similar. Em 15 de novembro de 1887 (o *baccalauréat* foi implantado em 1884), Francisque Sarcey declara em *La République Française*:

Já só se trabalha com os olhos fixos no programa e tendo em vista o programa. E esse infelizmente programa tem se distendido dia após dia: introduziram todos os conhecimentos humanos (...) este fantasma do *baccalauréat* que é preciso conquistar desorganizou todas as classes. Os professores proclamam-no: não obtém mais nenhuma atenção para os estudos desinteressados (...). Será isso válido para o *baccalauréat*? Os alunos não

conseguem se libertar disso. E em lugar de fazerem os trabalhos que lhes são dados, mergulham nos manuais onde a ciência é comprimida, amontoadada, e distribuída em tabelas de Liebeg. (Baudelot e Establet, 1994, p. 137)

Conhecemos muito bem o efeito “cursinho”, a ponto de não precisarmos comentá-lo. Mas vale registrar a antiguidade de seus antecedentes, como esse caso francês.

Essa avaliação acima comentada é, antes de mais nada, verificação de conhecimentos. Pretende medir e cobrar o conhecimento, até mesmo desconsiderando que muitos e importantes conhecimentos não resultam do trabalho escolar, nem sequer foram ensinados. Ademais, essas avaliações que se resumem a testes, tendem a absolutizar medidas e critérios cujo valor é sempre relativo. Além disso, os resultados finais não servem à melhoria do processo, que, aliás, não é considerado. Os exames são estáticos e utilizados em geral no fim de um processo ou tarde demais para orientar ações transformadoras. Por outro lado, são eficientes, isso sim, para sancionar noções de nível, estabelecer as hierarquias e oferecer consistência institucional às seleções, aos privilégios, à dualização e às segregações sociais.

No caso da verificação de conhecimentos, é preciso atentar cada vez mais à volatilidade e às transformações que ocorrem na ordem do conhecimento, que rompem com estabilidades e certezas habituais. Dentre as mudanças e rupturas, García-Guadilla (1995) destaca as seguintes:

No epistemológico: desintegração da ciência unificada e perda de consenso na natureza da racionalidade científica; no organizativo: tendência a conhecimentos integrados, com possibilidades de incluir aspectos que haviam estado marginalizados do conhecimento, por exemplo, o estético e o ético; no valorativo: novas formas de conhecimentos não necessariamente legitimados pelo até agora considerado “científico”; questionamento de noções-chaves como a objetividade, certeza, predição, quantificação; debilitamento do conhecimento abstrato e fortalecimento do conhecimento contextualizado; aumento de importância do conhecimento ordinário ou popular como fontes de sabedoria acerca da realidade social e do entendimento humano. (p. 90)

Em outras palavras, num cenário de transformações rápidas, se quisermos incrementar os valores de solidariedade e cooperação, devemos valorizar as dimensões subjetivas, éticas, políticas e mais complexamente humanas do conhecimento. As características de racionalização e de homogeneização das provas

ou exames pontuais ou nacionais de medida de conhecimentos não dão conta dessas dinâmicas.

A emergência da avaliação do domínio escolar está ligada às necessidades crescentes de ordenamento e seleção social que eclodiram a partir da Revolução Industrial. Por isso, a avaliação sempre carregou as características de mensuração e comparação. Os manuais estão cheios de avaliações que carregam esse peso semântico e sua utilidade ideológica. São muito conhecidos e dispensam maiores citações. Daremos só alguns poucos exemplos, servindo-nos de um documento não muito conhecido no Brasil.

Segundo o Manual de Avaliação, produzido pelo CIEES – Centro de Documentação do Sistema Educacional Tecnológico do México –, há uma grande diversidade conceitual em torno da avaliação. As concepções poderiam estar agrupadas da seguinte maneira:

a) Avaliação como julgamento de aspectos. Nesse caso, especialistas emitem juízos de valor a respeito de matéria de sua área, de forma inapelável, sem necessidade de conhecimentos especiais sobre avaliação.

b) Avaliação como medição. Com o desenvolvimento da psicologia experimental e diante do crescente anseio pela objetividade na verificação da aprendizagem, a avaliação passa a se reduzir à elaboração de instrumentos de medição e à interpretação quantitativa dos resultados. Uma vez que a medição não conduz à emissão de juízos de valor, ela não pode ser considerada avaliação em sentido pleno.

c) A avaliação como comprovação de logro de objetivos. Essa concepção leva a avaliação para além da questão da aprendizagem dos alunos. Envolve a avaliação dos docentes, dos planos e programas de estudo, dos materiais e recursos didáticos, em função de objetivos internos previamente estabelecidos.

Além desses três blocos, os autores do Manual de Avaliação apresentam uma série de outras concepções rapidamente mencionadas a seguir:

- Scriven, M. (1967). A avaliação determina o valor de algo por meio dos registros dos dados obtidos.
- Alkin, M. C. (1969). A avaliação seleciona informações, analisa e reporta dados úteis para a tomada de decisões.
- Stake e Demy (1969). A avaliação descreve e identifica os méritos e limitações de algo, ou seja, descreve e julga o avaliado.
- Guba e Stufflebam (1970). A avaliação permite obter informações úteis para tomada de decisões.

- Bloom (1971). A avaliação é entendida como a retroalimentação contínua que permite realizar as correções necessárias.
- Cronbach (1971). A avaliação é o procedimento mediante o qual se medem as conseqüências de uma ação que orienta o cumprimento de objetivos determinados.
- Ausubel (1982). A avaliação é a emissão de juízos de valor, tendo em vista resultados de um programa em função do cumprimento de metas ou objetivos.

Essas e outras concepções citadas pelos autores do Manual de Avaliação mexicano reduzem a avaliação a aspectos e funções definidos, apresentando em comum o seguinte: “Avaliação é um processo de indagação, de comparação, de obtenção de informação que permite a emissão de juízos e contribui para a tomada de decisões”. Os autores desse documento apontam carências e limitações naquele conceito, especialmente no que se refere à aplicabilidade a situações concretas e à falta de referências para a emissão de juízos de valor. Para eles, esses conceitos não chegam a ultrapassar plenamente a mensuração ou medição. A partir dessas observações, propõem uma conceituação mais abrangente: “avaliação é o processo sistemático e permanente que permite captar informação sobre o objeto avaliado para contrastá-lo com um marco de referência e a partir dessa contrastação emitir juízo de valor e propor alternativas para melhorar o referido objeto” (idem, p. 10).

A partir de Guba e Lincoln (Fourth Generation Evolution), uma síntese da evolução do conceito de avaliação na área educacional é apresentada por Bielschowsky e Penna Firme (1994, p. 134). A primeira geração, predominantemente técnica, não distinguia avaliação de medida. Tratava-se de medir, através de testes objetivos, o rendimento escolar. A segunda geração também se dedicou bastante à tecnologia de elaboração de testes, voltada à verificação do alcance dos objetivos e às diferenças individuais. “A distinção entre medida e avaliação foi marcante nessa fase, e o papel do avaliador, embora ainda técnico, foi muito mais o de descrever padrões e critérios em relação ao sucesso ou ao fracasso de objetivos.” (idem, p. 135). Para além da medida e da descrição, o que mais propriamente caracteriza a terceira geração é o julgamento de valor. Mas, os papéis técnicos, descritivos e de julgamento dos avaliadores tornaram-se muito complexos e, por vezes, de difícil conciliação, ante a pluralidade de valores e as dificuldades de gerenciamento. A quarta geração passa a ter, por isso, a negociação como elemento principal de integração. Para além de aspectos técnicos e informativos, a avaliação passa a envolver dimensões políticas, sociais e cul-

turais, o processo torna-se participativo e acolhe as preocupações, as percepções e os pontos de vista dos interessados.

A característica principal dessa abordagem é a de responder à multiplicidade de interesses, por sua flexibilidade e sua amplitude. Conhecida como abordagem *responsiva* (Stake, 1975) ela leva em consideração toda e qualquer indagação que se apresenta ao longo do processo avaliativo, detectando ou criando, para cada uma, os procedimentos metodológicos mais pertinentes. (Idem, p. 136)

Aos poucos, vamos chegando a um conceito de avaliação muito mais complexo e abrangente que descrever atributos e julgar os graus de aceitabilidade de um programa, de um procedimento de ensino ou do comportamento de um indivíduo ou de um grupo, como queriam, por exemplo, Thorndike e Hagen. Não só se amplia consideravelmente o objeto, como também o sujeito ou os atores da avaliação. Não só o avaliador individual, com seus procedimentos técnicos e inapeláveis, mas sim o sujeito coletivo, em atitudes técnicas e políticas discutidas e negociadas publicamente.

A avaliação que propomos é, pois, um processo bastante complexo, no qual os aspectos e etapas não se justapõem, mas se articulam de forma organizada. Mais do que uma simples medida episódica, a avaliação é um processo que requer a articulação de diferentes etapas e intrincados procedimentos em todas as atividades humanas. Tratando-se de processo formal e intencional, demanda a definição de critérios, campos, normas e referências bastante claros, que orientarão a produção de juízos de valor. Estes devem ter função instrumental de orientar as tomadas de decisão para a transformação da realidade avaliada.

A avaliação institucional, ao superar as dimensões individuais de verificação de conhecimento, de ensino e aprendizagem e as relações professor-aluno em situação de sala de aula, para abranger as múltiplas e até mesmo conflituosas dimensões e funções da universidade (esquemáticamente: científicas, pedagógicas e ético-políticas), envolvendo como atores ou sujeitos o conjunto da população universitária, faz transbordar o processo de avaliação do domínio mais estreitamente especializado e técnico para o complexamente social, ético e político. Das preocupações tecnicistas e das dificuldades instrumentais passamos ao campo dos debates, dos conflitos, das negociações, das interações, das construções coletivas e não isentas de contradições. Como assinala L. Branco,

as principais dificuldades não são técnicas, senão políticas, de vontade, de sensatez, de sentido comum, de sensibilidade pedagógica, de flexibilidade e de criatividade. Convém não perder de vista que a melhor avaliação não

é a mais técnica e precisa, senão a mais operativa; isto é, aquela que seleciona e se centra em obter dados que elabora e divulga para ajudar a tomar e melhorar as decisões. (In Gairén Sallán, 1993, p. 338)

A melhor avaliação não é necessariamente a que esclarece os problemas e cala as dúvidas e divergências. Antes, “a qualidade da avaliação pode ser julgada segundo a qualidade de sua contribuição ao debate crítico, informando-o e melhorando-o” (Kemmis, in Santos Guerra, 1993, p. 41).

Estamos num campo em que o mais importante não é verificar um estado (de conhecimento, por exemplo) ou um produto, e sim os processos, as trajetórias, as relações. A literatura da área educacional que trata da verificação de conhecimentos é bastante farta. No caso da medição de produtos científicos, o campo dos estudos sociais da ciência produziu várias orientações das quais uma das mais utilizadas, especialmente com motivações de políticas públicas, é a cientometria, que se dedica a medir, através dos assim chamados indicadores científicos, produtos ou resultados da ciência em face dos insumos (infra-estrutura e recursos). “A medição de insumos indica a atividade científica, e a medição de produtos indica a produtividade.” (Velho, 1994, p. 310).

Não é difícil operar os indicadores de insumo, como levantar os números referentes aos recursos humanos, aos recursos orçamentários, à organização física, etc. É uma questão basicamente de estabelecer as quantidades dos elementos e dos agentes que produzem ou fornecem as condições para a produção dos resultados da atividade científica. Também não é difícil contabilizar os produtos. Mas, os indicadores de produtos são sempre parciais e não dão conta do aspecto principal que são os conhecimentos e as idéias. É o que reconhece a Science Indicators, conforme destaca Lea Velho: “somos poucos sofisticados em nossa medição da ciência como um corpo de idéias e de suas conexões com a ordem social e econômica. Assim, nossos indicadores continuam sendo – por enquanto – principalmente medidas de aspectos da ciência” (Velho, 1994, p. 311). As limitações das medidas quantitativas são de ordem conceitual e metodológica. A cientometria, muito em voga desde os anos sessenta, principalmente entre sociólogos e historiadores da ciência norte-americanos (notadamente Garfield e Price do Institute for Scientific Information), embora fortemente incorporada às agendas de políticas científicas e tecnológicas de muitos países, se já é insuficiente para dar conta dos aspectos cognitivos, é muito mais ainda incapaz e inadequada para compreender e avaliar as relações sociais e políticas e as demais dimensões qualitativas do mérito científico. A questão básica é que essas análises quantitativas se dedicam mais à mensuração, de um ponto de vista externo e pretensamente neutro, do eixo insumos-produtos, o que obvia-

mente é um enfoque redutivo da ciência e está longe de dar conta do fenômeno do processo de produção científica. Além disso, essas análises e esses indicadores científicos são inadequados para países em desenvolvimento (Martinez, 1994).

A insuficiência que apontamos no caso específico das análises quantitativas da ciência ou cientometria é também pertinente a toda avaliação restrita ao levantamento de quantidades e à orientação normativa nos processos institucionais. A propósito, devemos aqui evitar as hipostasias e as posições extremas e unidimensionais, que acabam opondo quantidade e qualidade e anulando uma ou outra. Quantidade e qualidade são dimensões inseparáveis de uma mesma realidade. Portanto, enfatizar uma delas é uma opção metodológica que se justifica em face de escolha de objetivos. Assim, se queremos conhecer objetivamente o estado de uma determinada realidade, precisamos construir numerosos indicadores quantitativos, lembrando sempre que a dimensão qualitativa aí também se apresenta, pois todas as atividades humanas são orientadas por critérios sociais e escolhas pessoais ou intersubjetivas. Mas, se além disso, queremos avaliar o mérito, emitir juízo de valor sobre a qualidade e os efeitos públicos, se pretendemos compreender e valorar os processos, as condições de produção, as relações sociais, o clima institucional, a pertinência e enraizamento no contexto, os compromissos e a missão e se, sobretudo, entendemos a avaliação como um processo público e social de ação de transformação da realidade, então temos que adotar mais decididamente as orientações qualitativas, para além das quantitativas. É sempre uma opção metodológica que se justifica diante de situações complexas e densas de significação. Trata-se, contudo, de ênfase, não de exclusão.

A avaliação qualitativa não se reduz a níveis e muito menos a classificações numéricas, nem se limita temporalmente a episódios analíticos. Ela cria condições e situações para o contínuo desenvolvimento e transformação da realidade avaliada e dos indivíduos nela implicados, reconhecendo neles a prerrogativa de serem sujeitos ou atores desses processos em suas relações sociais. Por isso, a avaliação qualitativa não é intervenção externa, no sentido de que não invade a instituição com critérios e objetivos exteriores e muitas vezes estranhos. Não sendo exclusivamente uma auto-avaliação, também comportando julgamentos de pares externos (conforme a tradição da comunidade acadêmica e científica), a avaliação qualitativa é intrínseca aos processos de uma instituição, constituindo-se como cultura avaliativa permanente.

Essa prática tem lá seus vícios e comete muitos equívocos. Mas, ainda é a mais recomendável e rica. “Não se trata, mais uma vez, de uma panacéia. Sistemas de revisão por pares estão sujeitos à captura por oligarquias ou cor-

porações de interesse e trabalham muitas vezes com grandes margens de erro. Não existem, no entanto, mecanismos adequados que o substituam.” (Schwartzman, 1992, p. 25). No caso de uma instituição educativa, esses processos internos e externos combinados são, para dizer o mais evidente, administrativos, científicos e pedagógicos, ou seja, necessariamente formativos. A instituição educacional é um sistema de formação, não o único e suficiente, mas certamente é portadora de funções de formação socialmente reconhecidas e publicamente constituídas.

A avaliação, conforme estamos pensando, não será apenas um instrumento para fazer o ponto de uma situação ou de um estado em um determinado momento. Não será apenas uma medida, mesmo porque as técnicas de medição, os exames e os testes são considerados insuficientes e até mesmo perniciosos por especialistas respeitadas em todo o mundo.

Pensamos aqui na avaliação como parte constitutiva dos processos formativos da instituição. Por isso, ela é necessariamente contínua, interativa, proativa e global. Sendo contínua ou permanente, não se esgota em momentos ou aspectos determinados. Cada avaliação fragmentada e episódica adquire uma significação nova na medida em que se insere intencionalmente num tempo total, que é o tempo da educação. Da mesma forma, as medidas, quantificações e análises parcelares e específicas também se tornam compreensivas à medida que se relacionam com a análise do conjunto e se integram a visões de ampla agregação.

“As características dos fenômenos sociais e dos educativos em particular desbordam as rígidas limitações das exigências do modelo experimental de investigação”, diz Pérez Gómez (1995). Referindo-se à compreensão da relação ensino-aprendizagem na escola, ele continua:

o objeto de investigação é um complexo sistema de comunicação, que tem lugar em um espaço institucional determinado, onde se intercambiam, espontânea e intencionalmente, redes de significados que afetam ao conteúdo e às formas de pensar, sentir, expressar e atuar daqueles que participam de referido sistema.

Conclui afirmando a necessidade de afrontar a complexidade, a diversidade, a singularidade e o caráter evolutivo da realidade social. A aceitação desse princípio não supõe a recusa

de processos mais concretos e moleculares de investigação, enfoques mais minuciosos e restritivos, senão a convicção de que tais indagações não podem conduzir a interpretações autônomas e independentes da realidade

estudada, por serem necessariamente parciais e enviesadas e que requerem, portanto, sua integração permanente na perspectiva de conjunto que pretende abarcar a complexidade real do sistema concreto. (Idem, pp. 115-16 e p. 128)

Dizemos que essa avaliação é global não porque ela produza uma qualidade absoluta e final, coisa que certamente não faz, mas sim porque ela tem a globalidade como perspectiva e princípio heurístico. A interatividade diz respeito à natureza social, política e pública da educação, da formação e de seus valores. Por isso, é bom insistir que o cerne da avaliação não é de natureza técnica, mas, sobretudo, política. A universidade, instituição formalmente dedicada à formação e ao conhecimento, que são produções humanas públicas, tem ineludivelmente funções públicas. Por isso, suas atividades hão de ser visíveis e sempre alimentadas pelo debate público.

É o conjunto de atores da instituição que coloca os problemas, que levanta as interrogações a que a avaliação deve responder. É também de modo social e público que devem ser constituídos o objeto (o que avaliar), os objetivos e usos (para que avaliar), as razões e princípios (por que), os critérios, as normas, a metodologia e os modos de construir e desenvolver o processo de avaliação (como) e os atores sociais (quem). Se decidimos que é a melhoria da qualidade da instituição em todos seus processos de produção científica e de formação humana com forte sentido social e público o que pretendemos da avaliação, então ela é necessariamente proativa, ou seja, otimista, voltada para a idéia de um futuro melhor, que aprende com os erros e falhas e valoriza o que já existe de positivo. É, portanto, uma concepção dinâmica e criativa e não punitiva, controladora e segregacionista.

A avaliação deve ser, pois, uma construção certamente social, como é cultura já sedimentada na comunidade acadêmica e científica, e criativa, porque os indicadores exclusivamente quantitativos e as práticas normativas, classificatórias e sancionadoras, não são adequados às instituições educacionais. Não há modelos prontos e universalmente válidos. É preciso construí-los levando-se em conta as realidades históricas específicas. Como sublinha J. Ardoino e G. Berger (1986), “a avaliação, no que à questão do sentido se refere, consiste essencialmente em produzir, em construir, em criar um referente a que, simultaneamente, ela se refere. (...) Consiste, precisamente na recusa de se colar a um modelo pré-definido” (apud Albrecht, 1994, p. 79).

A avaliação, assim, põe em relevo o sentido mais central de sua etimologia: os valores. Não é simplesmente questão de controle e de competência em áreas específicas. É questão de críticas, de debates públicos, de conflitos, de negocia-

ção, de preocupação com o processo, com o contínuo e com o interesse pelo social, é busca de construção de uma visão de conjunto, “de uma forma integradora, onde os elementos constituintes já não são percebidos como tais, mas como significando um todo” (Albrecht, 1994, p. 166). A avaliação institucional de orientação qualitativa e formativa passa pela medida isolada, digamos, do nariz, dos lábios e dos olhos, mas a análise desses elementos constitutivos deve permitir a visão da forma integradora e concreta de um rosto. Analisa as “sete ou setenta e sete maravilhas” desse conjunto, mas isso como meio para obter o que realmente importa: “as respostas às nossas perguntas” ou, já que é um processo dinâmico e contínuo, “as perguntas que nos colocamos para nos obrigar a responder” (Calvino, 1990, 44).

As análises quantitativas, as medidas objetivas, a cientometria, os exames e testes pontuais e somativos, os gráficos, as estatísticas, as listas numéricas são insuficientes e até mesmo perniciosos, se entendidos isoladamente e fora do contexto de sua produção e de sua execução. Mas, como não pode haver qualidade sem quantidade, também não pode haver uma avaliação qualitativa consistente que não tenha por base concreta os dados quantitativos de realidade. Dados e pedras, isolados, podem ser só montes de números e pó. Organizados, intencional e socialmente, os elementos objetivos e quantitativos adquirem uma significação de cultura, de conjunto articulado, constituem uma forma concreta e integradora, tornam real a sua qualidade. Sem juízo de valor, não há avaliação, só medida. Sem o quantitativo, não há o qualitativo.

Marco Polo descreve uma ponte, pedra por pedra.

- Mas qual é a pedra que sustenta a ponte? Pergunta Kublai Khan.
- A ponte não é sustentada por essa ou aquela pedra – responde Marco – mas pela curva do arco que estas formam.
- Kublai Khan permanece em silêncio, refletindo. Depois acrescenta:
- Por que falar das pedras? Só o arco me interessa.
- Sem pedras o arco não existe. (Calvino, 1990, 79)

Bibliografia

- Albrecht, R. (1994). *A avaliação formativa*. Rio Tinto, Edições Asa.
- Baudelot, C. e Establet, R. (1994). *O nível educativo sobre*. Porto, Porto.
- Bielschowsky, C. E. e Penna Firme, T. (1994). Avaliação na Universidade Federal do Rio de Janeiro. *Educación Superior y Sociedad*, Vol. 5, nºs 1 e 2. Caracas, Cresalc – Unesco.
- Calvino, I. (1990). *As cidades invisíveis*. São Paulo, Companhia das Letras.

- CIEES (s/d). *Manual de evaluación*. México, mimeo.
- Dias Sobrinho, J. (1994). A avaliação institucional na Universidade Estadual de Campinas: considerações teóricas e processo. *Educación Superior y Sociedad*, vol. 5, nºs 1 e 2, Caracas, Cresalc-Unesco.
- _____. (org.) (1994). *Avaliação institucional da Unicamp: processo, discussão e resultado*. Campinas, Editora da Unicamp.
- _____. (1995). Avaliação institucional: A experiência da Unicamp. *Pro-Posições*, nº 16. Campinas, Unicamp.
- Dias Sobrinho, J. e Balzan, N. C. (orgs.) (1995). *Avaliação institucional: teoria e experiências*. São Paulo, Cortez.
- Gairén Sallán, J. (1993). La autoevaluación institucional como vía para mejorar los centros educativos. *Bordón*, vol. 45, nº 3. Madrid, Sociedad Española de Pedagogía.
- García-Guadilla, C. (1995). Globalización y conocimiento en tres tipos de escenarios. *Educación Superior y Sociedad*, vol. 6, nº 1. Caracas, Cresalc-Unesco.
- Martínez, E. (ed.) (1997). *Ciência, tecnologia y desarrollo: interrelaciones teóricas y metodológicas*. Caracas, Nueva Sociedad, Unesco.
- Pérez Gómez, A. I., in: Gimeno Sacristán, J. e Pérez Gómez, A. (orgs.) (1995). *Comprender y transformar la enseñanza*. 4ª edición, Madrid, Ediciones Morata.
- Santos Guerra, M. A. (1993). *La evaluación: un proceso de diálogo, comprensión y mejora*. Granada, Ediciones Aljibe.
- Schwartzman, S. (1992). O contexto educacional e político da avaliação. In Durham, Eunice e Schwartzman, S. (orgs.). *Avaliação do ensino superior*. São Paulo, Edusp.
- Unesco (1995). *Documento de política para a mudança e o desenvolvimento na educação superior*. Caracas, Venezuela.
- Velho, L. (1994). Indicadores científicos: aspectos teóricos y metodológicos. In Martínez, E. (ed.). *Ciencia, tecnologia y desarrollo: interrelaciones teóricas y metodológicas*. Caracas, Nueva Sociedad, Unesco.
- Torroba Arroyo, I. (1993). Evaluación cualitativa de centros en el marco de la LOGSE. *Bordón*, vol. 45, nº 3. Madrid, Sociedad Española de Pedagogía.