

AVALIAÇÃO DE PROFESSORES: UMA QUESTÃO DE ESCOLHAS

NATHALIA CASSETTARI

RESUMO

O artigo apresenta uma revisão da literatura sobre a avaliação de professores, destacando os diferentes objetivos, critérios e instrumentos que podem ser utilizados na construção de novas propostas. As alternativas disponíveis refletem concepções muitas vezes divergentes, o que pode repercutir em enfrentamentos durante o processo de tomada de decisões. Ainda assim, é grande o consenso em relação à necessidade desse processo ser o mais participativo possível, sendo ressaltada a importância do envolvimento dos professores. Outro consenso encontrado na bibliografia refere-se à superioridade das iniciativas que empregam mais de um instrumento avaliativo, possibilitando uma análise mais completa do trabalho docente.

PALAVRAS-CHAVE AVALIAÇÃO DO DESEMPENHO DOCENTE • POLÍTICA EDUCACIONAL • AVALIAÇÃO EDUCACIONAL • METODOLOGIA DA AVALIAÇÃO.

RESUMEN

El artículo presenta una revisión de la literatura sobre la evaluación de profesores, destacando los distintos objetivos, criterios e instrumentos que es posible utilizar en la construcción de nuevas propuestas. Las alternativas disponibles reflejan concepciones muchas veces divergentes, lo que se puede convertir en enfrentamientos durante el proceso de toma de decisiones. Aun así, es grande el consenso en relación a la necesidad de que este proceso sea lo más participativo posible, subrayándose la importancia del involucramiento de los profesores. Otro consenso encontrado en la bibliografía se refiere a la superioridad de las iniciativas que emplean más de un instrumento evaluativo, posibilitando un análisis más completo del trabajo docente.

PALABRAS CLAVES EVALUACIÓN DEL DESEMPEÑO DOCENTE • POLÍTICA EDUCACIONAL • EVALUACIÓN EDUCACIONAL • METODOLOGÍA DE LA EVALUACIÓN.

ABSTRACT

This paper presents a literature review on teacher evaluation, highlighting the different objectives, criteria, and tools that can be used in elaborating new proposals. The available alternatives will oftentimes reflect diverging conceptions, which may result in conflicts during the decision-making process. Despite that, there is considerable consensus as to the necessity for such process to be as participative as possible, the teachers' involvement being paramount. Another consensus found in the bibliography refers to the superiority of ideas that make use of more than one evaluation tool, allowing for a more thorough analysis of the teachers' work.

KEYWORDS TEACHER'S PERFORMANCE EVALUATION • EDUCATIONAL POLICY • EDUCATIONAL EVALUATION • EVALUATION METHODOLOGY.

1 O levantamento dos textos foi feito principalmente na internet nos seguintes *sites*: Banco de Teses da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes (<http://capesdw.capes.gov.br/capesdw/>); *Scientific Eletronic Library Online* – SciELO (<http://www.scielo.br/>); *Education Resources Information Center* – Eric (<http://www.eric.ed.gov/>); Sistema Integrado de Bibliotecas da Universidade de São Paulo – SiBi (<http://www.usp.br/sibi/>); e Google Acadêmico (<http://scholar.google.com>). As buscas foram feitas, sempre que possível, em três idiomas (português, inglês e espanhol) e evidenciaram a existência de um grande número de trabalhos no idioma inglês sobre a avaliação de professores, em especial de origem norte-americana, e um número ainda muito restrito de publicações nacionais sobre o tema.

INTRODUÇÃO

Existem inúmeras possibilidades para a criação de programas de avaliação de professores; diversas alternativas que podem ser combinadas de diferentes maneiras, gerando os mais variados resultados. Como não há nenhum formato que possa ser considerado ideal para todas as circunstâncias, a definição das características de cada programa depende de uma sequência de escolhas que são influenciadas por diferentes fatores, entre eles: as concepções e os valores dos envolvidos no processo de tomada de decisões, a correlação de forças entre os programas e a realidade em que estão inseridos (legislação vigente, grau de organização e mobilizações das associações de professores, recursos disponíveis etc.). De acordo com Duke e Stiggins (1990, p. 127), “sistemas de avaliação não existem no vácuo [...] um ambiente avaliativo é um reflexo do sistema de ensino e até mesmo da comunidade em que ele se localiza”.

O objetivo deste trabalho é apresentar as alternativas disponíveis na literatura¹ para a elaboração de programas de avaliação de professores. Conhecer tais alternativas pode

contribuir tanto para a formulação de novas propostas, quanto para a análise de iniciativas em curso.

O artigo foi organizado a partir das seguintes questões: por que avaliar os professores? – referindo-se às justificativas e aos propósitos da avaliação; o que avaliar? – referindo-se aos critérios utilizados; e como avaliar? – referindo-se aos instrumentos disponíveis. Para melhor compreensão do texto, é preciso ter em mente que essas questões estão interligadas, ou seja, que as decisões tomadas influenciam as escolhas subsequentes. Nesse sentido, alguns dos instrumentos disponíveis serão mais apropriados para determinados objetivos da avaliação e outros instrumentos para outros objetivos.

Cabe ressaltar que, apesar do amplo reconhecimento de que as escolhas sobre os sistemas de avaliação de professores geram disputas, em nenhum dos textos analisados houve a defesa de procedimentos autoritários para a tomada de decisões. Apesar de alguns autores não se posicionarem sobre o assunto, é grande o número dos que destacam as vantagens decorrentes do estabelecimento de um processo participativo de elaboração da avaliação, envolvendo os professores em todas as suas etapas. Entre as vantagens elencadas, destacam-se: a melhor adequação das propostas à realidade das escolas, a maior facilidade na obtenção do apoio dos participantes e a possibilidade de reflexão sobre o trabalho docente decorrente da discussão sobre a sua avaliação (STRONGE; TUCKER, 2003; DANIELSON; MCGREAL, 2000; BACHARACH et al., 1990).

Também não foram encontrados na bibliografia posicionamentos totalmente contrários à avaliação de professores. Mesmo nos textos que fazem fortes críticas às políticas vigentes predomina o entendimento de que são as características dos programas em questão que apresentam problemas, e de que, em outro formato, a avaliação de professores poderia trazer inúmeras contribuições para esses profissionais, bem como para as escolas e os sistemas de ensino.

POR QUE AVALIAR OS PROFESSORES?

A busca pela qualidade docente como condição primordial para uma educação de qualidade, aparece como a principal

justificativa para a criação de sistemas de avaliação de professores. A ideia é que essas avaliações não apenas ajudem a verificar a qualidade do corpo docente, mas contribuam com a melhoria de tal qualidade.

Entretanto, partindo de concepções distintas sobre a “qualidade docente”, diferentes objetivos podem ser atribuídos à avaliação de professores, entre os quais: orientar decisões referentes à carreira (contratação, efetivação, promoção e demissão); reconhecer e premiar contribuições meritórias; avaliar a qualidade dos serviços prestados e dos processos em que eles são desenvolvidos; estimular reflexões sobre a prática profissional; identificar pontos fortes e fracos dos avaliados; prescrever políticas que ajudem a remediar os problemas encontrados; orientar programas de treinamento e formação em serviço; providenciar *feedback* para professores e gestores etc. (JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION, 2009; SHINKFIELD; STUFFLEBEAM, 1995).

Tais objetivos podem ser agrupados em duas categorias: a que privilegia a função formativa da avaliação e o desenvolvimento profissional dos professores por meio de avaliações com fracas consequências (*low stakes*); e a que privilegia a função somativa² da avaliação e a responsabilização dos professores, por meio de avaliações com altas consequências (*high stakes*).

² Cabe o esclarecimento que, neste trabalho, a oposição entre a avaliação formativa e a somativa não se refere ao momento em que a avaliação é realizada – durante ou ao final do processo –, mas sim aos seus usos, conforme definição utilizada por Gardner *et al.* (2010, p. 6), segundo a qual: “o termo ‘formativo’ se aplica unicamente aos usos da avaliação que apoiam a aprendizagem”.

Nas avaliações que privilegiam a função formativa, existe o entendimento de que todos os professores podem aprender algo novo que os ajude a aprimorar sua prática profissional, ainda que se reconheça que alguns (iniciantes e com maiores dificuldades) precisem de mais apoio do que outros. Entre os objetivos dessas avaliações, Bauer (2013) cita os seguintes: melhorar e assegurar a qualidade do ensino; fornecer informações a pesquisas avaliativas sobre ensino; proporcionar o desenvolvimento de equipe – informação para basear as propostas de desenvolvimento profissional/formação continuada; buscar plano de desenvolvimento pessoal – identificar pontos fracos e áreas em que cada docente precisa investir; propiciar a autorreflexão do professor sobre o seu trabalho (autoavaliação); acompanhar a prática docente visando a sua melhoria; fornecer informações para conselhos escolares.

O fato de essas avaliações não servirem para a tomada de decisões com grandes consequências para os professores as diferencia dos processos aqui denominados somativos e provoca intensos debates na área. Por um lado, a ausência de grandes consequências associadas à avaliação gera questionamentos sobre como lidar com professores que demonstraram *performances* muito abaixo ou muito acima do esperado; por outro, essa ausência é entendida como essencial para que os professores não se sintam ameaçados pela avaliação e, assim, colaborem com os processos avaliativos e com as iniciativas de desenvolvimento profissional deles decorrentes.

Já as avaliações de cunho somativo podem apresentar os seguintes propósitos: seleção de docentes/contratação; concessão de licenças; certificação; aprimoramento de indicadores de gestão; distinção e premiação dos melhores profissionais; decisões sobre penalidades (redução salarial, afastamentos, demissão, rebaixamento de cargo etc.); seleção para formação continuada; decisões sobre benefícios (aumento salarial, promoção, pagamento por mérito etc.); classificação de docentes para fins de *accountability* (BAUER, 2013).

A maior crítica em torno desses propósitos refere-se à dificuldade de se estabelecerem processos justos que orientem a tomada de decisões com grandes consequências para os professores. Nesse sentido, existe o risco de as avaliações acabarem punindo ou premiando os professores por resultados que estão fora do seu controle (relacionados, por exemplo, às condições de trabalho ou ao nível socioeconômico dos seus alunos) ou, ainda, de estimularem comportamentos questionáveis por parte dos professores. Esse último fator tem como base a lei formulada por Campbell (1979, p. 85), que atesta: “Quanto mais um indicador social quantitativo é utilizado para tomada de decisão social, mais sujeito estará a pressões da corrupção e mais propenso a distorcer e corromper os processos sociais que pretende monitorar”.³

Existe ainda a discussão sobre a (im)possibilidade de combinação dos objetivos formativos e somativos em um único processo avaliativo. Para os que acreditam que tal combinação não apenas é possível como também desejável, há o entendimento de que o desenvolvimento

3 Todas as citações de obras estrangeiras presentes no texto foram traduzidas livremente pela autora deste artigo.

profissional e a responsabilização não são incompatíveis, mas sim complementares (DANIELSON; MCGREAL, 2000; STRONGE; TUCKER, 2003), ainda que admitam que esse tipo de avaliação apresente desafios complexos, em especial quanto ao reconhecimento por parte dos professores de que, ao buscarem seu aperfeiçoamento profissional, estarão também alcançando os objetivos do sistema de ensino (SHINKFIELD; STUFFLEBEAM, 1995).

Entretanto, alguns autores ressaltam que as duas vertentes partem de concepções distintas sobre o desenvolvimento profissional dos professores. Enquanto a avaliação somativa pressupõe que a busca das premiações ou o medo das punições estimula o desenvolvimento profissional, a avaliação formativa trata de identificar (e depois resolver) possíveis deficiências dos professores, que teriam resistência em expor suas deficiências se isso pudesse prejudicá-los (DUKE; STIGGINS, 1990; POPHAM, 1988). Esses autores reconhecem a importância tanto dos propósitos formativos quanto dos propósitos somativos da avaliação, mas acreditam que eles não podem ser combinados em um único processo, nas palavras de Popham (1988, p. 269): “essas duas funções da avaliação são esplêndidas se separadas, mas contraproduativas quando combinadas”.

Desse modo, é preciso considerar que a avaliação de professores pode ter muitos objetivos diferentes, sejam eles mais formativos, somativos, ou uma combinação dos dois, e que cada uma dessas alternativas tem potencialidades e limitações próprias.

O QUE AVALIAR?

A discussão sobre os critérios que podem ser utilizados nas avaliações de professores guarda fortes relações com o debate sobre a qualidade docente. Nesse sentido, sem uma clara definição do que se espera dos professores (ou do que se entende por qualidade docente) não é possível estabelecer critérios e padrões avaliativos. Nas palavras de Darling-Hammond (1990, p. 26): “Qualquer método avaliativo é acompanhado de uma concepção do que é a docência, e mais especificamente, uma concepção do que é uma docência boa e eficaz”.

O problema é que não é fácil definir o que constitui a qualidade docente, uma vez que o trabalho dos professores tem várias dimensões que podem ser mais ou menos valorizadas de acordo com o contexto em que está inserido e com as concepções de educação e sociedade dos envolvidos. Contudo, reconhecer essa dificuldade não é negar a possibilidade (e necessidade) de os sistemas de ensino determinarem o que esperam de seus profissionais e o que consideram uma boa prática docente.

Os critérios de avaliação encontrados na literatura podem servir como início de discussão nas escolas e nas redes de ensino, pois a partir deles podem ser debatidos quais são mais interessantes e quais adaptações são necessárias. Danielson e McGreal (2000) ressaltam que esse processo deve gerar oportunidades para amplas discussões e customizações, para que exista um senso de pertencimento e comprometimento com os critérios escolhidos. Cabe ainda ressaltar que essas discussões são políticas e não apenas técnicas, conforme expressa Bauer (2013, p. 32):

[...] a discussão a ser realizada não é eminentemente técnica, mas é política, à medida que indica a necessidade de discutir o que é um bom professor e quais são as características esperadas de um ensino de qualidade. É essa a discussão que deveria estar na base da definição dos critérios de avaliação dos docentes [...]. Apesar de essa lógica poder parecer óbvia, o que se tem observado atualmente é que a discussão sobre o bom professor e o bom ensino fica subsumida ao debate sobre o que se pode mensurar, considerando, inclusive, os custos do processo, invertendo a ordem lógica da avaliação.

Em geral, os critérios e padrões utilizados para a avaliação de professores se baseiam em listas dos componentes da atividade docente. Uma das principais listas foi elaborada por Scriven (1994) e contém cinco domínios que, segundo o autor, devem ser considerados igualmente importantes: conhecimento da matéria de ensino; competência de instrução; competência de avaliação; profissionalismo; outros deveres para com a escola e a comunidade. Cada um des-

ses domínios contém uma série de elementos que são apresentados com clareza pelo autor, que também sugere uma maneira de classificar o desempenho dos professores em: competente (quando a maioria dos requisitos é preenchida de maneira satisfatória); boa (quando a maioria dos requisitos é preenchida para além do satisfatório); e excelente (quando a maioria dos requisitos é preenchida de maneira excepcional). Para Scriven (1994, p. 152), uma *performance* excelente "é uma meta à qual nós professores não só devemos aspirar, mas também obter progressos significativos".

Mais uma lista foi desenvolvida por Danielson (1996) e contempla quatro domínios: componentes da prática profissional; ambiente da sala de aula; instrução; responsabilidades profissionais. Apesar de apresentar vários pontos em comum com os elencados por Scriven, essa lista é bem mais detalhada e se desdobra em 22 componentes e 76 elementos. A autora ainda estabelece níveis de *performance* para cada um dos domínios e componentes da lista, que vão desde o insatisfatório, passando pelo básico e pelo proficiente, até o distinto. Essa lista é utilizada como base para a avaliação de professores de diversos sistemas educacionais norte-americanos e é uma forte inspiração para o *Marco para la buena enseñanza* (CHILE, 2008), matriz de referência do Sistema de Evaluación del Desempeño Profesional Docente do Chile.

Destacam-se ainda os critérios utilizados pelo Interstate Teacher Assessment and Support Consortium (InTASC) e pelo National Board for Professional Teaching Standards (NBPTS), ambas instituições norte-americanas responsáveis pela avaliação de milhares de professores anualmente. O InTASC tem como base dez padrões essenciais de *performance*: 1. Entendimento do desenvolvimento do estudante; 2. Compreensão das diferenças de aprendizagem; 3. Criação de um ambiente favorável à aprendizagem; 4. Conhecimento do conteúdo; 5. Aplicação do conteúdo; 6. Avaliação; 7. Planejamento da instrução; 8. Estratégias de instrução; 9. Desenvolvimento profissional e conduta ética; 10. Liderança e colaboração (COUNCIL OF CHIEF STATE SCHOOL OFFICERS, 2013).

Já o NBPTS toma como base cinco proposições fundamentais para definir seus padrões de sucesso no ensino:

1. Professores estão comprometidos com seus alunos e com a aprendizagem; 2. Professores conhecem a matéria que ensinam e sabem como ensiná-la para seus alunos; 3. Professores são responsáveis pela gestão e pelo monitoramento da aprendizagem dos alunos; 4. Professores pensam sistematicamente sobre sua prática e aprendem com a experiência; 5. Professores são membros de comunidades de aprendizagem (NATIONAL BOARD FOR PROFESSIONAL TEACHING STANDARDS, 2002).

No Brasil, uma primeira iniciativa de estruturação de padrões nacionais para a avaliação de professores encontra-se nos “Referenciais para o Exame Nacional de Ingresso na Carreira Docente: documento para consulta pública”, elaborados pelas equipes técnicas do MEC/Inep (BRASIL, s/d). Nesse documento são listados vinte fatores que compõem o perfil de um “bom professor”, com base em padrões docentes estabelecidos nos sistemas educacionais de países: Austrália, Canadá, Cingapura, Chile, Cuba, Estados Unidos e Inglaterra.

A listagem brasileira contempla desde fatores ligados ao conhecimento dos professores (tanto da matéria de ensino, da didática e do desenvolvimento dos alunos, quanto do sistema educacional e da legislação vigente), passando por fatores relacionados à instrução (planejamento das atividades; seleção e aplicação de recursos de aprendizagem e estratégias de avaliação apropriados; manutenção de um clima na sala de aula favorável à aprendizagem, com normas de convivência e altas expectativas para os alunos; boa organização do tempo; utilização de estratégias de aprendizagem desafiantes; entre outros), até fatores ligados ao profissionalismo e à boa relação com a comunidade escolar (constante reflexão e aprimoramento do seu trabalho, trabalho em equipe, comunicação com os pais de alunos etc.).

Com base no exposto, temos diferentes critérios e padrões de *performance* disponíveis na literatura, que partem de definições do que é esperado do trabalho docente e o desdobram em aspectos ou atividades considerados de maior relevância. É preciso reconhecer que, apesar das ênfases diferenciadas, muitos dos tópicos levantados se repetem em várias listagens – por exemplo, o conhecimento da matéria de ensino e das estratégias de instrução, a capacidade de

avaliar a aprendizagem dos alunos, a disposição para o trabalho em equipe, entre outros – e, conforme exposto anteriormente, podem iniciar as discussões sobre quais critérios melhor se adequam à realidade de cada sistema de ensino e aos objetivos de cada programa de avaliação de professores.

COMO AVALIAR?

A revisão da literatura mostra diferentes instrumentos disponíveis para a avaliação de professores, sendo alguns mais adequados para alcançar certos objetivos do que outros, e alguns mais apropriados para determinadas concepções da qualidade docente do que outros.

Uma primeira discussão sobre o tema refere-se à superioridade das avaliações que combinam diferentes instrumentos, possibilitando uma melhor análise do trabalho docente. Peterson (1990, p. 110) afirma que “nenhuma linha de evidência fornece, sozinha, uma imagem total do que os professores fazem”. Darling-Hammond (1990, p. 26) diz que “qualquer método é mais útil em combinação com outros que complementem as informações por ele fornecidas” e que “o conjunto de métodos mais adequados pode variar de acordo com os seus propósitos ou funções”. E Barber (1990, p. 224) acrescenta:

As fraquezas de algumas técnicas podem ser amenizadas ou equilibradas ao combiná-las com outras técnicas. O propósito é permitir aos avaliados uma sensação de conforto e confiança, juntamente com a melhoria do ensino e, ao mesmo tempo, atender às demandas de administradores e do público por objetividade.

Assim, seria desejável que os sistemas combinassem diversas fontes de informação para a construção de melhores avaliações de professores, contudo, esse nem sempre é o caso, não sendo raras as vezes em que a avaliação se restringe a um único instrumento.

Cada um dos instrumentos disponíveis para a avaliação de professores possui características específicas que fornecem diferentes informações sobre o trabalho docente e

exigem uma quantidade variada de recursos (tempo para elaboração, aplicação e análise dos instrumentos; contratação de pessoal; treinamento dos avaliadores; cópia e distribuição de materiais impressos etc.). Por isso, eles serão apresentados separadamente a seguir.

OBSERVAÇÕES EM SALA DE AULA

Uma das particularidades do trabalho docente é que, na maioria das vezes, ele é realizado dentro de salas de aula e, portanto, os dados a ele referentes limitam-se às pessoas que têm acesso cotidiano a esse espaço: em geral, o próprio professor e seus alunos.

As observações permitem que o avaliador entre nas salas de aulas (fisicamente ou por meio de filmagens) a fim de recolher informações essenciais sobre o trabalho docente que não podem ser obtidas de outra forma, entre as quais: a maneira como os professores apresentam os conteúdos aos seus alunos; como distribuem o tempo das suas aulas; como organizam o espaço físico; como se relacionam com seus alunos (atentando se há diferença no tratamento destes de acordo com sexo, etnia, classe social e desempenho acadêmico); como os alunos se relacionam entre si; como participam das aulas etc. Danielson e McGreal (2000, p. 47) afirmam: “Um observador astuto pode notar como o professor estrutura o ambiente físico, como envolve os alunos na aprendizagem e como estabelece e mantém padrões de conduta”.

Ainda que as observações possibilitem o acesso a importantes informações sobre o trabalho docente, é preciso ter consciência de suas limitações. Nesse sentido, ressalta-se a dificuldade do observador de apreender o que de fato acontece na sala de aula, e não uma situação artificial, criada justamente por sua presença física no ambiente ou quando é representado por um equipamento de filmagem. Nas palavras de Stronge e Tucker (2003, p. 54), “observações formais da sala de aula podem trazer uma visão artificial do que ocorre em seu dia a dia”.

Stodolsky (1990) afirma, por sua vez, que as observações são bons instrumentos para avaliar comportamentos e ações dos professores, mas são limitados para avaliar seus

pensamentos e sentimentos. Assim, as observações podem indicar como os professores se comportam diante dos alunos, mas não o porquê de se comportarem dessa maneira e nem o que pensam sobre esse comportamento.

Outra limitação das observações diz respeito a aspectos que não podem ser percebidos nas visitas às salas de aula, por exemplo, o planejamento dos professores, a análise e as modificações de materiais didáticos, a escolha e a adaptação de metodologias de ensino e a relação dos professores com os colegas, os pais de alunos e outros membros da comunidade escolar (STODOLSKY, 1990).

Existe ainda o risco de os observadores apresentarem análises enviesadas do que acontece nas salas de aula. Strong (2011) destaca as operações cognitivas que podem afetar as observações, entre elas: o viés de confirmação, a cegueira não intencional, a racionalidade limitada; a percepção influenciada etc. Entretanto, a formação adequada dos observadores pode amenizar esse problema.

As observações em sala de aula variam de acordo com quem as realiza (em especial se o avaliador é externo ou interno); com o tipo de inferência exigida (pode-se cobrar dos avaliadores informações sobre comportamentos fáceis de identificar e mensurar, como também outros que exigem um maior nível de julgamento); com o foco das observações (podendo variar desde uma análise mais global do trabalho docente até a verificação de alguns comportamentos e atividades descritos em um *checklist*); e com o número de observações realizadas.

Em alguns casos, os professores são informados com antecedência das observações, em outros não. A notificação dos professores tem a vantagem de permitir que o observador realize reuniões prévias com o professor nas quais são discutidos os objetivos e o formato da observação, e o objetivo e o formato da aula, permitindo que o observador colete informações adicionais sobre o trabalho docente. Já as observações realizadas sem nenhum tipo de notificação evitam que os professores preparem um ambiente artificial para suas aulas.

Em estudo realizado como parte do Measures of Effective Teaching Project (2012), são definidos seis requisitos

mínimos para observações de alta qualidade: 1. Instrumentos de observação devem ter claras expectativas – definição das competências que serão observadas com exemplos de diferentes níveis de *performance*; 2. Observadores devem demonstrar precisão antes de taxarem o desempenho dos professores – necessidade de treinamento e de confirmação periódica da habilidade dos observadores; 3. Necessidade de múltiplas observações e de alta confiabilidade delas quando decisões de grandes consequências são tomadas; 4. Monitoração do nível de confiabilidade do sistema deve ser constante, por meio de observações imparciais de avaliadores externos, sem relações com os professores; 5. Combinação das observações com outras fontes de informações, como *feedback* dos alunos e o desempenho destes em avaliações externas; 6. Acompanhamento regular dos professores mais bem avaliados para verificar se suas turmas são diversificadas ou se são compostas apenas de alunos com os melhores desempenhos.

PESQUISAS DE OPINIÃO DOS ESTUDANTES

Os estudantes vivenciam diariamente o trabalho dos seus professores e, portanto, podem trazer importantes contribuições para a avaliação desses, sobretudo no ensino superior, no qual é bastante utilizado. Contudo, a validade das respostas dos alunos é questionada, especialmente na educação básica. Segundo Strong (2011, p. 78),

Ainda que os estudantes tenham o conhecimento mais íntimo e consistente sobre o que os professores fazem na sala de aula, a sua opinião é a procurada com menor frequência com o propósito de avaliar os professores. A avaliação de alunos é mais usada no ensino superior, onde existem diversas razões para considerá-las válidas.

Estudos como os de Peterson, Wahlquist e Bone (2000) e Wilkerson *et al.* (2000) indicam que os alunos da educação básica podem fornecer informações válidas e confiáveis para a avaliação de professores. Os padrões de resposta verificados nos instrumentos aplicados, sugerem que eles respondem às questões com razão, intenção e valores consistentes.

O primeiro estudo mostra ainda que eles conseguem diferenciar, em suas respostas, os professores de que mais gostam daqueles com quem aprendem mais.

A discriminação nas estruturas fatoriais entre um professor como fonte de aprendizagem e como uma pessoa que demonstra respeito e cuidado corrobora a validade da fonte de dados: os alunos de diferentes idades distinguem entre essas duas dimensões. Pesquisas com estudantes não são apenas concursos de popularidade, os alunos distinguem entre apenas gostar de um professor e reconhecer aquele que favorece a sua aprendizagem. Enquanto os estudantes podem distinguir entre um professor que apoia a aprendizagem e um que os trata bem, este estudo indica que o primeiro é mais importante para os alunos mais velhos, enquanto o último é mais importante para os mais novos. (PETERSON; WAHLQUIST; BONE, 2000, p. 148)

Entretanto, Goe (2007, p. 40) ressalta que “a confiabilidade e a validade das pesquisas com estudantes dependem, em certa medida, do instrumento utilizado, de como ele foi desenvolvido, de como é administrado e do nível de detalhe que pretende medir”.

Deve-se considerar que os alunos não são boas fontes de informação de todas as dimensões da qualidade docente; por isso, é necessário ter muita atenção ao formular as questões a eles direcionadas. Não é recomendável, por exemplo, perguntar-lhes a respeito do conhecimento dos professores sobre a matéria ensinada, mas sim, sobre certos eventos que ocorrem na sala de aula e sobre a interação entre os professores e os alunos.

É preciso também que as questões direcionadas aos estudantes sejam apropriadas à sua idade. Por isso, em geral, são elaborados instrumentos diferentes para alunos de idades diferentes. Quanto maior a criança, mais complexas podem ser as questões e mais abertas podem ser as possibilidades de resposta. Para crianças não alfabetizadas, as questões devem ser bem explicadas, e elas podem expressar sua opinião por meio de desenhos ou da indicação de símbolos

(por exemplo: ☺ para sim/sempe, ☹ para às vezes e ☹ para não/nunca).

Peterson, Wahlquist e Bone (2000) defendem que os questionários incluam um item global do tipo “o meu professor é um bom professor?”. Em sua pesquisa, eles descobriram que a resposta a esse item sintetiza as demais respostas dos questionários, expressando a opinião dos estudantes sobre os seus professores.

É recomendável, ainda, que os professores não participem nem da aplicação do instrumento e nem da contabilização das respostas dos alunos para que não haja conflitos de interesse e perda da credibilidade da avaliação.

De maneira geral, a opinião dos estudantes é importante porque eles analisam o trabalho docente de um ponto de vista particular, diferente de qualquer outro grupo. Nessa perspectiva, Stronge e Tucker (2003, p. 59) afirmam: “O *feedback* dos estudantes pode ser uma fonte viável de informação que aumenta o nosso entendimento do ensino e aprendizagem e, portanto, pode servir como uma adição valiosa para o quadro de avaliação de professores”.

PESQUISAS DE OPINIÃO DOS PAIS DE ALUNOS

O argumento de que a opinião dos pais deve fazer parte da avaliação de professores baseia-se na percepção de que eles têm interesse direto na qualidade do trabalho docente e de que analisam esse trabalho de uma maneira singular, podendo oferecer informações diferentes (e complementares) das encontradas em outras fontes.

Além disso, entende-se que a comunicação com os pais faz parte do trabalho docente e que o envolvimento deles na escola traz muitos benefícios para a educação (MASTER, 2013), devendo ser estimulado, inclusive, nos processos de avaliação de professores. Apesar de desejável, esse envolvimento nem sempre acontece, ainda que estimulado pelos professores e pelas escolas, o que levanta questionamentos sobre a capacidade dos pais de opinarem sobre o desempenho dos professores.

Existem também questionamentos quanto à validade e confiabilidade das respostas. Contudo, estudos como os de

Peterson *et al.* (2003) e Master (2013) mostram que os pais oferecem uma medida estável e distinta da qualidade dos professores, sendo possível estabelecer diferenciações entre eles (dentro de uma mesma escola e entre escolas diferentes) que se mantêm estáveis ao longo dos anos (em avaliações dos mesmos professores realizadas por um novo conjunto de pais).

Uma recomendação recorrente na literatura (PETERSON *et al.*, 2003; STRONGE; TUCKER, 2003; DANIELSON; MCGREAL, 2000; PETERSON, 1990) refere-se ao tipo de questões que devem ser colocadas aos pais de alunos. Por exemplo, eles podem responder sobre a comunicação que estabelecem com os professores, sobre as lições de casa dos alunos e o suporte oferecido pelos professores para a aprendizagem em casa, mas não são as pessoas mais adequadas para opinarem sobre o conhecimento dos professores acerca da matéria de ensino, sobre o que os professores fazem em sala de aula, etc. Para Peterson *et al.* (2003, p. 327), é preciso considerar os seguintes fatores na elaboração dos questionários:

Os pais: a) não estão na sala de aula; b) não conhecem importantes dados comparativos entre os alunos; c) são especialistas nos seus filhos, mas não no ensino e aprendizagem que acontecem na sala de aula; d) são influenciados pela reação emocional ao sucesso de uma criança ao invés do sucesso global da sala de aula; e e) são excessivamente influenciados pela sua própria história escolar.

A análise da literatura indica que a opinião dos pais pode trazer benefícios para a avaliação de professores, mas é preciso considerar suas limitações. Peterson *et al.* (2003, p. 327) afirmam que um “alto índice na avaliação dos pais não necessariamente significa a mesma coisa que uma boa docência. Talvez, a melhor interpretação seja que um alto índice dos pais, em conjunto com vários outros indicadores positivos, é um bom indicador da qualidade docente”.

RELATÓRIOS ELABORADOS POR PARES

Os professores têm muito a dizer sobre o desempenho de seus colegas, pois conhecem as metodologias de ensino, o

conteúdo curricular e, principalmente, a realidade das salas de aula. Fernandes (2008, p. 23) afirma:

São os pares que estão em melhor posição para se pronunciar com conhecimento de causa acerca da competência e do desempenho dos seus colegas. São os pares que estão mais familiarizados com os contextos concretos e específicos em que os professores trabalham e com as exigências a que têm de dar resposta. Por isso mesmo podem formular sugestões específicas e práticas que apoiem os processos de mudança e de melhoria.

Além disso, a participação dos pares no processo avaliativo cria uma oportunidade singular para que os professores conheçam o trabalho de seus colegas e, assim, possam compartilhar e reproduzir boas práticas.

Para a elaboração dos relatórios, os professores podem realizar reuniões com seus colegas, observar as salas de aula ou ainda analisar os materiais didáticos por eles utilizados. Não são todos os professores que devem avaliar o desempenho de seus colegas, mas apenas aqueles que têm uma *performance* reconhecida como equivalente ou superior à de seus pares. De outro modo, corre-se o risco de os professores “não reconhecerem competência dos avaliadores e, por isso, não ouvirem ou não aceitarem as suas opiniões” (FERNANDES, 2008, p. 28).

Em alguns casos são os próprios professores que escolhem por quem serão avaliados; em outros, essa decisão parte de órgãos superiores. De qualquer maneira, para alguns professores avaliar os colegas implica um novo encargo, que pode ou não ser recompensado financeiramente.

Em alguns sistemas, professores experientes saem da sala de aula para exercer a função de avaliadores. Nesse caso, questiona-se se esses professores podem continuar a ser considerados “pares”. Para McCarthy e Peterson (1988, p. 260), “um professor que está fora da sala de aula por mais de seis semanas não é mais um par por conta das diferentes pressões sociológicas do papel de observador”.

Uma preocupação constante na literatura diz respeito à possibilidade de as relações pessoais, profissionais e

políticas entre os professores influenciarem os resultados das avaliações. De maneira geral, os professores prezam o bom relacionamento com os colegas e entendem que a formalização de opiniões negativas sobre o desempenho de seus pares pode comprometer esse relacionamento. Para amenizar o problema, alguns programas sugerem que os professores deem opinião sobre o desempenho de pares com quem não têm contato direto, por exemplo, professores de outras escolas. Nesse caso, existe o risco de que os professores não reconheçam a legitimidade desses outros professores, ou não os encarem como pares.

Por fim, destaca-se mais uma vez a recomendação de que a opinião dos pares não seja a única fonte de informação sobre o desempenho dos professores, mas que seja combinada com outros instrumentos avaliativos.

AUTOAVALIAÇÕES

A autoavaliação consiste em um processo no qual os professores são seus próprios avaliadores e, assim, responsáveis por estabelecer juízos de valor sobre a própria *performance*. O principal objetivo é fazer com que os professores tomem consciência das suas práticas, por meio da reflexão e interpretação de dados e, com isso, possam identificar o que precisam fazer para melhorar seu desempenho. Vale dizer que, para algumas concepções sobre o trabalho docente, a autoavaliação deve fazer parte das atribuições cotidianas de todos os professores.

Mais do que qualquer outro instrumento, a autoavaliação pode explorar as “intenções dos professores, processos de pensamento, conhecimentos, e crenças” (GOE et al., 2008, p. 38). Para Strong (2011, p. 77), a maior vantagem dessa prática é “ser capaz de lidar com as intenções dos professores e com seus conhecimentos e crenças sobre o ensino” .

Já Barber (1990, p. 218) acredita que “o maior valor da autoavaliação é a melhoria da instrução, que resulta de um maior conhecimento sobre os seus pontos fortes e fracos”. Esse autor acredita que “a autoavaliação, como parte da avaliação de professores, pode ser a força mais poderosa disponível a nós, enquanto educadores, em nossa busca incessante

para melhorar comportamentos, estratégias e técnicas de instrução” (BARBER, 1990, p. 227). Danielson e McGreal (2000, p. 24) por sua vez, sustentam que “poucas atividades são mais poderosas para a aprendizagem profissional do que a reflexão sobre a prática. [...] nós não aprendemos tanto da experiência, mas das nossas reflexões sobre a nossa prática”.

Como instrumento de avaliação de professores, a autoavaliação tem uma característica particular: ela só funciona como parte de avaliações formativas, ou seja, daquelas que têm como principal objetivo o desenvolvimento profissional e não são utilizadas nem para premiar e nem para punir os professores. Essa particularidade resulta da necessidade de que os professores sejam honestos em suas autoavaliações, o que pode não acontecer quando existem grandes consequências em jogo.

Carvalho (2011, p. 67) acrescenta a isso que é como fundamental para o sucesso das autoavaliações que os professores estejam abertos à crítica sobre sua prática, pois “se não assumir que a humildade é um elemento indispensável para crescer profissionalmente e não reconhecer que pode aprender com os outros, não é capaz de levar a cabo uma autoavaliação válida e precisa para melhorar a sua docência”.

A autoavaliação não está imune a críticas. Barber (1990) lista as principais limitações desse instrumento: 1. Falta de objetividade; 2. Falta de precisão e confiabilidade; 3. Tendência de os indivíduos se considerarem proficientes; 4. Possibilidade de a avaliação se tornar um mecanismo de autojustificação (apoia uma visão prévia e transfere a responsabilidade da prática docente a fatores extrínsecos); 5. Propensão de os professores “mediócras” serem menos precisos em suas autoavaliações do que os professores mais competentes; 6. Tendência de os professores dependerem de consultores, caso eles sejam utilizados; 7. Potencial inerente de autoincriminação, se os resultados forem usados de forma somativa; 8. Dificuldade de quantificar a avaliação; 9. Propensão de os professores se concentrarem em questões periféricas, não em questões centrais, quando analisam sua *performance*.

Goe *et al.* (2008) ressaltam que as respostas dos professores podem ser suscetíveis ao que chamaram de “desejabilidade social” (*social desirability*), isto é, uma tendência dos

professores apresentarem resultados que seriam esperados, mas que não necessariamente condizem com a realidade.

Afirmam as autoras:

Esse fenômeno incluiria tanto a deturpação consciente das práticas docentes para que elas sejam bem vistas, como a deturpação não intencional devido à percepção do professor de que ele ou ela estaria implementando corretamente a prática quando, na verdade, ela não está sendo implementada com fidelidade. Esse viés em potencial pode levar tanto ao sobrerregistro, quanto à subnotificação de práticas, tornando os dados difíceis de interpretar. (GOE et al., 2008, p. 39)

Entretanto, para Carvalho (2011) existem algumas condições que, quando atendidas, podem minimizar as limitações das autoavaliações: 1. Criação de um ambiente em que os professores se sintam seguros e à vontade para participar da autoavaliação – para tanto, a garantia de anonimato e confidencialidade é essencial; 2. Formação dos professores – para que eles aprendam a refletir sobre a sua prática; 3. Disponibilização dos recursos (inclusive, e talvez principalmente, de tempo) para que os professores possam realizar as autoavaliações; 4. Estabelecer, em conjunto com os professores, noções claras sobre o que consiste a qualidade docente; 5. Fornecer padrões de *performance* com base nas definições de qualidade docente, para que os professores tenham com o que comparar as suas práticas; 6. Garantir que os professores tenham uma pessoa que possa tirar suas dúvidas e orientá-los sobre o processo; 7. Providenciar os meios necessários para que os professores possam aprimorar as práticas que consideram deficitárias.

PORTFÓLIOS

Portfólios são instrumentos parecidos com as autoavaliações no sentido de que são preparados pelos próprios professores e exigem algum nível de reflexão sobre o trabalho realizado. A diferença é que eles exigem um conjunto de registros e documentos que comprovem as experiências e realizações de cada profissional.

Como o trabalho docente é complexo, com múltiplos objetivos difíceis de evidenciar, os portfólios de professores podem assumir diferentes características, de acordo com suas finalidades e com as concepções de qualidade docente que representam. Nas palavras de Wolf e Dietz (1998, p. 10),

O que é chamado “portfólio” pode variar de uma ampla coleção de produtos personalizados a um restrito conjunto de materiais padronizados. Apesar de todas essas versões caberem no rótulo “portfólio de ensino”, essas coleções de informações variam consideravelmente no que contêm, no como são construídas e nas maneiras como são avaliadas.

É grande a variedade dos registros que podem compor os portfólios de professores. Entre eles podem ser encontrados: planos de aulas, exemplos de atividades dos alunos, filmagens de aulas, descrição de projetos desenvolvidos, cartas de recomendações de colegas e supervisores, documentos que comprovem o envolvimento na gestão escolar, registros de comunicação com os pais de alunos, certificados de participação em atividades de formação continuada, autoavaliações etc. Na maioria dos casos, tais conteúdos são acompanhados de reflexões dos professores sobre o contexto em que foram desenvolvidos e sobre sua contribuição para a aprendizagem dos alunos. Essas reflexões permitem que os professores examinem sua prática, e podem ser uma oportunidade para o desenvolvimento profissional.

Na seleção dos conteúdos que irão compor o portfólio deve-se ter em mente os objetivos da avaliação, as concepções de trabalho docente adotadas e a garantia da viabilidade da sua execução – ou seja, os professores devem dispor do tempo e dos recursos necessários para organizar as informações indicadas. Nesse sentido, existe a ressalva para que os portfólios não sejam nem tão longos – o que demandaria muito tempo e dinheiro para sua elaboração e avaliação –, nem tão curtos – o que restringiria a análise do trabalho docente.

Para o sucesso da confecção dos portfólios, Bird (1990, p. 248) entende ser importante lembrar que o objetivo não é criar portfólios excelentes, mas sim cultivar um ensino e uma aprendizagem de excelência. Essencial o treinamento

em conjunto de professores, mentores, supervisores e avaliadores, para que todos “compartilhem seus conhecimentos e compartilhem seus conhecimentos e criem as relações de trabalho que necessitarão”. Afirma o autor:

Sem uma tradição refinada para documentar o seu trabalho, é de se esperar que os professores produzam portfólios que entediem a eles mesmos e seus supostos avaliadores. Aprender a estudar e revelar o seu próprio ensino pode ser um dos desafios mais formidáveis da ideia dos portfólios. (BIRD, 1990, p. 249)

Os portfólios de professores também podem ser diferentes com relação ao seu nível de estruturação. Os mais estruturados são organizados de acordo com critérios e padrões específicos, e exigem a realização de tarefas (muitas vezes predeterminadas) que ajudem a ilustrá-los. Já nos portfólios mais abertos (ou menos estruturados), a organização é deixada a cargo dos próprios professores, a partir de algumas questões norteadoras. Para Delandshere e Petrosky (2010), a estruturação dos portfólios é “uma faca de dois gumes”, pois, por um lado, os portfólios muito estruturados podem fazer com que os professores se sintam presos a um único modelo, inibindo sua criatividade, por exemplo, para apresentar outros aspectos do seu trabalho que julguem merecer destaque; por outro lado, os portfólios muito abertos podem gerar confusões, e os professores podem ficar perdidos, sem saber exatamente o que devem fazer. Os autores ressaltam que portfólios mais estruturados são mais adequados para avaliações com fortes consequências, enquanto os portfólios menos estruturados são mais adequados para avaliações formativas.

De maneira sintética, Wolf e Dietz (1998, p. 13) apresentam os portfólios de professores como uma “coleção estruturada do trabalho do professor e de seus alunos criada em diversos contextos ao longo do tempo, emoldurada pela reflexão e enriquecida pela colaboração, que tem como seu principal objetivo o avanço do professor e a aprendizagem do aluno”. Os autores enfatizam ainda que os portfólios deveriam ser entendidos como um processo a ser realizado na companhia de mentores e colegas de trabalho.

Por fim, cabe a ressalva apresentada por Wolf (1996, s.p.): “é importante lembrar que o objetivo não é a criação de portfólios excelentes, mas sim a cultivação de um ensino e aprendizagem excelentes”.

PROVAS

A ideia de que os professores precisam comprovar que possuem conhecimentos e habilidades necessários para a realização do trabalho docente não é nova. As provas são adotadas, principalmente, em processos de seleção e contratação, mas também em alguns sistemas de avaliação de professores em serviço.

O modelo mais utilizado consiste em provas de múltipla escolha que verificam o conhecimento da matéria de ensino, de noções gerais de pedagogia e da organização e do funcionamento dos sistemas de ensino. Como argumentos a favor desse modelo, Madaus (1990, p. 258) ressalta:

O conhecimento é um aspecto particularmente importante da competência dos professores [...] Existe uma grande quantidade de estudos que demonstram que os itens de múltipla escolha podem mensurar o conhecimento. Testes convencionais podem contemplar um amplo domínio, ajudando a garantir que nenhum dos conhecimentos necessários deixe de ser testado. Testes convencionais são eficientes para administrar e pontuar.

Entretanto, o autor reconhece que esse modelo apresenta limitações, entre elas: algumas pessoas apresentam dificuldades para demonstrar seus conhecimentos em provas de múltipla escolha; assinalar a alternativa correta não necessariamente significa que a pessoa consiga elaborar uma resposta para o problema colocado; o formato desses testes produz resultados que não são adequados para avaliar a *performance* dos professores. Sobre esse último ponto, Madaus (1990) enfatiza o tipo de inferência que pode ser feita com base nos resultados desses testes. Ele atenta para o fato de que um bom desempenho na prova não necessariamente garante uma boa prática docente, mas que um mau desempenho muito provavelmente levará a uma prática insatisfatória.

O argumento é que a prova consegue avaliar o conhecimento dos professores e não sua *performance* em sala de aula, todavia, sem o conhecimento necessário; essa *performance* será comprometida.

Wolf e Dietz (1998) também destacam o fato de os testes de múltipla escolha apresentarem pouca ou nenhuma semelhança com o trabalho realizado pelos professores em sala de aula. A fim de amenizar esse problema, em alguns casos são incluídas nas provas para os professores questões mais complexas, que requerem a realização de tarefas mais próximas ao exigido cotidianamente dos professores, como a correção e a discussão de um exercício realizadas por alunos, o planejamento de uma atividade para o ensino de determinado conteúdo etc. Nesse caso, as respostas não podem se limitar à escolha de uma alternativa, o que tem implicações para a sua correção tanto no que diz respeito à padronização desta quanto aos seus custos.

Outro problema apontado é que os professores podem adequar as suas respostas ao que consideram como “correto” ou “esperado”, mesmo que isso não corresponda ao que acreditam e/ou realizam. Mais do que isso, ainda que as provas mostrem o que os professores são capazes de fazer, elas não revelam o que os professores tipicamente fazem (HAERTEL, 1990).

Desse modo, a literatura aponta a importância dos testes para a verificação dos conhecimentos dos professores, indicando suas limitações, principalmente no que se refere à análise do desempenho deles em sala de aula.

DESEMPENHO DOS ALUNOS

Um dos temas que provoca maiores debates no campo da avaliação de professores refere-se à pertinência da utilização dos resultados dos alunos em sistemas de avaliação em larga escala como indicadores da qualidade dos professores. Em geral, existe um consenso de que a aprendizagem dos alunos é uma das tarefas, senão a principal dos professores. Entretanto, existem inúmeras divergências quanto à capacidade das avaliações em larga escala de mensurar ou não o quanto os professores contribuíram com a aprendizagem dos alunos.

Entre as críticas à utilização dos resultados das avaliações em larga escala para a avaliação de professores destacam-se: o entendimento de que esse instrumento não foi elaborado para avaliar o desempenho dos professores, mas sim dos seus alunos (GOE, 2007); a limitação dos conteúdos contemplados pelas avaliações em larga escala, que as impede de analisar toda a aprendizagem que acontece na sala de aula (DANIELSON; MCGREAL, 2000; INGERSOL, 2003; STRONG, 2011); a existência de outros fatores, além da qualidade docente, que influenciam o desempenho dos alunos, sendo difícil isolar a contribuição de cada professor para os resultados dessas avaliações (GOE, 2007; DANIELSON; MCGREAL, 2000).

Para amenizar esses problemas, são realizados controles estatísticos das variáveis que influenciam o desempenho dos estudantes, em especial do seu nível socioeconômico. Contudo, persiste o questionamento de se esse tipo de controle é suficiente para isolar a contribuição de cada professor, especialmente quando não são utilizados modelos de valor agregado – que realizam duas testagens da proficiência dos alunos, permitindo a verificação dos ganhos de aprendizagem entre as avaliações.⁴

Há autores que questionam até mesmo a utilização dos modelos de valor agregado, por acreditarem que os professores não são os únicos responsáveis pela variação do desempenho dos estudantes. Por exemplo, Danielson e McGreal (2000, p. 43) afirmam que “as vantagens que alguns estudantes desfrutam em seus ambientes familiares afetam não apenas os seus níveis absolutos de desempenho, mas também a taxa de sua aprendizagem”. Goe (2007, p. 15) complementa que “muitas variáveis entram na composição de uma escola ou sala de aula e é difícil imaginar que os professores são os únicos responsáveis pelos resultados nos testes dos alunos após realizar o controle estatístico do desempenho pretérito dos estudantes”.

As críticas aos modelos de valor agregado referem-se, ainda, a problemas na estabilidade dos resultados – significativo movimento dos professores nos quintis de distribuição da qualidade docente (AARONSON; BARROW; SANDER, 2007; GOE, 2007; STACY et al., 2013); ao fato de não considerarem

⁴ Nesse modelo as testagens iniciais são consideradas como o controle das outras variáveis, além do trabalho docente, que afetam o resultado dos estudantes (SANDERS; RIVERS, 1996).

as diferenças nos ritmos de aprendizagem dos alunos, em especial após longos períodos de convivência com as famílias, como férias e feriados (DESIMONE; LONG, 2010; DANIELSON; MCGREAL, 2000), e ao fato de não levarem em conta que a aprendizagem não acontece de maneira crescente e contínua. Sobre esse último ponto, Strong (2011, p. 50) afirma:

Outro fator que pode ter consequências para o valor agregado são as pesquisas que mostram que a aprendizagem dos alunos nem sempre acontece em uma linha reta [...] crianças entre 7 e 10 anos de fato pioram antes de melhorarem, mostrando uma clara curva de aprendizagem no formato de U. Esse fenômeno pode ter implicações para o cálculo da pontuação de um professor com a utilização do valor agregado, dependendo de onde está a sua turma na curva de aprendizagem.

O posicionamento contrário à utilização dos resultados dos alunos em avaliações em larga escala para a avaliação de professores é sintetizado por Danielson e McGreal (2000, p. 41) da seguinte maneira:

Os modelos de avaliação de professores que incorporam medidas da aprendizagem dos alunos exigem técnicas válidas para a aferição dessa aprendizagem. Isto é, se nós assumimos que a aprendizagem dos alunos é um indicador da qualidade docente, então esperamos que medidas válidas da aprendizagem estejam disponíveis. Infelizmente, medidas válidas estão longe de estarem asseguradas.

Os que defendem a utilização dos resultados dos alunos para a avaliação de professores reconhecem a existência de condições que precisam ser atendidas para esses resultados representarem medidas válidas da contribuição dos professores para o desempenho dos alunos. Eles ressaltam a importância de que os testes sejam administrados por avaliadores, e não por professores; que seja feito o controle do desempenho prévio dos alunos (valor agregado); que se considere que não existem testes para mensurar a contribuição de alguns professores (por exemplo, de artes e educação física); entre outras questões (PETERSON, 2004; STRONGE; TUCKER, 2003).

Ao contrário do posicionamento anteriormente apresentado, eles acreditam que é possível atender a essas condições e que alguns sistemas de avaliação em larga escala já o fazem, possibilitando que essas medidas sejam incorporadas na avaliação de professores. Além disso, enfatizam as limitações dos outros instrumentos de avaliação de professores (subjetividade, risco da ênfase na popularidade dos professores, pouca relação com o trabalho desenvolvido em sala de aula etc.), indicando que, assim como os demais instrumentos, a incorporação dos resultados dos alunos na avaliação de professores apresenta problemas potenciais que precisam ser considerados, mas que não impedem a sua utilização, visto que esta representa mais uma fonte de informação sobre a qualidade docente.

CONSIDERAÇÕES FINAIS

Conforme o exposto, fica claro que existem várias alternativas para a construção de modelos de avaliação de professores, tanto no que se refere aos seus objetivos, quanto aos critérios e instrumentos utilizados.

As escolhas realizadas ao longo do processo estão sujeitas a disputas entre concepções diferentes. Tais disputas também se fazem presentes na literatura, com autores argumentando a favor de determinados formatos para a avaliação de professores e outros apresentando fortes críticas a essas formulações.

Em um cenário de tantas possibilidades e disputas, constatou-se que existe consenso apenas em relação à necessidade de envolvimento de diversos atores – principalmente os professores – no processo de tomada de decisões e à superioridade dos modelos que combinam diferentes instrumentos avaliativos.

Seria interessante que a elaboração de programas brasileiros de avaliação de professores considerasse as alternativas disponíveis na literatura e, particularmente, os consensos na área. Infelizmente, não parece que esses preceitos estejam sendo observados até o presente momento.

REFERÊNCIAS

- AARONSON, D.; BARROW, L.; SANDER, W. Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics*, v. 1, n. 25, p. 95-135, 2007.
- BACHARACH, S. B.; CONLEY, S. C.; SHEDD, J. B. Evaluation teachers for career awards and merit pay. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin Press, 1990.
- BARBER, L. W. Self-Assessment. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.
- BAUER, A. Avaliação do desempenho de professores: pressupostos e abordagens. In: GATIL, B. (Org.). *O trabalho docente: avaliação, valorização, controvérsias*. Campinas: Autores Associados; São Paulo: Fundação Carlos Chagas, 2013.
- BIRD, T. The Schoolteacher's portfolio: an essay on possibilities. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais. *Referenciais para o Exame Nacional de Ingresso na Carreira Docente*: Documento para Consulta Pública. Brasília: MEC, s/d.
- CAMPBELL, D. Assessing the impact of planned social change. *Evaluation and Program Planning*, v. 2, n. 1, p. 67-90, 1979.
- CARVALHO, E. M. R. *Autoavaliação e desenvolvimento profissional docente: estudo exploratório*. 2011. 183f. Dissertação (Mestrado em Ciências da Educação) – Instituto de Educação, Universidade do Minho, Braga. 2011.
- CHILE. Ministerio de Educación. Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas. *Marco para la buena enseñanza*. 2008. Disponível em: <<http://www.docentemas.cl/docs/MBE2008.pdf>>. Acesso em: jan. 2014.
- COUNCIL OF CHIEF STATE SCHOOL OFFICERS. *InTASC Model Core Teaching Standards and Learning Progressions for Teachers 1.0: A Resource for Ongoing Teacher Development*. Washington, DC: Council of Chief State School Officers, 2013.
- DANIELSON, C. *Enhancing professional practice: a framework for teaching*. Alexandria: Association for Supervision and Curriculum Development, 1996.
- DANIELSON, C.; MCGREAL, T. L. *Teacher evaluation to enhance professional practice*. Alexandria: Association for Supervision and Curriculum Development, 2000.
- DARLING-HAMMOND, L. Teacher evaluation in transition: emerging roles and evolving methods. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

DELANDSHERE, G.; PETROSKY, A. R. The use of portfolios in pre-service teacher education. In: KENNEDY, M. (Ed.). *Teacher assessment and the quest for teacher quality: a handbook*. San Francisco, CA: Jossey-Bass/Wiley, 2010.

DESIMONE, L. M.; LONG, D. Teacher effects and the achievement gap: do teacher and teaching quality influence the achievement gap between black and white and high- and low-SES students in the early grades?. *Teachers College Record*, v. 112, n. 12, Dec. 2010.

DUKE, D. L.; STIGGINS, R. J. Beyond minimum competence: evaluation for professional development. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

FERNANDES, D. *Avaliação do desempenho docente: desafios, problemas e oportunidades*. Córdoba: Texto, 2008.

GARDNER, J. et al. *Developing Teacher Assessment*. Berkshire: Open University Press, 2010.

GOE, L. *The link between teacher quality and student outcomes: a research synthesis*. Washington: National Comprehensive Center for Teacher Quality, 2007.

GOE, L.; BELL, C.; LITTLE, O. *Approaches to evaluation teacher effectiveness: a research synthesis*. Washington: National Comprehensive Center for Teacher Quality, 2008.

HAERTEL, E. H. Performance tests, simulations and other methods. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

INGERSOLL, R. M. *Who controls teachers' work? Power and accountability in America's schools*. Cambridge: Harvard University Press, 2003.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. *The personnel evaluation standards: how to assess systems for evaluation educators*. 2. ed. Thousand Oaks: Corwin Press, 2009.

MADAUS, G. Conventional tests for licensure. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

MASTER, B. *What can parents tell us about teacher quality? Examining the contributions of parent perspectives in comparison to a portfolio of alternative teacher evaluation measures*. 2013. (Dissertation paper) – Stanford University, Stanford. 2013.

MCCARTHEY, S. J.; PETERSON, K. D. Peer review of materials in public school teacher evaluation. *Journal of Personnel Evaluation in Education*, v. 1, n. 3, p. 259-267, 1988.

MEASURES OF EFFECTIVE TEACHING PROJECT. *Gathering feedback for teaching: combining high quality observations with student surveys and achievement gains*. Seattle: Bill and Melinda Gates Foundation, 2012.

NATIONAL BOARD FOR PROFESSIONAL TEACHING STANDARDS – NBPTS. *What teachers should know and be able to do*. Arlington: NBPTS, 2002.

PETERSON, K. D. Assistance and assessment for beginning teachers. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

_____. Research on school teacher evaluation. *NASSP Bulletin*, v. 88, n. 639, p. 60-79, 2004.

PETERSON, K. D.; WAHLQUIST, C; BONE, K. Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education*, v. 14, n. 2, p. 135-153, 2000.

PETERSON, K. D. et al. Parent surveys for teacher evaluation. *Journal of Personnel Evaluation in Education*, v. 17, n. 4, 2003.

POPHAM, W. J. The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education*, v. 1, n. 3, p. 269-273, 1988.

SANDERS, W. L.; RIVERS, J. C. Cumulative and residual effects of teachers on future student academic achievement. *Research Progress Report*. Knoxville: University of Tennessee Value-Added Research and Assessment Center, 1996.

SCRIVEN, M. Duties of the teacher. *Journal of Personnel Evaluation in Education*, v. 8, n. 2, p. 151-184, 1994.

SHINKFIELD, A. J.; STUFFLEBEAM, D. L. *Teacher evaluation: guide to effective practice*. Norwell: Kluwe Academic, 1995.

STACY, B. et al. Does the Precision and Stability of Value-Added Estimates of Teacher Performance Depend on the Types of Students They Serve? In: AEFPP ANNUAL CONFERENCE, 38., 2013, New Orleans. *Anais...* New Orleans, Louisiana: AEFPP, 2013.

STODOLSKY, S. S. Classroom Observation. In: MILMAN, J.; DARLING-HAMMOND, L. (Org.). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park: Corwin, 1990.

STRONG, M. *The highly qualified teacher: what is teacher quality and how do we measure it?*. New York: Teachers College Press, 2011.

STRONGE, J. H.; TUCKER, P. D. *Handbook on teacher evaluation. Assessing and improving performance*. Larchmont: Eye on Education, 2003.

WILKERSON, D. J. et al. Validation of student, principal, and self-ratings in 360° feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, v. 14, n. 2, p. 179-192, 2000.

WOLF, K. Developing an effective teaching portfolio. *Educational Leadership*, v. 53, n. 6, p. 34-37, Mar. 1996.

WOLF, K.; DIETZ, M. Teaching portfolios: purposes and possibilities. *Teacher Education Quarterly*, v. 25, n.1, p. 9-22, 1998.

NATHALIA CASSETTARI

Doutoranda em Educação na Faculdade de Educação da
Universidade de São Paulo (FE/USP)

nathalia.cassettari@gmail.com

Recebido em: MARÇO 2014

Aprovado para publicação em: ABRIL 2014