

EIXO 1 - Políticas públicas, financiamento, avaliação e gestão da educação.

POR QUE SECRETARIAS DE EDUCAÇÃO INVESTEM EM SISTEMAS PRÓPRIOS DE AVALIAÇÃO EM LARGA ESCALA?

Vandré Gomes da Silva
FCC e UNISANTOS
vgomes@fcc.org.br

Nelson A. S. Gimenes
FCC
ngimenes@fcc.org.br

Lisandra Marisa Príncipe
Faculdade Sumaré
lisandraprincipe@uol.com.br

Paula Louzano
FEUSP
paula.louzano@gmail.com

Resumo: O estudo analisa os objetivos e justificativas de quatro Secretarias de Educação ao investir em sistemas de avaliação em larga escala próprios ao invés de utilizar as avaliações nacionais já disponíveis, em especial a Prova Brasil, como instrumento precípua de sua política educacional. O foco recai sobre avaliações externas que lançam mão dos mesmos modelos adotados pelas avaliações nacionais em meio a contornos e desenhos distintos conforme o tipo de organização e política impressos às diferentes redes de ensino investigadas. Destacam-se as demandas que se impõem a essas avaliações, notadamente naquilo que podem conter de subsídio pedagógico voltado diretamente ao trabalho escolar.

Palavras-chave: avaliação de larga escala; avaliação externa; política educacional.

É notório o aumento e o peso atribuídos aos sistemas de avaliação de desempenho discente no Brasil, protagonizados tanto pelo governo federal, como por iniciativas de governos estaduais e municipais nos diferentes níveis de ensino da Educação Básica a partir do início dos anos 1990. Contudo, a utilização de seus resultados, em seu sentido mais pedagógico, por parte de gestores, coordenadores e professores é, ainda, assunto pouco explorado.

Em meio ao esforço de se investigar e buscar descrever eventuais tipos de apropriação de resultados desencadeados por sistemas de avaliação externa, a primeira impressão é a de que qualificar tais usos como *instrumentos* ou *políticas de gestão* no âmbito do processo decisório das Secretarias de Educação (Souza e Oliveira 2010; Bonamino e Souza, 2012; Brooke e Cunha, 2010) parece não abranger, em profundidade, outras formas de utilização dessas avaliações e seus resultados, sobretudo quando praticadas no cotidiano escolar.

De qualquer forma, alguns usos dessas avaliações se destacam no cenário educacional atual, ainda que seus eventuais efeitos sejam pouco conclusivos, para além do caráter polêmico de algumas ações gerenciais que lançam mão de seus resultados, em especial a bonificação de professores tendo como um de seus critérios o desempenho discente obtido. Abre-se, assim, outra perspectiva de análise e intervenção ao investigarmos a pertinência dessas avaliações em larga escala na prática cotidiana das escolas em seu sentido mais pedagógico.

Dessa forma, esse trabalho pretende analisar os objetivos e justificativas de sistemas de públicos de educação ao investir em sistemas de avaliação externa próprios. Reconhecendo a profusão de diversos tipos de avaliações externas protagonizadas por diversas secretarias municipais de educação (Souza, Pimenta e Machado, 2012), o presente estudo¹ volta-se especificamente para a análise de sistemas de avaliação externa que lançam mão dos mesmos modelos adotados pelas avaliações nacionais disponíveis, em especial a Prova Brasil, em meio a contornos e desenhos distintos conforme a organização e gestão de diferentes secretarias de educação públicas. Busca assim, analisar de forma mais detida quais seriam os objetivos e demandas que se impõe a essas avaliações, em especial naquilo que podem conter de subsídio pedagógico voltado diretamente ao trabalho escolar.

O escopo geral dessa pesquisa investigou os usos das avaliações externas no âmbito de quatro sistemas de ensino público do país: a rede estadual do Espírito Santo e as redes municipais de São Paulo (SP), Sorocaba (SP) e Castro (PR)². A escolha de diferentes perfis de Secretarias de Educação - estaduais e municipais de grande, médio e pequeno porte -

¹ Os dados parciais aqui apresentados são fruto de pesquisa maior realizada pela Fundação Carlos Chagas em parceria com a Fundação Itaú Social em torno dos usos que se tem feito dessas avaliações em larga escala, bem como dos próprios modelos de avaliação adotados em relação ao seu potencial pedagógico como subsídio direto ao trabalho realizado em âmbito escolar. A coleta de dados foi realizada em entre março de 2011 e maio de 2012.

² A orientação da escolha dessas Secretarias de Educação se deu em função do desenho de suas políticas de avaliação e não a partir de seus resultados em seus sistemas próprios ou mesmo nas avaliações nacionais. Como se sabe, resultados considerados bons ou a sua melhora gradativa ao longo de várias ondas de avaliação podem ser influenciados por diversos fatores extraescolares, não se constituindo necessariamente em um efeito direto ou principal, por exemplo, de uma política orientada pelos usos dessas avaliações.

proporcionou uma variabilidade interessante de modelos e arranjos de políticas educacionais vinculadas à avaliação externa.

Como opção metodológica no âmbito de uma pesquisa de caráter exploratório (Babie, 1999; Deslauriers e Kerisit, 2008) foram selecionadas três Secretarias de Educação que possuem sistemas próprios de avaliação em larga escala – o que de alguma forma denota uma preocupação institucional em relevar e avaliar o desempenho cognitivo de seus alunos. Como contraponto, foi investigada também uma Secretaria de Educação que não possui um exame próprio em larga escala nesses moldes (Sorocaba-SP), porém conta com uma política de avaliação externa a partir dos resultados disponíveis através de outras avaliações em larga escala como o SARESP – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo – sob a responsabilidade da respectiva Secretaria Estadual de Educação e, também, da própria Prova Brasil. Em distintos níveis de investigação (Gestão central, instâncias intermediárias e escolas) além da coleta de documentos, diversos atores foram entrevistados individual e coletivamente, como secretários de educação, técnicos da secretaria, coordenadores regionais, supervisores de ensino, gestores escolares, coordenadores pedagógicos e professores.

Em meio às limitações metodológicas impostas por pesquisas exploratórias como essa, não se pretende generalizar os resultados apresentados a partir dos objetivos específicos desse trabalho, mas indicar possíveis tendências em relação à esse tipo de política educacional, assim como destacar o significado e sentido específicos das experiências regionais aqui abordadas.

Nas diferentes redes de ensino pesquisadas foi possível identificar diversas justificativas para a elaboração dos sistemas próprios de avaliação externa, agrupadas em quatro grandes grupos: 1) o caráter censitário e anual das avaliações externas promovidas pelas redes de ensino estudadas; 2) a possibilidade de detalhar e trabalhar os dados nas escolas com maior rapidez e agilidade; 3) a amplitude dos anos/séries avaliadas pelo modelo de avaliação adotado pelas redes de ensino e 4) a possibilidade de identificação nominal dos resultados de cada aluno.

Tais justificativas podem ser mais bem compreendidas a partir do sentido que tomou as políticas de avaliação em larga escala nas duas últimas décadas no Brasil. A introdução maciça desse tipo de avaliação nos países, em especial a partir dos anos 1990, acompanhou uma agenda e tendência internacionais já consolidadas que enxergam nesse tipo de ação uma forma de se aferir a qualidade e a efetividade dos sistemas educacionais.

Ainda que o termo qualidade no campo educacional possua diversos significados e comporte diferentes dimensões (Dourado, 2007; Silva, 2008), o fato é que tais avaliações rapidamente se associaram à ideia de qualidade, limitando-a aos resultados obtidos pelos alunos. Assim, não é estranho o argumento de que essas avaliações seriam a própria tradução de uma política educacional (Oliveira, 2000; Souza, 2003) cujo alcance ainda carece de avaliação mais ampla e diversificada quanto ao seu potencial indutor de transformação e maior qualificação do trabalho escolar e da própria prática docente.

De qualquer forma, no caso brasileiro, as avaliações em larga escala apresentam em comum a pretensão de servir de diagnóstico educacional a partir do monitoramento dos resultados de aprendizagem dos alunos, dentre outros objetivos e contornos. Essa parece ser a vocação do Sistema de Avaliação da Educação Básica – SAEB -, iniciativa pioneira de avaliação realizada em âmbito nacional pelo Ministério da Educação implementada no início dos anos 1990.

Segundo os gestores das redes de ensino pesquisadas, o caráter censitário das avaliações por eles adotadas configura-se como uma justificativa relevante quanto se procura compreender os motivos pelos quais tais sistemas optam pela organização e elaboração de um sistema próprio de avaliação externa. Acompanhada de tal justificativa, podemos também identificar uma crítica comum dos gestores em relação à demora na divulgação dos resultados da Prova Brasil. Isto é, o grande intervalo de tempo entre a aplicação dos instrumentos dessa avaliação nacional e o acesso aos resultados pelas redes de ensino públicas em foco.

Já o formato original do SAEB – ainda em voga - não favorece objetivos mais abrangentes como, por exemplo, a utilização imediata de seus resultados por unidades escolares e professores. Os dados que o SAEB disponibiliza podem fornecer um retrato abrangente de sistemas de ensino e regiões brasileiras em relação aquilo que determinado conjunto dos alunos demonstra saber, nos termos do exame. Contudo, seu o desenho amostral impede que sejam disponibilizados os resultados por escola.

Avaliações como o SAEB podem ser compreendidas como de *primeira geração* conforme a classificação estabelecida pelas Souza & Bonamino (2012). Sua característica eminente seria a de fornecer um diagnóstico da qualidade da educação ofertada no Brasil, sem acarretar consequências diretas a escolas, seus profissionais ou mesmo em termos do currículo efetivamente ensinado aos alunos.

Ainda que os impactos do SAEB sejam pouco significativos no contexto escolar em função de seu desenho amostral, é importante destacar que essa avaliação e seu desenho

metodológico atual - introduzido a partir de 1995³ - continuam a exercer forte influência junto aos diversos sistemas estaduais e municipais criados ainda nos anos 1990 e, cada vez mais, presentes em diversas redes estaduais e municipais de ensino, em especial na última década. É nesse período, por exemplo, que surgem os sistemas de avaliação de Castro (2007), Espírito Santo (2001, 2004 e continuamente a partir de 2007) e São Paulo (2007), investigados nesse estudo.

De forma concomitante ao SAEB diversos sistemas de avaliações presentes no cenário educacional atual apresentam, além dessa característica diagnóstica inerente, a pretensão de, ao produzir resultados por escolas, propiciar algum tipo de consequência nesse âmbito e/ou a seus profissionais. Nesse sentido, as avaliações de *segunda geração* seriam vinculadas a consequências de caráter mais simbólico em função da divulgação e apropriação dos resultados pelos pais de alunos e demais segmentos da sociedade de modo geral e, evidentemente, pelos próprios agentes escolares (Sousa e Bonamino, 2012). Normalmente, nos estudos sobre o tema esse tipo de política de avaliação é chamada *low stakes* ou de *responsabilização branda* (Brooke, 2008). Segundo essa classificação a Prova Brasil – feita em âmbito federal e uma espécie de extensão do modelo SAEB, gerando resultados por escolas e secretarias de educação - e os sistemas de avaliação implementados em Castro-PR e São Paulo-SP seriam exemplos de avaliação de *segunda geração*.

Já as avaliações de *terceira geração* estariam relacionadas às chamadas políticas de *responsabilização forte* ou *high stakes* que envolveriam, além da divulgação dos resultados aferidos, também sanções ou recompensas estabelecidas a partir dos desempenhos obtidos por alunos e escolas, em grau e diversidade variados. Uma ação de gestão nesses moldes implicaria algum tipo de consequência material a escolas e professores (Brooke, 2008). No Brasil, o exemplo mais contundente desse tipo de política seriam os processos de bonificação ou premiação de escolas e professores mediante os resultados obtidos, como se verificou no caso da Rede Estadual de Ensino do Espírito Santo.

Evidentemente, determinados tipos de usos das avaliações podem gerar efeitos diversos a depender dos contornos e pressupostos das ações implementadas, o histórico e características do sistema de ensino em que são realizados, o sentido que adquirem tais ações

³Sem dúvida, o SAEB tem servido de parâmetro para a formulação de diversos sistemas de avaliação em larga escala, em especial em relação à apropriação de suas matrizes de avaliação e a utilização da metodologia denominada Teoria da Resposta ao Item – TRI.

ao longo do tempo e, em última instância, pelos tipos de resultados obtidos pelos alunos fortemente influenciados por suas condições sociais e econômicas de origem (Brooke & Soares, 2008).

De todo modo, independente das motivações que as políticas de avaliação em larga escala tenham, diferentes objetivos e expectativas atreladas às avaliações externas podem se apresentar como dois lados de uma mesma moeda, ou seja, melhorar progressivamente os resultados implica – de alguma forma - em tentar modificar e aperfeiçoar práticas de ensino e modos de organização escolares.

A tendência em aproximar a política de avaliação externa à apropriação de seus resultados por parte de escolas e professores parece se destacar em relação a outros objetivos pelos quais as Secretarias de Educação investigadas possam se orientar. Embora a efetividade ou efeitos de uma política educacional não possam ser avaliados apenas por seus objetivos declarados, é importante verificarmos o tipo de discurso empregado em favor da melhora ou aperfeiçoamento do trabalho escolar, e que parece ir além do desejável aumento dos resultados.

Na secretaria municipal de Educação de Castro-PR, é bastante presente nos documentos e falas de seus gestores uma concepção de avaliação que procura se distanciar do *ranqueamento* de escolas a partir de resultados em avaliações externas, ao mesmo tempo em que enfatiza um tipo de apropriação eminentemente pedagógica dos resultados. Ao expor os objetivos do IDEC (Índice de desenvolvimento educacional de Castro), o então secretário de educação procura ir além da realização de um diagnóstico educacional ao afirmar que é preciso

(...) oferecer para o professor condições de atuar sob aqueles dados que chegaram a nós (...) oferecer informações adequadas e reais para que o professor possa aprimorar seu planejamento e, muitas vezes, mudar sua prática em sala de aula.

Nesse mesmo sentido, o Programa de Avaliação Básica do Espírito Santo – PAEBES – embora apresente certa diversidade de objetivos, enxerga o trabalho escolar e a ação docente como seu alvo preferencial. Os objetivos do PAEBES podem ser divididos em dois grandes grupos, estreitamente ligados: o primeiro deles diz respeito à utilização da avaliação externa na elaboração e desenvolvimento de ações no nível da macro-política, ações de caráter mais geral e sistêmico sob a responsabilidade da Secretaria de Educação; o segundo grupo refere-se ao desdobramento dessa avaliação externa em ações e procedimentos a serem adotados em

cada uma das unidades escolares que compõem o sistema⁴. De qualquer forma, ambos os grupos são acompanhados de uma forte expectativa de apropriação e utilização dos resultados da avaliação pelas unidades escolares na fala de seus gestores.

De forma semelhante ao Espírito Santo, na secretaria municipal de educação de São Paulo a perspectiva de articulação entre uma ação mais sistêmica e o uso dos resultados pelas escolas também é explícita, destacando-se essa segunda dimensão como uma espécie de alvo principal de sua política de avaliação. Assim, o sistema de Avaliação de Aproveitamento Escolar dos Alunos da Rede Municipal de Ensino de São Paulo, apresenta dois grandes objetivos⁵:

- I - subsidiar a Secretaria Municipal de Educação nas tomadas de decisão quanto à Política Educacional do Município;
- II - fornecer ao sistema de ensino, às equipes técnico-pedagógicas das Coordenadorias de Educação e às Unidades Educacionais informações que subsidiem seu trabalho, mais especificamente no que diz respeito à (a) formação continuada dos professores, (b) reorientação da proposta pedagógica, (c) articulação dos resultados da avaliação com o planejamento escolar, a formação dos professores e o estabelecimento de metas para o projeto pedagógico de cada escola; (d) orientação para os trabalhos desenvolvidos nas Salas de Apoio Pedagógico -SAPs das unidades escolares com os alunos que necessitam de reforço na aprendizagem.

A articulação entre objetivos de caráter mais gerencial e do uso das avaliações em sentido mais pedagógico também é verificada em Sorocaba, cuja política de avaliação se pauta não em um modelo interno de prova, mas em uma política de avaliação que lança mão dos dados disponibilizados por outras instancias, como as avaliações nacionais (Prova Brasil e IDEB) e de forma especial o SARESP – Sistema de Avaliação do rendimento escolar do Estado de São Paulo. De maneira geral, seus gestores pontuam a necessidade de se garantir a equidade de aprendizagem bem como sua constante melhoria a partir da análise e apropriação dos dados da avaliação, tanto em relação às decisões de caráter mais geral efetuadas pela gestão central desse sistema, mas, fundamentalmente, na apropriação dos resultados por escolas e professores.

A coincidência de diversos objetivos e a preocupação comum com a apropriação efetiva dos resultados por agentes escolares ajudam a explicar o fato de os sistemas de ensino pesquisados buscarem desenhos de avaliação próprios ou, no caso de Sorocaba, da utilização

⁴ A esse respeito ver a *Revista do Gestor* e *Revista do Educador* - constituindo relatórios dos resultados do PAEBES referente aos anos 2008, 2009 e 2010 e o *Roteiro Básico para Discussão e Apropriação dos Resultados* e o *Direito ao Aprendizado: Guia de Orientações para a Intervenção Pedagógica* - Ensino Fundamental – Ano II.

⁵ Conforme Lei Municipal Nº 14.063, DE 14 DE OUTUBRO DE 2005, que institui o Sistema de Avaliação de Aproveitamento Escolar dos Alunos da Rede Municipal de Ensino de São Paulo.

do modelo Estadual de Avaliação (SARESP), ao invés de uma apropriação exclusiva dos modelos nacionais existentes.

Um denominador comum em relação às razões citadas por diversos gestores dos sistemas pesquisados seria a necessidade de uma avaliação externa ter, além de um caráter censitário, ser realizada de forma constante, pelo menos uma vez ao ano. Ou seja, para os gestores entrevistados, o uso dos resultados é fundamental, desde que eles tragam informações “efetivas”, isto é, informações periódicas e frequentes sobre toda a rede e que, de alguma forma, possam subsidiar as unidades escolares em seu trabalho cotidiano.

Uma perspectiva de avaliação nesses moldes destoa do modelo da Prova Brasil, cuja periodicidade se dá de dois em dois anos e cujos resultados, sob o controle do INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira -, não são disponibilizados em tempo considerado hábil e a contento por parte dos responsáveis pela gestão central de todas as secretárias investigadas.

De forma geral, os tipos de resultados disponibilizados pela Prova Brasil nitidamente não suprem as necessidades identificadas por esses sistemas no que deveriam ser seus objetivos. A percepção de um gestor do município de Castro é bastante ilustrativa em torno da questão:

Nós nos sentíamos bastante incomodados com os instrumentos que nós tínhamos, principalmente a Prova Brasil, porque a Prova Brasil não chega num nível de detalhe suficiente para que a gente possa tomar algumas decisões lá na sala de aula. O principal foco nosso é o aprendizado, tentar entender quais são os problemas que estão acontecendo respeitando todas as particularidades das escolas, mas tentar identificar onde estão os gargalos no sentido da aprendizagem, que a gente possa atacar. Precisamos desenvolver um instrumento que seja capaz de fornecer ao professor informações suficientes para ele saber onde estão os problemas e como ele vai poder solucionar esses problemas. Nós não tínhamos esse nível de detalhamento com as informações que vinham até nós com os outros instrumentos de avaliação que a gente conhecia: a Prova Brasil, a Provinha Brasil, Por exemplo, a Prova Brasil chega ao nível da proficiência da escola como um todo. Nós não conseguimos desdobrar esses dados para identificar quais as habilidades que estão faltando nos nossos alunos.

Essa fala sintetiza a maior parte das razões que orientariam a adoção de um sistema de avaliação próprio ou mesmo a adoção de uma política de avaliação pautada na avaliação externa elaborada pela Secretaria Estadual de São Paulo, como visto em Sorocaba. Em geral, todos os sistemas pesquisados apresentam discurso semelhante como uma das principais justificativas para o desenvolvimento de avaliações próprias ou políticas diferenciadas de avaliação, justamente pela perspectiva de uma apropriação detalhada desses resultados por escolas e professores e, também, por permitir aos gestores das secretarias de educação vislumbrar um panorama dinâmico e direto de suas redes e escolas.

Nos casos dos sistemas ou modelos de avaliação próprios chama à atenção a necessidade, citada por seus gestores, de obter resultados que, além de serem gerados por escola, possam ser discriminados por turma e, em alguns casos, até por alunos⁶, algo que não ocorre em relação à divulgação dos resultados da Prova Brasil. Essa possibilidade de controle das informações geradas pela avaliação externa e a relação entre o todo – sistema – e o indivíduo – o aluno – se configura como uma das razões que levariam as secretarias de educação a desenvolverem avaliações próprias.

Na medida em que as avaliações externas nacionais – em especial SAEB e Prova Brasil - se constituem na maior referência das políticas de avaliação dessas redes de ensino, assim como de diversas secretarias de educação estaduais e municipais, sua análise permite a compreensão mais ampla do movimento das Secretarias de educação ao propor alternativas que possam atender melhor suas expectativas e intenções.

De forma resumida, pode-se afirmar que os principais recursos oferecidos pelo Governo Federal para o uso dos resultados da Prova Brasil⁷, por parte dos gestores das Secretarias e equipes escolares, são os boletins das escolas, divulgados por meio do site do INEP para acesso irrestrito a toda a sociedade e também em forma de cartazes impressos enviados às escolas.

Os boletins de desempenho de cada escola que participou da edição de 2009⁸ da Prova Brasil, por exemplo, contêm as médias e a distribuição das notas em níveis de proficiência obtidas pelo conjunto de seus alunos de 5º e/ou 9º anos do ensino fundamental, nas provas de leitura e matemática. Além disso, os boletins contêm as metas da escola no IDEB nos anos de 2005, 2007 e 2009, bem como as notas do Índice de Desenvolvimento da Educação Básica que a escola de fato atingiu nesses anos, o que permite uma análise tanto do alcance das metas estipuladas como da evolução da rede nesse indicador, especificamente.

No caso da edição de 2009 da Prova Brasil, esses boletins foram divulgados quase dois anos após a aplicação, o que foi motivo de críticas por parte de diversos gestores das secretarias de educação entrevistados, como já anunciado anteriormente. Antes disso, mais

⁶Em relação aos sistemas pesquisados podemos citar como exemplos o PAEBES ALFA que avalia os primeiros anos do ensino fundamental e divulga as notas de cada aluno; a divulgação dos resultados do IDEC, feitos por aluno, turmas e escolas e, em São Paulo, os boletins individuais dos alunos entregues às famílias via correio.

⁷ Como se está tratando da apropriação dos resultados por parte das equipes escolares, será analisado mais especificamente o desenho da Prova Brasil, já que o SAEB abrange os estudantes de maneira amostral, apresentando resultados que chegam apenas ao nível das Unidades da Federação.

⁸ Trata-se aqui dos boletins em formato mais completo disponíveis até 2009 e disponíveis para acesso. A partir de 2011, em meio eletrônico, a divulgação dos dados se atém apenas às médias dos resultados pela escola em relação à media municipal (rural e urbana) e estadual (rural e urbana) em “língua portuguesa e Matemática”. A esse respeito ver <http://portal.inep.gov.br/web/prova-brasil-e-saeb/resultados>.

especificamente na metade do ano seguinte à aplicação, houve a divulgação apenas dos resultados preliminares – as médias obtidas pelo conjunto dos alunos de 5º e/ou 9º anos de cada escola. Esse acesso ficou restrito aos gestores educacionais no intuito de que as escolas e redes pudessem tomar conhecimento, analisar e solicitar eventuais correções ao INEP.

Pode-se constatar a partir da fala dos gestores, que a Prova Brasil tem um potencial limitado de utilização naquilo que poderia fornecer um *acesso mais imediato* dos resultados por parte das equipes escolares ao não favorecer um tipo de apropriação dos resultados capaz de servir como diagnóstico em relação aos próprios alunos testados nessa edição da Prova.

De modo geral, foi possível identificar de forma clara a busca de alternativas para utilização mais imediata e individualizada de avaliações em larga escala e que possam, ao mesmo tempo, dialogar e se pautar pelos mesmos critérios presentes na Prova Brasil/ SAEB, em especial pelo fato de esses sistemas adotarem o mesmo modelo metodológico de avaliação utilizado por essas avaliações nacionais, pautado pela *Teoria da Resposta ao Item*– TRI (Valle: 2000). Ainda assim, todas as redes investigadas buscaram avançar e se diferenciar, de alguma forma, em relação ao desenho da avaliação externa nacional, procurando favorecer uma maior aproximação dos diferentes atores institucionais com o tipo de prova aplicado e um uso mais imediato dos resultados por parte das equipes escolares.

Pode-se detectar diferenças significativas em relação aos diversos desenhos regionais de avaliação adotados - mesmo que a partir de um mesmo referencial metodológico inspirado nas avaliações nacionais. No caso da rede municipal de São Paulo, por exemplo, a Prova São Paulo avalia os alunos e/ou turmas compreendidos entre o 3º e 9º ano do ensino fundamental, de modo a oferecer, anualmente, diversos resultados por escola. Esta rede ainda proporciona um recurso adicional para esse tipo de *acesso mais imediato* às avaliações externas através da *Prova da Cidade*, elaborada de forma centralizada, mas aplicada e corrigida pelas próprias escolas e que, segundo a gestão central da secretaria, seria um instrumento pedagógico para que professores e gestores escolares possam realizar um diagnóstico mais detalhado da aprendizagem dos alunos em diferentes bimestres de um mesmo ano letivo. Nessa rede de ensino, além da questão do prazo e da oferta de resultados individualizados por aluno verificou-se a demanda, sobretudo por parte das equipes escolares, em conhecer melhor os instrumentos aplicados nessas avaliações o que poderia ter motivado, em parte, a proposição da Prova da Cidade associada à Prova São Paulo.

Na rede municipal de Castro o sistema próprio de avaliação externa conta com duas provas: uma no primeiro semestre e outra no segundo semestre, aplicadas anualmente para o 3º, 4º e 5º anos – especificamente, essa rede de ensino possui apenas os segmentos da pré-

escola e anos iniciais do ensino fundamental. Aqui também os resultados são divulgados por aluno às equipes escolares. A partir desse desenho, a divulgação dos resultados individuais dos alunos obtidos na prova do primeiro semestre também favorece o replanejamento do segundo semestre por parte das equipes escolares. Já os resultados da prova realizada no segundo semestre, por sua vez, podem ser utilizados pelos próximos professores de cada um desses alunos para o planejamento do ano letivo seguinte.

Já a rede municipal de Sorocaba passou a oferecer, como recurso adicional à sua política de avaliação, provas simuladas construídas a partir de itens divulgados da Prova Brasil e do SARESP, inspiradas em uma prática já recorrente em sua rede de ensino: a elaboração e aplicação de simulados de forma descentralizada pelas escolas. No caso do simulado da rede, a aplicação ficou por conta das escolas, mas as provas foram corrigidas no nível central – que tabulou os dados e devolveu os resultados para cada escola uma semana após a aplicação. Conforme os relatos obtidos nessa rede de ensino, na semana seguinte à essa prova os professores já estariam usando os dados referentes aos resultados para orientar e replanejar seu trabalho.

Na rede estadual do Espírito Santo, esse movimento distintivo foi percebido com relação ao PAEBES Alfa, por meio do qual foram aplicadas provas de leitura a todos os alunos dos 1º, 2º e 3º anos do ensino fundamental, com a devolução dos resultados por aluno às equipes escolares. Nos demais anos avaliados também anualmente – 5º e 9º anos do ensino fundamental e 1º e 3º do ensino médio – os resultados são divulgados somente no nível das escolas e do conjunto de alunos do ano avaliado, não chegando ao detalhe de resultados de turmas específicas e alunos individualmente, ainda que se verifique o desejo de uma avaliação externa nominal para todas as séries avaliadas por parte de diversos gestores da secretaria estadual.

Outro diferencial das políticas de avaliação externa observado nas redes pesquisadas diz respeito à forma como são apresentados os resultados das avaliações. O formato dos boletins de resultados obtidos pelas escolas na Prova Brasil, contendo a(s) nota(s) das turmas avaliadas, necessitam das escalas de desempenho das disciplinas avaliadas – divulgadas separadamente e com menos destaque no página eletrônica do Ministério da educação⁹. Nessa escala de proficiência os resultados obtidos podem ser interpretados pedagogicamente a partir de níveis crescentes de desempenho, contendo a descrição do que os alunos são capazes de

⁹ A esse respeito consultar <http://provabrasil.inep.gov.br/escalas-da-prova-brasil-e-saeb1>.

fazer a partir das notas divulgadas em uma escala que varia de 0 a 350 em língua portuguesa, e de 0 a 425 em Matemática.

Além dessas informações, fundamentais para apropriação e uso das avaliações externas, as secretarias de educação pesquisadas utilizam um sistema de classificação dos resultados, diferentemente da Prova Brasil. A partir de categorias arbitrárias como “Abaixo do Básico”, “Básico”, “Adequado” e “Avançado”, por exemplo, a divulgação dos resultados nesse tipo de classificação permite às equipes escolares um parâmetro claro do quanto os resultados obtidos se aproximam ou se distanciam das expectativas gerais de aprendizagem estipulados pela secretaria de educação em uma determinada série avaliada.

Verificou-se que o estabelecimento de um sistema de classificação dos desempenhos obtidos se converte em um recurso adicional que facilita a apropriação dos resultados pelas equipes escolares. Ao fornecer uma interpretação pautada em um juízo de valor, pôde-se observar que essa forma de divulgação chama a atenção, não só das equipes escolares, mas de toda a Secretaria de Educação para as escolas e turmas que estariam abaixo de um nível considerado mínimo, denominando na maior parte as redes como “Básico”. Assim, os alunos, turmas e/ou escolas que apresentam um desempenho abaixo desse nível mínimo tem, frequentemente, recebido tratamento prioritário, em maior ou menor grau nas quatro redes de ensino pesquisadas.

Para além dos tipos de resultado e interpretações que as secretarias de educação oferecem às suas redes, observava-se a preocupação e atenção crescentes em torno das estratégias de divulgação dos resultados e, também, em relação às formas de envolvimento das equipes escolares com as avaliações e seus resultados.

Todas as Secretarias pesquisadas produzem, ou adquirem junto às empresas terceirizadas que aplicam seus testes, algum tipo de material de divulgação explicativo dos resultados das avaliações para suas escolas. Em que pese as diferenças entre esses materiais verifica-se, além da divulgação de resultados por escola nos formatos já citados, algumas explicações acerca do sistema de avaliação adotado e análises pedagógicas de questões reunidas em um mesmo material. Em alguns casos são divulgadas análises do perfil do alunado e outros aspectos de interesse escolar com base nos dados coletados pelos questionários aplicados junto com as provas aos estudantes, professores e diretores. No caso da Secretaria de Educação de Castro – que configura uma exceção ao produzir exclusivamente uma avaliação em larga escala utilizando os parâmetros do SAEB/Prova Brasil por parte de sua própria equipe de técnicos e gestores – são divulgadas as questões utilizadas nas avaliações a professores e escolas.

No caso específico do município de Sorocaba, o fato de não se ter desenvolvido um sistema próprio de avaliação não desqualifica boa parte das críticas em relação aos limites de Prova Brasil – ainda que essa avaliação seja fortemente considerada e analisada no âmbito de sua política de avaliação naquilo que pode oferecer, de maneira correlacionada aos outros dados que a gestão dispõe sobre a aprendizagem dos alunos provenientes do SARESP. Nesse sentido, o tipo de apropriação desse conjunto de dados, embora provenientes de avaliações elaboradas fora do âmbito da secretaria, é reelaborado e interpretado de acordo com as demandas e intenções internas à gestão, incluindo análises específicas a partir das bases de dados de ambos os sistemas, gerando relatórios próprios para cada escola, diferentes daqueles divulgados pelos sistemas utilizados. Esse movimento, aliado ao fato de Sorocaba não ter que despende esforços e orçamento para elaborar uma prova própria, pode ser destacado como um elemento que ajuda a concatenar esforços para a tarefa prioritária de análise e interpretação dos dados disponíveis e sua articulação com as demais ações de gestão alegadamente motivadas pelo incremento da aprendizagem dos alunos em sala de aula.

Há ainda, um traço característico da diferença entre os modelos próprios de avaliação e a Prova Brasil, identificado no âmbito das Secretarias municipais de Castro e São Paulo¹⁰. Para alguns gestores desses sistemas, a construção de *rankings* de escolas a partir de resultados de avaliações externas – como é frequente em relação à divulgação dos resultados da Prova Brasil - é bastante criticada. Uma saída para superar esse aspecto e, de certa forma, se desprender desse tipo de classificação midiática, seria a realização de avaliações próprias que propiciariam, ao mesmo tempo, um maior controle dos resultados e um recurso que, destituído de outras conotações potencialmente prejudiciais, se pautasse como subsídio direto ao trabalho pedagógico. O conjunto de percepções desses gestores, especificamente, indica que um modelo de avaliação própria seria mais “justo” ou “adequado” ao levar em conta as especificidades de uma determinada rede e de seus alunos, impedindo ou amenizando o fato de as diferenças entre as escolas criarem abismos dentro de um *ranking* nacional ou regional. Para isto, tanto o município de Castro como o de São Paulo não divulgam publicamente os resultados individuais de escolas, turmas ou alunos, entregando essa informação de maneira individual a cada uma das escolas.

Entretanto, há sensíveis diferenças nas formas adotadas pelas diferentes secretarias para divulgar e apresentar os resultados obtidos por escolas e os diferentes materiais que

¹⁰ Diferentemente dessas Secretarias de educação, no caso da Rede Estadual do Espírito Santo, a divulgação ampla dos resultados de sua avaliação, desde a escola até os meios da imprensa local, é marcadamente enfatizada e valorizada.

possam ajudar em sua compreensão e comprometimento por parte das equipes escolares. A rede estadual do Espírito Santo, por exemplo, divulga os resultados das escolas de forma ampla para toda a sociedade – através dos diversos órgãos da imprensa local. A aposta aqui é a de que a partir dessa divulgação, a sociedade se mobilizaria para cobrar as equipes escolares, o que por sua vez poderia impulsionar a apropriação e o uso dos resultados nas escolas.

Contudo, tanto na rede do Espírito Santo quanto na rede de Sorocaba - que também divulga os resultados por escola de forma ampla -, não foram coletados relatos ou indícios de que atores externos às redes de ensino tenham feito algum tipo de cobrança às equipes escolares em relação aos resultados dos alunos nos testes ou tomado alguma atitude a partir disso. Nessas redes, a apropriação dos resultados por parte das equipes escolares, quando observada, parece estar atrelada a outras estratégias e motivações.

De forma geral podemos destacar alguns aspectos considerados como relevantes e que podem, de alguma forma, contribuir para o debate a respeito das políticas de avaliação externa, ao menos em relação aos sistemas de ensino participantes da pesquisa.

Embora tais avaliações possam ser interpretadas como subsídios importantes para a melhoria do trabalho escolar, a escolha de um desenho de política de avaliação por parte de Secretarias de Educação estaduais ou municipais se depara com alguns obstáculos, notadamente em relação à sobreposição de avaliações externas no ambiente escolar, uma vez que as unidades escolares já participam de avaliações federais como a Prova Brasil e, eventualmente, o SAEB. Desde modo, corre-se o risco de as escolas serem levadas a uma exaustiva rotina de avaliações externas, não raro subtraindo da escola uma quantidade de tempo útil para desenvolvimento de seu trabalho pedagógico e principalmente, favorecendo o próprio desgaste desse tipo de ação avaliativa.

Outro aspecto importante, quando são analisados os objetivos e justificativas de uma avaliação externa realizada pelos sistemas de educação estaduais e municipais em pauta, diz respeito às estratégias utilizadas para a divulgação dos resultados. Nos casos estudados, parece não haver um consenso quanto a melhor forma de divulgação dos resultados das avaliações externas. Por um lado, observamos secretarias que produzem ou disponibilizam os resultados de forma a propiciar a organização de rankings quanto aos resultados obtidos por cada uma de suas escolas na avaliação externa. Por outro, redes de ensino que optam por não divulgar ou favorecer a elaboração de *rankings* de escolas a partir de seus resultados. Nesse caso, há a pressuposição de que tal prática não teria muito a acrescentar à desejável melhoria

do trabalho pedagógico e provavelmente acirraria as reações contrárias ao uso dos resultados em seu sentido estritamente pedagógico.

Por fim cabe destacar que as eventuais implicações da avaliação externa e as expectativas nela depositadas demandam a verificação de sua *validade consequencial* (Viana, 2003) a partir dos efeitos que os resultados podem despertar no pensamento, nas atitudes e na ação dos diversos interessados, em especial dos agentes escolares¹¹. Nesse sentido, o uso de avaliações externas pode e deve ser mais abrangente, demandando um conjunto coerente de ações em diferentes níveis e direções a partir das características próprias de cada secretaria de educação.

Da mesma forma que diversos elementos presentes na estrutura e gestão das secretarias de educação podem determinar a efetividade de uma política de avaliação, a recíproca também é verdadeira. Ao eleger as avaliações externas como uma política educacional estratégica, as secretarias de educação investigadas estão cada vez mais atentas a diferentes aspectos que podem favorecer o seu uso em seus diversos segmentos, sobretudo no cotidiano escolar – tendo em vista esse tipo de política, relativamente recente no cenário educacional brasileiro. Parafraseando a perspectiva de Hannah Arendt (2007, p.17) em torno da condição humana, também as redes de ensino são entidades *condicionadas*: “tudo aquilo com o qual elas entram em contato torna-se imediatamente uma condição para sua existência”.

Bibliografia

ARENDRT, Hannah. *A Condição Humana*. Rio de Janeiro: Forense Universitária, 2007.

VIANNA, Heraldo. *Fundamentos de um Programa de Avaliação Educacional*. In *Estudos em Avaliação Educacional*, n. 28, jul-dez/2003. São Paulo: Fundação Carlos Chagas, p. 23-38.

¹¹ Como destaca Heraldo Vianna (2003, p.32): “As avaliações, além das características normais relacionadas a diversos tipos de validade (conteúdo, preditiva e de construto), devem ter, necessariamente, *validade consequencial*. A expressão pode determinar controvérsias, necessitando, portanto, ser plenamente esclarecida. A validade consequencial não se refere a distinções, prêmios e/ou bônus, e muito menos a *rankings* e menos ainda a comparações. É fundamental que os resultados das *avaliações cheguem aos alunos, aos pais, aos educadores e a toda a comunidade educacional*, não devendo ficar restrita apenas aos *policy-makers* da administração escolar”.

BABBIE, E. Métodos de Pesquisas de Survey. Tradução Guilherme Cezarino. Belo Horizonte: Ed. UFMG, 1999.

BONAMINO, Alicia e SOUSA, Sandra Zákia. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educ. Pesqui.* [online]. ahead of print, pp. 0-0. Epub 14-Fev-2012. ISSN 1517-9702. <http://dx.doi.org/10.1590/S1517-97022012005000006>.

BROOKE, Nigel; SOARES, José Francisco (Orgs). Pesquisa em eficácia escolar: origens e trajetórias. Belo Horizonte: Editora UFMG, 2008.

BROOKE, Nigel. Responsabilização Educacional no Brasil. *Revista Iberoamericana de Evaluación Educativa*. V.1, nº. 1, pp. 93-109, 2008.

BROOKE, Nigel; UNHA, Maria Amália de A. *A avaliação externa como instrumento da gestão educacional nos estados*. In Estudos & Pesquisas Educacionais – n. 2, novembro 2011 – Fundação Victor Civita, São Paulo.

DOURADO, L. F; SANTOS, C. A; OLIVEIRA, J. F. *A qualidade da educação: conceitos e definições*. Série Documental (INEP), Brasília, v. 24, n. 22, p. 05-34, 2007.

OLIVEIRA, Romualdo. Reformas educativas no Brasil na década de 90. In: CATANI, Afrânio; OLIVEIRA, Romualdo (orgs.). *Reformas educacionais em Portugal e no Brasil*. Belo Horizonte: Autêntica, 2000.

DESLAURIERS, J; KERISIT, M. O delineamento da pesquisa qualitativa In: A Pesquisa Qualitativa: enfoques epistemológicos e metodológicos. Petrópolis: Vozes, 2008. p.127-153.

SOUSA, Sandra Zákia; PIMENTA, Claudia; MACHADO, Cristiane. Avaliação e Gestão Municipal da Educação. *Est. Aval. Educ.*, São Paulo, v. 23, n. 53, p. 14-36, set/dez. 2012

SOUSA, Sandra Zákia e OLIVEIRA, Romualdo Portela de. Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. *Cadernos de Pesquisa*, vol.40, n.141, pp. 793-822, 2010.

SOUSA, Sandra Zákia L. Possíveis impactos das políticas de avaliação no currículo escolar: *Cadernos de Pesquisa*, n. 119, jul. 2003.

VALLE, Raquel da Cunha. Teoria de resposta ao item. *Est. Aval. Educ.* Nº21, Jun. 2000, p.07-92.

SILVA, Vandré Gomes. A narrativa instrumental da qualidade na educação. Estudos em Avaliação Educacional. *Estudos em Avaliação Educacional*.V.19. n.40, pp. 191-221, 2008.