

A avaliação externa como instrumento da gestão educacional nos estados¹

NIGEL BROOKE | MARIA AMÁLIA DE A. CUNHA

GAME/FAE/UFMG

1. Introdução

Estudo realizado em 2006 sobre os sistemas de avaliação de cinco estados brasileiros concluiu que houve melhoras significativas no desenho técnico dos instrumentos e na confiabilidade dos resultados. Entretanto, os pesquisadores observaram que os resultados eram pouco aproveitados pelas Secretarias de Educação (SEs) para a formulação de políticas de gestão (SOUSA & OLIVEIRA, 2007). Em documento de 2005, o próprio Conselho Nacional de Secretários de Educação (Consed) parecia confirmar essa conclusão ao afirmar que o principal uso dos resultados das avaliações externas era para a orientação da formação continuada de professores e destacou como desafio para os próximos anos o “melhor aproveitamento da avaliação em larga escala” (CONSED, 2005).

As explicações mais correntes para a aparente incipiência de políticas de gestão calcadas em informações sobre o desempenho dos alunos eram três: as Secretarias ainda não percebiam o potencial dessas informações para a formulação de políticas de gestão de recursos e pessoal; as Secretarias estavam constrangidas de formular políticas de gestão pelas resis-

1 Este estudo foi realizado pelo Game/FAE/UFMG sob encomenda da Fundação Victor Civita (FVC). Participou do planejamento desta pesquisa como consultora técnica, Paula Louzano.

tências encontradas entre os professores; e os instrumentos de avaliação usados pelas Secretarias nem sempre correspondiam aos usos almejados pelos gestores.

Levantamentos feitos pelos autores no ano de 2010 sobre os diferentes usos dos sistemas estaduais de avaliação mostram que essas conclusões estão desatualizadas. Pela expansão no número e sofisticação dos sistemas de avaliação durante os últimos anos e pela exploração de novas formas de aproveitar as informações geradas, constata-se a diversificação nas práticas de gestão educacional com base nos resultados dos alunos. De forma isolada ou em conjunto com outras informações, os resultados gerados pelos sistemas estaduais de avaliação educacional estão sendo usados para finalidades que variam entre a criação de indicadores estaduais de desenvolvimento educacional e a avaliação de desempenho individual de diretores escolares.

Percebe-se que as Secretarias não estão sentindo os mesmos constrangimentos do passado. Se ainda há alguma dificuldade no uso dos resultados para orientar a prática dos professores, ela não tem impedido sua aplicação em uma nova geração de políticas de gestão. Muitas vezes, essas novas políticas envolvem a identificação e classificação das escolas com base nos resultados de Português e Matemática dos alunos do 5º e 9º anos do Ensino Fundamental e do 3º ano do Ensino Médio, que são as informações disponibilizadas pela maioria dos sistemas estaduais de avaliação que seguem o padrão do Sistema de Avaliação da Educação Básica (Saeb). Mesmo não sendo propriamente uma avaliação da escola, por não testar todos os alunos em todas as disciplinas e anos e muito menos os campos não acadêmicos do currículo, as Secretarias presumem que os resultados das duas disciplinas básicas nos anos finais de cada ciclo de estudos possam oferecer uma amostra das atividades da escola que sirva de indicador do desempenho geral da instituição. Mesmo esparsa, há alguma evidência na literatura de que esse pressuposto seja válido, por causa da correlação positiva entre diferentes medidas de desempenho acadêmico dos alunos (RUTTER et al., 1979; MORTIMORE et al., 1988).

2. A classificação dos usos da avaliação

A necessidade de fazer distinções entre os novos usos das avaliações nos obriga a procurar um sistema classificatório que dê conta da variedade das políticas de gestão. Essa tarefa representa um desafio pelo fato de as classificações tradicionais se ocuparem principalmente dos propósitos genéricos dos sistemas de avaliação e não dos diferentes usos da informação. Por exemplo, é comum a distinção entre a avaliação formativa e somativa, sobretudo na descrição das avaliações de sala de aula, e entre a avaliação diagnóstica e certificatória. Nesses casos, é a natureza da tarefa que vai determinar a forma da avaliação e o uso posterior dos resultados. Ou seja, o desenho do instrumento é definido pelo tipo de informação que se precisa produzir de forma que, grosso modo, o uso da informação já é dado pela natureza e pelo propósito do instrumento. No caso do instrumento diagnóstico, por exemplo, talvez nem seja necessário aplicar os testes à totalidade dos alunos nem colocar todos os itens num único teste vultoso. Basta uma amostra de alunos e versões paralelas do teste para ter uma radiografia da aprendizagem das diferentes matérias. Esse é o modelo de avaliação seguido pelo Saeb. O instrumento certificatório, por outro lado, precisa ser aplicado a todos os alunos que querem a certificação e cada teste precisa dar conta da matéria inteira para oferecer uma medida fidedigna da aprendizagem de cada aluno. Ravela admite que uma avaliação possa ter mais de uma finalidade e que as diferentes finalidades não são necessariamente exclusivas, mas lembra que “cada opção requer um determinado desenho e tem diferentes exigências técnicas e custos” (RAVELA, 2008, p. 12).

Outro tipo de classificação, que reconhece a multidimensionalidade da avaliação, procura encontrar as categorias que descrevem os diferentes tipos de informação gerada. Shulha e Cousin, por exemplo, em estudo que resumia os usos da avaliação nos Estados Unidos no período entre 1986 e 1996, empregam três categorias de informação: instrumental, conceitual e simbólica (SHULHA & COUSIN, 1997). Em esquema parecido, descrito por Soares (2002), encontramos uma tipologia que fala das três funções genéricas dos sistemas de avaliação. A primeira é a função métrica, de medir onde se situam os sistemas de educação em comparação com algum parâmetro para responder perguntas sobre possíveis avanços. A segunda função

é a analítica, ou seja, oferecer subsídios aos pesquisadores e gestores no seu trabalho de investigar o funcionamento do sistema. A terceira, considerada a mais difícil, é a função pedagógica, que trata do uso da avaliação como instrumento para a melhoria do ensino (SOARES, 2002).

No caso brasileiro, fica evidente que os mesmos instrumentos estão sendo usados com objetivos diversos de modo que a finalidade precípua ou original do instrumento pouco esclarece sobre seus desdobramentos posteriores. Também fica evidente pelas declarações das autoridades que as finalidades dadas como razão pela criação dos sistemas de avaliação são muito abrangentes e, às vezes, até ultrapassam a própria capacidade da avaliação educacional. Essas declarações incluem, por exemplo, a finalidade de “melhorar a qualidade da educação” ou, no caso do Saeb, “reverter o quadro de baixa qualidade e produtividade do ensino caracterizado, essencialmente, pelos índices de repetência e evasão escolar” (SOUSA, 2003). A classificação com base nas finalidades oferecidas pelas autoridades como justificativas deixaria de captar as especificidades dos diferentes usos ao mesmo tempo que obrigaria a inclusão de categorias fora do escopo real da avaliação.

Nesse cenário, a nossa classificação dos usos dos sistemas precisa se basear em outra lógica, a das diferentes políticas efetivamente implementadas pelas autoridades uma vez construídas e aplicadas as diversas formas de avaliação em larga escala. Ao fazer essa distinção, estaríamos concordando com a diferença entre finalidade e utilidade explicitada no trabalho de Dantas sobre o uso para efeitos pedagógicos dos resultados do sistema de avaliação do estado da Bahia (DANTAS, 2009). Para essa autora, a finalidade é o objetivo expresso, enquanto a utilidade relaciona-se com a demanda de informações por parte dos *stakeholders* do sistema, incluindo os gestores.

Encontramos uma classificação de “utilidades” em termos de políticas de gestão no trabalho de Shepard em que são listados sete propósitos ou políticas governamentais (SHEPARD, 1977). O autor não objetiva definir quais propósitos podem ser alcançados com os mesmos instrumentos. Pelo contrário, parece subentendido que cada propósito significaria um instrumento diferente ou no mínimo um desenho para a avaliação que incorporasse as utilidades previstas. Eles são:

1. Diagnóstico da aprendizagem do aluno
2. Certificação do aluno
3. Avaliação de pessoal docente, diretores e outros
4. Avaliação de programas
5. Alocação de recursos
6. *Accountability*²
7. Pesquisa

Para as nossas necessidades, essa lista tem dois defeitos. O primeiro reside no desdobramento desnecessário de 4. Avaliação de programas e 7. Pesquisa. Em ambos os casos, os resultados podem ser usados como evidência da existência de efeitos causados por políticas, programas e até situações quase experimentais, em que há interesse em medir os impactos de mudanças no funcionamento das escolas e na aprendizagem dos alunos. As avaliações podem acontecer em nível estadual, municipal ou escolar e podem ser para efeitos tanto somativos, para determinar as consequências de determinadas políticas, como formativos, para corrigir os rumos de programas em andamento. A essas utilidades pode-se acrescentar as funções de planejamento e monitoramento, que também pressupõem a projeção e análise dos resultados ao longo de um período de tempo. Conjuntamente, todas essas diligências expressam o propósito de usar os resultados para avaliar e orientar a ação do próprio governo em nível mais sistêmico.

-
2. Neste texto usaremos o termo responsabilização como tradução dessa palavra. A tradução não é perfeita, já que a palavra *accountability* expressa dois conceitos, um relacionado com a prestação de contas e o outro com a ideia da responsabilização. Quando aplicada especificamente à escola, a noção de prestação de contas significa a cobrança legítima de um serviço público de qualidade em troca da manutenção da escola com recursos públicos oriundos dos impostos pagos pela população. O segundo significado da palavra é inseparável e consequência do outro, mas expressa a ideia que o público tem o direito de demandar que a escola assuma sua responsabilidade direta pelos resultados de seus alunos. *Accountability*, portanto, significa uma cobrança por bons resultados e a demanda de que cada um dos atores envolvidos assumam a sua responsabilidade na produção desses resultados (BROOKE, 2006; CORVALÁN & MCMEEKIN, 2006; FERRER, 2006).

O segundo defeito reside no uso da categoria *accountability* para classificar determinadas avaliações de forma exclusiva. Os autores Klinger, DeLuca e Miller, na descrição da evolução da avaliação educacional no Canadá, e Bauer, no caso dos sistemas nacionais de avaliação na América Latina, também empregam o termo dessa forma ao criar uma distinção entre os objetivos das avaliações relacionados à *accountability* e os objetivos relacionados com a seleção, classificação, premiação ou certificação (KLINGER; DELUCA & MILLER, 2008; BAUER, 2010). Porém o que se pode mostrar é que *accountability* no contexto educacional significa acima de tudo a associação entre resultados de aprendizagem e consequências para o distrito educacional, a escola, os professores ou os alunos. O grau de pressão que a autoridade educacional deseja aplicar ao processo de indução pode ser calibrado de acordo com o tipo ou a seriedade das consequências. Nesse contexto, é comum rotular os níveis de pressão/indução mais altos de *high-stakes* e os mais baixos de *low-stakes*, usando a palavra inglesa *stakes*, que significa “o valor colocado em jogo pelo apostador”. Por isso, pode haver uma política de gestão com aparente utilidade única, mas cujas consequências são variáveis de acordo com o contexto, as intenções das autoridades e o nível dos *stakes* associados a elas. Dependendo desses condicionantes, as consequências de resultados de aprendizagem baixos podem ser insignificantes ou nulas ou podem ser extremamente prejudiciais à vida futura do aluno ou da escola. Para dar conta dessas diferentes realidades, em que o objetivo da avaliação pode estar associado ou não a diferentes medidas de *accountability*, observa-se a necessidade de conceituar a *accountability* como se fosse a segunda dimensão de um sistema classificatório bidimensional. Com isso, para todas as utilidades da avaliação, potencialmente existe a possibilidade de agregar consequências de impacto maior ou menor. Na realidade, como será mostrado, há categorias de uso em que o nível de *accountability* costuma ser único, sem que isso invalide a bidimensionalidade das outras categorias.

Outra característica do esquema de Shepard, que não chega a ser um defeito, é a inclusão do propósito “certificação de alunos”. Diferentemente de outros sistemas, como o dos Estados Unidos, os sistemas brasileiros de avaliação nunca foram planejados para determinar a aprovação dos alunos

do ensino básico, salvo uma breve e polêmica experiência no estado de São Paulo em 2001, quando o Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp) foi usado para definir a aprovação ou reprovação de alunos ao final dos ciclos do Ensino Fundamental (4ª e 8ª séries). A criação do Exame Nacional do Ensino Médio (Enem), que aos poucos assume as características de um exame certificatório de saída do Ensino Médio na medida em que se torna requisito e critério para o ingresso à universidade, alterou o panorama da avaliação federal, mas ainda não há sinal de interesse nem cobertura legal para a criação de avaliações certificatórias equivalentes por parte dos estados.

No entanto, a função certificatória não pode ser eliminada do esquema classificatório devido ao possível uso dos resultados da avaliação pelas SEs na sua capacidade reguladora como critério para autorizar ou reautorizar o funcionamento de escolas particulares. O uso de resultados de avaliações externas como componente da nota semestral ou bimestral atribuído ao aluno pelo professor também poderia ser considerado um caso de certificação se essa nota influenciasse a decisão de promover ou reter o aluno.

Com as modificações sugeridas, e a inclusão de três categorias novas, que logo serão explicadas, chega-se a uma nova matriz classificatória para a tarefa de categorizar todas as políticas dos estados que fazem uso das informações geradas pelos seus sistemas próprios de avaliação padronizada. A discussão que segue centra-se na explicitação dessa matriz certificatória mediante o emprego de exemplos levantados pela pesquisa.

Quadro 1. Classificação dos usos da avaliação externa como instrumento da gestão educacional nos estados

	Uso dos resultados da avaliação (Política de gestão)	Sem consequências ou consequências menores (<i>low-stakes</i>)	Com consequências importantes (<i>high-stakes</i>)
1	Para avaliar e orientar a política educacional	<ul style="list-style-type: none"> • Edurural¹ (MEC²) • Avaliação PDE³ (Ceará) • Avaliação Escolas de Referência (PE) • Idesp⁴; Idepe⁵; Ideam⁶; IDE-Rio⁷ etc. • Gestão para resultados (MG) 	<ul style="list-style-type: none"> • PAR⁹ (MEC) • Prog. Escolas-Padrão (SP)
2	Para informar as escolas sobre a aprendizagem dos alunos e definir as estratégias de formação continuada	<ul style="list-style-type: none"> • Boletim de Resultados (PE) • Relatório Pedagógico (DF) • PIP⁸ (MG) 	
3	Para informar ao público	<ul style="list-style-type: none"> • Boletim da escola (Paraná) • Propaganda política (MG) 	<ul style="list-style-type: none"> • Escolha de escola (Chile)
4	Para a alocação de recursos	<ul style="list-style-type: none"> • Aprender Mais (PE) • Escola Referência (MG) 	<ul style="list-style-type: none"> • Escola Nota 10 (CE) • Índice Municipal de Qualidade Educacional (CE)
5	Para políticas de incentivos salariais		<ul style="list-style-type: none"> • Acordo de Resultados (MG) • Escola Nota 10 (CE) • Bonificação por Desempenho (SP)
6	Como componente da política de avaliação docente		<ul style="list-style-type: none"> • ADI¹⁰ (MG) • Recondução de diretores (DF)
7	Para a certificação de alunos e escolas		<ul style="list-style-type: none"> • Promoção de alunos de 4^a e 8^a series, 2001 (SP) • Aprovação de funcionamento de escolas (DF)

1. Educação Rural 2. Ministério da Educação 3. Plano de Desenvolvimento da Educação 4. Índice de Desenvolvimento da Educação do Estado de São Paulo 5. Índice de Desenvolvimento da Educação do Estado de Pernambuco 6. Índice de Desenvolvimento da Educação do Amazonas 7. Índice de Desenvolvimento da Educação do Estado do Rio de Janeiro 8. Plano de Intervenção Pedagógica 9. Programa de Ações Articuladas 10. Avaliação de Desempenho Individual

3. Pesquisa

Pela natureza da investigação, que foi mais um levantamento para fins descritivos do que um retrato da diversidade das políticas de gestão, nosso intuito não foi o de oferecer uma medida da frequência dessas políticas nos diferentes cantos do país. Por ter um propósito mais elucidativo, enfatizamos as características e qualidades das políticas encontradas, e não sua importância relativa ou distribuição em termos geográficos. Dessa forma, o estudo foi planejado em duas partes. A primeira se concentrou na coleta, de todas as fontes secundárias possíveis, de todas as informações disponíveis sobre a avaliação educacional em nível subnacional. Para essa primeira etapa, julgamos que a fonte mais importante para a coleta das informações era a internet e os sites patrocinados pelos próprios governos estaduais e municipais.

A segunda fase se concentrou no aprofundamento dos padrões que surgiram na primeira fase e na elucidação de exemplos de políticas de gestão cujas informações tinham chegado de forma incompleta ou que suscitavam interesse especial. Nessa fase foram realizadas 20 entrevistas semiestruturadas com gestores, técnicos e especialistas em avaliação em sete estados diferentes. O objetivo da abordagem empírica, na investigação desse objeto, era caracterizar melhor as políticas de gestão e descobrir indicações da sua eficácia.

4. Exemplos dos usos da avaliação educacional

4.1. Para avaliar e orientar a política educacional

A história do desenvolvimento das avaliações padronizadas em larga escala no Brasil está associada à avaliação de programas educacionais com financiamento federal. O primeiro exemplo foi a avaliação do Projeto Edu-rural, que utilizou instrumentos padronizados criados sob medida baseados em exemplos de exercícios e trabalhos de sala de aula colhidos nas escolas de três dos estados do Nordeste onde o projeto se desenvolvia – Piauí, Ceará e Pernambuco. A avaliação acompanhou a implementação e o desenvolvimento do projeto de 1982 a 1986, medindo o rendimento dos alunos, analisando as formas de gerenciamento do projeto e acompanhando profes-

sores, organizações municipais de ensino, alunos e famílias. A avaliação do Edurural se tornou um marco justamente por ela fornecer as bases técnicas e mostrar a relevância da avaliação educacional em larga escala para a identificação mais nítida dos problemas das escolas. Até o final da mesma década, os alicerces de um sistema nacional de avaliação já haviam sido criados (GATTI, 2002).

Apesar da conexão histórica entre a avaliação educacional e a avaliação de programas federais, os exemplos de uso dos resultados dos sistemas estaduais brasileiros para essa finalidade são limitados. Introduzir a aprendizagem dos alunos como variável dependente em estudo de impacto de determinada política ou programa de intervenção estadual ainda é incomum em comparação com os estudos descritivos dos métodos e processos de implantação desses mesmos programas. Não somente no Brasil, mas na América Latina como um todo, poucos programas voltados para a melhoria dos resultados têm sido acompanhados de modo a determinar sua eficácia (GTEE/PREAL, 2011).

Uma exceção recente é o trabalho da Secretaria de Educação do Ceará (SEE-CE), em que foram usados indicadores baseados nos resultados do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaee) para avaliar três programas, o Prêmio Escola Nota 10, o Plano de Desenvolvimento da Educação (PDE) e o Prêmio Aprender pra Valer.

Por outro lado, o uso das avaliações estaduais para fins de um monitoramento genérico da evolução geral do sistema se torna cada vez mais frequente, como evidenciado pela criação recente de indicadores que tentam captar em um único número uma medida da qualidade de cada fase de educação em cada escola. Influenciados e incentivados pela criação do Índice de Desenvolvimento da Educação Básica (Ideb) em nível nacional, alguns estados, incluindo Pernambuco (Idepe), Amazonas (Ideam), São Paulo (Idesp), Ceará (IDE-Alfa³), Rio de Janeiro (Iderj), e os municípios do Rio de Janeiro (IDE-Rio), São Paulo (Indique⁴) e o Distrito Federal (IDDF) já criaram indicadores desse tipo.

3. Indicador de Desempenho Escolar em Alfabetização.

4. Índice de Qualidade da Educação.

O Índice de Desenvolvimento da Educação do Distrito Federal (IDDF) se apresenta como exemplo tanto da necessidade de criar um indicador compreensível como do valor do indicador como um elemento estruturador dos planos da Secretaria. Segundo o material usado para explicar o IDDF às escolas, o desafio primordial da Secretaria de Estado de Educação do Distrito Federal (SEE-DF) é garantir que os alunos aprendam sem desperdiçar tempo com reprovações e sem abandonar a escola. Para cumprir essa tarefa, a SEE-DF precisa de “indicadores objetivos” para acompanhar a evolução das escolas e definir “metas individuais para cada instituição educacional, a serem alcançadas em curto prazo”. O objetivo, em nível da escola, é que as metas “sirvam como guia para seus gestores e se traduzam em esforços pedagógicos capazes de elevar o desempenho dos estudantes, garantindo que uma proporção cada vez maior de alunos domine um sólido conhecimento dos conteúdos e habilidades esperados para o seu estágio escolar”. Visando esse objetivo, com base em um indicador que junta os dois critérios para uma avaliação da qualidade da escola, o IDDF é composto de uma medida da aprendizagem dos alunos, advinda dos exames do Sistema de Avaliação de Desempenho das Instituições Educacionais do Sistema de Ensino do Distrito Federal (Siade), e uma medida da movimentação escolar derivada das taxas de aprovação, em fórmula análoga à do Ideb.

O tema dos indicadores estaduais ainda não ultrapassou os círculos de planejamento no interior das Secretarias para se tornar um assunto de discussão acadêmica. Uma exceção é a discussão do indicador do estado de São Paulo – Saresp. Seguindo uma metodologia mais completa que o Ideb, o indicador de São Paulo capta também a equidade das estratégias educacionais usadas naquele sistema pela inclusão de uma medida da distribuição dos alunos por faixa ou categoria de desempenho (SOARES, 2009). Ao incentivar a progressão de todos os alunos, e não somente a melhoria da média, o Saresp consegue corrigir uma falha do Ideb e promover uma equidade maior nas estratégias de melhoria das escolas. O indicador mais recente, o Indique da cidade de São Paulo, vai um passo além ao criar um indicador de esforço da escola com base em uma medida do nível socioeconômico da sua clientela derivada de informações fornecidas pelos alunos nos questionários contextuais aplicados junto com a Prova São Paulo.

A utilização dos resultados das avaliações educacionais como peça central do planejamento estratégico tem como caso emblemático o estado de Minas Gerais. Inspirado nas iniciativas do Governo Federal e nas experiências internacionais, Minas Gerais foi um dos primeiros estados a incorporar o modelo de gestão por resultados e oferece um exemplo da afinidade entre os propósitos da avaliação educacional e as exigências dos novos padrões da administração pública. Na primeira fase, iniciada em 2003, os planejadores produziram o Plano Mineiro de Desenvolvimento Integrado (PMDI), que serviu como a visão do futuro para o período até 2020. Os cenários futuros evidenciaram a necessidade da reorganização e modernização do setor público estadual por meio do chamado Choque de Gestão. Para criar a ponte entre os cenários futuros e as ações do secretariado, foram criados 30 projetos estruturadores e depois os programas setoriais para compor o plano plurianual do governo para o período até 2006. Para definir os objetivos de seus programas para o Ensino Fundamental e Médio, a SE optou por uma bateria de indicadores, entre os quais as proficiências médias dos alunos de 3º, 5º e 9º anos do Ensino Fundamental e 3º ano do Ensino Médio. Derivados do Sistema Estadual de Avaliação (Simave), essas proficiências ofereceram uma forma concreta de fixar metas com base em resultados finalísticos, quando a maioria das Secretarias só oferecia indicadores de processo.

Na segunda fase de implantação do novo modelo de gestão, chamada Estado para Resultados, ficou ainda mais evidente o uso do Simave para atender aos requisitos dos planejadores. Como estratégias para a fixação de critérios comuns para a reorganização e modernização do serviço público, foram aprimorados o sistema de avaliação de desempenho de todos os servidores e o chamado Acordo de Resultados. Esse último, uma metodologia para estabelecer metas de desempenho para todas as instâncias das Secretarias e pagar uma bonificação em caso de cumprimento dessas metas, encontrou na SE o melhor exemplo de como calcular com exatidão os resultados esperados. A orientação política a favor da melhoria dos resultados dos serviços públicos achou no sistema de avaliação educacional as respostas desejadas. Ao mesmo tempo, o sistema de avaliação se desdobrou e penetrou em todos os programas da Secretaria em resposta a um ambiente de planejamento francamente propício ao uso de objetivos e metas quantificáveis.

Outro estado a criar uma conexão entre as metas do governo e os resultados da escola foi o Espírito Santo. O Plano Estratégico Espírito Santo 2025 reafirmou a prioridade absoluta a ser dada à Educação pelo governo que tomou posse em 2003. Posteriormente, a parte do plano que cabia à SE recebeu um detalhamento para o período 2008-2011 e assumiu o nome Plano Estratégico Nova Escola. O objetivo maior era “o alcance de escolas vivas”, com capacidade criativa e inovadora, modernas e atrativas, que na essência “tenham como resultados de suas ações a aprendizagem dos alunos”, segundo o secretário da época, Haroldo Corrêa Rocha (ESPIRITO SANTO, 2008, p.5). Alinhado com a ênfase nos resultados de aprendizagem, o plano assumiu como sua Visão de Futuro 2011 um cenário em que o Espírito Santo figuraria entre os estados mais bem colocados “nas avaliações de resultados” e formulou diversas metas pensadas para levar o estado a essa posição. Entre as metas assumidas, constam a elevação das médias do 3º ano do Ensino Médio do Saeb para 350 pontos em Português e 400 em Matemática, a elevação do Ideb das séries iniciais e finais do Ensino Fundamental de, respectivamente, 3,7 e 3,5, em 2005, para 5,9, em 2021, e o Ideb no Ensino Médio de 3,1 para 4,5. Se o Plano Estratégico Nova Escola ainda não expressava suas metas em termos dos resultados do seu próprio sistema de avaliação, Programa de Avaliação Básica do Espírito Santo (Paebes), a inclusão do projeto de “avaliar sistematicamente as condições de funcionamento das escolas e de desempenho dos alunos visando a obtenção de melhores resultados da aprendizagem” deixa claro quanto o plano oferecia de respaldo às atividades de avaliação educacional e quanto o espírito da avaliação estava presente no pensamento dos planejadores.

Não obstante a importância do planejamento estratégico, talvez o aspecto mais inovador da conexão entre a avaliação e o planejamento seja a relação, ainda em construção, entre os indicadores estaduais de desenvolvimento e a fixação de metas em nível escolar. Com uma medida confiável do desempenho dos alunos e um índice que capta a qualidade em um único número compreensível, a tarefa de fixar e cobrar metas se tornou mais fácil. Isso tem significado o avanço na utilização de metas no planejamento das atividades das Secretarias e também na formulação de políticas de incentivos que dependam do cumprimento das metas fixadas.

4.2. Para informar as escolas sobre a aprendizagem dos alunos e definir as estratégias de formação continuada

A preocupação com a qualidade da Educação Básica está no cerne dos programas subnacionais de avaliação. Apesar de usarem instrumentos descendentes do Saeb, que são fundamentalmente de monitoramento e não de diagnóstico, todos têm como objetivo explícito fornecer *feedback* aos professores de sala de aula e ajudar na busca de soluções para superar as dificuldades de aprendizagem dos alunos. Sem exceção, as Secretarias fundamentam seu apoio pedagógico aos professores na produção de materiais informativos que tentam traçar um diagnóstico do nível de aprendizagem dos alunos. Esses materiais são apresentados basicamente em duas versões: a) Boletins Pedagógicos, que seguem um formato próprio para o consumo do professor e disponibilizam, em meio a informações padronizadas, resultados específicos de cada escola; e b) Relatórios, destinados aos gestores, que resumem de forma mais técnica os resultados da rede. Juntamente com essas práticas de divulgação, observam-se diversas estratégias para a formação continuada de professores com base nos resultados da avaliação para facilitar a compreensão dos Boletins ou remediar problemas comuns. Ambas as políticas têm evoluído de forma marcante ao longo dos anos, acompanhando a crescente sofisticação dos instrumentos e a importância atribuída ao processo de devolução dos resultados. Além de dar destaque ao conteúdo, os gestores também têm dado mais importância à apresentação gráfica do Boletim, procurando tornar mais fácil a compreensão dos resultados.

Uma parte do crédito pela melhoria progressiva dos materiais “devolutivos” deve ser atribuída às próprias agências de avaliação, como o Centro de Políticas Públicas e Avaliação da Educação (CAEd), Centro de Seleção e Promoção de Eventos (Cespe) e Centro de Seleção de Candidatos ao Ensino Superior do Grande Rio (Cesgranrio). Incentivadas pelas SEs que são seus clientes, essas organizações experimentaram diferentes conteúdos e formatos ao longo dos anos e puderam verificar por meio dos contatos diretos com professores e técnicos o grau de acerto das diferentes maneiras de apresentar as informações. Quando chamadas para atender a novos clientes ou retomar

avaliações suspensas temporariamente, essas agências puderam oferecer um produto que incorporava a aprendizagem de muitos anos, permitindo que as novas avaliações principiassem em patamares mais sofisticados.

Mesmo com a evolução na qualidade dos Boletins Pedagógicos e, em alguns casos, a criação de um dia específico para o estudo coletivo do material devolutivo pela escola, a divulgação dos resultados não redundava automaticamente na sua apropriação por parte dos professores. As críticas dos professores sugerem que os resultados não são imediatamente consumíveis pela comunidade escolar, não obstante todo o empenho das Secretarias na apresentação de um material com cunho mais didático e instrutivo. Além disso, os Boletins Pedagógicos costumam chegar às escolas com certo atraso, dificultando a conexão entre os resultados apresentados e o tipo de intervenção necessário.

De modo geral, as Secretarias acreditam que o problema principal reside nas inúmeras dificuldades dos professores para entender as matrizes de referência dos testes, associá-las aos conteúdos ministrados e trabalhá-las em sala de aula. Observam-se tentativas de criar um diálogo que possa superar essas dificuldades. No Ceará, os professores participam de oficinas de elaboração de itens para difundir o conhecimento das matrizes e houve até a publicação das matrizes na capa dos diários de classe para garantir sua compreensão pelos docentes.

Uma forma de superar esse desafio tem sido a criação de programas estruturados de formação continuada com base nos resultados das escolas, que vão além do trabalho individual dos analistas ou supervisores pedagógicos, na tentativa de atingir um número mais significativo de professores. Nesses casos, as estratégias de formação ultrapassam as oficinas voltadas para a explicação dos resultados das avaliações para se tornarem programas permanentes de intervenção junto aos professores. Na cidade de Belo Horizonte, por exemplo, foi implantado o programa de Monitoramento da Aprendizagem. Com base nos resultados das avaliações externas realizadas na rede municipal de Educação (Proalfa, Provinha Brasil e Avalia BH), a Secretaria Municipal de Educação (SME) iniciou em 2009 um acompanhamento sistemático e individualizado de todos os alunos do Ensino Fundamental, chamado de Monitoramento da Aprendizagem. Além de ações

voltadas para os alunos, como o reforço escolar, são também desenvolvidas ações de formação docente, em que os professores recebem formação específica e acompanhamento constante da SME.

Talvez o exemplo mais acabado de uma política de formação continuada formulada com os resultados de avaliações externas se chame Plano de Intervenção Pedagógica (PIP), da Secretaria da Educação do estado de Minas Gerais (SEE-MG). Com base nos resultados das avaliações do Programa de Alfabetização, Documentação e Informação (Proalfa), educadores e gestores elaboram um PIP cuja finalidade reside na construção de estratégias para melhorar o desempenho dos alunos. O objetivo é fazer com que as escolas com os piores resultados recebam visitas de uma equipe pedagógica competente para discutir e trabalhar pontualmente as lacunas observadas com os resultados da avaliação.

O PIP teve início em 2008 com a implantação do programa chamado Alfabetização no Tempo Certo⁵, cuja meta era garantir as competências de leitura e escrita de toda criança até os 8 anos de idade até 2010. O PIP entra como suporte a esse programa a fim de oferecer subsídio para que o estado pudesse alcançar a meta de alfabetização e letramento de todas as crianças na idade certa, corrigindo problemas de fluxo de desempenho entre as escolas das diferentes regiões do estado, evidenciado pelos resultados da avaliação do Proalfa.

A atenção do PIP se volta para o atendimento das chamadas “escolas estratégicas”, isto é, as escolas com os piores resultados na avaliação externa. Essas escolas estão distribuídas por todo o estado, mas com uma concentração mais evidente na Região Metropolitana de Belo Horizonte e na Região Norte. Ao acompanhar as dificuldades dos professores no domínio de certos conteúdos, a equipe desenvolve um trabalho de capacitação individual dos professores baseado no resultado da avaliação dos alunos. As metas de alfabetização para o estado, ultrapassadas com folga, demonstram o êxito do PIP:

5. Modelado em programa quase homônimo do estado de Ceará, chamado Programa de Alfabetização na Idade Certa.

“Então na verdade quando você vê, por exemplo, que tínhamos uma meta de 74% [de alunos com] proficiência de alfabetização, vamos assim dizer, e nós atingimos 86%, alguma coisa melhorou, alguma coisa eles estão indo atrás. A gente não sabe se realmente é a intervenção do PIP, isso aí, a gente ainda não sabe qual que é a causa dessa melhora. Mas de 74 que era a nossa meta e nós atingirmos 86, é muita coisa né?” (técnico/SEE-MG)

As nossas investigações junto às outras Secretarias evidenciaram uma heterogeneidade de práticas, métodos e concepções teóricas sobre o uso dos resultados da avaliação na política de formação do professor. Se o PIP em Minas Gerais fornece um exemplo de um programa bem estruturado e aparentemente bem-sucedido, ele não representa um fenômeno isolado. Essa dimensão do uso dos resultados das avaliações parece estar em expansão. Na medida em que os instrumentos usados começam a fornecer informações sobre alunos individuais e vir acompanhados de materiais curriculares que procuram estabelecer um consenso sobre os conteúdos a serem ensinados, pode-se prever a diversificação ainda maior de iniciativas dessa natureza.

4.2.1. O currículo básico

O constante aprimoramento dos boletins pedagógicos foi uma aposta que, ao aprofundar cada vez mais a compreensão do professor em relação às matrizes de referência das provas e às habilidades que se pretendiam formar nos alunos, o aproveitamento dos resultados das avaliações seria cada vez melhor. Nesse sentido, foram sendo eliminadas informações que não diziam respeito à aprendizagem dos alunos, como as médias do Saeb no caso dos Boletins do Programa de Avaliação da Rede Pública de Educação Básica (Proeb) em Minas Gerais, e enfatizados exemplos de itens para ajudar o professor a entender a natureza das habilidades cobradas pela matriz. Os Boletins se expandiram com visualizações da escala de proficiência, mostrando a faixa de desempenho em que se situava a média da turma em relação a todas as principais competências testadas, e com exemplos de itens usados

nas provas para fazer a conexão entre a matriz, a escala e as habilidades identificadas pelos descritores.

Ao mesmo tempo, e possivelmente como uma indicação da frustração com a falta de impacto dos Boletins e os outros meios tradicionais de informar o professor, acendeu-se o interesse de algumas Secretarias em implantar um currículo oficial composto dos elementos considerados mínimos para cada série. Chamado de Currículo Básico Comum (CBC), em Minas Gerais, a Base Curricular Comum (BCC), em Pernambuco, e Proposta Curricular, em São Paulo, esses currículos parecem advir da constatação de que os esforços na busca por melhores resultados podem ser inócuos se não houver uma política bem focada na definição e distribuição dos conteúdos ensinados.

Nossa investigação demonstrou uma dinâmica de construção de um currículo comum oficial bastante recorrente: após a implantação dos sistemas de avaliação, chega-se à conclusão de que pouco vale investir em formação, oficinas pedagógicas ou na divulgação dos resultados por meio de boletins se não existir um currículo básico comum, que permita construir um diálogo com o professor a respeito daquilo que está sendo cobrado pelo sistema de avaliação.

No Espírito Santo, por exemplo, as próprias matrizes de referência foram transformadas em conteúdos curriculares básicos. Nesse exemplo, a linguagem das competências e habilidades usada pelas matrizes está ainda mais próxima aos objetivos do ensino:

“A avaliação puxou o currículo. Nós tínhamos a avaliação já avançada e não tínhamos o currículo [...] assim fica difícil, não tem uma referência para o professor, o professor não pode ficar usando matriz de avaliação como currículo. Aí começou a confecção do currículo com competências [...] e eles partiram da matriz para montar o currículo.”
(gestora/SEE-ES)

No caso de Ceará, após a decisão da Secretaria de imprimir a lista de competências e habilidades contidas nas matrizes de referência nos próprios diários dos professores, constatou-se que a carga horária dedicada efetivamente ao ensino das habilidades era mínima e que se precisava partir para a produ-

ção de um currículo básico calcado em conteúdos. Em Pernambuco, já houve uma evolução com a BCC em direção a uma especificação ainda mais detalhada dos conteúdos que os professores precisam ensinar.

Na maioria dos casos, há indicações de que os currículos em construção nos estados contam com a participação dos professores e refletem, portanto, um esforço comum para estipular os objetivos do ensino. Essa construção coletiva cria a esperança de uma implantação mais tranquila. Mesmo assim, a implantação e adoção gradual dos currículos estaduais é um processo lento. O domínio insuficiente dos conteúdos deixa os gestores frustrados e indicam problemas de fundo em relação à formação inicial:

“Eu diria que mesmo com o CBC ainda tem um bom espaço para andar [...] principalmente porque o professor não sabe o CBC. Esse é o problema, ele não sabe.” (especialista/MG)

Essas reflexões sobre a evolução das estratégias das Secretarias em direção a um currículo de conteúdos básicos servem para indicar que há uma vertente da relação entre avaliação e política de gestão que a presente pesquisa não teve condições de aprofundar. Essa relação não é a do uso dos resultados do sistema de avaliação para estabelecer os critérios para uma política de gestão. Em vez disso, ela reflete uma interpretação da dificuldade em fazer uso mais produtivo dos resultados como consequência da ausência de uma política curricular. Portanto, é mais a compreensão da fragilidade da avaliação externa em servir de instrumento pedagógico do que a compreensão da sua utilidade. O movimento em direção ao currículo básico expressa a conclusão de que, na ausência de um consenso compartilhado com os professores a respeito dos conteúdos curriculares e de um alinhamento entre os objetivos do ensino, o currículo da escola e as matrizes do sistema de avaliação, a contribuição deste para a melhoria da qualidade será seriamente prejudicada.

Pela natureza federativa do sistema educacional brasileiro e pela aparente impossibilidade de um currículo nacional que vá além dos seus parâmetros filosóficos, o Saeb foi criado com base em pesquisas que, na média, tentaram captar o currículo real usado pelas escolas sem que ele fosse o oficial ou comum a todas as escolas. Quando os estados começaram a adotar sistemas de

avaliação modelados no Saeb, eles também adotaram as mesmas matrizes e a mesma orientação no sentido da especificação de objetivos em termos de competências de diferentes níveis de sofisticação. Isso não significa que não haja um amplo grau de convergência entre as matrizes do Saeb e aquelas produzidas pelos estados que já passaram pelo processo de especificação do currículo básico. A diferença reside na explicitação e implantação de um currículo que possa legitimamente servir de foco para todos os processos de formação continuada, sobretudo quando expressado em termos de expectativas e quando fruto de uma construção coletiva. Usar as matrizes definidas externamente para essa finalidade não só restringe o currículo a um subconjunto daquilo que as escolas deveriam ensinar como também obriga a utilização de uma linguagem de especialistas que não é a dos livros didáticos e muito menos a das atividades que o professor costuma desenvolver.

Uma pesquisa sobre essa vertente da relação entre a avaliação e a política educacional mostraria provavelmente que, se não fosse pela dificuldade em comunicar os resultados da avaliação aos professores, o progresso na definição dos currículos estaduais teria sido menor. Mesmo atrasados, os currículos básicos dos estados e o progresso na definição de expectativas de aprendizagem para as diferentes séries são avanços na direção certa, que em algum grau podem ser atribuídos ao fato de os sistemas de avaliação terem saído na frente.

4.3. Para informar ao público

A lógica do sistema federal de avaliação no Brasil é de dar ampla divulgação aos resultados dos testes com o intuito de apoiar as instâncias subnacionais nas suas tarefas constitucionais e de envolver a população na discussão da qualidade da educação.

Por muito tempo, a disseminação dos resultados dos sistemas estaduais seguia outra lógica. Na sua maioria, esses sistemas eram vistos pelos seus criadores exclusivamente como instrumentos internos de gestão, e não como uma oportunidade para a prestação de contas a um público com direito a informações transparentes sobre a qualidade da educação. As políticas de divulgação se restringiam aos boletins para os professores e a outros poucos documentos internos.

A falta de transparência pode refletir mais a atitude de autopreservação, comum a todos os setores governamentais, do que uma característica própria das SEs. Mesmo assim, é surpreendente a ausência de informações detalhadas sobre os resultados das avaliações nos sites das Secretarias na internet. Em vez de acesso irrestrito a todos os dados, desagregados por escola e agregados por rede, região, município, nível e disciplina, o comum é um acesso liberado exclusivamente para as escolas, com informações que permitem poucas comparações. Em nenhum lugar se encontram as séries históricas do tipo publicado nos relatórios do Saeb para permitir que a população faça suas próprias análises das tendências mostradas pelos sistemas estaduais.

Uma das possíveis razões pela falta de informação nos sites das Secretarias tem a ver com a relação terceirizada entre o órgão público e a agência responsável pela avaliação. No Brasil, são quatro ou cinco as principais agências e empresas que oferecem o serviço de produção, aplicação e análise dos testes previstos pela legislação estadual que institui os sistemas de avaliação. De longe a maior dessas agências, o CAEd, ligado à Universidade Federal de Juiz de Fora, oferece alguma informação no seu próprio site sobre as avaliações realizadas. A impressão, no entanto, é que a atividade de divulgação ocupa um lugar secundário para o CAEd, mesmo nos casos em que há o pedido explícito da SE. Se as próprias Secretarias fossem as responsáveis diretas pela análise dos resultados e pela produção dos diversos textos que são distribuídos para as escolas, é provável que haveria uma melhora na qualidade e diversidade das informações nos seus sites.

As exceções que começam a aparecer, juntamente com o foco nos resultados, cada vez mais explícito, incluem os estados de São Paulo e Minas Gerais. No caso de São Paulo, encontram-se os dados do Saesp de 2007 a 2010 no site da Secretaria, com os documentos legais, os questionários aplicados à equipe escolar e a descrição das matrizes de referência. Os resultados são apresentados ano a ano por rede, nível, disciplina, coordenadoria da capital e interior e também de acordo com os diferentes níveis de desempenho, para ver a distribuição dos alunos em termos das competências adquiridas. Mesmo na ausência de gráficos e tabelas para facilitar o processo de comparação, os interessados podem encontrar todas as informações básicas, incluindo cópia dos boletins remetidos para as escolas.

No caso de Minas Gerais, são oferecidas informações completas sobre os últimos três anos de aplicação das provas do Proeb, juntamente com a matriz curricular, o boletim pedagógico e a “revista contextual”, como é chamado o documento contendo dados sobre o nível socioeconômico da escola e seu grau de eficácia. Não são mostrados nenhuma média por ano, disciplina, Superintendências Regionais de Ensino (SRE) ou qualquer outro parâmetro de agregação. Como no caso do Rio Grande do Sul e outros estados, a divulgação via site se dirige à escola e pressupõe que o único interesse dela seja ver seus próprios resultados, isolados de quaisquer comparações.

Não há dúvida de que na ausência de informações que permitam controlar a influência do nível socioeconômico dos alunos e outros fatores externos à escola, a comparação de escolas com base nos dados brutos dos sistemas estaduais não é recomendável. Pela impossibilidade de identificar claramente a contribuição da escola no resultado geral dos alunos, a comparação de dados brutos pode levar a conclusões errôneas tanto no sentido de atribuir qualidade maior a escolas com alunos de nível socioeconômico mais alto como no sentido de achar de pior qualidade escolas com alunos mais pobres. Perante esse risco, e o possível constrangimento das escolas, as Secretarias fazem bem em não permitir a divulgação de informações que possam facilitar o ranqueamento de escolas.

A única exceção que encontramos a essa regra geral foram os Relatórios de Desempenho Geral da Rede Estadual de Ensino produzidos pela Secretaria de Educação de Mato Grosso do Sul. Nesses relatórios, sobre os resultados dos alunos no Sistema de Avaliação da Educação Básica do Mato Grosso do Sul (Saems) em Língua Portuguesa, Produção de Texto e Matemática em 2009, encontram-se ranqueados escolas e municípios, abrindo os flancos para o efeito de comparação. No entanto, como a divulgação dos boletins pedagógicos atinge somente os diretores das escolas e não há política explícita de disseminação externa, deve-se supor que os ranqueamentos não faziam parte de uma estratégia premeditada de divulgação de comparações junto à comunidade.

Não obstante a circulação maior de informações no Brasil, não foi encontrada evidência de interesse em influenciar a decisão dos pais e, assim, implantar políticas de “escolha de escolas”, como existem nos Estados Uni-

dos e no Chile. O único caso documentado de um canal oficial de comunicação com os pais dos alunos mediante a produção de um informativo contendo os resultados da escola é o caso do Boletim de Escolas da Secretaria de Educação do Paraná, descrito por Ayres (2003). Criado como parte de uma estratégia para fomentar um mínimo de responsabilização para as escolas e dar poder aos pais para demandar resultados, o Boletim também tinha o propósito menos divulgado de fortalecer as associações de pais e contrabalancear o poder do sindicato de professores, que na época estava em desavença com o governo do estado. A teoria por trás dessa estratégia parece ser a de que as escolas se sentiriam mais pressionadas em demonstrar bons resultados se os pais estivessem mais bem informados a respeito da sua escola em comparação com as outras.

Com a mudança da administração estadual em 2003 e a nomeação de um secretário da educação com fortes laços com os sindicatos dos professores, o sistema de boletins foi abandonado. As dificuldades em treinar os pais na interpretação e no uso do boletim, a falta de comunicação real entre pais e professores e a falta de confiança geral dos professores e diretores com relação ao propósito do Boletim levaram ao seu abandono precoce. O fracasso do Boletim da Escola do Paraná mostra que uma estratégia de divulgação aparentemente *low-stakes* e sem consequências para os professores pode criar desconfianças e ser combatida se não for implementada de forma adequada. Pode-se supor que se o Boletim tivesse usado uma estratégia *high-stakes*, incorporando comparações entre as escolas e atribuindo alguma consequência aos indicadores empregados, a sua vida teria sido ainda mais curta.

Uma política similar, o Informativo de Contexto e Desempenho para as Escolas Municipais, foi implementada em Belo Horizonte em 2005 sem provocar resistências. De vida curta por força de mudanças na equipe de gestão e visivelmente mais direcionada às escolas do que aos pais dos alunos, o Informativo mostrou que é possível sugerir um nível moderado de responsabilização sem provocar a reação dos professores e sem prejudicar o objetivo de promover uma autoavaliação fundamentada em dados objetivos.

Outra modalidade de divulgação diz respeito ao uso dos resultados das avaliações como um elemento de propaganda política. Foram encontradas

peças dessa natureza em Minas Gerais, onde, no período anterior às eleições estaduais de 2010, foram distribuídos documentos .pdf contendo informações positivas sobre os avanços da educação no estado.

Essas peças de divulgação chamam a atenção justamente por serem tão incomuns. Enquanto as notícias sobre a educação eram sempre tão infelizes e os resultados dos sistemas de avaliação tidos como ruins, não havia o que celebrar. Os governantes que cumpriam seu papel de dar ao público o conhecimento dos resultados das avaliações o faziam mais por obrigação do que por querer a publicidade. Além do mais, na ausência de uma relação entre a política educacional e a popularidade do governo, não havia razão para o governante se expor. A divulgação dos resultados das avaliações não era uma prática calculada para ganhar votos.

A mudança desse cenário está associada à consolidação de instrumentos de avaliação que permitem fazer comparações – ao longo dos anos e entre os estados – e à criação de políticas educacionais voltadas para os resultados que pelo menos em alguns estados parecem estar produzindo efeito. Nesses lugares, começam a surgir melhorias que podem ser atribuídas à qualidade da gestão pública e também sinais de que a população está pronta para priorizar a qualidade da educação e disposta a premiar os governantes que mostram avanços nessa direção. No Ceará, o sucesso da política educacional do município de Sobral, disseminado por todos os meios de comunicação, contribuiu para a eleição do governador e continua fazendo escola⁶. Nesse estado, o governador procura fazer um evento político da entrega dos prêmios dos sistemas estaduais de incentivos.

A estratégia de divulgação do governador cearense pode ser atribuída ao desejo de dar prestígio à política educacional e angariar apoios entre os municípios para os esforços da SE. O evento pomposo da Escola Nota 10 seria, então, uma estratégia de mobilização para entusiasmar os dirigentes e provocar melhorias nos resultados dos municípios. Por outro lado, pode também ser um evento com fins políticos de caráter eleitoral. De qualquer forma, o cálculo é que o assunto da premiação renderá frutos tanto para o

6. No município de Sobral, houve 34,7% de aumento no Ideb nas séries iniciais entre 2007 e 2009.

estado como para os municípios se houver uma melhoria nos resultados das escolas.

A conclusão a que se chega é que a avaliação educacional faz parte de uma nova maneira de conduzir a educação que, ao promover a melhoria dos resultados da escola, está propiciando ganhos políticos para os gestores dos sistemas públicos. Nesse cenário, é a avaliação que fornece as informações sobre os avanços do sistema que ao serem divulgadas angariam os apoios desejados tanto em termos de colaboração como em termos eleitorais. Se a educação nos estados continuar a melhorar e se o interesse da população no assunto da qualidade da educação continuar a crescer, pode-se esperar um volume cada vez maior de informação nos sites das Secretarias, em inserções na televisão e em publicações de diversos tipos. Não há dúvida de que a circulação de maiores informações e a divulgação de resultados educacionais com propósitos políticos trarão benefícios para a qualidade da educação pública.

4.4. Para a alocação de recursos

O Plano de Ações Articuladas (PAR), implantado pelo MEC em 2007 e que contemplava, originalmente, os municípios considerados prioritários em função do Ideb alcançado em 2005, pode ser classificada aqui como política de alocação de recursos federais com base em resultados da avaliação externa. Com a extensão do plano para todos os municípios, a política adquiriu conotações de *accountability*. Pelo menos, é nessa luz que Adrião e Garcia analisam os objetivos e métodos do Ministério (ADRIÃO & GARCIA, 2008). Composto de programas e ações com o objetivo declarado de promover a melhoria da qualidade da Educação Básica brasileira, o plano é operacionalizado por meio de convênios entre os municípios e a União e da elaboração local de um PAR.

A existência de um documento legal que estabelece um acordo público na forma de metas implica, necessariamente, algum grau de responsabilização subjetiva. Dependendo da mobilização da população, o não-cumprimento das metas acordadas pode trazer um custo político para os gestores municipais. Resta saber se o MEC se interessa por condicionar a continui-

dade do financiamento ao cumprimento das metas e dessa forma elevar conscientemente os *stakes* dessa responsabilização.

Apesar de recente no Brasil, o uso das informações sobre o desempenho escolar para a definição de políticas de alocação de recursos já deve estar bastante difundido. Essa impressão deriva da existência em quase todos os estados visitados de exemplos de políticas de pequeno e de grande porte para a distribuição de materiais, treinamentos e outros recursos físicos e financeiros com base em critérios que levam em consideração o desempenho da escola ou da sua região no sistema de avaliação.

No caso de políticas compensatórias (também chamadas de discriminação positiva), definidas para favorecer escolas com desempenho menor, podem ser citados o projeto Aprender Mais, da Secretaria de Educação de Pernambuco (SEE-PE), que garante reforço escolar em Português e Matemática dado por 3.275 professores temporários para todos os estudantes de 4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio, um “kit boletim” e cursos para as escolas que não atingem as metas do Idepe; uma política veiculada pela Secretaria de Educação do Rio Grande do Sul (SEE-RS) em 2010 para a identificação de escolas com resultados “insuficientes” para receber apoio do poder público e replanejamento de sua gestão e ação pedagógica; uma distribuição de recursos e ações pela SEE-MG para favorecer as escolas da Região Grande Norte do estado, identificada como detentora dos piores resultados no Sistema Mineiro de Avaliação da Educação Pública (Simave); e outra política de Pernambuco chamada Certificação de Padrões Base. Nesse caso, trata-se de um processo de implantação e manutenção de um padrão físico de funcionamento das 221 escolas com o pior resultado no Idepe de 5ª a 8ª série. Com dinheiro do Banco Mundial e por meio da gerência geral de engenharia da própria Secretaria, propõe-se que as escolas atinjam um padrão de ventilação, iluminação, salas, carteiras, material etc. e recebam os recursos e o apoio para a manutenção desse padrão.

Há também políticas que favorecem escolas ou regiões com base na lógica contrária. Nesses casos, os beneficiários são escolas ou regiões com os melhores resultados. A Lei 14.023/07 do Ceará, que distribui a cota estadual do Imposto sobre Circulação de Mercadorias e Prestação de Serviços

(ICMS) para os municípios de acordo com os resultados do Spaece, seria um exemplo. Nesse caso, foi criado o Índice Municipal de Qualidade Educacional (IQE) com base nos resultados de Português e Matemática do 2º e do 5º ano do Ensino Fundamental e todo ano o governo calcula o repasse de ICMS aos municípios de acordo com os critérios estipulados no Decreto 29.306, de 5 de junho de 2008. Nessa legislação, em nome de “uma gestão pública por resultados”, se estabelece que 18% dos 25% da arrecadação do ICMS serão distribuídos “em função do Índice Municipal de Qualidade Educacional de cada município, formado pela taxa de aprovação dos alunos do 1º ao 5º ano do Ensino Fundamental e pela média obtida pelos alunos de 2º e 5º ano da rede municipal em avaliações de aprendizagem” (Art. 1º, Parágrafo Único, II).

Pelo componente de responsabilização inerente à distribuição de recursos com base no desempenho, as consequências dessa legislação têm sido a mobilização dos prefeitos e uma concorrência para melhorar os resultados nos municípios:

“Veja bem, o prefeito que nunca ligava para a educação hoje ele está louco aqui, ligando pra gente o tempo todo, por quê? Porque tem o ICMS atrelado ao resultado do 2º e 5º ano também.” (técnica/SEE-CE)

Outra consequência, menos interessante, tem sido a perda de receita da capital por causa das características da população, que confirma que a distribuição da receita do ICMS com base em um critério aparentemente meritocrático é um jogo de soma zero. Mesmo com as dificuldades provocadas pela perda de receita de alguns municípios, a política começa a fazer escola e em Pernambuco o governo já distribui 3% da receita do ICMS dos municípios com base em critérios parecidos.

Mais comum em se tratando da alocação de recursos de acordo com os resultados são os prêmios pagos diretamente às escolas em reconhecimento aos resultados alcançados nas avaliações estaduais. Um exemplo seria o Prêmio Escola de Valor, criado pelo governo do Amazonas, para “estimular a obtenção de melhores índices educacionais”. Anualmente são oferecidos prêmios de 30 mil reais a cada uma das escolas que obtiverem os melhores

índices no Ideb. O valor precisa ser empregado em benefício da comunidade escolar e administrado pelas Associações de Pais, Mestres e Comunidade (APMC). No primeiro ano de aplicação, 39 escolas foram premiadas e, no segundo ano, 180. Segundo o site da SE, o prêmio já ajudou a modificar a realidade dessas escolas. Conforme as regras, foram aplicados 40% para a melhoria da estrutura física, 40% para a compra de materiais pedagógicos e equipamentos tecnológicos e 20% para a realização de eventos educacionais.

Existem exemplos de políticas que fazem as duas coisas simultaneamente. Mais complexas, e potencialmente mais eficazes, essas políticas partem da premissa de que, além de merecerem seus prêmios, as escolas mais bem colocadas nas avaliações também têm alguma coisa a oferecer às escolas com o desempenho mais fraco. O que essas políticas promovem, portanto, é uma relação de assistência técnica entre as escolas premiadas e as escolas adotadas. Essa combinação dá um novo significado à responsabilização. Além de serem responsabilizadas pelos seus próprios resultados, as escolas premiadas também assumem uma nova responsabilidade para com as escolas adotadas ou apoiadas.

O município de Rio de Janeiro abraçou uma política desse gênero ao definir em 2010 que as 25 melhores escolas de cada ano da Prova Rio, avaliação externa daquela cidade, vão apadrinhar as 25 piores escolas, “dando orientações e mostrando algumas estratégias que deram certo em suas unidades”. No entanto, a versão mais completa dessa política de apadrinhamento foi lançada no Ceará em 2009 pela Lei 14.371, de 19 de junho 2009, e se chama Prêmio Escola Nota 10. Nesse caso, os critérios para a seleção das escolas premiadas são de que tenham média no IDE-Alfa entre 8,5 e 10, percentual de participação nas provas de Spaece de 85% e pelo menos 20 alunos de Ensino Fundamental na época da avaliação. As escolas que atendem a essas condições têm o direito de receber R\$ 2.500,00 por aluno matriculado no 2º ano do Ensino Fundamental. A contrapartida é a adoção de uma das escolas que obtiveram os menores resultados no IDE-Alfa. A partir da sua adoção, essas escolas recebem uma contribuição financeira do Estado no valor de R\$ 1.250,00 por aluno avaliado e ações de cooperação técnico-pedagógica, vindas das escolas premiadas.

A premiação recebida pela Escola Nota 10 é dividida em duas parcelas. Para receber a segunda, a escola precisa manter sua própria média na

avaliação seguinte e a escola assistida mostrar que houve avanço no desempenho ao ponto de alcançar a nota 5 no IDE-Alfa. Esse critério é para garantir que haja de fato uma relação de apoio entre as duas escolas. Outra questão de importância diz respeito ao processo de escolha das escolas adotadas e aos usos do dinheiro. Na opinião de um gestor, é importante obrigar a escola premiada a se relacionar com escolas distantes e não se acomodar com escolas mais próximas. Na opinião de um técnico, a política melhoraria se a premiação fosse menor, porém estendida a um número maior de escolas.

Outra questão é a confiabilidade das avaliações para operar um sistema de premiação com base em uma única medida. Com informações fornecidas por um técnico, descobriu-se que muitas escolas podem mudar de categoria entre adotadas e premiadas de um ano para outro, o que levanta a suspeita de que a categorização deixa a desejar ou que a margem de erro é grande. Uma escola que passa das piores para as melhores em período tão curto desafia a própria lógica dos sistemas de apoio.

Mesmo não tendo sido criado como uma política de incentivo ou premiação, o Projeto Escolas-Referência de Minas Gerais se situa nessa categoria do sistema classificatório por promover investimentos de diversos tipos nas escolas beneficiadas e também por obrigar que elas adotem outras escolas de desempenho mais fraco. Um dos programas estruturantes da política educacional de Minas Gerais, o Projeto Escolas-Referência busca a reconstrução da excelência na rede pública e a superação do fracasso escolar por meio de uma educação de qualidade (PEREIRA, 2009). Iniciado em 2004, com a seleção de 223 escolas pela sua tradição, número de alunos nos ensinos Fundamental e Médio e resultados positivos no Simave, o projeto visa transformar as escolas escolhidas em focos irradiadores de educação de qualidade no estado.

Como comentário final, convém destacar a diferença entre as políticas de alocação de recursos com base em resultados de acordo com o nível de impacto das consequências ou responsabilização. A distribuição de recursos em favor de escolas com desempenho menor não pode ser considerada política de responsabilização por não existir nenhum propósito de estimular as escolas a reduzir seu desempenho para se tornarem escolas

beneficiadas. No caso das políticas de premiação, por outro lado, em que as escolas com melhor desempenho são aquinhoadas com recursos até vultosos, como no caso do Ceará, a mensagem é claramente de competição e de incentivo a todas as escolas para que melhorem seus resultados. A não-premiação da escola acarreta prejuízos pela perda de melhores condições de trabalho e pelo não-reconhecimento público da sua qualidade. Por colocar os resultados nas mãos das escolas e dos seus profissionais, a política é de responsabilização, mesmo sendo uma premiação para a escola, e não para os profissionais. Na seção seguinte, serão tratados os incentivos dirigidos aos profissionais das escolas.

4.5. Para políticas de incentivos salariais

A bibliografia brasileira sobre o tema da responsabilização baseada no pagamento de bônus ou prêmios aos membros da equipe escolar não é extensa, mas permite traçar o histórico dessa política desde o princípio da década passada. Os exemplos mais bem documentados da primeira geração de políticas dessa natureza são os casos da Nova Escola, do estado do Rio de Janeiro, o sistema de bônus salarial do estado de São Paulo, o prêmio Escola do Novo Milênio, do estado do Ceará, e o bônus salarial para professores alfabetizadores do município de Sobral, todos criados no ano de 2001 (MEC/INEP, 2005; BROOKE, 2006, 2008; RODRIGUES, 2007; ANDRADE, 2008). O que essas políticas tiveram em comum foi a formulação de critérios para o pagamento coletivo de incentivos salariais aos profissionais das escolas com base nas médias de desempenho dos alunos do ano anterior.

A segunda geração de políticas de responsabilização tenta superar as críticas mais comuns. Criados nos últimos dois ou três anos e exemplificados pelos sistemas de incentivos coletivos de Pernambuco, Minas Gerais, São Paulo e Espírito Santo, esses novos sistemas incorporam procedimentos mais sofisticados para a definição de critérios e a avaliação de escolas. É o que veremos com a análise dessa política em quatro dos estados investigados.

4.5.1. Investigações em loco

Quatro dos estados escolhidos para a pesquisa de campo têm políticas de bonificação regidas pelos seus sistemas de avaliação – São Paulo, Minas Gerais, Pernambuco e Ceará. Mais um estado, Espírito Santo, e o Distrito Federal já prepararam as metodologias para o cálculo do bônus, mas ainda não fizeram o primeiro pagamento⁷. Essa demora se deve à decisão de deixar passar o período eleitoral de 2010 e aguardar as orientações dos novos secretários de Educação. As incertezas desse período, que se estendeu pelo primeiro semestre de 2011, dificultou a pesquisa de campo e tornou necessário restringir as entrevistas à discussão das políticas dos governos cujos mandatos tinham terminado. As mudanças já efetuadas ou ainda em curso nas políticas educacionais de diversos estados com a chegada dos novos governantes levam a crer que algumas das observações dos pesquisadores descritas a seguir possam estar já desatualizadas.

As observações não repetirão os detalhes sobre as metodologias dos sistemas de bonificação que possam ser encontradas na bibliografia sobre esse tema. Elas se restringirão aos detalhes do funcionamento desses sistemas não disponíveis na bibliografia e àqueles aspectos que esclarecem melhor os princípios e problemas dessa modalidade de política de gestão.

4.5.1.1. Ceará

No caso do Ceará, cabe esclarecer que o prêmio Escola Nota 10, já apresentado na categoria *alocação de recursos*, também será estudado aqui pelo fato de parte do prêmio ser destinada aos professores. Conforme o Manual de Orientações para Elaboração, Execução e Prestação de Contas do Plano de Aplicação dos Recursos Financeiros do Prêmio, documento de 63 páginas, dos 75% do valor total do prêmio recebidos na primeira parcela, as escolas podem destinar até 20% para “bonificar os professores do 2º ano do Ensino Fundamental; profissionais do núcleo gestor e demais profissionais lotados

7. No caso de Espírito Santo, o bônus foi regulamentado em 31 de maio de 2011 e o primeiro pagamento, referente ao ano de 2010, estava programado para julho/agosto de 2011.

na escola, no ano de referência da premiação”. Do resto do prêmio, até 70% é destinado à aquisição de materiais de apoio pedagógico, material permanente, formação continuada e adequações dos ambientes físicos da escola e “outras ações que contribuam para a melhoria da qualidade do ensino e da aprendizagem dos alunos”, e até 20% destinados a deslocamento, alimentação e hospedagem para a equipe escolar com o objetivo de desenvolver ações de cooperação técnico-pedagógicas junto à escola apoiada. Dentre os limites fixados pela legislação, a escola fica livre para seguir seus próprios critérios, até na forma de calcular e fazer a distribuição da bonificação.

A SEE-CE também patrocina o prêmio Aprender pra Valer, voltado para escolas de Ensino Médio, conforme a Lei 14.484, de 2008. Nesse caso, são premiados os professores de até 45 escolas que atingem os seguintes índices na soma das médias de Língua Portuguesa e Matemática em cada uma das três séries do Ensino Médio:

- Evolução de 7 a 10% sobre a média de proficiência obtida no ano anterior.
- Não apresentar média de proficiência no padrão Muito Crítico.
- Percentual de participação mínima de 80% (no Spaece) com base na matrícula inicial informada no Educacenso.

O prêmio, destinado aos gestores, professores e servidores efetivos e temporários, é concedido no valor mensal da remuneração de cada servidor, proporcional ao acréscimo na média de proficiência dos alunos. Ou seja, no limite inferior de 7% de crescimento na média da escola, os professores ganham uma bonificação de 70% do salário. No limite superior de 10% ou mais, a bonificação é de 100%. Em 2010, a premiação representou um gasto de cerca de R\$ 2 milhões e, como no caso da Escola Nota 10, houve solenidade para a entrega às escolas de placas de homenagens pelo desempenho alcançado. A ideia de trabalhar várias políticas de incentivos simultaneamente não é por acaso:

“No sistema de motivação precisa perceber é o seguinte, que você precisa de mais de um modelo. Você não pode trabalhar só com uma

linha de força, você tem que ter várias linhas de força coerentes e convergentes. Então com Escola Nota 10 eu motivo os professores e diretores, com o ICMS eu motivo o prefeito, e com o Aprender para Valer, eu motivo diretores, gestores, professores, funcionários, e aí, eu motivo independente do contrato de trabalho, se é cargo efetivo ou se é temporário ou se é terceirizado, eu motivo todos. Nenhuma pessoa da escola pode ficar fora do prêmio.” (gestor/SEE-CE)

A convicção do estado do Ceará na eficácia dos prêmios como incentivo para os professores tem uma década de história que, para as escolas de ensino regular, começa com a criação do prêmio Escola do Novo Milênio pela Lei 13.203, de fevereiro 2002⁸. Alterado em 2004, quando o estado institui o Programa de Modernização e Melhoria da Educação Básica e cria o Selo Escola Destaque do Ano, o prêmio continua ininterrupto até 2007. Nesse ano, o atual governador, Cid Gomes, toma posse. Prefeito de Sobral de 1997 a 2005, quando aquele município implantou o primeiro exemplo de um sistema de bonificação individual para professores alfabetizadores, o novo governador dá continuidade às políticas de incentivos ao criar o Índice Municipal de Qualidade Educacional como base para a distribuição dos 25% do ICMS pertencentes aos municípios e estabelece as bases para a retomada da premiação de escolas em 2008 por meio da Escola Nota 10. No ano seguinte, são criados o Prêmio Aprender pra Valer e também o Prêmio para os Alunos do Ensino Médio, também pioneiro, que distribui um laptop a todos os alunos que conseguem média de 325 em Língua Portuguesa e 350 em Matemática na escala do Spaece⁹. A própria continuidade

-
8. Em 2001, o projeto chamado Tempo de Avançar, que utilizou o telecurso 2000 da Fundação Roberto Marinho para implantar 2 mil turmas de alfabetização com alunos fora de idade escolar, foi a primeira experiência com a política de premiação. De acordo com os resultados na Prova de Êxito, ao final do curso foram premiados os professores dos alunos com o melhor desempenho.
 9. Política similar instituída na cidade do Rio de Janeiro para “homenagear e valorizar os alunos que obtiveram as melhores notas em Português e Matemática” vai premiar os cinco melhores alunos de cada disciplina dos dois anos avaliados com uma bicicleta.

da política de incentivos sugere que os gestores locais estão convictos da sua eficácia.

Nas entrevistas, foi constatado que a opinião muito positiva dos gestores cearenses não vem só do efeito cumulativo da experiência mas também de um raciocínio que deposita fé no efeito real do incentivo tanto na melhoria da competência como na motivação do professor. Para um dos gestores entrevistados, a questão de se haveria outras formas de abordar os problemas da educação ou de organizar a própria política de incentivos não é uma preocupação. A política de premiação é vista como uma opção pragmática que chega ao cerne da preocupação pelos resultados da escola e que qualquer política nesse contexto é melhor do que nenhuma.

Isso não significa que a política de incentivos salariais não oferece problemas na hora da aplicação prática das regras. No caso da Escola Nota 10, por exemplo, foi necessário criar o IDE-Alfa, que leva em consideração a proporção de alunos matriculados que efetivamente participam do Spaece porque as escolas nem sempre aplicavam as provas aos alunos mais fracos para não prejudicar suas médias. Além desse tipo de *gaming*, em que as escolas aprendem a manipular as regras, também se observa a tendência das escolas de se preocuparem menos com as séries que não são levadas em consideração no cálculo do IDE-Alfa.

Outros problemas incluem a “transferência branca” de alunos para outras escolas para tentar se livrar de alunos de desempenho mais baixo e a insatisfação criada pela pressão irreal do prêmio Aprender pra Valer, que estipula um crescimento de no mínimo 7% ao ano na média de desempenho da escola, independentemente do patamar já alcançado.

Na opinião dos técnicos da Secretaria, há indicações de que o programa Escola Nota 10 produz resultados positivos no conjunto das escolas atingidas. Em comparação com 2009, as médias tanto das escolas premiadas como das apoiadas cresceram em 2010 e, das 148 apoiadas, 107 chegaram a receber a segunda parcela do apoio financeiro em 2010, mostrando que suas médias já alcançaram 5 no IDE-Alfa.

No entanto, o efeito parece não ser pelo recebimento do dinheiro. Uma evidência nesse sentido é o fato de as escolas apoiadas, que recebem menos, terem avançado mais que as escolas premiadas. Outra evidência é fornecida

pelo fato de os recursos terem demorado a chegar às escolas e o avanço no desempenho ter acontecido mesmo na ausência dos recursos. A explicação alternativa tem a ver com o efeito psicológico da premiação tanto para as premiadas como para as apoiadas:

“O efeito da avaliação e a divulgação do resultado por si só já é uma dinâmica de motivação. A escola saber que ela está entre as melhores. Aí eu fico perguntado: por que as apoiadas evoluíram mais do que as premiadas? As premiadas não conseguiram manter a média de desempenho. Então talvez é o efeito psicológico do prêmio, do fato de ter o prêmio e ter a publicização do resultado e a escola ser conhecida que está numa situação crítica e tudo mais. Talvez isso tenha gerado um movimento interno na escola de tentar rever o seu projeto e tentar rever suas práticas. Talvez é muito mais do que o efeito do dinheiro que chegou a essa escola ou mesmo o efeito da cooperação técnica da apadrinhada que chegou a essa escola.” (técnico/SEE-CE)

Mesmo assim, nem todas as escolas premiadas conseguem manter suas médias para justificar o recebimento da segunda parcela da premiação, equivalente a 25% do total. Das 148 premiadas em 2009, 75 não receberam a segunda parcela por não manterem o IDE-Alfa da escola no mesmo nível e mais 12 pelo fato de suas escolas apoiadas não alcançarem o índice de 5 nessa mesma escala.

No caso do Ensino Médio, houve também uma melhoria expressiva nos resultados. Mas nesse caso o efeito pode ter sido pelo novo entusiasmo dos alunos, provocado pelo Prêmio para os Alunos do Ensino Médio, e refletir, portanto, a melhoria causada pela medição mais exata da proficiência dos alunos.

A taxa de participação dos alunos matriculados, criada para contornar o problema da “transferência branca” e a falta de comparecimento dos alunos no dia das provas, tem criado seus próprios problemas. Um deles tem sido a não-premiação de escolas que por motivos alheios têm perdido alunos por abandono ou transferência. Nesses casos, escolas com desempenho inferior podem chegar a receber a premiação por não ter sofrido a mesma perda de alunos. Há relatos de funcionários das escolas viajando até os novos domi-

cílios de alunos transferidos para receber a documentação oficial da transferência e de um diretor que teve que falar com o conselho tutelar do Rio de Janeiro para garantir que o aluno transferido tivesse sido de fato matriculado em escola daquela cidade. Um técnico da SEE/CE sugere que uma forma de contornar o impacto da perda inevitável de alunos seria estipular um percentual mínimo de participação e fazer a distribuição dos prêmios entre os que satisfazem esse mínimo.

No fundo, essa dificuldade diz respeito à aparente “injustiça” dos critérios estipulados para o pagamento do incentivo. Em se tratando de um acréscimo importante à renda do professor, equivalente a um mês de salário, não é trivial a questão de se as regras da competição são percebidas como legítimas e justas. A experiência de Sobral, e os ajustes sucessivos nas regras do bônus do professor alfabetizador para o aperfeiçoamento da política de incentivos mostra o quanto os gestores precisam fazer as adaptações de acordo com as definições locais de justiça. Numa primeira versão dessas regras, as cinco primeiras escolas na prova de alfabetização levaram o prêmio, independentemente do tamanho da escola e do número de turmas. No ano seguinte, a competição foi subdividida em categorias pelo número de turmas das escolas em reconhecimento à dificuldade maior em melhorar os resultados em escolas maiores. No terceiro ano, houve nova mudança pelo fato de algumas escolas não receberem prêmios mesmo tendo desempenho melhor que escolas premiadas em outros grupos por só existir um prêmio para cada grupo. Neste ano, foi eliminado o ranqueamento e criado no seu lugar um sistema de metas com base no desempenho da própria escola no ano anterior. Nesse formato, a escola competia consigo mesma para alcançar pelo menos 90% da meta. A lição aprendida é que o sistema de incentivo nunca será perfeito.

4.5.1.2. Espírito Santo

O sistema de bonificação do Espírito Santo, que prevê pagamento a toda a equipe da escola que atingir determinado índice de desenvolvimento da escola, é de elaboração recente, o que permite o resgate mais pormenorizado dos diferentes objetivos perseguidos pelos gestores. Mesmo não havendo ainda uma aplicação prática da metodologia, o raciocínio por trás da sua

criação ainda está fresco na memória dos envolvidos. O primeiro objetivo é levar para dentro da escola a mesma preocupação com os resultados que aquela expressa no planejamento estratégico do estado.

Uma vez aceita a relação entre qualidade e resultados e estabelecida a preocupação da escola em melhorar seus resultados, parece que haverá naturalmente a compreensão de por que a escola ganhou ou deixou de ganhar o bônus.

“O que a gente queria com o sistema de bonificação é levar essa discussão da qualidade para dentro da escola. Ou seja, o Paebes, que é o nosso sistema de avaliação, passa a ganhar uma nova cara porque hoje ele não vai para a escola simplesmente como uma avaliação. [...] Com a bonificação ele coloca uma coisa nova, então você cria um estímulo para que as pessoas pelo menos passem a entender por que é que eu ganhei mil e você ganhou 2 mil, por que a escola A é assim e a B é assado”. (gestora/SEE-ES)

O que não fica explicitado nesse raciocínio é como o professor pode de fato entender as diferenças entre as escolas e, com base nessa compreensão, alterar de alguma forma a sua prática. As informações que circularão entre as escolas podem explicitar as diferenças nos índices, mas dificilmente identificarão porque uma escola atingiu determinado patamar e outra não.

O segundo objetivo diz respeito ao propósito de incentivar os professores a se verem como responsáveis pelos resultados da escola. Esse raciocínio parece completo, já que o pagamento de uma bonificação aos professores das escolas com os melhores resultados de fato sinaliza a crença de que esses professores têm pelo menos uma parte da responsabilidade por esses resultados, sobretudo na presença de uma metodologia de cálculo que leva em consideração o nível socioeconômico dos alunos. O que não fica garantido é que os professores sintam as mesmas convicções a respeito da sua contribuição ou que entendam o que precisam fazer para ganhar o prêmio no futuro.

O terceiro objetivo é retribuir o esforço dos professores. Como o bônus é classificado como prêmio, e não como remuneração, ele se justifica como um reconhecimento da dedicação adicional que os professores mostraram

ao produzirem resultados acima da média. O incentivo esperado é no sentido de suscitar níveis de esforço maiores no futuro, mesmo sabendo que o nível de esforço pode variar entre os professores de uma mesma escola e que o incentivo pode não afetar a todos igualmente.

O quarto objetivo, ao condicionar o pagamento do bônus ao cumprimento de um número mínimo de dias de trabalho, é coibir as faltas dos professores, tal como na política de bonificação do estado de São Paulo. Como será discutido, há evidência de que o efeito maior da bonificação é justamente no sentido de promover a assiduidade.

4.5.1.3. Minas Gerais

O sistema de premiação de Minas Gerais merece uma análise mais detalhada por seguir um modelo diferente dos outros estados. Chamado Acordo de Resultados, a metodologia implantada em 2008, e repetida anualmente desde então, se estende a todas as Secretarias do governo. No caso da SE, o acordo diz respeito aos resultados da Secretaria, das SREs e de todas as escolas. No caso das escolas, a nota recebida é composta em 65% da nota da Secretaria, 10% da nota da SRE e 25% da avaliação dos “indicadores finalísticos” da escola. Por meio dessa composição, mostra-se claramente que a responsabilidade pelos resultados finais da escola é compartilhada entre as três instâncias. Essa característica do Acordo de Resultados a diferencia das demais políticas de incentivos estudadas que tendem a identificar os professores, os dirigentes e demais membros das equipes escolares como os únicos integrantes do sistema de ensino responsáveis pelos resultados de aprendizagem, apesar da conhecida importância das condições materiais fornecidas pelo órgão central e os diversos serviços de apoio das instâncias regionais.

Os indicadores finalísticos da escola dizem respeito a um conjunto de dez metas, expressas tipicamente em termos da proficiência média dos alunos por ano e matéria. Como exemplo, a sexta meta, com peso valendo 2,14% do total da nota da escola, diz respeito à proficiência média dos alunos do 9º ano do Ensino Fundamental em Português. A escola ganha a nota máxima se o valor observado para a meta for maior ou igual ao Nível Recomendável para o ano, de acordo com a classificação publicada pelo sistema de avalia-

ção do estado. No caso de a meta da escola ser menor que o Nível Recomendável, precisa-se empregar uma equação para calcular o seu desempenho de 0 a 1 com base na diferença entre a proficiência alcançada no presente ano com a proficiência do ano anterior dividido pela diferença entre a meta e a proficiência do ano anterior. No final, a nota do indicador = Desempenho \times Peso do Indicador \times 10. Quando somadas todas as notas dos indicadores, mais as notas da Secretaria e da SRE a que a escola pertence, a nota final da escola mostrará a proporção da meta global alcançada e, portanto, o percentual do prêmio de produtividade a que terá direito.

Essa matemática é necessária devido à decisão de beneficiar a escola com alguma recompensa mesmo no caso de não alcançar 100% da meta acordada. Essa flexibilização torna o processo mais brando, permitindo a diluição do significado das metas ao ponto de elas se tornarem indicadores de resultados, o que permite levar em consideração a realidade das escolas que nem sempre conseguem melhorar no ritmo desejado, apesar dos esforços. O raciocínio da Secretaria que levou a essa flexibilização é mostrado na seguinte fala:

“Meta tem que ser flexível por que meta rígida é meta boba. Meta é uma coisa importante? É importante, mas ela não foi feita para ser sustentada a ferro e fogo; de qualquer jeito você tem que ganhar. Se for a ferro e fogo o cara toma anabolizante para não sei o que lá nas olimpíadas, se for a ferro e fogo as pessoas usam tudo quanto é tipo de artifício para chegar lá. Passa menino sem saber, altera nota do menino, ensina o menino a fazer prova. Acontece esse tipo de coisa por quê? Porque se eu não conseguir, eu vou perder salário e ninguém está em condições de perder salário.” (gestor/SEE-MG)

O aspecto mais interessante do Acordo de Resultados é o procedimento para a fixação das metas. Esse processo culmina na assinatura de um termo de compromisso por parte das SREs e das escolas após a consulta ao sistema online, em que as metas estão listadas escola por escola. No caso de a escola não concordar, existe um procedimento para a sugestão de nova meta e para o julgamento do pleito da escola. Ou seja, mesmo sendo um processo centralizado, a possibilidade de a escola fazer uma contraproposta torna mais

legítimas as metas fixadas e justifica a assinatura do termo de compromisso que pactua os resultados esperados.

Outro aspecto relevante, que remete à discussão sobre o impacto do modelo de planejamento estratégico assumido por diversos governos estaduais, diz respeito à relação entre as metas dos planos governamentais e as metas colocadas para as escolas. No caso de Minas Gerais, as metas do PMDI 2003-2023 foram desdobradas em metas plurianuais por meio dos Planos Plurianuais de Ação Governamental (PPAG). No caso do PPAG 2008-2011, coube à Secretaria planejar suas metas de acordo com o percentual de crescimento fixado pelo estado.

No entanto, nem sempre os planejadores ouvem os gestores antes de fixar as metas globais do estado. Na primeira edição do Acordo de Resultados, os planejadores dividiram as escolas em cinco faixas de acordo com as proficiências do ano de 2007 e apesar de esse ano mostrar uma evolução acima da média, usaram a taxa de crescimento para fixar as metas até 2010.

Não é trivial a questão de como estabelecer as metas de um sistema de incentivos quando não se quer usar um ranqueamento competitivo nem fixar as metas exclusivamente com base no histórico do crescimento anual da proficiência. O que está em jogo é uma combinação de fatores que também leva em consideração a vontade dos governantes de ir além da meta.

O sistema de acordo de resultados tem vários méritos metodológicos. Também tem uma deficiência criada pelo uso de médias para a fixação de metas e o cálculo do desempenho da escola. Pelo fato de a escola poder melhorar suas médias com mais facilidade mediante atenção especial aos alunos mais capazes, existe o risco de a escola esquecer-se dos alunos mais atrasados. Entretanto, a Secretaria já percebeu esse problema e pensa resolvê-lo ao criar requisitos para melhorar a distribuição dos alunos entre as faixas de desempenho, de forma análoga à de São Paulo e Espírito Santo:

“A gente acha que a proficiência às vezes engana um pouco. Porque às vezes a escola cresce na proficiência média, mas reduz o número de alunos no nível recomendado, o que não é desejável. Então talvez a gente tenha que alocar esses dois indicadores.” (técnico/SEE-MG)

Outra possível deficiência diz respeito à dificuldade de a escola entender o cálculo do percentual do bônus a que tem direito e a dificuldade de aceitar as diferenças nas metas de cada escola:

“O acordo avalia a escola em relação a ela mesma. Então teve um problema que no primeiro acordo foi muito claro. A escola que tinha uma proficiência maior e outra que tinha menor em municípios iguais. Aquela que tinha proficiência menor tinha um prêmio de produtividade maior por que ela cresceu mais que a outra escola [...] Então assim ela não entendia e a gente teve que explicar: olha, o acordo com para você com você mesma.” (técnico/SEE-MG)

O fato de os indicadores estarem relacionados às matérias avaliadas pelo Simave no 3º, 5º e 9º anos do Ensino Fundamental e 3º ano do Ensino Médio também traz consequências para a maneira de pensar dos professores de outras matérias e anos. Mesmo pago a todos os professores da escola, a capacidade de o bônus incentivar os professores não incluídos nos indicadores deve ser menor, o que abre a possibilidade do *free riding*, encontrado na bibliografia:

“Muitos funcionários e outros professores de outras matérias não se sentem parte do acordo de resultados, eles acham que não vão contribuir com nada, eles têm esse sentimento. Então assim, eles não se sentem parte desse pacto. Eles recebem o prêmio e “que bom, mais um prêmio”, mas eles não sentem que o seu trabalho foi reconhecido, que a nota que está ali traduz o trabalho que eles desenvolveram durante aquele ano. Muitos. E eu acho que eles têm completa razão.” (técnico/SEE-MG)

Apesar das deficiências, tanto os técnicos como os gestores da Secretaria atribuem uma parte do avanço nos resultados das escolas durante os últimos anos à incorporação de um regime de metas e incentivos.

4.5.1.4. Pernambuco

O sistema de incentivo salarial introduzido pela SEE-PE em 2009 era pensado como um dos elementos de um conjunto de ações voltados para a revisão da situação funcional e salarial do professor. Entre essas ações, figurava a reformulação do contrato de trabalho e do plano de carreira de modo que o professor pudesse se dedicar à docência em uma única escola, em vez de acumular até 70 horas aula/semana em diferentes escolas e redes, como permitido pela legislação. Portanto, o incentivo era visto como um complemento para um salário que já resolveria as necessidades do professor. Além disso, a visão da Secretaria de Pernambuco atrelava o aumento salarial da nova carreira a um esquema de supervisão bem mais rigoroso. O incentivo salarial, pago com base nos resultados da escola, seria então um dos elementos desse esquema de supervisão.

Para enfatizar a política de fixar os professores em uma única escola, o bônus salarial é pago aos professores que estiveram na escola por um período superior a seis meses. De forma parecida à de Minas Gerais, o valor do bônus é calculado com base no percentual de cumprimento da meta fixada para a escola, a partir de 50%. As metas são determinadas para cada escola com base na faixa de desempenho em que se encontra – baixo, médio ou alto – de acordo com os resultados do Saepe do ano anterior.

O que também chama a atenção do sistema de Pernambuco é a maneira de estipular o valor do bônus pago aos professores. Comum em outros sistemas é o pagamento a uma proporção fixa de escolas ou, como no Chile, às escolas que atendem a uma proporção fixa de alunos, no caso 25%. Em Pernambuco, é criado todo ano um fundo representando a folha de pagamento de um mês, o que permitiria pagar um bônus equivalente a um salário a todos os professores caso todas as escolas atingissem 100% da meta. Em 2008, 51,8% das escolas tiveram acesso ao bônus, o que permitiu pagar 2,4 vezes o salário-base dos beneficiados no caso de escolas que atingiram 100% da meta. Em 2009, a proporção de escolas atingindo pelo menos 50% da meta subiu para 82%. Com isso, o fator usado no pagamento do bônus caiu para 1,6 salário nas escolas que alcançaram 100% da meta.

Nesse mesmo ano, a Secretaria estendeu o pagamento do bônus às gerências regionais com base na média ponderada da taxa de cumprimento das metas das suas escolas, sinalizando a responsabilidade também atribuída aos órgãos de apoio e supervisão, como em Minas Gerais. No próximo passo, pretende-se criar uma sistemática para o pagamento de bônus aos funcionários do órgão central.

No entanto, o fato de as políticas referentes à nova carreira dos professores não terem sido implantadas e os salários não terem sido reformulados reduz bastante a credibilidade da política de incentivos salariais de Pernambuco. Sem essas mudanças, colocadas como pré-requisitos pela própria Secretaria, será mais penoso cobrar as atividades esperadas pelo sistema de supervisão e mais difícil concretizar os avanços esperados do sistema de incentivos.

4.5.1.5. São Paulo

Sendo o sistema estadual de bonificação mais contestado por integrantes do magistério e o que mais aparece no noticiário, a política implantada pela Secretaria Estadual de Educação de São Paulo (SEE-SP) merece uma discussão mais detalhada. Foram analisadas em outros lugares as questões técnicas sobre a fórmula usada para avaliar a qualidade e a equidade da escola (SOARES, 2009). Nesta seção, precisa-se investigar o status da Bonificação por Desempenho como o mais problemático dos diversos exemplos em estudo das novas políticas de incentivos salariais.

O primeiro problema da Bonificação por Desempenho, lançada em maio de 2008, é que ela tem antecedentes. Entre 2001 e 2006, a Secretaria já pagava um bônus anual aos professores e diretores. Entre os anos de 2001 e 2003, esse bônus foi calculado de acordo com a assiduidade média da equipe, o desempenho global da escola, medido durante o ano escolar anterior através do Saesp, e outros indicadores de qualidade, principalmente o nível de retenção de alunos pelas escolas. Em 2004, os critérios foram modificados para, em primeiro lugar, incluir o mínimo de 200 dias de trabalho efetivo e o desenvolvimento de projetos envolvendo a comunidade da escola, e, em segundo lugar, para excluir o uso dos resultados de desempenho do Saesp.

Em 2006, o Saresp foi suspenso e os resultados do desempenho dos alunos novamente retirados do cálculo do bônus. Isso se deveu, em grande medida, à aplicação no ano anterior da nova Prova Brasil.

A resistência aos bônus anuais pagos pela Secretaria foi contínua. A oposição mais vociferante, do Sindicato dos Professores do Ensino Oficial do Estado de São Paulo, expressou uma rejeição à “transferência abusiva da lógica de mercado” para a educação (CAVALIERE, 2003, p. 30) e à competição entre as escolas e os diretores das escolas, que seria fomentada por essa política (CDEP, 2006). Outras críticas tiveram mais a ver com problemas de trapanças, havendo meios para os professores ajudarem os alunos a preencher as folhas de respostas ou alterarem as respostas dos alunos, uma vez que eram os professores que aplicavam e corrigiam as provas. Supostamente esses casos eram “bem conhecidos” entre os professores das escolas estaduais (DISSIDÊNCIA, 2006), bem como os problemas burocráticos associados ao cálculo e ao pagamento individual dos professores. Outro argumento, aparentemente bem justificado, dizia respeito ao uso de resultados de desempenho médio dos alunos sem nenhuma tentativa de controlar por diferenças socioeconômicas entre as escolas ou pelo impacto de diferenças no nível anterior de aprendizagem.

Com a eliminação dos resultados dos alunos, o programa de bônus se tornou mais uma maneira de justificar um benefício natalino. Porém o uso contínuo de medidas de assiduidade e até de taxas mínimas de comparecimento em cursos de treinamento para os professores, tal como incluído entre os critérios de 2005, é uma indicação do problema perene de absentismo e da vontade da Secretaria de cobrar a frequência dos professores mesmo na ausência de um indicador de qualidade da escola.

Com esses antecedentes, é difícil saber se a resistência atual de uma parte dos professores se deve mais às cobranças de frequência, ao sistema de avaliação em si ou aos efeitos negativos provocados pela associação *high-stakes* entre os testes e o pagamento da bonificação. Há evidências de que ainda há restrições ao Saresp, no qual a bonificação está fundamentada. Sem uma atitude tranquila em relação à legitimidade e confiabilidade do próprio sistema de avaliação, dificilmente será recebida de braços abertos uma política alicerçada nesse instrumento. A forte politização do tema da avaliação, classificada

como produto das reformas neoliberais de recorte gerencialista, e a falta de simpatia do principal sindicato de professores para a política de avaliação do estado, impulsionado e consolidado por um governo do Partido da Social Democracia Brasileira (PSDB), ajudaram a criar um ambiente ainda um tanto hostil ao próprio tema da avaliação.

No entanto, os argumentos publicados pelo Sindicato dos Professores do Ensino Oficial do Estado de São Paulo (Apeoesp) para sua oposição atual à política de bonificação não incluem nenhuma restrição explícita ao sistema de avaliação. Os argumentos empregados dizem respeito, primeiro, à falta de “uma verdadeira política salarial, que assegure reajustes e aumentos reais para todos os professores, da ativa e aposentados”; segundo, à falta de evidência de que o sistema de bonificação possa elevar a qualidade do ensino ou o nível de satisfação dos professores; terceiro, à falta de inclusão de outros fatores, além dos resultados dos alunos; e quarto, ao valor irrisório do bônus pago (APEOESP, 2011). Surpreendentemente, as últimas duas críticas apontam para a necessidade de melhorias na política, e não para sua abolição.

Ao mesmo tempo, o sindicato sinaliza que pretende contestar na Justiça a Lei Complementar 1.078/2008 por não ter considerado todas as justificativas de faltas previstas no Estatuto do Servidor Público do Estado na hora de calcular a bonificação de cada servidor. A reivindicação da Apeoesp é de que a Secretaria deixe de descontar as faltas que têm justificativa legal. A posição da Secretaria é de que a bonificação não é salário, mas uma forma de premiar os professores e outros servidores que mostrem dedicação ao deixar de aproveitar todas as faltas que a lei faculta.

O argumento do sindicato de que o sistema de bonificação é incapaz de elevar a qualidade do ensino deriva de uma leitura sobre a experiência de outros países. Dependendo da bibliografia consultada, essa leitura fornece uma série de argumentos sobre os efeitos colaterais dos testes *high-stakes* usados por políticas de *accountability*. Recém-sumarizados pelo Movimento Contra Testes de Alto Impacto em Educação, criado no âmbito da Associação Nacional de Pós-graduação e Pesquisa em Educação (Anped), em 2010, os efeitos colaterais dos testes de alto impacto (*high-stakes*) incluem a capacidade de falsear os dados das avaliações, induzir à desonestidade, substituir os esforços de ensino pelo de preparação para testes, estreitar o

currículo e desestruturar os profissionais do magistério, além de provocar extrema ansiedade em alunos, pais, professores e administradores (MOVIMENTO, 2010).

Além dos possíveis efeitos colaterais negativos, o sindicato pode também apontar a falta de evidência de um impacto positivo na aprendizagem e até de evidência pelo contrário. O incentivo pago em 2011 causou constrangimento ao revelar um recuo no total pago e a queda no número de escolas aptas para receber o bônus, em comparação com o ano anterior. O valor pago a 190 mil funcionários, de 3.778 escolas – 75% das 5.019 unidades estaduais –, foi de R\$ 340 milhões. Em 2010, foram concedidos R\$ 655 milhões. O número de escolas que não receberam o bônus triplicou. Em 2010, 510 escolas (9,9% do total) ficaram sem a bonificação, enquanto em 2011 foram 1.474, o que representa 29,1% da rede pública do estado. As causas principais foram as variações negativas do Saresp entre 2009 e 2010 nos anos finais do Ensino Fundamental e no Ensino Médio. Com isso, só houve aumento no valor do Idesp nos anos iniciais do Ensino Fundamental, que subiu de 3,86 para 3,96, em uma escala que vai até 10. Nos anos finais do Ensino Fundamental, houve o recuo de 2,84 para 2,52 e, no Ensino Médio, o Idesp caiu de 1,98, em 2009, para 1,81.

4.5.2. Discussão

Para alguns especialistas ligados à universidade, a queda na bonificação de São Paulo permite inferir que a política como um todo é um fracasso e que não são os incentivos financeiros que motivam os professores (*O Globo*, 23/4/2011). No entanto, as atitudes plenamente favoráveis entre os gestores dessa mesma política em outros estados e o interesse de mais estados como Pará (*O Liberal*, PA, 3/4/2011) e a cidade de São Paulo em começar a experimentar com o pagamento de bônus podem sugerir que o impacto dependa do contexto político e das relações entre as partes. A literatura também recomenda cautela antes de chegar a conclusões gerais devido à grande diversidade na configuração das políticas de incentivos e nos resultados efetivamente alcançados (HOUT & ELLIOTT, 2011). Deixaremos essa discussão para a última seção.

Do ponto de vista das metodologias usadas, é inegável que as políticas de incentivos de segunda geração, que utilizam indicadores de desempenho construídos de diversos elementos para levar em consideração o fluxo, o nível socioeconômico dos alunos, o esforço da escola, a distribuição dos alunos por faixa de desempenho e nível de ensino etc. são mais sólidas e confiáveis do que seus antecessores. No entanto, até a construção de indicadores que se baseiem no valor agregado e que permitam a avaliação da escola ao longo de um período de anos, sempre haverá dúvidas sobre a influência de movimentos aleatórios no desempenho dos alunos. Os resultados da Escola Nota 10 do Ceará, em que a metade das escolas premiadas em 2009 não conseguiu mostrar o mesmo desempenho no ano seguinte, leva a crer que uma parte dos prêmios se deve a flutuações aleatórias. O fato de algumas das escolas identificadas como de desempenho mais baixo em 2009 conseguirem melhorar, a ponto de receberem prêmios como escolas nota 10 no ano seguinte, também levanta suspeitas sérias. Dificilmente uma escola cumpre essa trajetória em um único ano escolar, mesmo com a assistência técnica de uma escola de excelência.

Os embates ideológicos que caracterizam o campo, e que também explicam por que a bibliografia continua incapaz de decantar algum consenso, indica que a formulação de novas políticas e a reformulação de políticas existentes continuará pelo futuro próximo mais uma questão de fé do que de ciência. Sendo assim, nós nos permitimos duas suposições. A primeira é que, quanto mais se avança no sistema em termos do ano escolar e idade dos alunos, maior a dificuldade em relacionar a atuação dos professores aos resultados dos alunos e mais tênues os sistemas de incentivos. Por isso, será sempre mais difícil criar prêmios que façam sentido para os professores do Ensino Médio ou os últimos anos do Ensino Fundamental. Pelo mesmo raciocínio, será mais provável encontrar evidência de algum impacto de um sistema de premiação nos primeiros anos do Ensino Fundamental. Nesses anos, é mais fácil o professor sentir que sua atuação fará alguma diferença no nível de desempenho absoluto dos alunos.

A segunda suposição é que, quanto menor o piso salarial e maior a insatisfação dos professores com seus salários, menor a probabilidade de o incentivo salarial produzir os efeitos desejados. Enquanto o incentivo for

visto como um remendo ou uma tentativa disfarçada de oferecer o salário que o professor reivindica, a sua lógica como prêmio fica prejudicada, e sua capacidade de provocar o tipo de atuação colaborativa pretendida, ainda mais remota. Isso significa que não é propriamente o tamanho do incentivo que faz a diferença, mas a distância entre o salário recebido e o salário visto como merecido. Independentemente do nível salarial efetivamente recebido, o fato de os professores de Espírito Santo estarem mais satisfeitos com seus contracheques proporcionará um futuro mais promissor à política de incentivo naquele estado do que em São Paulo.

4.6. Como componente da política de avaliação docente

As políticas descritas na seção anterior são baseadas na avaliação da escola como um todo e o pagamento coletivo dos incentivos. No Brasil, não há registro atual de políticas de incentivo salarial que façam distinções entre os professores de uma mesma escola. Não é de surpreender, portanto, que há pouco interesse por parte dos gestores em usar a avaliação externa para efeitos da avaliação docente e de tomar decisões administrativas sobre a situação funcional do professor com base nos resultados dos seus alunos.

As únicas exceções encontradas foram uma política de curta duração do Distrito Federal para decidir sobre a recondução de diretores escolares com base na média de desempenho da escola e o sistema de avaliação de desempenho individual dos professores de Minas Gerais.

No caso do Distrito Federal, o uso da média da escola no Siade foi uma decisão *ad hoc* para facilitar a recondução de uma parte dos diretores e vice-diretores de escola. As equipes gestoras eleitas em 2007 que ainda permaneciam em seus cargos poderiam ganhar mais dois anos de mandato em 2009 se as comunidades escolares os reelegessem. Para isso, as escolas dirigidas por elas precisavam ter um bom desempenho do Siade 2009, realizado em outubro desse ano. Se o desempenho dos alunos não fosse melhor em relação aos resultados de 2008, essas equipes obrigatoriamente teriam que participar de novo processo seletivo no final de 2010. Portanto, os resultados dos alunos foi um critério definitivo para determinar a duração do período de gestão do diretor.

A partir da implantação em 2009 de um sistema de avaliação da Gestão Compartilhada e da Gestão Escolar Regimental, como parte do sistema de avaliação das instituições educacionais do Distrito Federal, a Secretaria fica de posse de *checklist* extenso de indicadores sobre o funcionamento de todas as escolas com base nas informações prestadas online pelos diretores. Esses indicadores vão desde a implantação da proposta pedagógica da escola até os processos participativos das instâncias colegiadas, passando por uma bateria de “parâmetros para a gestão compartilhada e desempenho escolar”. Nessa última categoria, estão incluídas as taxas de aprovação, promoção, evasão, distorção idade-série e também a variação do Ideb de 4ª e 8ª séries e a relação entre a avaliação escolar dos alunos e a avaliação do Siade (indicador 2B23). A presença desses indicadores sinaliza a importância dos resultados no conjunto de critérios da boa gestão e provavelmente dispensa o uso específico da média no Siade em decisões futuras sobre a recondução da equipe escolar.

No caso de Minas Gerais, por outro lado, o uso de informação sobre o desempenho dos alunos na avaliação anual dos docentes tem todas as características de uma política permanente. Mesmo sendo só mais uma informação entre muitas nas mãos dos avaliadores, valendo no máximo 10% no conjunto de critérios empregados, fica evidenciado o interesse das autoridades em colocar o foco nos resultados e responsabilizar os professores pela aprendizagem dos alunos.

A Avaliação de Desempenho Individual (ADI) é realizada todo ano por uma comissão avaliadora de cinco integrantes. Os avaliadores levam em consideração cinco critérios ao analisar o professor regente: desenvolvimento profissional, relacionamento interpessoal, compromissos profissional e institucional e habilidades técnicas e profissionais. Cada critério é composto de dois a nove indicadores, tendo peso 1 ou 2, o que permite aumentar ou diminuir a relevância do indicador na composição da pontuação final. O último indicador do quarto critério especifica que “[O professor] obtém bons resultados de aprendizagem de seus alunos evidenciados pelas avaliações internas, realizadas pela escola, e por avaliações externas”. Valendo até 5 pontos, mas com peso 2, esse indicador pode contribuir com até 10% da pontuação total do professor, dependendo da opinião dos avaliadores.

Essa avaliação do professor é *high-stakes* na medida em que a progressão horizontal ou vertical na carreira só é permitida para quem recebeu uma avaliação igual ou superior a 70%. Por exemplo, quem obteve um diploma de pós-graduação, mas não atingiu a nota mínima no ADI, fica impedido de ocupar o nível correspondente ao novo título. E, com duas avaliações seguidas inferiores a 70% ou com quatro em dez avaliações abaixo desse limite, o professor é exonerado do cargo. Até o presente, mais de 400 professores já foram exonerados por esse motivo.

Esse sistema de avaliação docente formaliza um critério usado pela Secretaria cinco anos atrás para exonerar um número grande de professores alfabetizadores. Uma resolução publicada em 2006 impediu a renovação dos contratos de professores eventuais em escolas com mais de 30% de alunos não alfabetizados ao final do 3º ano de acordo com a avaliação do Proalfa. O gestor responsável pela política não tem dúvida de que levar em consideração os resultados dos alunos na hora da avaliação individual contribui para incentivar os professores:

“O resultado do aluno é o negócio que mais pesa na pontuação e tem várias fontes. A comissão pode olhar várias coisas e o professor pode dar a sua explicação. [...] E o que que isso tem a ver com estímulo? Tem a ver o seguinte; a condição para você ser promovido na carreira é que você tem avaliação de desempenho positivo.”
(gestor/SEE-MG)

Uma gestora da SEE-PE mostrou entusiasmo com uma política de avaliação docente para efeitos de promoção na carreira que também incorporaria critérios relacionados ao desempenho dos alunos:

“[...] vai ter [...] um sistema de informação. Ele vai ser em cima do resultado do aluno, do tempo de ensino, do cumprimento do programa de ensino e do resultado do Saepe. [...] o projeto já está desenhado. E por sua vez o plano de carreira vai prever, porque ele já prevê avaliação de desempenho. A avaliação de desempenho vai ter

também um link com o resultado do desempenho do aluno. Tudo nosso é em função do desempenho do aluno.” (gestora/SEE-PE)

A política ainda não foi implantada e não se sabe se haverá tentativas por parte da Secretaria de Pernambuco de incorporar medidas de valor agregado para torná-la mais palatável. Mesmo nesse caso, se existe qualquer reflexão por parte dos professores sobre as imprecisões e instabilidades inerentes à avaliação docente, pode-se prever um grau merecido de resistência à nova política.

4.7. Para a certificação de alunos e escolas

Diferentemente de outros países, os responsáveis pela criação dos sistemas estaduais de avaliação não cogitam usar os resultados como um critério para a aprovação ou diplomação dos alunos. Mesmo na série final do Ensino Médio, quando os alunos concluem o Ensino Básico e saem do âmbito do sistema estadual de ensino, a aprovação dos alunos é determinada exclusivamente pela escola, com base em avaliações realizadas pelos professores.

Em outros países, é comum o emprego de avaliações externas para dar uma nota nas diferentes matérias que compõem o currículo do Ensino Médio ou para decidir se o aluno está apto para receber o diploma desse nível de ensino. A vantagem sempre citada para a criação de parâmetros externos para a avaliação dos alunos diz respeito ao estabelecimento de um padrão único e a garantia de equivalência do significado do diploma em diferentes lugares e tempos. Na ausência desse padrão, podem existir diferenças significativas entre os critérios de diferentes escolas e entre os critérios da mesma escola em diferentes anos. As autoridades ficam sem parâmetro para a comparação das escolas e as escolas sem uma medida clara dos seus resultados. Os esforços recentes de promover mudanças na qualidade da educação tanto nos Estados Unidos como na Inglaterra têm como critério-chave os resultados das escolas baseados nas taxas de diplomação dos alunos ou nas notas dos alunos nas matérias que compõem o certificado de Ensino Médio. Nesses países, o aluno se submete às avaliações externas por ser o único

meio de garantir o diploma ou certificado de Ensino Médio e de se habilitar para o processo de seleção ao Ensino Superior.

Há diversas razões históricas e estruturais para a não-utilização dos sistemas estaduais de avaliação implantados a partir da década de 1990 para a certificação de alunos ao final do Ensino Médio, incluindo a autonomia dos sistemas estaduais e municipais e a autonomia das próprias escolas. Há também restrições de natureza mais técnica devido à incapacidade dos instrumentos de avaliação externa usados nesses sistemas de oferecer uma visão completa do rendimento de cada aluno. Sem um instrumento de avaliação capaz de aferir o nível alcançado por cada aluno em cada uma das competências e habilidades que compõem o currículo do Ensino Médio, torna-se indispensável o processo avaliativo interno respaldado pelo conhecimento do aluno pelos seus professores.

O único caso encontrado em que foram usados os resultados de uma avaliação externa para determinar a progressão de alunos individuais foi em São Paulo no ano de 2001. Nesse caso, porém, não foram os alunos do final do Ensino Médio, mas alunos de 4ª e 8ª séries que foram promovidos ou retidos com base nos resultados do Saresp. A experiência foi polêmica por diversas razões, incluindo o problema de usar um instrumento não desenhado para a avaliação de indivíduos. Outro problema foi o fato de os alunos não saberem de antemão que os resultados seriam usados com essa finalidade. Pelo fato de o Saresp ser um instrumento de monitoramento, sem consequências para os alunos, estes nem sempre levam os testes a sério e podem ter sido prejudicados na ocasião por não responderem às questões com o devido cuidado.

O fato de professores estarem recentemente liberados para aproveitarem notas individuais produzidas por diferentes componentes dos sistemas de avaliação externa em Rio de Janeiro e Minas Gerais na composição da última nota bimestral levantam algumas das mesmas preocupações que no caso de São Paulo de 2001. Pode-se prever que os testes externos usados, as Provas Bimestrais do 4º bimestre no caso do Rio de Janeiro e a da 2ª aplicação do Programa de Avaliação da Aprendizagem Escolar (PAAE) no caso de Minas Gerais, logo assumirão uma importância que antes não tinham caso a nota bimestral seja suficiente para alterar a decisão sobre a retenção

ou promoção e, por consequência, a trajetória escolar do aluno. Pelas suas novas consequências como parte de um processo certificatório, esses testes se tornam mais *high-stakes* que antes, podendo tanto incentivar o bom desempenho quanto provocar atitudes que solapam seus propósitos didáticos.

A criação e consolidação do Enem como um exame certificador para os alunos que querem continuar seus estudos mostra que algumas das restrições históricas sobre a avaliação externa ao final do Ensino Médio foram superadas. Ao mesmo tempo, o Enem mostra que, quando a matriz de referência da prova é definida mais com base em competências e menos em termos de conhecimentos curriculares, é possível criar um instrumento capaz de medir o nível alcançado por cada aluno. O que o Enem nunca fornecerá é uma medida da qualidade da escola onde os alunos estudaram. Apesar de ser muito usada com essa finalidade, a média dos resultados dos alunos de determinada escola não indica a qualidade da escola porque os alunos que fazem a prova não representam nem a totalidade nem uma amostra representativa dos egressos daquela instituição. A falta de correspondência exata entre a matriz de referência e o currículo das escolas e o fato de estas atenderem a populações bastante diferenciadas seriam outras razões para não confiarem nos resultados do Enem como um indicador de qualidade da instituição.

Mesmo incomum, também há exemplo de uso dos resultados dos alunos para efeitos de certificação de escolas. Trata-se da certificação das escolas privadas do Distrito Federal. Essa situação se deriva do fato de a SEE-DF ser o órgão responsável pela autorização de funcionamento das escolas privadas e pela sua recertificação para continuarem em funcionamento depois de cinco a dez anos, dependendo do tipo de escola. No ano anterior ao vencimento da sua autorização, as escolas privadas candidatas à recertificação são automaticamente inscritas entre as escolas a serem avaliadas pelo Siade. Outras escolas privadas podem inscrever-se voluntariamente para essa mesma avaliação mediante o pagamento dos seus custos. A ideia da Secretaria era comparar os resultados da escola privada durante dois anos seguidos para determinar se ela mantinha o mesmo nível de qualidade:

“Quando ela for recredenciar, ela vai ser avaliada antes para [poder] comparar se ela melhorou mesmo ou não, até para dizer se ela vai ser recredenciada ou não, em cima de um resultado do Siade.” (técnico/SEE-DF)

Não havia informações disponíveis para verificar o impacto dessa política em termos do número de escolas privadas avaliadas nem se havia escolas impedidas de funcionarem com base nos seus resultados. Suspeita-se que a política foi divulgada mais para alertar as escolas privadas do interesse da Secretaria em promover a melhoria dos resultados do que para fechar escolas com crescimento de desempenho aparentemente insatisfatório. Essa conclusão deriva do fato de que tomar a decisão sobre o funcionamento de uma escola com base na comparação da média dos alunos entre um ano escolar e outro seria uma temeridade. Por não levar em consideração as flutuações naturais nessa média, nem as taxas diferenciadas de crescimento de acordo com o nível absoluto de desempenho da escola, a medida não oferece o rigor necessário para uma decisão *high-stakes* de tamanha importância.

5. Considerações finais¹⁰

Este artigo teve como propósito descrever as diferentes políticas de gestão empreendidas pelas SEs dos estados e municípios maiores com base em seus sistemas próprios de avaliação educacional e refletir sobre elas. Para essa finalidade, foi elaborado um esquema classificatório e levantadas informações primárias e secundárias sobre diversos exemplos das políticas para permitir uma discussão dos seus prováveis impactos no novo cenário da gestão educacional.

10 Especialistas que participaram do painel desse estudo, com objetivo de debater os resultados encontrados e elencar proposições com foco em políticas públicas, em ordem alfabética: Antônio Bara Bresolin (Fundação Itaú Social), Bernardete Gatti (Fundação Carlos Chagas), Clara Etienne Lima de Souza (Inep), Mozart Neves Ramos (Todos pela Educação), Ruben Klein (Fundação Cesgranrio), Sandra Maria Zakia Lian Sousa (USP e Unicid), Tufi Machado Soares (Universidade Federal de Juiz de Fora) e Wanda Engel (Instituto Unibanco).

A primeira constatação da pesquisa é da diversificação e progressiva sofisticação nas metodologias empregadas pelas Secretarias Estaduais e Municipais na utilização dos resultados das avaliações para a tomada de decisões. O aumento no número e na competência técnica dos métodos dos sistemas estaduais, concomitante com a melhoria nos sistemas de informação, tem contribuído para essa disseminação. Outras contribuições advêm da incorporação pelos governos estaduais de um novo modelo de gestão pública, voltado para o controle dos resultados dos serviços oferecidos à população, e da liderança principalmente do MEC na criação de um consenso nacional sobre a relevância dos resultados das avaliações externas como medida da qualidade do ensino e parâmetro para o progresso da educação brasileira. A criação de um indicador próprio, tipo Ideb, por um número cada vez maior de estados é indicativo dessa nova abordagem para a política de monitoramento dos serviços públicos e da disposição de empregar os resultados da avaliação na formulação de novas políticas, como a da responsabilização.

Confirmando a diversificação nos usos da avaliação, foram identificadas sete categorias de políticas de gestão que se fundamentam nos resultados dos alunos e das escolas. Em determinadas condições, a maioria dessas políticas pode ser entendida como contribuições importantes para a melhoria da qualidade da educação. Em outros casos, porém, notadamente no uso dos resultados dos alunos como um componente da política de avaliação docente e para a certificação de alunos e escolas, observa-se o emprego desavisado das avaliações para a tomada de decisões que precisam de informações mais confiáveis, de mais de uma fonte e por períodos maiores de tempo.

Em relação às políticas de responsabilização que atingem a remuneração dos professores, é necessário fazer algumas distinções para evitar a adoção apressada de conclusões derivadas da literatura norte-americana. No contexto brasileiro, os incentivos empregados não dizem respeito ao futuro imediato dos alunos e não envolvem os pais. Também não há riscos no sentido de fechamento de escolas nem da perda de emprego dos membros da equipe. Nas palavras de Afonso, aqui são “formas parcelares de *accountability*” e, portanto, incompletas (AFONSO, 2009). Não deixa de ser verdade que,

quanto melhores os resultados do aluno, melhores suas chances na vida, mas essa equação não tem a mesma urgência em um sistema que diploma todos os alunos que sobrevivem até o final do ciclo e que jamais discriminaria as escolas com os piores resultados. No caso brasileiro, as avaliações dos alunos que são usadas para determinar se as escolas cumpriram suas metas ou alcançaram um patamar de desempenho considerado satisfatório representam muito pouco para os alunos. Pelo contrário, há evidências de que os alunos fazem tão pouco caso dos testes que, quando se criam consequências, como os laptops e as bicicletas para os alunos de Ensino Médio do Ceará e do Rio de Janeiro, respectivamente, a mudança no seu empenho é tão radical que altera por completo a base de comparação.

No contexto das políticas brasileiras de incentivos, os critérios empregados não obrigam o professor a assimilar o drama dos alunos nem requer que transfira para eles a sua ansiedade em relação aos resultados dos testes. O que está em jogo nesses casos não são as consequências para os sujeitos principais da educação, os alunos, mas o nível de recompensa dos professores. Em vez de ser a própria definição do objeto de seu trabalho, pelas consequências para os alunos, a motivação para a melhoria nos resultados se torna uma decorrência do nível de preocupação do professor em melhorar sua remuneração.

Talvez por esse motivo, ainda não existam evidências claras para os efeitos colaterais das políticas de responsabilização no Brasil. Apesar dos temores, mesmo em estados como Ceará, com dez anos de experiência com sistemas de incentivos, não se registra qualquer pesquisa que relate os efeitos prejudiciais dessa política. Ou seja, em lugares como os Estados Unidos, onde a pressão da responsabilização é alta, pelo número de políticas que estabelecem consequências para as escolas e pela natureza muitas vezes negativa dessas consequências, talvez seja compreensível o alto grau de preocupação com os efeitos colaterais dos testes *high-stakes*. O nível do risco de não se alcançar a meta é muito alto, o que pode provocar os comportamentos vistos como destrutivos tanto ética quanto pedagogicamente. No contexto brasileiro, por outro lado, há razões para acreditar que esses danos colaterais talvez não sejam tão graves justamente pela natureza da política de responsabilização em jogo. Quando a política de responsabilização se fixa no pagamento de

prêmios que possam atingir até a maioria dos professores, a consequência negativa mais óbvia da política é o não-recebimento da totalidade do bônus.

O nosso estudo da implantação e dos primeiros passos da nova geração de políticas de incentivos salariais também não trouxe evidências objetivas de benefícios em termos de ganhos na aprendizagem dos alunos provocados pelos efeitos motivadores dos incentivos. Pelo contrário, os resultados do sistema de bonificação de São Paulo de 2010, em que o número de professores beneficiados e o volume pago de bônus foram menores que em 2009, já foram interpretados como evidência de que o sistema de bonificação não é uma ação suficiente para promover a melhoria dos resultados dos alunos em curto prazo.

A pesquisa brasileira para medir os impactos dos incentivos ainda é incipiente, só havendo uma pesquisa finalizada (RODRIGUES, 2007) e outra em andamento, no estado de Pernambuco. A pesquisa norte-americana, por outro lado, é abundante. No entanto, quando essa pesquisa é submetida a um crivo metodológico exigente, como em metaestudo recente patrocinado pela Comissão sobre Incentivos e Responsabilização baseada em testes do *National Research Council* dos Estados Unidos, descobre-se que os impactos dos incentivos em termos de ganhos de aprendizagem naquele país são mínimos (HOUT & ELLIOTT, 2011). Em outros países, como Israel, Índia e Quênia, os autores atestam que o impacto é maior, sendo mais pronunciado para Matemática no Ensino Fundamental. Mesmo assim, as conclusões do estudo apontam para benefícios pequenos e a necessidade de investimentos simultâneos em outras políticas voltadas para a melhoria da qualidade do ensino.

Com base em quatro argumentos, chega-se à conclusão de que no contexto brasileiro é provável que as políticas de responsabilização formuladas em termos de incentivos salariais sejam ou inócuas ou só marginalmente eficazes no seu objetivo primordial de promover ganhos nos resultados dos alunos. O primeiro argumento deriva da pesquisa estrangeira, que mostra que os esquemas de incentivos coletivos podem impactar a assiduidade dos professores, mas dificilmente alteram sua forma de ensinar (GOODMAN & TURNER, 2010). Essa conclusão não é de difícil compreensão dado o fato de que o pagamento coletivo sai ou para todos os membros da equipe ou para

ninguém da equipe. Isso dificulta o trabalho do professor que quer receber o bônus, mas que não tem com quem aprender para mudar seu ensino. E, mesmo motivado, o professor só receberá o bônus se a escola inteira se interessar pelo mesmo objetivo. Pesquisa chilena tende a mostrar que as escolas que não estão próximas de receberem o bônus dificilmente passam a recebê-lo. Ou seja, o bônus chileno só premia quem já tem resultados melhores e não influencia a quase metade das escolas que nunca foram premiadas (MIZALA & ROMAGUERA, 2005).

O segundo argumento é a falta de evidência sobre a relação entre a assiduidade do professor e o desempenho dos alunos. Se for verdade que os sistemas de incentivos podem influenciar positivamente a assiduidade dos professores, no sentido de coibir as faltas excessivas que caracterizam os sistemas públicos, seria natural pressupor um impacto positivo sobre a aprendizagem dos alunos. No entanto, essa relação é relativizada por diversos elementos, incluindo o desempenho dos professores substitutos, o desempenho do professor titular quando presente e a falta de critério no controle das faltas. Ainda não sabemos se a melhoria na assiduidade conduz de fato a uma melhoria nos resultados dos alunos.

De toda forma, observa-se a necessidade de clareza maior sobre o tipo de impacto desejado pelos sistemas de incentivos brasileiros. A falta de evidência de que o prêmio incentive o professor a procurar se capacitar ou modificar sua prática de ensino impõe a necessidade de explicitar que o propósito seja muito mais o de bonificar a assiduidade. Nesse caso, talvez nem se precisasse de indicadores de aprendizagem. Outros usos de incentivos, como para trazer professores para áreas menos atraentes, como é o caso das escolas rurais, ou para atrair professores para disciplinas curriculares em que há falta de professores, ainda não constam do elenco de políticas dos governos estaduais, mesmo havendo uma conexão bem mais clara entre o pagamento de bônus para essas finalidades e os resultados alcançados.

O terceiro argumento tem relação com o currículo. Para o *accountability* funcionar, precisa-se de quatro elementos, segundo Corvalán: informação (sobre o desempenho dos alunos), autoridade (para definir os critérios), consequências (em termos de prêmios e castigos) e padrões (em termos do significado pedagógico das metas) (CORVALÁN & McMEEKIN, 2006).

No caso brasileiro, os padrões estão faltando. O professor não saberia qual a diferença entre 220 e 240 na escala de proficiência usada pelo Boletim Pedagógico em termos dos conteúdos que ele precisa ensinar para chegar ao nível de proficiência requerido. Essa falta de correspondência entre os critérios empregados pelos sistemas de incentivos e a compreensão do seu significado em termos da aprendizagem dos alunos está na base da falta de progresso da maioria dos estados no uso dos resultados das avaliações como material de apoio pedagógico.

Na hipótese de os sistemas de incentivos serem marginalmente eficazes, seu impacto será a consequência da ênfase colocada nas metas e nos resultados do ensino. Essa contribuição, que diz respeito a um novo ideário em relação à ordem de importância dos objetivos do sistema educacional e do trabalho do professor, é a expressão de uma filosofia de gestão que produz as políticas de incentivo salarial, mas que não precisa delas para existir. Em outras palavras, a importância cada vez maior atribuída aos resultados dos alunos é uma mensagem inerente às políticas de responsabilização que pode ter sua própria influência na maneira de os professores enxergarem seu papel.

A razão por que essa mensagem possa ter algum impacto no Brasil, diferentemente dos Estados Unidos, é justamente pela sua novidade e pela falta de outras políticas de responsabilização competidoras. Nos Estados Unidos, as políticas de incentivos salariais não acrescentam nenhuma novidade a um ambiente já crivado por políticas de responsabilização, que transmitem a preocupação com os resultados por meio de várias exigências, incluindo o progresso anual aceitável em direção às metas do NCLB¹¹. Nesse ambiente, é mais difícil imaginar que os incentivos salariais agreguem alguma coisa às preocupações já estabelecidas dos professores. No Brasil, por outro lado, os

11. Legislação federal nos Estados Unidos, aprovada pelo congresso em 2001, recebeu o nome de Nenhuma Criança Deixada para Trás, ou NCLB (No Child Left Behind). O NCLB aumentou significativamente as obrigações dos estados na aplicação de avaliações, estipulou metas de *accountability* para escolas, distritos e estados e legislou sobre o “progresso anual aceitável” em direção às metas de 2014 para todos os alunos e subgrupos de alunos por nível socioeconômico, raça, língua materna e necessidade especial.

incentivos são os mensageiros de uma nova ordem de prioridades que ainda está para ser assimilada plenamente.

Referências

- ADRIÃO, T. & GARCIA, T. “Oferta educativa e responsabilização no PDE: o plano de ações articuladas”. *Cadernos de Pesquisa*, v. 38, n. n. 135, p. 779-796, set./dez. 2008. Disponível em: <<http://www.scielo.br/pdf/cp/v38n135/v38n135a12.pdf>>. Acesso em: 03/11/10.
- AFONSO, A. J. “Políticas avaliativas e accountability em educação: subsídios para um debate ibero-americano”. *Sísifo. Revista de Ciências da Educação*, v. 09, p. 57-70, 2009. Disponível em: <<http://sisifo.fpce.ul.pt/?r=23&p=57>>. Acesso em: 08/10/10.
- ANDRADE, E. D. C. “‘School accountability’ no Brasil: experiências e dificuldades”. *Revista de Economia Política*, v. 28, n. 3(111), p. 443-453, julho-setembro 2008. Disponível em: <<http://www.scielo.br/pdf/rep/v28n3/a05v28n3.pdf>>. Acesso em: 05/10/10.
- APEOESP. “Fax urgente 2011”. Disponível em: <http://apeoespsub.org.br/teste/Fax/Fax_2911.pdf>. Acesso em: 06/06/2011.
- AYRES, S. “Boletim da escola: diagnóstico multidimensional das escolas do Paraná”. *Ensaio*, v. 11, n. 38, p. 33-44, 2003.
- BAUER, A. “Usos dos resultados das avaliações de sistemas educacionais: iniciativas em curso em alguns países da América”. *Revista Brasileira de Estudos Pedagógicos*, v. 91, n. 228, p. 315-344, maio/ago. 2010. Disponível em: <<http://www.rbep.inep.gov.br/index.php/RBEP/article/viewFile/1690/1341>>.
- BROOKE, N. “As perspectivas para as políticas de responsabilização educacional no Brasil”. *Cadernos de Pesquisa*, v. 36, n. 128, p. 377-401, 2006. ISSN 0100-1574. Disponível em: <http://www.scielo.br/scielo.php?script=sci_issuetoc&pid=0100-157420060002&lng=en&nrm=iso>.
- BROOKE, N. “Responsabilização educacional no Brasil”. *Revista Iberoamericana de Evaluación Educativa*. v. 1, pp. 93-109, 2008. Disponível em: <http://www.rinace.net/riee/numeros/vol1-num1/art7.pdf>. Acesso em 20/10/10.

- CAVALIERE, A. M. “Entre o pioneirismo e o impasse: a reforma paulista de 1920”. *Educação e Pesquisa*, v. 29, n. 1, p. 27-44, 2003.
- CDEP. *Comitê em defesa da escola pública*. <<http://www.adusp.org.br/PEE/PEE.rtf> São Paulo 2006>.
- CONSED. *Estudo sobre avaliação da educação e valorização do magistério*. Consed, Brasília, 2005.
- CORVALÁN, J. & MCMEEKIN, R. W. O. “Accountability educacional: rendición de cuentas más responsabilización por los procesos y resultados de la educación”. In: MCMEEKIN, J. C. R. W. (Ed.). *Accountability educacional: posibilidades y desafíos para América Latina a partir de la experiencia internacional*. Santiago: PREAL – CIDE, 2006.
- DANTAS, L. M. V. *As contribuições das políticas de avaliação educacional em larga escala: o caso da avaliação de aprendizagem na Bahia*. Salvador, 2009. Salvador, Tese (Doutorado). Faculdade de Educação, Universidade Federal da Bahia.
- DISSIDÊNCIA. “Gilberto Dimenstein e a bonificação dos professores”. *Dissidência*. Disponível em: <<http://www.andrekenji.com.br/weblog/?p=264>>. Acesso em: 10/04/2006.
- ESPÍRITO SANTO. *Plano Estratégico Nova Escola*. Vitória: Secretaria da Educação, 2008.
- FERRER, G. *Sistemas de evaluación de aprendizajes en América Latina: balance y desafíos*. Março 2006, p. 174, 2006
- GATTI, B. A. “Avaliação educacional no Brasil: pontuando uma história de ações”. *EccoS revista científica*, v. 4, n. 1, p. 17-41, junho 2002. Disponível em: <<http://www4.uninove.br/ojs/index.php/eccos/article/viewFile/291/280>>.
- GOODMAN, S. & TURNER, L. *Teacher incentive pay and educational outcomes: evidence from the NYC bonus program*. Harvard, June, 2010.
- GTEE/PREAL. “Evaluación y políticas educativas: ¿Dos caminos paralelos?” In: *Observatorio Regional de Políticas de Evaluación Educativa*. Boletín n. 28, 2011. Disponível em: <http://www.preal.org/NoticiaGNN.asp?Id_Noticia=345&Id_Grupo=3 2011>.
- HOUT, M. & ELLIOTT, S. W. (Eds.). *Incentives and test-based accountability in Education*. National Research Council Committee on Incentives and

- Test-Based Accountability in Public Education. Washington: The National Academies Press, 2011.
- KLINGER, D. A.; DELUCA, C. & MILLER, T. “The evolving culture of large-scale assessments in canadian education”. *Canadian Journal of Educational Administration and Policy*, n. 76, 3 jul. 2008. Disponível em: <<http://www.umanitoba.ca/publications/cjeap/articles/klinger.html>>.
- MEC/INEP. *Vencendo o desafio da aprendizagem nas séries iniciais: a experiência de Sobral/CE*. Brasília: MEC/INEP, 2005. Disponível em: <www.sic.inep.gov.br/index.php?option=com_docman&task...>.
- MIZALA, A. & ROMAGUERA, P. “Teachers’ salary structure and incentives in Chile”. In: VEGAS, E. E. (Ed.). *Incentives to improve teaching: lessons from Latin America*. Washington: The World Bank, 2005.
- MORTIMORE, P. et al. *School matters*. Berkeley: University of California Press, 1988.
- MOVIMENTO. *Movimento contra testes de alto impacto em Educação*. 2010. Disponível em: <<http://sites.google.com/site/movimentocontratestes/home>>. Acesso em: 04/06/2011.
- PEREIRA, M. C. “As avaliações externas mineiras: contexto de políticas de responsabilização”. In: XXIV Simpósio Brasileiro de Política e Administração da Educação Universidade Federal do Espírito Santo (UFES), 2009, Vitória: Associação Nacional de Política e Administração da Educação (ANPAE), 2009.
- RAVELA, P. A.; VALVERDE, Gilbert; WOLFE, Richard; FERRER Guillermo; MARTÍNEZ RIZO, Felipe; AYLWIN, Mariana & WOLFF, Laurence. *Las evaluaciones educativas que América Latina necesita*. Preal. Santiago: 2008.
- RODRIGUES, J. R. S. *Responsabilização e resultados escolares no Rio de Janeiro*. Rio de Janeiro, 2007. Dissertação (Mestrado). Departamento de Educação, PUC-Rio.
- RUTTER, M. et al. *Fifteen thousand hours: secondary schools and their effects on children*. London: Open Books, 1979.
- SHEPARD, L. *A Checklist for evaluating large-scale assessment programs*. The Evaluation Center, Western Michigan University, Kalamazoo, 1977.
- SHULHA, L. M. & COUSIN, J. B. “Evaluation use: theory, research and practice since 1986”. *American Journal of Evaluation*, v. 18, p. 195-208, 1997.

- SOARES, J. F. “Índice de desenvolvimento da Educação de São Paulo (IDESP): bases metodológicas”. *Revista São Paulo em Perspectiva*, v. 23, n. 1, p. 29-41, jan.-jun. 2009.
- SOARES, S. “Avaliação educacional como instrumento pedagógico”. *Trabalho e Sociedade*, ano 2, n. 4, p. 23-25, agosto, 2002.
- SOUSA, S. M. Z. L. “Possíveis impactos das políticas de avaliação no currículo escolar”. *Cadernos de Pesquisa*, v. 119, p. 175-190, 2003. Disponível em: <<http://www.scielo.br/pdf/cp/n119/n119a09.pdf>>. Acesso em: 05/12/10.
- SOUSA, S. M. Z. L. & OLIVEIRA, R. P. D. *Sistemas de avaliação educacional no Brasil: características, tendências e uso dos resultados*. Universidade de São Paulo, Faculdade de Educação, Centro de Estudos e Pesquisas em Políticas Públicas de Educação (CEPPPE), 2007.

