AVALIAÇÃO DE POLÍTICAS PÚBLICAS: O CASO DA POLÍTICA DE PAGAMENTO POR BÔNUS NA EDUCAÇÃO NA CIDADE DE NOVA YORK E O CENÁRIO BRASILEIRO

DISCUSSING ASSESSMENT: THE CASE OF TEACHER INCENTIVE POLICY IN THE CITY OF NEW YORK AND THE BRAZILIAN SITUATION

EVALUACIÓN DE POLÍTICAS PÚBLICAS: EL CASO DE LA POLÍTICA DE PAGO MEDIANTE BONOS EN EL SECTOR DE LA EDUCACIÓN EN NUEVA YORK Y LA SITUACIÓN BRASILEÑA

RESUMO

Este artigo trata da avaliação de políticas educacionais, especificamente da política de pagamento de bônus na educação pública. Além da análise da política implementada em Nova York, foi estabelecido um panorama da situação brasileira na avaliação de políticas dessa natureza tendo o caso dessa cidade como um balizador. Para tal, utilizou-se o subsídio da literatura a respeito da avaliação de políticas públicas e realizou-se uma análise documental dos relatórios e dos artigos científicos produzidos com base na experiência de Nova York e dos estudos acerca das experiências brasileiras. Assim, verificou-se que o caso segue os procedimentos descritos na literatura e, dessa forma, cumpre os objetivos ao qual se destina. Em relação ao Brasil, nota-se a necessidade de que as avaliações sejam consideradas pelos gestores públicos, o que permitirá que os resultados contribuam para o processo de planejamento e aprimoramento da política.

PALAVRAS-CHAVE: Avaliação, políticas públicas, políticas educacionais, pagamentos de bônus na educação, tomada de decisões.

Antonio Bara Bresolin¹
tombresolin@gmail.com

Caio César de Medeiros Costa²
caiocosta@unb.br

Ronan Pereira Capobiango³
ronan.capobiango@ufv.br

Submetido 22.05.2017. Aprovado 08.04.2018

Avaliado pelo processo de *double blind review*.

DOI: <http://dx.doi.org/10.12660/cgpc.v23n74.66777>

¹Instituto Ayrton Senna, São Paulo, SP, Brasil

²Universidade de Brasília, Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas, Brasília, DF, Brasil

³Universidade Federal de Viçosa, Viçosa, MG, Brasil

Abstract

This paper discusses the assessment of educational policies, particularly teacher bonus pay in public education policy. We analyze the assessment of the policy implemented in the City of New York and draw a parallel with the Brazilian situation concerning the assessment of this kind of policy. Based on the literature on public policy evaluation and using documentary analysis for the New York experience and the situation in Brazil, this paper presents the evaluative experiment in New York, identifying and analyzing its design and purpose and the use of information from the process as a resource for government decision making. In addition, we show that in Brazil the assessment of this policy requires greater support of policy makers so that assessment results can be used in policy planning and improvement.

Key words: Assessment, public policy, educational policies, education bonus payments, decision making.

Resumen

Este trabajo trata sobre la evaluación de políticas educativas, específicamente, sobre la política de pago de bonos en el sector de la educación pública. Además del análisis de la política implementada en la ciudad de Nueva York, se ha establecido un paralelo con la situación brasileña en la evaluación de políticas de esa naturaleza. Para ello, se recurrió a la literatura sobre evaluación de políticas públicas y se realizó un análisis documental de los informes y artículos científicos producidos con base en la experiencia de Nueva York y de Brasil. Así, se verificó que ambos casos siguen los procedimientos descritos en la literatura y cumplen los objetivos propuestos. Con relación a Brasil, se observa la necesidad de que las evaluaciones realizadas sean consideradas por los gestores públicos a los efectos de permitir que sus resultados puedan contribuir al proceso de planificación y mejora de políticas.

Palabras clave: Evaluación, políticas públicas, políticas educativas, pagos de bonos en la educación, toma de decisiones.

INTRODUÇÃO

Após a promulgação da constituição federal brasileira em 1988, principalmente ao longo das décadas de 1990 e 2000, uma série de reformas políticas foram realizadas. Durante esse período, algumas agendas implementadas em contexto nacional e subnacional seguiram modelos adotados por outros países. Entre elas, destaca-se a criação de instrumentos de avaliação em larga escala do sistema público de educação (Schwartzman, 2013).

A realização de provas padronizadas desenvolvidas como instrumento para mensurar a aprendizagem dos alunos tem crescido ao longo dos anos (Souza & Oliveira, 2010). A criação desse modelo de avaliação permite a orientação de diferentes etapas da gestão da política educacional por esse instrumen-

to e seus resultados, como demonstra Brooke (2011) ao identificar as seguintes possibilidades de utilização dos resultados: 1) Avaliar e orientar a política educacional; 2) Informar as escolas a respeito da aprendizagem dos alunos e definir estratégias de formação continuada; 3) Informar o público sobre o progresso dos alunos; 4) Alocar recursos; 5) Elaborar políticas de incentivos salariais; 6) Avaliar o docente; 7) Certificar alunos e escolas.

As discussões desta pesquisa concentram-se no item número 5, comumente chamado de políticas de pagamento de bônus a professores. O uso de provas como instrumento de incentivo para melhoria do desempenho de professores ganhou força e teve sua prática ampliada principalmente na última década. O desenho de programas de bônus para professores com desempenho diferenciado visando estimular a melhoria na aprendiza-

gem dos alunos foi adotado por diferentes governos, incluindo sete estados brasileiros que já implantaram esse tipo de política: Amazonas, Ceará, Espírito Santo, Minas Gerais, Pernambuco, Rio de Janeiro e São Paulo (Brooke, 2011).

Além das diferentes experiências iniciadas no Brasil, outros países fazem ou fizeram uso semelhante desse tipo de instrumento como incentivo para professores na área pública. Apesar da sua larga utilização ao redor do mundo, evidências internacionais a respeito do impacto do bônus sobre desempenho não são conclusivas, ainda que apresentem sinais positivos em contextos específicos (Figlio & Loeb, 2011). Experiências avaliadas em países como Índia (Muralidharan & Sundararaman, 2009), Inglaterra (Atkinson et al., 2009) e Israel (Lavy, 2009) mostram que o programa de bônus para professores baseado no desempenho pode ter resultados em provas padronizadas feitas pelos estudantes. Entretanto, as discussões sobre o assunto ainda estão longe de um consenso e carecem de estudos avaliativos conclusivos e generalizáveis.

Em contextos como esses, a avaliação das políticas governamentais com base em métodos e modelos científicos é necessária. De acordo com Serpa (2010), são raros os casos sobre uso da avaliação na área pública no Brasil e existem poucos estudos avaliativos com desdobramentos e consequências práticas no aprimoramento da gestão de políticas públicas.

A política de pagamento por bônus a professores no Brasil segue a mesma lógica de escassez de estudos avaliativos explicitado pela baixa quantidade de informações e

análises sobre o tema.

Os únicos trabalhos publicados até o momento foram feitos por Oshiro e Scorzafe (2011), Bruns e Ferraz (2011) e Bresolin (2014). O primeiro avalia a política do Estado de São Paulo a respeito do efeito do pagamento de bônus aos professores sobre a proficiência escolar. O segundo também analisa o impacto do programa, todavia tendo como recorte as escolas básicas do estado de Pernambuco. O terceiro averigua o impacto do programa de pagamento por bônus em diferentes unidades da federação sob a ótica de três dimensões: a. práticas pedagógicas dos professores; b. absenteísmo e rotatividade ou visão do professor em relação à sua assiduidade, rotatividade; c. interlocução com familiares para garantir a frequência dos alunos.

Apesar do baixo número de trabalhos publicados, cumpre destacar que a realização da avaliação das políticas públicas é fundamental para o desenvolvimento e a adaptação das formas e dos instrumentos da ação pública como aponta Frey (2000), para isso a avaliação deve subsidiar o processo de tomada de decisões com informações corretas e úteis que permitam o aprendizado organizacional.

Considerando tais apontamentos, este artigo tem como foco a avaliação das políticas educacionais, especificamente a política de pagamento de bônus na educação pública, tomando como referência empírica a avaliação da política de pagamento por bônus na educação realizado por Nova York. Trata-se de um exemplo bem-sucedido de avaliação. Os resultados encontrados foram utilizados como base para que o governo tomasse

seus resultados como subsídios para a tomada de decisões. Além de ter abordado diferentes nuances da política, forneceu grande quantidade de informações e propiciou a aprendizagem organizacional. Experiências como essas mostram a necessidade e a pertinência da avaliação de políticas públicas que permitam o emprego consequente dos processos avaliativos no aprimoramento das ações desenvolvidas.

Levando em conta o baixo uso da avaliação das políticas públicas no Brasil, no que se refere ao pagamento de bônus por desempenho e à necessidade de se analisar a contribuição potencial de tal política de incentivo (Bresolin, 2014), este estudo pretende contribuir para a reflexão teórica, assim como estimular uma mudança de prática ao apresentar e discutir um caso que poderá servir como referencial para concepção e realização de avaliações futuras.

Considerando ainda o escasso recurso na administração pública aliado ao custo de uma avaliação que ultrapasse a percepção de qualidade e da mensuração de impacto (Bresolin, 2014), torna-se cada vez mais importante a utilização de benchmarkings a fim de dirimir possíveis erros e garantir uma avaliação que forneça as informações necessárias aos gestores públicos e à população.

Aliados à importância da avaliação, esses argumentos depõem a favor do objetivo deste estudo que é discutir a avaliação de um programa de bônus para professores, com base em uma experiência internacional de sucesso que preencheu tal lacuna. O caso de Nova York certamente poderá orientar, inspirar e referenciar processos avaliativos

de diferentes políticas que estão em elaboração ou já foram implementadas no Brasil nesse âmbito.

AVALIAÇÃO DE POLÍTICAS PÚBLICAS

Ala-Harja e Helgason (2000, p. 8) definem avaliação assegurando que “o termo compreende a avaliação dos resultados de um programa em relação aos objetivos propostos”. No entanto, os autores advertem que não existe um consenso quanto ao que seja avaliação de políticas públicas, já que o conceito admite múltiplas definições. Uma definição comumente empregada seria a de Weiss (1998): avaliação é a estimativa sistemática da operação e/ou dos resultados de um programa ou de uma política, comparada com um conjunto de parâmetros implícitos ou explícitos, como forma de contribuir para o aprimoramento do programa ou política.

Apesar do importante papel realizado pela avaliação, ela não deve ser julgada como sendo autossuficiente para a definição dos rumos da política pública (Cohen & Franco, 2007), mas sim como uma das etapas do planejamento que pode ser apresentada como uma “produtora de informações, que possibilitam novas escolhas; e analisa resultados que podem sugerir a necessidade de reorientação das ações para o alcance dos objetivos traçados” (Souza, 2015, p. 48).

Assim, deve-se entender avaliação como uma etapa que auxilia na tomada de decisões. Arretche (2001) destaca que a avaliação deveria ser uma etapa da política posterior à implementação de programas e projetos com o intuito de influenciar a sua reestruturação, durante ou posterior a sua execução, como preconiza o modelo do ci-

clo político. Tornando-se fundamental para o progresso e a adaptação contínua dos instrumentos de ação pública (Frey, 2000).

Além da importância do entendimento dos conceitos, da finalidade e dos tipos de utilização da avaliação, a metodologia para sua realização é cada vez mais relevante no debate acadêmico. A pesquisa de Trevisan e Van Bellen (2008) se utiliza de diferentes trabalhos e discute a metodologia da avaliação seguindo a classificação apresentada por Cotta (2001), isto é, avaliação classificada de acordo com o *timing* (antes, durante ou depois da implementação da política ou do programa), com a posição do avaliador em relação ao objetivo avaliado (avaliador interno ou externo) e com a natureza do objetivo avaliado (avaliar a eficácia, a eficiência e/ou a efetividade da política ou do programa).

Apesar da variada gama de características e possibilidades destacadas pela literatura no que se refere à avaliação das ações governamentais, algumas dificuldades, além das já expostas, são atreladas à sua realização e seu uso subsidiando o processo decisório. Nessa direção, Cohen (2001) expressa que as dificuldades podem ser advindas tanto da natureza das políticas que impedem a obtenção de indicadores, como da falta de vontade de gestores e burocratas em realizar a última etapa do ciclo político. No que se refere à falta de interesse, o autor argumenta que o Estado tem grande interesse em controlar os insumos e procedimentos atrelados à gestão pública do que propriamente os efeitos da atuação estatal; adicionam-se as dificuldades trazidas por Cohen (2001) e a argumentação de Weiss (1998) acerca das dificuldades políticas e organizacionais para que a avaliação tenha os seus resultados

considerados pelos governos.

Os problemas ocorridos nos processos avaliativos e, principalmente na utilização dos resultados dessa avaliação, enfatiza a pertinência da análise dos processos de avaliação que cumpriram os objetivos de fornecer informações. Para isso, destaca-se o caso de Nova York que permite a análise das questões que envolvem a avaliação de políticas sociais, mais especificamente da educação, cujos *outcomes* sofrem influência de outros aspectos, demandando mais cuidado na etapa de avaliação.

AVALIAÇÃO DE POLÍTICAS PÚBLICAS E SUA APLICAÇÃO NO CONTEXTO EDUCACIONAL

A avaliação de políticas públicas tem múltiplos objetivos e atende a uma variada audiência, a qual no Brasil, de acordo com Calmon e Serpa (2009), tem ocorrido de forma tardia e sem uma perspectiva estratégica, predominando a existência de grupos de acadêmicos ou de alguns gestores que reconhecem as contribuições potenciais da avaliação para o aprimoramento das ações desenvolvidas.

Na área da educação pública, tal questão é similar às demais. Todavia, ao atentar para as especificidades de cada política, convém transpor tais particularidades também para o processo avaliativo.

Como pontua Fernandes e Gremaud (2009), ao analisar o caso brasileiro, as avaliações das políticas educacionais podem ser vistas como casos particulares das análises de políticas e programas atrelados aos resultados da política buscando aferir resultados pas-

sados ou inferir acerca de resultados futuros.

Ampliando essa discussão, os autores apontam que atualmente as avaliações das políticas educacionais devem incorporar o debate sobre *accountability*, ou seja, ir além da sala de aula e da análise do desempenho dos alunos, levando em conta também os responsáveis pela performance dos estudantes, escolas, professores, diretores e gestores.

De forma semelhante, o estudo de Ceneviva e Farah (2012) analisa a relação entre a responsabilização dos agentes públicos e as características do sistema de avaliação implementado. O trabalho de Falsarella (2011), por sua vez, expõe os pressupostos políticos, teóricos e metodológicos da avaliação de políticas públicas no campo da educação. Como enfoque nessa análise, a autora explora as mudanças necessárias para os atores ligados à política.

Os dois estudos supracitados apresentam motivação científica e visão restrita acerca da avaliação e seguem em direção aos trabalhos de Fernandes e Gremaud (2009), reiterando que as avaliações das políticas educacionais devem considerar outros atores e instituições educacionais aos resultados das políticas educacionais.

Assim, pode se inferir que os estudos avaliativos das políticas de educação demonstram uma tendência à responsabilização e à visão de que as políticas educacionais são arenas nas quais cabem a atuação de múltiplos atores.

Essa preocupação pode ser ilustrada pelo

trabalho de Abrucio (2016) que realiza um diagnóstico da formação docente no Brasil e destaca estrategicamente, com base nesse processo, perspectivas para a modificação do quadro atual da educação básica pública no Brasil.

Ao considerar o papel dos demais atores e especificamente dos professores no desempenho das estudantes, a dificuldade de se avaliar a política pública e principalmente de obter uma relação causal entre a política e os seus resultados se torna ainda mais difícil. Tal dificuldade decorre principalmente de um problema de agência, como destacam Shapiro e Stiglitz (1984), uma vez que o empregador (Estado – Principal) não obtém informações completas acerca da atuação do seu empregado (Professor – Agente) devido às dificuldades de monitoramento do esforço e à sua dedicação (Bresolin, 2014).

A questão relatada no parágrafo anterior vai ao encontro das preocupações atuais dos *policy makers* da área educacional no Brasil, que buscam a melhoria dos indicadores educacionais. Além de analisar comparativamente as condições educacionais entre os entes federados, subsidiam outras políticas balizadas por esses indicadores, entre elas a política de premiação por bônus (Bresolin, 2014), que se faz necessária principalmente devido à dificuldade de monitoramento de determinadas características do professor, conforme mencionado anteriormente.

O uso de testes como indicador de resultados de aprendizagem dos estudantes é parte crucial dos sistemas de premiação por desempenho, pois representa um aspecto bastante controverso no processo avaliativo de programas e políticas. Por se tratar de

um instrumental restrito e passível de erros, muito se reflete e discute sobre essa questão, o que ainda causa discordância.

Todavia, a avaliação das políticas educacionais com base em indicadores de desempenho oriundos de provas padronizadas, de modo a contribuir com o ciclo da política pública, deve ir além das verificações de impacto e avaliar aspectos tais como a modificação dos comportamentos e das atitudes das equipes, verificando problemas que possam ser visualizados na implementação da política (Marsh et al, 2011).

Semelhantemente, Falsarella (2011) discute que uma concepção totalizante de avaliação educacional deve incorporar uma gama de procedimentos, acompanhando o ciclo da política pública e permitindo uma visão sistêmica, de acordo com o que argumenta Weiss (1998). A mesma autora aponta que a instalação de uma cultura de avaliação em detrimento de ações de caráter pontual e com efeito reduzido poderia propiciar o aprimoramento da educação.

METODOLOGIA

Este estudo faz uma análise descritiva do processo de avaliação da política de pagamento de bônus por desempenho em Nova York, ressaltando os diversos critérios considerados e a relevância desse tipo de processo. Quanto ao método, a pesquisa caracteriza-se como estudo de caso. A escolha dessa estratégia baseia-se nos destaques de Yin (2001), de que se trata de uma estratégia adequada em situações nas quais o foco da pesquisa se encontra em fenôme-

nos contemporâneos sob os quais o pesquisador detém baixo ou nenhum controle.

Realizou-se uma análise de documentos governamentais, relatórios de avaliação e artigos científicos que abordaram tal temática. A pesquisa pode ser classificada como documental, uma vez que se pauta em trabalhos já publicados. A principal fonte de informações foi o relatório RAND Corporation, considerada a avaliação mais completa sobre a política de pagamento por bônus em Nova York.

A fim de verificar os desdobramentos do resultado dessa avaliação, foram realizadas pesquisas sobre o tema tanto no Google, Google Scholar, Scopus, Web of Science, Science Direct e no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), buscando encontrar informações veiculadas na mídia acerca dessa avaliação específica, além de outras avaliações e análises do desdobramento dessa avaliação que foram alvo de problematização por parte da academia. Por esse motivo, alguns filtros foram aplicados, analisando aqui somente as informações publicadas pelos dez principais jornais norte-americanos, tendo por base a tiragem. (Mais informações podem ser encontradas em: <www.cision.com/us/2014/06/top-10-us-daily-newspapers>). Em relação aos artigos analisados, somente foram fonte de dados para este estudo aqueles que abordaram especificamente a política avaliada. No quadro 1 foram sistematizadas as informações acerca dos procedimentos metodológicos realizados.

Quadro 1. Síntese dos procedimentos para a coleta de dados

Informação	Descrição
Documentos que compuseram a análise	Relatórios de avaliação realizados por instituições de pesquisa ou consultorias especializadas. Notícias e artigos científicos publicados sobre a temática
Fonte	Google, Google Scholar, Web of Science, Scopus, Science Direct e todas as bases que compõe o Portal de Periódicos da Capes
Filtros	Foram analisadas as notícias publicadas nos principais jornais publicados nos EUA no período que vai desde janeiro de 2010 a julho de 2017. Os artigos científicos foram analisados durante o mesmo intervalo de tempo. Foram verificados os artigos científicos no que se refere a avaliação das políticas de bônus educacionais (e suas derivações) em português para dar subsídio as análises do contexto brasileiro; Acerca do caso de Nova York foi buscado nas bases o <i>termo Evaluation of bonus payment policy for education in New York</i> (suas derivações). Visando evitar a repetição de informações foram priorizados os trabalhos que possuíam um maior número de citações.

As fontes de dados do caso aqui analisado foram os trabalhos de Gall e Guedes (2009), Lasagna (2010), Marsh et al. (2011), Goodman e Turner (2010, 2011, 2013) e o trabalho de Fryer (2013), além do material não acadêmico de autoria do NYT (2011). Ressalta-se que a base principal dos dados foi o relatório da avaliação realizada pela RAND Corporation de autoria de Marsh et al. (2011). Os demais artigos analisados serviram para ampliar a análise da avaliação da política de

bônus para professores em Nova York.

A partir da obtenção das informações, elas foram analisadas de maneira descritiva de modo a embasar a análise das discussões sobre a realização das avaliações nas políticas de pagamentos de bônus aos professores, permitindo, com base no referencial teórico, realizar reflexões e incitar questionamentos acerca das avaliações para o setor aqui analisado, destacando a avaliação rea-

lizada pela RAND Corporation referente ao caso de Nova York.

RESULTADOS E DISCUSSÕES

Seguindo os procedimentos metodológicos apresentados, analisou-se, de acordo com a teoria de avaliação de políticas públicas, a avaliação contratada pelo Fundo da Cidade de Nova York para Escolas Públicas realizada pela RAND Corporation em parceria com a Universidade Vanderbilt e com o Centro Nacional de Incentivo em Performance (NC-PIN).

Optou-se por sua análise em razão de sua amplitude e por ter sido avaliada por avaliadores independentes e, principalmente, por ser uma política implementada em larga escala no Brasil, entretanto, carente de estudos avaliativos, daí ser utilizada como *benchmarking* e incentivador da realização de avaliações. Além disso, o aprofundamento realizado neste estudo permite analisar a avaliação pela ótica da teoria e, dessa forma, avançar no estado da arte do campo.

Programa de bônus por performance escolar

A Reforma Educacional de Nova York, chamada *Children First*, foi estruturada com o intuito de superar diversos obstáculos em relação ao aprimoramento do ensino e da aprendizagem em instituições educacionais públicas, após décadas de tentativas inócuas no sentido de reverter uma cultura de fracasso nas escolas da cidade.

Essa reestruturação, liderada pelo prefeito de Nova York, Michael Bloomberg, e pelo secretário de Educação, Joel Klein, foi ini-

ciada em 2002. Entre os diferentes temas abarcados no processo da reforma educacional, que produziu mudanças na rede pública de ensino da cidade, foram incluídos a descentralização e autonomia escolar, o monitoramento dos processos de aprendizagem e a responsabilização pelos resultados. Além do processo de formação e apoio aos professores em sala de aula, aproximação entre a escola e a família, entre outros (Gall & Guedes, 2009).

Após debates que se estenderam por um longo período, a Secretaria de Educação de Nova York (NYCDOE – New York City Department of Education) e a Federação Unida de Professores (UFT) implantaram conjuntamente o Programa de Bônus por Performance Escolar (SPBP) no ano escolar de 2007-2008. O desenho do projeto contemplou a elaboração, a implementação e a avaliação de uma política de incentivos por desempenho aos profissionais do setor, com duração inicial prevista para dois anos, 2007-2008 e 2008-2009, que acabou sendo prolongado por mais um período, de 2009-2010, como resultado de negociações adicionais posteriores entre as duas instituições.

Concebido para escolas extremamente carentes e vulneráveis, com baixo nível de proficiência dos alunos, indicadores elevados de pobreza e situações sociodemográficas preocupantes, o Programa de Bônus foi implantado como uma iniciativa piloto, em pequena escala, de modo a possibilitar a avaliação da pertinência de sua expansão. Em um segundo momento, para toda a cidade de Nova York (Lasagna, 2010).

O bônus foi idealizado como recompensa e incentivo para motivar comportamentos de-

sejados por parte dos professores. Criou-se, assim, um instrumento supostamente poderoso que permitia relacionar performance e pagamento.

Lasagna (2010) ressalta que o Programa teve como objetivos prioritários: aumentar a aprendizagem dos alunos por meio da criação de incentivos e recompensas para professores de escolas em situação de elevada vulnerabilidade e baixa performance, instituir um sistema de recompensa profissional para os professores, criar um clima de colaboração entre os docentes e estimular uma mudança de comportamento nas práticas de ensino, além de atrair e reter educadores qualificados.

O foco central foi dado à melhoria da aprendizagem dos estudantes. O sucesso da política seria mensurado de maneira objetiva, considerando as notas obtidas pelos estudantes em provas padronizadas de matemática e língua inglesa, o nível de frequência às aulas e a adequação do ambiente escolar à aprendizagem. Consequentemente, objetivos intermediários também seriam atingidos: retenção de professores altamente qualificados, aumento de colaboração e cooperação entre os funcionários da escola e melhoria do nível motivacional dos professores.

A adesão ao programa era feita de forma voluntária pela unidade escolar desde que pelo menos 55% de seus representantes na UFT escolhessem participar. No primeiro ano, 427 escolas que apresentavam as condições para elegibilidade ao programa se inscreveram e aproximadamente metade foi selecionada aleatoriamente para participar. Após desistências e ajustes na seleção, o

programa foi implantado com a participação de 205 escolas. No decorrer do tempo, esse número caiu para 198 escolas no segundo ano e para 196 no terceiro.

Em contrapartida, cada escola selecionada para participação criou um Comitê de Compensação (CC) com a atribuição de determinar a maneira como o valor total de bonificação recebido pela instituição seria distribuído entre os participantes. O Comitê era composto pelo diretor da escola, por um integrante indicado pelo diretor e por outros dois participantes eleitos pela equipe escolar que fossem representantes da UFT, e teria autonomia completa para a definição do critério de divisão do bônus recebido pela escola. A única restrição imposta foi não usar critério de senioridade para a distribuição.

Fundamentado em medidas fixas de performance, o bônus foi concedido com base no estabelecimento de níveis referenciais – metas – para desempenho da escola no Boletim de Progresso Escolariv (School Progress Reports – RP) da Secretaria de Educação, elaborado para cada distrito escolar. O índice criado para a medida do desempenho foi composto pela performance e pelo progresso dos estudantes, pelo ambiente escolar e por um critério de pontuação adicional por exemplaridade. O cálculo final da pontuação no RP foi baseado na pontuação total da escola em relação a todas as outras instituições com população estudantil semelhante.

De acordo com Marsh et al. (2011), o número de escolas que receberam o bônus cresceu substancialmente entre os dois primeiros anos, passando de 62% das instituições participantes, sendo contempladas com bônus integral ou parcial 84%. Já no último ano,

como resultado de uma calibragem das metas nos testes de proficiência dos alunos, as notas das escolas caíram generalizada-mente, provocando uma redução de 13% no percentual de participantes que receberam o bônus no terceiro ano.

A avaliação – desenho geral e principais resultados

Em 2009, o Fundo da Cidade de Nova York para Escolas Públicas contratou a RAND Corporation para, em parceria com a Universidade Vanderbilt e com o Centro Nacional de Incentivo em Performance (NCPI), avaliar de maneira independente a implementação e a efetividade do SPBP. Com esse propósito, diversos procedimentos metodológicos foram adotados para a coleta de informações, tais como: revisão de documentos relacionados ao programa; análise de dados administrativos da secretaria municipal de educação; pesquisas com os integrantes dos CC, com os professores das escolas participantes do programa (grupo de tratamento), com os professores de escolas inscritas não contempladas pelo sorteio (grupo de controle) e também com membros do NYCDOE e líderes da UFT, e ainda entrevistas em 14 escolas escolhidas para representarem estudos de caso.

A avaliação dos efeitos do programa no desempenho escolar dos alunos foi feita a partir de diversas análises e estimativas dos dados do PR, incluindo os resultados acadêmicos dos estudantes e os dados demográficos dos dois grupos de escola desde o ano escolar 2006-2007 até 2009-2010.

O relatório final da avaliação do programa, intitulado *A Big Apple for Educators – New*

York City's Experiment with Schoolwide Performance Bonuses foi publicado em julho de 2011 (Marsh et al., 2011). Utilizando uma combinação de métodos quantitativos e qualitativos, a avaliação foi dividida em três grandes blocos temáticos conforme detalhamento a seguir:

Em relação ao *timing*, a avaliação foi feita após a implementação do programa iniciado em 2002. Nesse caso, foi realizado uma avaliação *ex-post*, a qual enfatiza a credibilidade e a objetividade dos achados e estudam-se a eficácia e os valores do programa como mostram Ala-Harja e Helgason (2000) e Trevisan e Van Bellen (2008). A posição do avaliador segue o que preconiza Arretche (1998): essas avaliações precisam ser realizadas por órgãos independentes dos que implementaram a política contribuindo para garantir uma avaliação sem vieses.

Aprendizagens sobre o processo avaliativo e o emprego dos resultados em Nova York

A avaliação do programa, documentada no estudo de Marsh et al. (2011), foi realizada por uma instituição externa contratada para conduzir esse processo. Constatou-se a ausência dos efeitos inicialmente esperados tanto para variáveis intermediárias, relacionadas ao comportamento dos professores, quanto para a meta principal, referente à performance acadêmica dos estudantes.

As conclusões a respeito da implementação do programa indicam tanto aspectos promissores quanto problemas incidentes. Alguns fatores importantes para o sucesso de um programa de pagamento por performance não foram devidamente encontrados em todas as escolas participantes em decorrên-

cia de problemas relacionados à implementação que podem ter enfraquecido o efeito do bônus e do programa em geral, o que reduziu o estímulo motivacional esperado. Entre os principais itens, foram ressaltados o entendimento equivocado a respeito do programa – especialmente no que diz respeito ao desentendimento sobre a mensuração dos resultados –, a incerteza com relação à capacidade dos professores de modificar suas práticas profissionais ou atuar de maneira a aumentar as chances de atingir as metas propostas, os questionamentos a respeito dos processos decisórios, a preocupação com os tempos de implementação e a duração do programa.

Melhorias no processo de implementação do programa poderiam ser realizadas, mas isso não garantiria um resultado efetivo. Motivar indivíduos, mais especificamente no ambiente escolar público, com recompensa financeira, é um desafio complexo, por haver uma resistência natural ao reconhecimento do estímulo e pela possível ausência motivacional gerada por um instrumento parcial e imperfeito.

Em termos de efeitos intermediários, identificados pelos próprios participantes de acordo com seus comportamentos, a avaliação constatou que o programa não os afetou pessoalmente, nem modificou suas performances profissionais, o nível de conhecimentos, as motivações e a maneira de se relacionar com os profissionais envolvidos. O bônus não foi compreendido como uma variável capaz de influenciar tais aspectos. Diferentes demandas por informações públicas a respeito do trabalho nas escolas têm contribuído para que gestores, professores e outros integrantes da equipe escolar se-

jam colocados como corresponsáveis pelo nível do desempenho educacional dos estudantes. Uma forte cobrança gerada por diversas iniciativas governamentais pode ter se sobreposto ao programa de bônus no caso de Nova York, reduzindo a sua importância, pois resultados parecidos podem decorrer de outras formas de estímulos e pressões políticas, também presentes nas escolas de controle.

É provável que algumas percepções sobre mudança motivacional tenham sido equivocadamente atribuídas ao programa por decorrerem de outros estímulos e não em razão do bônus. Atingir uma pontuação elevada no programa, por exemplo, pode ser uma variável motivacional em si, também presente nas escolas não participantes. É possível que mudanças na prática dos professores levem mais do que três anos para produzir melhoria no desempenho dos alunos.

No entanto, a ausência de impacto significativo no comportamento dos professores, concebida como etapa intermediária do efeito final almejado, reduz a plausibilidade desse argumento. Sendo assim, as evidências com base na avaliação estruturada não suportaram a continuidade do programa de bônus por performance, uma vez que não atingiu os objetivos a que se propunha.

É possível verificar a relação entre os objetivos e metas da política e os resultados auferidos pela avaliação, mostrando a realização de uma avaliação absoluta (Costa & Castanhar, 2003). Apesar de ter realizado um estudo dos três anos do programa, uma análise com esse intervalo temporal pode ser insuficiente para uma mudança que impacta a cultura organizacional, principalmente para

as estruturas de recompensa e motivação dos professores apresentarem efeitos em relação aos resultados educacionais.

Apesar dessas questões, a avaliação foi utilizada segundo a classificação de Faria (2005). Tendo como referência as principais conclusões e recomendações fundamentadas nesse processo avaliativo, foi anunciado em julho de 2011 que o programa seria descontinuado (Carvalho, 2011).

Segundo reportagem do The New York Times (2011), o estudo realizado pela RAND Corporation foi essencial para a decisão de encerrar o programa: “A decisão foi tomada à luz de um estudo que descobriu que os bônus não tinham nenhum efeito positivo no desempenho dos alunos ou nas atitudes dos professores em relação aos seus empregos” (NYT, 2011, p. 1, tradução nossa).

O que confirma a importância da avaliação, de acordo com o que aponta Trevisan e Van Bellen (2008), é que, na etapa da avaliação, se determina, de acordo com o impacto, o seguimento a ser dado ao ciclo político: ser suspenso caso não haja atendimento dos objetivos ou levar ao início de um novo ciclo. Todavia, os apontamentos de Cohen e Franco (2007) mostram que a avaliação não deve ser julgada como autossuficiente, dado que faz parte de um processo complexo de planejamento da política. Dessa forma, o emprego da avaliação para se definir a continuidade ou não de uma política deve ser feito com parcimônia, como no caso de Nova York, segundo o porta-voz do departamento de Educação de Nova York na matéria supracitada veiculada pelo The New York Times (2011, p. 1, tradução nossa): Este estudo confirma que foi a decisão certa e nos

fornece informações importantes [...]. Esse recorte da fala demonstra o papel de produtora de informação para as decisões governamentais da avaliação como conceituado por Souza (2015).

A avaliação nova-iorquina e os insights para as políticas de bônus no Brasil

Diversos governos estaduais brasileiros já elaboraram projetos e leis para o estabelecimento de programas de bônus voltados aos docentes e profissionais ligados à educação fundamentados em metas de desempenho dos alunos da rede. É conhecida a existência de programa de bônus para servidores na área educacional nos estados brasileiros: Amazonas, Ceará, Espírito Santo, Minas Gerais, Pernambuco, Rio de Janeiro e São Paulo (Brook, 2011).

As políticas de remuneração variável por desempenho em educação no Brasil apresentam desenhos variados quanto às formas de medir o desempenho, seja do professor, seja da escola, ao critério de desempenho dos alunos, à natureza dos indicadores escolares considerados, à distribuição da bonificação, para o indivíduo ou para a escola, para todos da equipe ou excluindo o pessoal de apoio, à maneira de contabilizar faltas previstas e não previstas legalmente, à forma de pagamento, anual ou mensal, ao estabelecimento do valor pago, entre outras variáveis (Pontual 2008).

Será que essa estratégia de incentivo é adequada para estimular a melhora no desempenho escolar dos alunos? Sob quais condições esse mecanismo pode, ou não, funcionar? Existe um modelo que seja mais efetivo ao objetivo que se pretende? Qual o

desenho que gera mais resultados? Sobre a infinidade de questionamentos possíveis que venham a surgir na implementação de uma política pública ou de um programa, as questões consideradas na avaliação devem se atentar aos objetivos, já que, como aponta Arretche (2001), seria fora de propósito que o avaliador se balizasse em objetivos ou metodologias externas ao próprio programa. Todavia, a situação brasileira não reflete tal apontamento, pois, como destaca Cohen (2001) e Arretche (2001), em alguns cenários, como no caso do Brasil, a vontade do gestor e o ambiente eleitoral influenciam na forma como a avaliação será feita. Tais questões podem ser responsáveis pelos aspectos que discutiremos a seguir sobre o panorama brasileiro envolvendo os fenômenos analisados neste estudo.

A escassez de avaliações verificadas no Brasil não permite saber se os programas desenhados para estimular a melhoria no desempenho dos alunos estão trazendo de fato uma mudança nos indicadores educacionais, ou seja, no impacto do programa. Para verificar isso, seria necessário compreender se a oferta de uma recompensa financeira adicional para servidores públicos que trabalham na área da educação, condicionada a variáveis preestabelecidas, como notas dos alunos em provas padronizadas, contribui para promover um desempenho diferenciado do professor a ponto de os alunos aprenderem mais.

Assim sendo, não é possível saber se a política de pagamentos por bônus na educação tem sido efetiva. A análise de efetividade, segundo destaca Carvalho (2003, p. 186): “é aquela que focaliza os efeitos ou impactos produzidos sobre a sociedade e, portanto,

para além dos beneficiários diretos da intervenção pública, avaliando-se sua efetividade social”. Essa análise foi feita no caso de Nova York, tendo por base a nota apresentada no RP. Todavia, os recortes geográficos estabelecidos na avaliação de Nova York, bem como as características homogêneas das regiões analisadas, favorecem esse tipo de análise. No Brasil, algumas variações dificultam esse tipo de avaliação, dado que nem tem todas as variáveis que interferem no desempenho dos alunos são mensuradas.

Destacam-se entre os trabalhos publicados no Brasil até o momento, com foco no impacto gerado por esse tipo de política, a pesquisa de Oshiro e Scorzafava (2011) e a de Bresolin (2014). O baixo número de avaliações realizadas vai ao encontro do que foi abordado por Calmon e Serpa (2009), como destacam os autores, esse tipo de estudo é feito de forma tardia e tem na academia um grupo de interesse na sua realização.

O estudo de Oshiro e Scorzafava avaliou o programa implementado no estado de São Paulo em 2008, que visava remunerar funcionários e professores da rede que cumpriam metas predeterminadas pela Secretaria de Educação do Estado de São Paulo na 4a série do ensino fundamental, na 8a série do ensino fundamental e no 3o ano do ensino médio. No caso dessa política, as metas foram estabelecidas fundamentadas nos indicadores sobre o fluxo escolar – taxa média de aprovação em cada ciclo da escolarização – e o desempenho dos alunos – nota na proficiência em Matemática e Português do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp).

A avaliação utilizou métodos econométricos

(*propensity score matching* e diferenças em diferenças) para mensurar o impacto do programa no desempenho dos alunos em Matemática e Português na Prova Brasil na 4a e 8a séries do ensino fundamental. Com o intuito de atribuir causalidade entre a política e o impacto nas notas que ela busca promover, foi considerado como grupo de tratamento as escolas estaduais de São Paulo submetidas à política, enquanto o grupo de controle foi composto por escolas com características semelhantes, mas pertencentes à outra rede (municipal paulistas) em que políticas dessa natureza não haviam sido implementadas.

O estudo identifica um impacto positivo e estatisticamente significativo na proficiência dos alunos da 4a série mas não encontra impacto na proficiência dos alunos da 8a série. Os próprios autores reforçam a necessidade de estudos complementares para compreender os motivos da diferença de impacto entre as séries e suprir a carência de investigação de como os agentes envolvidos avaliam a política. Elaborado em ambiente acadêmico, a avaliação foi feita com base em dados secundários e superou dificuldades decorrentes de não ter sido planejada antes da implementação da política, nem em parceria com a área pública.

A pesquisa de Bresolin (2014) soma-se aos estudos relatados nos parágrafos anteriores e adiciona alguns componentes na avaliação das políticas de pagamento por bônus na educação no Brasil. Ao analisar o efeito da política em um número maior de Estados brasileiros, o autor permite conhecer seu efeito em distintos contextos.

O estudo analisa o efeito da política em três

dimensões: “Práticas pedagógicas dos professores ou ações ligadas ao ensino”, “Absentismo e comprometimento de professores ou visão do diretor e professor sobre a assiduidade e rotatividade do professor” e “Interlocução com familiares para garantir a frequência escolar dos alunos ou ações ligadas à permanência dos alunos na escola”. Os resultados do estudo demonstram que a política gera mudança em alguns indicadores intermediários, porém, a ausência de impacto em um grande número de indicadores, como destaca o autor, deixa sem resposta a questão caracterizada como fundamental:

O estímulo financeiro proposto pelas políticas analisadas é suficiente para estimular mudanças nos comportamentos dos professores e diretores que proporcionem resultados intermediários na melhora da aprendizagem dos alunos? (Bresolin, 2014 p. 66)

Ao não responder a esse questionamento, torna-se impossível elucidar a principal questão relacionada à política de incentivo: “A política tem sido suficiente para contornar os problemas de agência na relação entre governos e professores?”. Questão de vital importância considerando o papel do professor para o melhor aprendizado do aluno, como comprovaram Rivkin, Hanushek e Kain (2005).

As avaliações apresentadas anteriormente tiveram sua concepção e execução sem qualquer articulação ou interlocução com a secretaria executora da política analisada, de tal forma que os resultados encontrados não tiveram qualquer desdobramento ou uso prático por parte dos gestores públicos, o que também pode ter contribuído para a

dificuldade de encontrar dados que permitissem elucidar a totalidade das questões.

Diferentemente do caso de Nova York, as avaliações realizadas no Brasil, ficaram restritas a fins acadêmicos. A falta de notícias acerca da sua realização e de seus efeitos nas decisões nos permite inferir que as avaliações não cumpriram seus objetivos principais. No caso de Nova York, a avaliação obteve ampla divulgação, permitindo inclusive que os cidadãos fizessem uso das informações nos processos democráticos, além de fornecer subsídio ao processo de tomada de decisões, dado que os resultados foram os principais responsáveis pela descontinuidade do programa. Resgatando o que aponta Arretche (2001), a respeito da necessidade de a avaliação ser realizada com o intuito de buscar a reestruturação do programa/política seja durante sua execução, seja posteriormente, no caso brasileiro, essa função não foi cumprida, situação oposta à verificada em Nova York.

A avaliação realizada em Nova York antes mesmo de seu início já demonstrava a preocupação dos gestores em avaliar o andamento da política, não somente sobre a ótica do impacto, mas também sobre os determinantes que levaram a seu fracasso ou sucesso. No Brasil, as avaliações encontradas foram feitas por demanda de atores com pouca possibilidade de interferir no ciclo político. A avaliação nos moldes das realizadas no Brasil não atentaram aos determinantes do sucesso ou fracasso, ficando restrita à análise da efetividade, desviando de aspectos como a efetividade e a eficiência, questões que, como defende Arretche (1998), devem estar contidas nas avaliações.

Cumprir destacar, porém, que esse cenário tem evoluído, de acordo com duas pesquisas discutidas nessa seção. O estudo de Bresolin (2014) apresenta os resultados intermediários da política e faz uma análise mais ampla, buscando uma visão sistêmica da política e permitindo inferir em certa medida acerca de fatores que podem estar contribuindo para os resultados verificados. Como essas pesquisas analíticas foram realizadas após a implementação da política que não contempla a possibilidade de avaliação, ela se torna um limitador dos resultados. Para corrigir tal gargalo, o destaque de Falsarella (2011) acerca da necessidade de se criar uma cultura de avaliação ganha ainda mais força.

Para que as avaliações realizadas no Brasil alcancem o efeito do caso aqui estudado, é fundamental que elas sejam realizadas levando em conta o contexto político, suas inspirações e os propósitos ao longo de sua implementação (Weiss, 1998; Souza, 2015). O caso de Nova York pode trazer uma ampla gama de insights que serão fundamentais para a consolidação de uma cultura avaliativa no Brasil, uma vez que começará a modificar as questões avaliadas (Cohen, 2001), além do interesse que move as avaliações ou a falta delas no país (Arretche, 2001; Cohen, 2001).

CONCLUSÕES

Tem sido cada vez mais consensual a legitimidade em reconhecer, prestigiar e premiar profissionais que apresentam resultados positivamente diferenciados, assim como a necessidade de apoiar quem não consegue avançar, ainda que se esforce para isso, ou punir os que não têm motivação nem se em-

penham para aprimorar-se. Porém, é preciso encontrar a melhor maneira de fazer isso, uma vez que não há uma fórmula que aumente a probabilidade de atingir esses objetivos.

É fundamental que diversas questões avaliativas sejam formuladas, medidas, planejadas, acompanhadas, mensuradas e analisadas para se chegar a uma conclusão a respeito dos resultados atingidos e dos impactos gerados por um programa de incentivo financeiro na área da educação pública. Considerando a importância da avaliação na política de pagamento por bônus e com o intuito de contribuir com as reflexões teóricas e práticas acerca do tema, mais especificamente na educação, foi realizada uma análise profunda do caso de Nova York. Nessa direção, verificamos que a experiência sobre política de bônus, como a promovida em Nova York, mostra diversos aspectos relacionados à avaliação de políticas públicas dessa área. Nesse caso, destaca-se que o programa começou com uma articulação política diferenciada estabelecida em parceria com a secretaria municipal de educação e com o sindicato de professores, de tal forma que essa parceria em si representa uma situação diferenciada, dada as inúmeras dificuldades e a baixa frequência desse tipo de colaboração entre governo e sindicato.

Essa experiência contribuiu para reforçar a importância de projetos pilotos menores, de caráter experimental, associados ao processo de desenho, implementação, monitoramento e avaliação sistemáticos capazes de permitir correções de rumo durante o processo e uma rigorosa análise dos resultados antes de se disseminar a inovação com resultados desconhecidos em larga escala.

As escolhas metodológicas feitas para a condução do processo avaliativo possibilitaram a análise de diversos aspectos referentes à implementação do programa, a seus efeitos intermediários na motivação e no comportamento dos professores e ao objetivo prioritário de melhoria na proficiência dos alunos.

A combinação de diferentes processos e metodologias, com a devida ponderação a respeito das dificuldades e limitações enfrentadas, gerou resultados relevantes para a gestão do programa. Apesar de também não haver consenso nem fórmulas preestabelecidas a respeito da maneira ideal de avaliar políticas de incentivo por performance. O caso do SPBP apresenta pontos de reflexão relevantes.

Verificamos ainda que os aspectos destacados no referencial teórico acerca dos possíveis vieses ocorridos no processo de avaliação, principalmente os de cunho eleitoral, não foram verificados no caso de Nova York. O direcionamento do estudo, principalmente no que se refere à relação entre os objetivos da política e a forma como o programa/política, foi avaliado em consonância com o que literatura estabelece como sendo o correto nesse processo.

No Brasil, temos vivenciado uma proliferação de políticas de bônus implementadas por diversas secretarias estaduais de educação. A maneira pela qual tais programas foram desenhados não demonstra qualquer preocupação com um processo de monitoramento e avaliação. Raras são as experiências avaliativas que, quando existem, utilizam métodos específicos, com interpretações limitadas, e que não foram deman-

dadas pelos gestores, de tal forma que os resultados não proporcionam qualquer uso para aprimoramento ou reflexão a respeito da política vigente.

É passada a hora dos governantes brasileiros incluírem a avaliação das políticas na agenda, para que essa ferramenta efetivamente contribua para o aprimoramento das ações desenvolvidas.

A impossibilidade de entrar em contato com os atores envolvidos no caso aqui analisado inviabilizou uma análise mais apurada, contudo, as análises e discussões apresentadas possibilitaram o cumprimento do objetivo estabelecido.

Uma pesquisa de maior profundidade que, entre outros pontos, busque analisar a relação desses atores com o processo de avaliação se caracteriza como sendo um novo ponto a ser posteriormente explorado.

REFERÊNCIAS

Abrucio, F. (2016). *Formação de professores no Brasil*: diagnóstico, agenda de políticas e estratégias para a mudança. São Paulo: Moderna.

Ala-Harja, M., & Helgason, S. (2000). Em direção às melhores práticas de avaliação. *Revista do Serviço Público*, Brasília, 51 (4), pp. 5-59.

Arretche, M.T.S. (1998). Tendências nos Estudos sobre Avaliação, In: RICO, E. M. *Avaliação de Políticas Sociais*: uma questão em debate. São Paulo: Cortez.

_____. (2001). Uma contribuição para fa-

zermos avaliações menos ingênuas. In: Moreira, M. C. R., & Carvalho, M. C. B. (Org.). *Tendências e Perspectivas na Avaliação de Políticas e Programas Sociais*. São Paulo: IEE/PUCSP, 1, pp. 7-224.

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., & Wilson, D. (2009). Evaluating the impact of performance-related pay for teachers in England. *Labour Economics*, 16 (3), pp. 251-261.

Bresolin, A. B. (2014). *Análise de resultados intermediários das políticas de bônus em escolas públicas estaduais brasileiras*. Dissertação – Mestrado em Administração Pública e Governo – EAESP/ FGV. São Paulo

Brooke, N. (2011). *A Avaliação Externa como Instrumento da Gestão Educacional nos Estados*. São Paulo: Fundação Victor Civita.

Bruns, B.; Ferraz, C. (2011) *Paying teachers to perform*: the effects of bonus pay in Pernambuco, Brazil. World Bank power point presentation, 7 de dezembro. Washington, DC. Recuperado em 12 jul. 2017, de http://siteresourcesworldbank.org/EDUCATION/Resources/ppt_paying_teachers_Dec.7.2011.pdf

Calmon, P. C. Du P., & Cerpa, S. M. H. C. (2009). Afinal de contas, por que avaliar? Resgatando o debate sobre utilização na avaliação de programas governamentais. In: 33o Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração (EnANPAD), 2009, São Paulo... *Anais 33 Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração*.

Carvalho, R. (2015). Política. *Carta Capital*.

Recuperado em 17 abr. 2018, de www.cartacapital.com.br/politica/copia-fiel

Ceneviva, R., & Farah, M. F. S. (2012). Avaliação, informação e responsabilização no setor público. *Revista de Administração Pública*, 46 (4), pp. 993-1016.

Cohen, E., & Franco, R. (2007). *Avaliação de Projetos Sociais*. (2. ed.). Petrópolis: Vozes.

Cohen, E. (2001). Reforma del Estado, modernización administrativa y evaluación del desempeño de la gestión. En: *Los desafíos de la reforma del Estado en los programas sociales: tres estudios de caso-LC/L. 1469-P-2001-p. 9-22*.

Costa, F. L. da, & Castanhar, J. C. (2003). Avaliação de programas públicos: desafios conceituais e metodológicos. *Revista de Administração Pública*, Rio de Janeiro, 37 (5), pp. 962-969.

Carvalho, S. N. de. (2003). Avaliação de programas sociais: balanço das experiências e contribuição para o debate. *São Paulo em Perspectiva*, São Paulo, 17 (3-4), pp. 185-197.

Cotta, T. C. (2001). Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb). *Revista do Serviço Público*, Brasília, 52 (4), pp. 89-110.

Falsarella, A. M. (2011). Avaliação das políticas públicas educacionais no Âmbito escolar. In: 25o Simpósio Nacional da Associação Nacional de Política e Administração na Educação - ANPAE, 2011... *Anais 25 Simpósio Nacional da Associação Nacional de Po-*

lítica e Administração na Educação.

Faria, C. A. P. (2003). Idéias, conhecimento e políticas públicas: um inventário sucinto das principais vertentes analíticas recentes. *Revista Brasileira de Ciências Sociais*, São Paulo, 18, (51), pp. 21-29.

Fernandes, R., & Gremaud, A. P. (2009). *Qualidade da educação: avaliação, indicadores e metas*. Educação básica no Brasil: construindo o país do futuro. Rio de Janeiro: Elsevier, 1, 213-238.

Fernandes, C. do C. P., Silveira, S. de F. R., Vieira, L. H. S., & Quirino, S. das G. (2008). Representações Sociais e Políticas Educacionais no Universo da Gestão Pública Municipal: Um estudo de Multicasos. In: Encontro de Administração Pública e Governança, 3, Salvador, *Anais*. EnAPG, CD.

Figlio, D., & Loeb, S. (2011). School Accountability. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, (edt): *Handbooks in Economics*, v. 3, The Netherlands: Ludger Woessmann.

Frey, K. (2000). *Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no Brasil*. Planejamento e Políticas Públicas, Brasília, (21), pp. 211-259.

Fryer R. (2013). Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. *Journal of Labor Economics*, 31 (2), pp.373-427.

Gall, N., & Guedes, P. (2009). *A Reforma Educacional de Nova York: possibilidades para o Brasil*. São Paulo, Instituto Fernand

Braudel.

Goodman, S., & Turner, L. (2011). Does whole-school performance pay improve student learning?. *Education Next*, 11(2).

Goodman, S., & Turner, L. (2010). Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program. Program on Education Policy and Governance Working Papers Series. PEPG 10-07. *Program on Education Policy and Governance*, Harvard University.

Goodman, S. F., & Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the New York City bonus program. *Journal of Labor Economics*, 31(2), 409-420.w

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annu. Rev. Econ.*, 6(1), 801-825.

Kolbe, T., & Strunk, K. O. (2012). Economic incentives as a strategy for responding to teacher staffing problems: A typology of policies and practices. *Educational Administration Quarterly*, 48(5), 779-813.

Lasagna, M. (2011). *New York City's School-Wide Performance Bonus Program*: Center for Educator Compensation Reform. U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Lavy, V. (2009). Performance Pay and Teacher's Effort, Productivity, and Grading Ethics, *American Economic Review*, 99 (5), pp.1979-2011.

Leigh, A. (2012). The economics and politics of teacher merit pay. *CESifo Economic Studies*, 59(1), 1-33.

Marsh, J., Springer, M., Mccaffrey, D., Yuan, K., Epstein, S., Koppich, J., Kalra, N., & Dimartino, C. (2011). *A Big Apple for Educators*: New York City's experiment with schoolwide performance bonuses. Final Evaluation Report. Santa Monica, CA: RAND Corporation.

Muralidharan, K., & Sundararaman, V. (2009). Teacher Performance Pay: Experimental Evidence from India. NBER *Working Paper* N. 15323. National Bureau of Economic Research.

Oshiro, C., & Scorzafeve, L. (2011). Efeito do pagamento de bônus aos professores sobre a proficiência escolar no estado de São Paulo. In: 39o Encontro Nacional de Economia, 2011, Foz do Iguaçu... *Anais 39 Encontro Nacional de Economia*.

Pontual, T. (2008). *Remuneração por mérito, desafio para a educação*. São Paulo: Fundação Lemann.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.

Schwartzman, S. (2013). Uses and abuses of education assessment in Brazil. *Prospects*, 43(3), 269-288.

Secchi, L. (2010). *Políticas públicas: conceitos, esquemas de análise, casos práticos*. São Paulo: Cengage Learning.

Serpa, S. M. H. C. (2010). *Para que Avaliar?* Identificando a tipologia, os propósitos e a

utilização das avaliações de programas governamentais no Brasil. 191f. Dissertação de Mestrado em Administração. Universidade de Brasília. Brasília.

Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review*, 74(3), 433-444.

Slavin, R. E. (2010). Can financial incentives enhance educational outcomes? Evidence from international experiments. *Educational Research Review*, 5(1), 68-80.

Souza, L. G. D. (2015). *Avaliação pública de políticas educacionais: concepções e práticas avaliativas dos organismos internacionais no Brasil* Tese de Doutorado. Universidade Federal da Bahia. Programa de Pós-Graduação em Educação. Bahia, Brasil.

Souza, S. Z., & Oliveira, R. P. de. (2010). Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. *Cadernos de Pesquisa*, 40 (141), pp. 793-822.

Trevisan, A. P., & Van Bellen, H. M. (2008). Avaliação de políticas públicas: uma revisão teórica de um campo em construção. *Revista de Administração Pública*, 42 (3), pp. 529-50.

Tymes, The New York. (2011). *New York City Abandons Teacher Bonus Program*. New York. Recuperado em 17 abr., 2018, de www.nytimes.com/2011/07/18/education/18rand.html?_r=1&

Yin, R. (2001). *Estudo de caso: planejamento e métodos* [Case study: planning and methods]. Porto Alegre: Bookman Ed.