

PÁGINA ABERTA

Indicadores para avaliação de programas educacionais: desafios metodológicos

Adriana Bauer^{a,b}
Sandra Zákia Sousa^b

Resumo

Este artigo discute as estratégias metodológicas utilizadas para enfrentar o desafio de estabelecer indicadores para compor a avaliação de impacto de um programa de formação de professores em serviço, o Programa Letra e Vida. Este programa foi implementado na rede de ensino do Estado de São Paulo, no período de 2003 a 2006, com o propósito de aprimorar a prática de professores alfabetizadores atuantes nos primeiros anos do Ensino Fundamental. Antecedendo a discussão procedimental, é apresentada, com base em contribuições da literatura, a noção de indicador, seus tipos e propriedades, bem como os aspectos políticos e técnicos – que estão subjacentes à definição de indicadores – que possam subsidiar a avaliação de programas educacionais.

Palavras-chave: Indicadores educacionais. Avaliação de programas. Formação de professores.

1 Introdução

Os programas educacionais, em uma perspectiva de consecução responsável, demandam ter a avaliação como atividade inerente à sua execução, vindo a subsidiar decisões aos seus proponentes e participantes, com vistas ao seu contínuo aprimoramento.

Entretanto, a construção de propostas de avaliação, até mesmo pelos condicionantes de ordem política, educacional, financeira, material e humana, não é uma

^a Fundação Carlos Chagas. São Paulo, São Paulo, Brasil.

^b Faculdade de Educação da USP. São Paulo, São Paulo, Brasil.

Recebido em: 25 Nov., 2013

Aceito em: 25 Ago., 2014

tarefas simples. Falamos em construção, pois, embora contando com relevantes contribuições, nacionais e internacionais, de teóricos da área, em realidade, cada situação a ser avaliada demanda um delineamento próprio, capaz de responder ao propósito da avaliação, em dado contexto e com dadas especificidades e características (SOUSA, 2006).

No Brasil, além da pouca tradição de avaliação de programas educacionais, quando esta ocorre, muitas vezes, reveste-se de um caráter formal, não se observando a valorização de seus resultados para revisão/reformulação das propostas e ações que constituem objeto da avaliação. Aproxima-se do que Barry MacDonald (1982, p. 16) denomina de avaliação burocrática, ou seja:

Avaliação burocrática constitui-se num serviço incondicional prestado às agências governamentais que têm maior controle sobre a alocação de verbas educacionais. O avaliador aceita os valores daqueles que o contratam e fornece-lhes informações que os ajudam a atingir seus objetivos políticos. Ele age como um consultor administrativo e seu critério de sucesso é a satisfação do cliente. Suas técnicas de investigação devem merecer crédito por parte dos administradores e não deixá-los expostos a críticas. O avaliador não tem nenhuma independência e nenhum controle sobre o uso que será feito das informações por ele fornecidas. [...] O relatório avaliativo pertence à burocracia e vai para seus arquivos.

Além dessa característica, assinala-se que não é incomum, no âmbito da administração pública, que programas sejam implantados, reformulados ou até mesmo extintos sem a presença de evidências empíricas que apoiem as decisões, por vezes referenciadas em opiniões e concepções de atores individuais.

Assim, é preciso avançar no sentido da maior institucionalização das avaliações não no plano meramente administrativo, fiscalizatório ou burocrático-formal, mas no plano formativo, que busque incorporar o uso dos resultados obtidos para eventuais mudanças de rumo, reajustes e melhorias das políticas públicas, e o controle social das intervenções realizadas.

Além disso, concretizar uma perspectiva de avaliação que venha a se constituir numa ação intrínseca à execução de um programa supõe o enfrentamento de limitações de natureza técnica, que abrangem desde a opção por uma abordagem

avaliativa adequada aos propósitos e ao objeto de avaliação, até a construção de instrumental válido e fidedigno.

Integrando as decisões relativas ao delineamento da proposta avaliativa, está o desafio de se estabelecerem indicadores capazes de sintetizar dimensões de qualidade do objeto em análise, temática que é foco deste artigo.

O processo de definição de indicadores é complexo não só porque condensa a noção de qualidade que se está tomando como referência para avaliação, mas também por dificuldades de operacionalização, em particular, quando se trata de investigar objetos multifacetados, usualmente focalizados na pesquisa social, âmbito em que se inserem as iniciativas em educação.

Como relata Bottani (1998), ao discorrer sobre o processo de produção de indicadores educacionais, conduzido pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), ter claros os parâmetros conceituais e teóricos que orientarão a definição de indicadores é o passo inicial para sua operacionalização. Portanto, a definição de indicadores supõe movimentos de explicitação de constructos teóricos em interação com sua tradução em evidências observáveis, supondo articulação conceitual e metodológica.

Este artigo organiza-se em torno de dois propósitos. Busca sistematizar, com base em contribuições da literatura, elementos que integram a noção de indicador, atentando-se para a discussão de aspectos políticos e técnicos inerentes ao seu delineamento. Em seguida, registra como se lidou com o desafio de estabelecer indicadores para compor a avaliação de impacto de um programa educacional, qual seja o Programa Letra e Vida. Ressalta-se que o foco do artigo são as informações relativas à metodologia utilizada para avaliação do programa, não se registrando, em seu âmbito, os resultados obtidos na pesquisa.

Este Programa foi implementado na rede de ensino do Estado de São Paulo, no período de 2003 a 2006¹, com o propósito de aprimorar a prática de professores alfabetizadores atuantes nos primeiros anos do Ensino Fundamental. Os objetivos anunciados na documentação do programa são: melhorar significativamente os resultados da alfabetização no sistema de ensino estadual, tanto quantitativa quanto qualitativamente; contribuir para uma mudança de paradigma no que se refere tanto à didática da alfabetização quanto à metodologia de formação

¹ A partir de 2007, o Programa Letra e Vida passa a denominar-se Ler e Escrever; embora mantendo características do Programa original, sofre alterações em seu delineamento. O presente artigo volta-se para a avaliação do Programa em suas características originais.

dos professores; contribuir para que se formem, na base do sistema estadual de educação, quadros estáveis de profissionais capazes de desenvolver a formação continuada de professores alfabetizadores; contribuir para que tanto as diretorias de ensino quanto as unidades escolares sintam-se responsáveis pela aprendizagem de todos os seus alunos; favorecer a ampliação do universo cultural dos formadores e dos professores cursistas, principalmente no que se refere ao seu letramento (SÃO PAULO, 2011).

Inicia-se o texto com elementos que visam a conceituar a noção de indicador, destacando os tipos e as propriedades que lhe são característicos. A seção seguinte trata da elaboração de indicadores, qualitativos e quantitativos, estabelecidos para a avaliação do Programa Letra e Vida. Tecem-se, ainda, considerações sobre desafios inerentes à elaboração e ao uso desse tipo de informação estruturada para a avaliação de programas na área educacional.

2 Indicadores – significado na avaliação de políticas

O uso de indicadores para planejamento, monitoramento e avaliação de ações tem se intensificado como suporte na formulação e implementação das políticas sociais, nas diferentes esferas de governo. Essa tendência, como assinala Jannuzzi (2005), reflete mudanças no processo de gestão pública, que valorizam o planejamento plurianual, como ferramenta de gestão, e o aprimoramento de mecanismos de auditoria e monitoramento de programas governamentais, realçando que o desenvolvimento das várias modalidades de tecnologia da informação permite a transparência e a rapidez no fluxo de informações, propiciando o controle social. Ao conceituar indicadores, o autor destaca seu papel nas políticas públicas na atualidade:

[...] os indicadores sociais são medidas usadas para permitir a operacionalização de um conceito abstrato ou de uma demanda de interesse programático. Os indicadores apontam, indicam, aproximam, traduzem em termos operacionais as dimensões sociais de interesse definidas a partir de escolhas teóricas ou políticas realizadas anteriormente. Prestam-se a subsidiar as atividades de planejamento público e a formulação de políticas sociais nas diferentes esferas de governo, possibilitam o monitoramento das condições de vida e bem estar da população por parte do poder público e da sociedade civil e permitem o aprofundamento da investigação acadêmica sobre a mudança social e sobre os

determinantes dos diferentes fenômenos sociais. (JANNUZZI, 2005, p. 138)

Em outros termos, como afirmam Ferreira e Tenório (2010, p. 72), construir indicadores supõe:

[...] transformar conceitos e relações entre conceitos, que constituem as teorias com seus diferentes graus de generalidade e abstração, em categorias e proposições capazes de avançar na direção da explicitação das suas configurações e aplicações particulares. Esse processo revela a sua importância e necessidade quando garante a demarcação de evidências concretas que permitam a compreensão ou explicação dos fenômenos que estão sendo investigados.

No entanto, o debate sobre a necessidade de produção de informações que possam ser “úteis” ou “aplicáveis” às políticas públicas não é novo – ocorre pelo menos desde meados do século XX e coincide com o desenvolvimento da avaliação de programas, principalmente nos Estados Unidos, a partir da década de 1960, como destacam diversos autores (CARO, 1971; MADAUS; STUFFLEBEAM, 1983; WORTHEN; SANDERS; FITZPATRICK, 2004; VIANNA, 2005).

Em retrospectiva que trata da origem e do uso de indicadores sociais no planejamento do setor público, Jannuzzi (2002) realça a intensificação de seu aprimoramento metodológico a partir dos anos 1960, motivada pelo interesse na avaliação do impacto das políticas sociais. Assim, buscaram-se, por exemplo, formas de mensurar os efeitos sociais da taxa de desemprego, dos índices de criminalidade, das estimativas de expectativa de vida, dos índices de saúde, do número de matrículas escolares e dos resultados de desempenho em testes, dentre outros indicadores que servissem a um sistema social de *accountability*².

Acompanhando esse movimento, a compilação de dados e a produção de relatórios de estatísticas educacionais foram expandidas durante o início de 1980, nos EUA; no início de 1990, nos países da OCDE, e, mais recentemente, nos países em desenvolvimento. Como destaca Morduchowicz (2006, p. 01):

² O conceito de *accountability* é aqui associado ao uso de informações obtidas e sua disponibilização como instrumento de ação e controle sociais sobre os serviços prestados e recebidos.

En particular, los indicadores educativos, destinados a comparar y juzgar el contexto y el funcionamiento de la enseñanza y sus resultados, han experimentado un gran desarrollo. El proceso de globalización, que propicia las comparaciones, ha provocado la preocupación de los organismos internacionales y nacionales por crear sistemas de información adecuados para evaluar diferentes aspectos de la escolarización de la población. La diversidad de estadísticas utilizadas en los distintos sistemas de indicadores educativos existentes, exige su análisis al momento de realizar su comparación.

Além de indicadores estatísticos de educação inter e intranações, outras estatísticas têm sido produzidas, resultando em indicadores complexos, tais como investimentos financeiros em educação, retenção e evasão, e desempenho de alunos em matemática, leitura, escrita e ciências, muitas vezes tomados como indicadores em relatórios de avaliação.

Fitz-Gibbon (2002) esclarece que a intensificação no uso de indicadores ocorreu em paralelo ao desenvolvimento de sistemas de informação que se utilizam da tecnologia para coletar e armazenar dados. Segundo a autora, os indicadores, processados a partir das informações disponíveis nos diversos sistemas de dados, têm sido desenhados para acompanhar o desempenho dos programas por meio de medições regulares.

A tradução de informações disponíveis em indicadores requer uma análise que articule os dados disponíveis aos propósitos da avaliação. Nesse sentido, Soares e Alves (2007) e Gatti (2004) comentam que, apesar da crescente disponibilidade de dados estatísticos na área da educação, há ainda pouca exploração destes, por parte da comunidade acadêmica, devido à pouca familiaridade com metodologias de análise de dados quantitativos.

Além desse desafio de ordem técnica, a seleção e/ou a construção de indicadores para avaliação de resultados de um programa envolvem opções valorativas, que expressam os posicionamentos políticos e ideológicos que irão pautar os julgamentos do pesquisador, seja na análise da formulação de um programa, de sua implementação, ou de seus efeitos, etapas que se inter-relacionam no sentido de possibilitar a interpretação do programa em avaliação.

A especificação de indicadores, quantitativos ou qualitativos, com base em informações disponibilizadas em bancos ou dados ou coletadas pelo avaliador,

requer, portanto, que se esteja atento à diversidade de dimensões de análise – insumos, processos, produtos – que podem estar presentes na avaliação de políticas e programas sociais, como alerta Morduchowicz (2006, p. 2):

[...] los indicadores de insumo y de proceso son indicadores de contexto. Si bien lo que interesa, en última instancia, es el desempeño de un sistema educativo o una política que se está llevando adelante, esos indicadores son necesarios porque ayudan a interpretar ese desempeño. Aun más, los mismos indicadores de insumo y de proceso contribuyen a explicar el desempeño de algunos de los que pueden conformar esos grupos: por ejemplo, el gasto educativo por alumno puede estar influido por la zona (urbana o rural) en que se imparte el servicio, los años de antigüedad de los docentes, el tamaño de la escuela, que puede contribuir a absorber (o, mejor dicho, distribuir) los costos fijos entre más alumnos, etc.

Tendo em conta os desafios inerentes à produção de indicadores, no caso do estudo aqui referenciado – relativo ao Programa Letra e Vida – procedeu-se a coleta, seleção e organização de informações, que foram traduzidas em indicadores quantitativos e qualitativos para a avaliação de possíveis impactos do Programa Letra e Vida no desempenho dos alunos e em práticas dos professores alfabetizadores da 1.^a Série do Ensino Fundamental³. Para evidenciar empiricamente esses impactos, o estudo recorreu à metodologia não experimental, com aplicação do método misto, integrando complementarmente instrumentos e dados quantitativos e qualitativos (STUFFLEBEAM; CORYN, 2012; CRESWELL, 2007), o que será detalhado nas próximas seções.

Creswell (2007), um dos pioneiros no tema, esclarece que adotar métodos mistos vai além de coletar dados quantitativos e qualitativos na mesma pesquisa, envolvendo a combinação das duas formas de dados em desenhos ou metodologia de pesquisa distintos. Em sua definição de método misto, o autor explica:

Método misto é um desenho de pesquisa que envolve pressupostos filosóficos e métodos de pesquisa. A abordagem parte de pressupostos filosóficos que orientam a coleta e a análise de dados e articula análises quantitativas e qualitativas em muitas

³ Quando da realização do estudo, o Ensino Fundamental estava organizado em oito anos de escolaridade e não em nove anos, como atualmente. Manteve-se, assim, a nomenclatura da época.

fases da pesquisa. O método se concentra na coleta, análise e mescla dos dados quantitativos e qualitativos em um único estudo ou em uma série de estudos.⁴ (CRESWELL; PLANO CLARK, 2007, p. 5, tradução nossa).

Essa perspectiva resultou no delineamento de indicadores qualitativos e quantitativos para a avaliação do Programa Letra e Vida, que apoiaram análises descritivas e estatísticas de evidências de impactos do referido Programa.

3 Indicadores qualitativos para análise do Programa

Como um dos procedimentos de pesquisa⁵ utilizados para se evidenciar eventuais impactos do Programa Letra e Vida na atuação docente, recorreu-se à observação de sala de aula e de cadernos dos alunos. Para estabelecer dimensões e critérios de análise, o passo inicial foi apreender, com base na análise dos materiais de formação docente do Programa Letra e Vida, pressupostos teórico-metodológicos por este difundidos. As principais fontes consultadas foram as apostilas utilizadas no curso de formação dos professores alfabetizadores e o Questionário para o Professor de 1.^a ou 2.^a Série do Saresp 2007, que contém itens relativos ao processo de alfabetização.

A interpretação dessas fontes resultou em referências de qualidade para análise de práticas pedagógicas dos professores e subsidiou a definição de dimensões a serem estudadas. Sem pretensões de exaurir as expectativas do Programa, as dimensões analisadas foram: 1) organização espacial das salas de aulas e 2) características do trabalho pedagógico e de atividades que integram a rotina dos alunos. Em seu conjunto, as informações coletadas foram tratadas como indicadores da implementação – ou não –, pelos professores, de um processo

⁴ “[A] mixed method is a research design with philosophical assumptions as well as methods of inquiry. As a methodology it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies.” (CRESWELL; PLANO CLARK, 2007, p. 5).

⁵ Para apreender possíveis impactos do Programa Letra e Vida, utilizou-se uma metodologia que combinou análises quantitativas e qualitativas em relação de dependência. A partir de análise descritiva dos resultados do SARESP 2007, foram escolhidas três escolas para a pesquisa de campo, que abrangeu observação e entrevistas com profissionais envolvidos com o Programa Letra e Vida. Com base nos resultados obtidos, nova análise quantitativa foi realizada, levando-se em conta as informações obtidas no campo. Nos estudos quantitativos, foram utilizados três procedimentos distintos: a comparação entre grupos semelhantes de escolas; a utilização de procedimentos de *data mining*, usando como recurso as árvores de decisão, e a regressão linear múltipla, sendo que os resultados obtidos foram triangulados (triangulação metodológica), a fim de procurar conferir maior robustez às análises.

de alfabetização compatível com a proposta disseminada pelo Programa Letra e Vida (SÃO PAULO, 2009, 2011).

Com base nas dimensões de análise supracitadas, foram detalhados os critérios que orientaram a elaboração do roteiro de observação:

✓ Quanto à organização de sala de aula

- Presença das carteiras (em fileiras ou agrupadas); espaço para atividades livres ou de leitura.

✓ Quanto ao trabalho pedagógico

- Existência de plano semanal, com objetivos claramente definidos.
- Realização de atividades diversificadas, contemplando atividades permanentes, sequenciadas, de sistematização, independentes e de projetos.
- Realização de atividades propostas na Coletânea de Textos e Vídeos fornecidos pela Secretaria Estadual de Educação aos alunos e professores.
- Utilização de materiais diversificados de leitura: textos informativos variados, letras de música, poesias/poemas, cantigas, jornais, gibis, parlandas, textos publicitários (cartazes/placas), rótulos e clássicos da literatura infantil.
- Avaliação dos alunos em uma perspectiva formativa, por meio de instrumentos que permitam levantar hipóteses sobre o processo de aprendizagem dos alunos e registrar sistematicamente o desempenho dos alunos (por exemplos: mapa da classe e tabulações).
- Presença, no caderno dos alunos, de atividades com nomes próprios, estudo de palavras retiradas de textos, produção e/ou revisão coletiva de textos, desenho a partir de histórias lidas pelo professor, atividades de reflexão ortográfica, ditado de palavras não estudadas, produção de texto a partir de sequência de gravuras e reescrita de textos.

- Execução de atividades de leitura em voz alta e de leitura silenciosa pelos alunos de diferentes tipos de texto: com e sem palavras, de prosa, de poesia; revistas; gibis; suplementos infantis de jornais; cartelas com nomes dos alunos; produções das próprias crianças e dicionários.
- Realização de rodas de leitura.
- Leitura em voz alta de histórias e textos não literários, pelo professor para os alunos.

Essas características da organização do ambiente da sala de aula, do trabalho pedagógico e das atividades presentes na rotina dos alunos, interpretadas como desejáveis por expressarem pressupostos teórico-metodológicos do Programa, integraram a estrutura do roteiro de observação.

O roteiro contemplou ainda práticas pedagógicas não disseminadas pelo Programa, mas que poderiam estar presentes nas salas de aula e evidenciarem, então, a adoção, pelo professor, de outras abordagens de alfabetização. Frente aos limites de extensão deste texto, citam-se apenas alguns exemplos dessas práticas: uso de cartilhas e livros didáticos, cartazes com letras cursivas, treinos mecânicos de escrita – cópia de sílabas, palavras, frases ou textos –, exercícios de associação/correspondência de letras/palavras a imagens e atividades de treino ortográfico.

Além de indicadores de prática pedagógica que focalizam o desempenho do professor, individualmente, em sala de aula, foram definidos indicadores que permitissem apreender evidências do acolhimento dos pressupostos teórico-metodológicos do Programa no contexto escolar. Para tanto, depoimentos de professores e gestores das escolas foram coletados por meio de entrevistas, abarcando relatos de práticas e opiniões sobre o Programa.

Foram estabelecidas como dimensões de análise: 1) Liderança da equipe técnica para a concretização do Programa Letra e Vida; 2) Estrutura e organização da escola para a implementação do programa; 3) Formação do professor 4) Dificuldades para a concretização dos pressupostos do curso na sala de aula, na ótica dos professores e gestores. Essas dimensões foram definidas com base na hipótese de que o sucesso na implementação do Programa extrapola a adesão individual do professor e supõe adesão coletiva e condições para sua realização no contexto escolar, o que é fartamente explorado na literatura (SAMMONS; HILLMAN; MORTIMORE, 1995; BROOKE; SOARES, 2008).

Essas dimensões foram detalhadas em critérios a seguir apresentados:

1. Liderança da equipe técnica para a concretização do Programa Letra e Vida: presença do diretor nas reuniões pedagógicas; parceria entre diretores e coordenadores para a consolidação do programa e o acompanhamento dos professores; motivação da equipe para a realização do trabalho, traduzida pelo suporte material e pela tentativa de manutenção de boas normas de convivência na escola; apoio pedagógico às práticas dos professores, e evidências de organização de estudos dentro dos pressupostos do Programa.
2. Estrutura e organização da escola para a implementação do programa: presença de professores-pesquisadores e depoimentos sobre a existência de materiais básicos do programa.
3. Formação em nível local, sob supervisão do coordenador, que contemplou os depoimentos dos professores sobre o conhecimento do coordenador dos pressupostos do Programa e suas iniciativas para acompanhar o trabalho dos professores e apoiá-los em suas dúvidas, prestando esclarecimentos e subsidiando sua prática pedagógica.
4. Compromisso do corpo docente com a realização do Programa, que considerou o depoimento dos professores sobre a aplicação dos pressupostos metodológicos do Programa.
5. Reflexão sobre a prática e estudo dos resultados obtidos pelo professor e pela equipe de gestão quanto à incorporação dos princípios do Programa à prática.
6. Influência do Programa na formação do professor, considerando-se depoimentos sobre mudanças na prática e na filosofia de ensino havidas após a participação no curso, no que se refere a: atividades em sala de aula; concepção de aluno e de processo de aprendizagem; organização e planejamento do trabalho; gestão de sala de aula, e leitura pessoal.
7. Dificuldades apontadas para a concretização dos pressupostos do curso na sala de aula, considerando-se as dificuldades extraescolares (nível socioeconômico das famílias e/ou falta de estrutura familiar; falta de preparo dos alunos na etapa escolar anterior; falta de apoio da comunidade ao trabalho de alfabetização; falta de suporte da SEE

e da Secretaria de Saúde, para lidar com necessidades dos alunos que influem na alfabetização, como dislexias, hiperatividade etc.), e intraescolares (falta de suporte pedagógico da gestão, dos colegas de trabalho ou da SEE; descrença no potencial dessa prática de alfabetização; necessidade de atendimento de demandas da SEE que implicam outro tipo de trabalho; rotatividade do corpo docente, e nível de formação insuficiente do corpo docente para aplicação dos pressupostos do Programa).

Após definição das dimensões que balizariam as observações, procedeu-se à seleção das escolas, considerando-se os seguintes critérios: escolas que concentrassem o maior número de professores de 1.^a e 2.^a Séries que participaram do programa Letra e Vida⁶ e, ao mesmo tempo, atendessem a uma clientela composta, majoritariamente, por alunos provenientes das camadas mais pobres da população. Foram visitadas três escolas, denominadas no trabalho original de Alfa, Gama e Beta, com o intuito de preservar seu anonimato.

A interpretação dos resultados obtidos com as observações e as entrevistas permitiu inferir a utilização de práticas de alfabetização e letramento disseminadas pelo Programa Letra e Vida pelos professores que o cursaram. Essas informações, obtidas na pesquisa de campo, subsidiaram, também, a definição de variáveis para as análises quantitativas.

4 Análise do Programa à luz de indicadores quantitativos

O recurso às análises quantitativas, complementando o estudo qualitativo, foi empregado para possibilitar que se respondesse sobre eventuais impactos do programa.

Abordagens metodológicas afeitas à avaliação de impactos de programas indicam a pertinência de análises quantitativas, com recurso às técnicas de regressão,

⁶ Diversas foram as condições utilizadas na seleção das escolas. Para essa seleção, trabalhou-se com as respostas dos professores alfabetizadores ao questionário do SARESP 2007. Procurou-se isolar as escolas com professores que tivessem sua formação continuada em alfabetização apenas pelo programa Letra e Vida, havendo cursado este de forma integral, ou seja, os três módulos. Consideraram-se, também, professores que estivessem na mesma escola desde 2005. Procedeu-se ao cálculo da quantidade de professores, por escola, que obedeciam a estas condições. As escolas que concentravam o maior número desses professores e atendiam os alunos mais pobres foram escolhidas para a pesquisa de campo. Inicialmente, seis escolas satisfaziam essas condições. No entanto, no momento da realização do trabalho campo, várias delas estavam fechadas, devido à epidemia de gripe suína que assolou São Paulo em 2009. Assim, o estudo ficou limitado a três escolas, que puderam ser visitadas no período disponível para a concretização do trabalho de campo.

*matching*⁷, experimentos randômicos controlados e modelagem hierárquica multinível, por exemplo, visto que é inerente, a esse tipo de avaliação, estabelecer relações causais entre o fenômeno em estudo e as variáveis que podem estar associadas aos resultados obtidos⁸.

Para avaliação de eventuais impactos do Programa Letra e Vida no desempenho dos alunos foram explorados e reorganizados dados provenientes do Sistema de Avaliação do Rendimento Escolar de São Paulo – Saresp, disponibilizados pela Fundação para o Desenvolvimento da Educação (FDE), de modo a ensejar análises distintas: a aplicação da regressão linear múltipla, o uso da mineração de dados por meio das técnicas de árvores de decisão e a comparação de médias de alunos de dois grupos de escolas com características equivalentes (contrafactual). A variável dependente utilizada nas análises foi o nível de proficiência dos alunos de 1.ª Série na prova de Língua Portuguesa do Saresp aplicado em 2007.

Os produtos obtidos com a aplicação dessas três técnicas foram triangulados para analisar se os resultados dos alunos variaram segundo a proporção de professores da escola que fizeram o Programa Letra e Vida, e a utilização, nessas escolas, de práticas de alfabetização e letramento disseminadas pelo Programa.

Os indicadores quantitativos utilizados no estudo foram processados a partir de respostas dos professores dos anos iniciais ao Questionário para Professores de 1.ª e 2.ª Séries do Saresp de 2007⁹. O universo de respondentes do instrumento totalizava 14.974 sujeitos, mas desejava-se trabalhar com as respostas dos docentes que haviam participado do programa em tela. Assim, procedeu-se a uma seleção dos questionários respondidos com os seguintes critérios:

1. Professores que participaram do Programa entre 2003 e 2004.
2. Professores que completaram os três módulos do Programa.

⁷ *Matching* é uma técnica estatística utilizada para avaliar o efeito de um tratamento por meio da comparação do objeto tratado com as unidades não tratadas em estudos observacionais ou em quase experimentos (quando o tratamento não é atribuído aleatoriamente). O objetivo do *matching* é encontrar, para cada unidade de tratamento, uma (ou mais) unidade(s) que não receba(m) o tratamento e que tenha(m) características observáveis semelhantes àquelas, tornando possível avaliar o efeito do tratamento, pois a suposição é de que apenas esta diferença as unidades tratadas das não tratadas.

⁸ Neste artigo, não serão tratados aportes teóricos relativos à avaliação de impactos. Para informações sobre tipo de avaliação, consultar Bauer (2011, 2010) e Mohr (1992).

⁹ O questionário aplicado aos professores pode ser acessado por meio do link <http://saresp.fde.sp.gov.br/2007/Arquivos/Questionarios_Gestao/FCC_Questionario_Professor_1e2_serie.pdf>. Acesso em: 01/12/2013

3. Professores que participaram do Programa como cursistas, visto que alguns dos professores que estavam em sala de aula, à época da pesquisa, tiveram oportunidade de participar do programa como formadores.
4. Professores que estavam na mesma escola, em 2007, havia quatro anos ou mais, ou seja, tiveram seus alunos avaliados pelo SARESP em 2005 e em 2007.

Aplicando-se estes critérios, chegou-se a um total de 9.289 (78,4%) professores da rede paulista que haviam cursado o programa completo até o final de 2007. Destes, 3.426 participaram do curso em 2003 e em 2004, sendo, respectivamente, 1.593 e 1.833 sujeitos.

Dos 3.426 professores que participaram do curso nos anos mencionados anteriormente, 144 sujeitos haviam atuado como formadores do curso e 3.282 foram cursistas; destes, 3.252 estavam havia mais de quatro anos na mesma escola, satisfazendo, portanto, as condições pré-determinadas para a composição do grupo cujas respostas seriam analisadas. Com esse recorte, as informações consideradas abrangeram docentes vinculados a 2.166 escolas estaduais.

As informações provenientes do questionário, composto por 163 questões objetivas, abrangiam características das escolas em que os docentes atuavam, dados de identificação, formação, experiência e trajetória profissional dos respondentes, bem como informações sobre suas práticas pedagógicas, atitudes e opiniões, suporte da equipe gestora à sua prática pedagógica; o questionário compôs-se, também, de um conjunto específico de nove questões a serem respondidas apenas pelos docentes que tivessem cursado o Programa Letra e Vida.

Diante desse vasto conjunto de dados, adotaram-se os seguintes procedimentos, visando à elaboração de indicadores que pudessem iluminar o entendimento sobre a influência do Programa na prática do professor e nos resultados de desempenho dos alunos no Saresp:

- ✓ Seleção de questões de interesse: foi realizado um estudo descritivo do banco de dados proveniente do Questionário para professores de 1.^a e 2.^a Séries do Saresp 2007, com o propósito de identificar quais seriam as questões, ou bloco de questões, de interesse. Optou-se por selecionar aquelas que se relacionavam ao perfil profissional dos professores –considerando-se os objetivos anunciados pelo Programa e as dimensões suscitadas pela literatura que trata da influência do perfil docente sobre o desempenho dos

alunos – e às suas práticas pedagógicas – buscando-se selecionar as que tratassem de aspectos potencialmente capazes de explicar o desempenho dos alunos, ou seja, respostas dos professores que pudessem evidenciar sua interação com o Programa Letra e Vida. Foram escolhidas ainda questões sobre a gestão da escola, no que se refere ao favorecimento de um ambiente propício para a realização de um trabalho coerente com os pressupostos do Programa Letra e Vida.

- ✓ Análise descritiva das respostas dos professores às questões selecionadas: realizou-se uma análise das frequências das respostas dos professores aos itens e foi possível verificar que algumas concentravam as respostas em uma única alternativa, denotando um baixo poder discriminativo. Tais questões foram desconsideradas, tendo como critério de descarte a concentração de mais de 70% das respostas numa única alternativa.
- ✓ Reorganização dos dados: a análise acurada do questionário e das respostas dadas pelos professores alfabetizadores revelou imprecisões na definição de escala que permitiam questionar a fidedignidade das respostas. Por exemplo, as questões sobre práticas pedagógicas utilizavam uma escala de três pontos – “frequentemente”, “de vez em quando”, “raramente ou nunca”. Desse modo, não há parâmetros que auxiliem o respondente a distinguir “de vez em quando” de “raramente ou nunca” e, conseqüentemente, tais alternativas suportam diferentes interpretações. Para superar essa limitação do instrumento, as escalas foram dicotomizadas, reorganizando-se as respostas “de vez em quando” e “raramente ou nunca” em um mesmo agrupamento.
- ✓ Definição das dimensões de análise: identificaram-se, com base nos indicativos presentes na literatura de referência e nos resultados obtidos durante a pesquisa de campo¹⁰, quais informações seriam importantes para constituir os indicadores a serem utilizados nas análises estatísticas. O estabelecimento de dimensões – ou aspectos característicos que se quer quantificar por meio de indicadores – é crucial para que se possam obter, direta ou indiretamente, as respostas desejadas (TRZESNIAK, 1998). Inicialmente, foram selecionados itens do questionário que forneciam informações sobre algumas dimensões do trabalho e da formação dos

¹⁰ Na definição de quais constructos se intencionava medir, foram consideradas as informações obtidas na pesquisa de campo, por meio da entrevista com diversos coordenadores pedagógicos e professores, que indicaram aspectos relativos ao Programa Letra e Vida e à prática dos professores que seriam importantes de mensurar e considerar nas análises estatísticas.

professores, a saber: leitura do professor, material utilizado para o ensino de leitura, indicativos de prática pedagógica coerente com o Programa Letra e Vida (designada na pesquisa prática pedagógica A), indicativos de prática pedagógica baseada em métodos “tradicionais” (denominada na pesquisa prática pedagógica B) e indicativos de suporte à prática pedagógica construtivista na escola.

- ✓ Processamento dos dados: os dados foram processados com o auxílio do *software* científico *Statistical Package for the Social Science* (SPSS)¹¹.
- ✓ Elaboração dos indicadores: escolhido o conjunto de questões que representasse as dimensões que se pretendia mensurar, fez-se uma Análise dos Componentes Principais (ACP), visando a reduzir o número de questões utilizadas na composição de cada uma das dimensões supracitadas. Considerando os objetivos do Programa e a forma como ele foi desenvolvido, foram escolhidas as seguintes dimensões: hábitos de leitura do professor; prática pedagógica de alfabetização; materiais usados para o ensino de leitura, e suporte à prática pedagógica.

Os Quadros 1 e 2 indicam de modo sintético as questões, respondidas por professores de 1.^a e 2.^a Séries do EF/Questionário Saresp 2007, utilizadas na composição dos indicadores.

O Quadro 1 explicita como se construíram os indicadores de hábitos de leitura do professor, os materiais utilizados para o ensino de leitura e o suporte à prática pedagógica disponibilizada aos professores no âmbito da escola, com base em 19 das questões. O Quadro 2 especifica as questões consideradas para construção dos indicadores de práticas pedagógicas A e B. O recurso à organização de dois quadros visa apenas a facilitar a apresentação das informações, pois ambos tratam de registrar as questões que foram tomadas como referência na construção dos indicadores.

Escolhidas as questões que comporiam cada dimensão da análise e, portanto, cada indicador, passou-se a enfrentar o desafio de quantificar os resultados. Para a quantificação de cada aspecto, considerou-se que a força do indicador como variável explicativa dos resultados obtidos pelos alunos estaria na possibilidade de

¹¹ O Pacote Estatístico para as Ciências Sociais é um programa estatístico científico muito utilizado nas ciências sociais; é também usado por pesquisadores de mercado, na pesquisa relacionada com a saúde, no governo, na educação e em outros setores. Foi criado em 1968 e, atualmente, pertence à empresa IBM.

Quadro 1 - Questões utilizadas na composição de três dos indicadores utilizados nas análises

Indicador	Número das questões utilizadas	Justificativa para elaboração de um indicador sobre essa dimensão	Conteúdo das alternativas
Hábitos de leitura do professor	42 a 47	Um dos objetivos do Programa era propiciar que os professores lessem mais e se tornassem melhores leitores.	Tipos de materiais lidos pelos professores: Q42 - Histórias infantis Q43 - Outros livros (literatura em geral) Q44 - Revistas especializadas na sua área de atuação Q45 - Jornais ou revistas de informação geral Q46 - Revistas em quadrinhos Q47 - Outras revistas
Materiais utilizados para o ensino de leitura	73,74, 77, 78	O Programa objetiva prover cada sala de aula de 1.ª e 2.ª Séries com diversos livros de literatura infantil e outros materiais de leitura, visando à aplicação, pelos professores, dos pressupostos de ensino de leitura aprendidos no curso, como a utilização de diversos materiais impressos, como jornais, rótulos, cartazes e peças publicitárias em sua prática de ensino de leitura.	Tipos de materiais usados pelos professores para ensinar leitura: Q73 - Gibis Q74 - Jornais Q77 - Textos publicitários (cartazes/placas) Q78 - Rótulos

Fonte: BAUER (2011).

Quadro 1 - Continuação...

Indicador	Número das questões utilizadas	Justificativa para elaboração de um indicador sobre essa dimensão	Conteúdo das alternativas
Suporte à prática pedagógica	140, 141, 142, 145, 146, 147, 149, 150, 152	Importância dos professores poderem contar com suporte às suas ações por parte da equipe gestora e de seus pares.	Tipos de suporte à prática fornecidos ao professor nas escolas: Q140 - Auxílio na utilização de recursos didáticos Q141 - Estímulo na adoção de estratégias didáticas que enfatizem a interação com o aluno na construção de sua aprendizagem Q142 - Orientações para o planejamento da aula Q145 - Assessoria na busca de novas maneiras de lidar com os alunos que têm dificuldades de aprendizagem Q146 - Estímulo à troca de experiências entre os professores do Ciclo I do EF Q147 - Análise das orientações apresentadas em atividades de capacitação para selecionar aquelas que serão implementadas na sala de aula Q149 - Análise dos índices de aproveitamento dos alunos, procurando, junto com os professores e a coordenação pedagógica, estratégias para melhorar esses índices Q150 - Valorização do trabalho dos professores e do professor coordenador por meio da divulgação de suas realizações Q152 - Orientações da oficina pedagógica para a utilização de diferentes materiais no seu trabalho em sala de aula

Fonte: BAUER (2011).

Quadro 2 - Questões utilizadas na composição dos indicadores de prática pedagógica

Indicador	Número das questões utilizadas	Justificativa para elaboração de um indicador sobre essa dimensão	Conteúdo das alternativas
Indicador de prática pedagógica A	95, 96, 100, 103, 104, 106, 107, 108, 111, 112, 113	Analisar em que medida os pressupostos ensinados durante o curso foram incorporados à prática dos professores	Práticas pedagógicas sugeridas no âmbito do Programa Letra e Vida Q95 - Atividades de reflexão ortográfica Q96 - Formação de frases com palavras aprendidas Q100 - Estudo de palavras retiradas do texto Q103 - Correspondência entre palavras e gravuras Q104 - Produção de texto à vista de gravuras Q106 - Produção de texto coletivo na lousa Q107 - Produção de texto individual Q108 - Reescrita de texto Q111 - Interpretação de textos a partir de questões Q112 - Dramatização das histórias lidas Q113 - Desenho a partir de histórias lidas
Indicador de prática pedagógica B	86, 87, 92, 93, 94, 97, 98	Analisar que outras estratégias de alfabetização e letramento estavam ainda presentes nas práticas dos professores	Práticas pedagógicas não indicadas pelo Programa Letra e Vida Q86 - Atividades com vogais e encontros vocálicos Q87 - Atividades de coordenação motora Q92 - Separação em sílabas Q93 - Atividades de gramática Q94 - Treino ortográfico Q97 - Cópia da lousa, da cartilha ou de livros Q98 - Ditado de palavras ou frases já estudadas

Fonte: BAUER (2011).

essa medida ilustrar práticas pedagógicas esperadas de um professor alfabetizador que se baseasse nos pressupostos aprendidos durante o Programa.

Como as questões que se relacionavam com cada dimensão de análise eram dicotômicas – ou seja, as possibilidades de resposta eram 1 ou 0, sendo que a resposta 1 indica uso frequente, pelo professor, de uma estratégia de ensino ou de material indicado no Programa Letra e Vida, e a resposta 0 denota ausência ou pouca utilização dessa mesma prática ou material –, optou-se por uma somatória simples dessas respostas, conforme indicado no Quadro 3.

Quadro 3- Padrões de codificação das respostas ao questionário para professores de 1.^a e 2.^a Séries do EF do Saresp 2007

Escala do questionário		Valor
Padrão I	Frequentemente	1
	De vez em quando	0
	Raramente ou nunca	0
	Concordo	1
	Concordo em parte	0
	Discordo	0
Padrão II	Suficiente	1
	Suficiente em parte	0
	Insuficiente	0
	Não recebo apoio nesse aspecto	0

Fonte: BAUER (2011).

Portanto, para cada indicador, havia uma pontuação máxima, oriunda da somatória simples do número de questões que o compuseram, conforme indicado na Tabela 1.

Finalmente, analisou-se a fidedignidade desses indicadores com o índice alfa de Cronbach¹². No estudo em pauta, foram considerados adequados os indicadores com alfa superior a 0,75, conforme parâmetros estabelecidos por DeVellis (1991). Os valores obtidos para cada indicador são apresentados na Tabela 2.

Considerando-se estes resultados e o parâmetro assumido no estudo, foram utilizados, nas análises quantitativas, os indicadores “suporte à prática pedagógica”, “material utilizado para o ensino de leitura” e “prática pedagógica B”.

¹² O coeficiente de fidedignidade Alfa de Cronbach tem sido a estatística mais usada para medir a confiabilidade de um questionário. Seu valor vai de 0 a 1 e a confiabilidade das respostas é maior se o valor obtido estiver mais perto de 1. Os valores de Alfa são afetados pelo aumento das intercorrelações entre itens de um questionário. Ou seja, se as questões medem um mesmo constructo, os valores de Alfa serão maiores e, portanto, a medida de confiabilidade de resposta será mais acurada. Mais informações podem ser obtidas em Maroco e Garcia-Marques (2006).

Tabela 1 - Quantificação dos indicadores

Nome do indicador	Número de Questões que compõem o indicador	Valor máximo do indicador
Suporte à prática pedagógica	9	9,00
Leitura do professor	6	6,00
Material utilizado para o ensino de leitura	4	4,00
Prática pedagógica A	11	11,00
Prática pedagógica B	8	8,00

Fonte: BAUER (2011).

Tabela 2 - Parâmetros utilizados na composição de variáveis para o modelo de regressão

Nome do indicador	Número de Questões que compõem o indicador	Alfa de Cronbach
Suporte à prática pedagógica	9	0,913
Leitura do professor	6	0,700
Material utilizado para o ensino de leitura	4	0,785
Prática pedagógica A	11	0,734
Prática pedagógica B	8	0,872

Fonte: BAUER (2011).

Com base nestes indicadores de contexto, que mensuraram o uso do material de leitura indicado no curso e o apoio da equipe gestora aos professores, bem como a produção da variável prática pedagógica B – indicativa da não utilização dos pressupostos do Programa Letra e Vida –, analisou-se sua correlação com os resultados dos alunos nas provas do Saresp. Como os resultados das correlações foram significativos, tais indicadores compuseram os estudos de regressão e as análises calcadas na técnica de árvores de decisão.

Além dos indicadores compostos que foram descritos, utilizaram-se indicadores simples para aquilatar possíveis impactos do Programa Letra e Vida no desempenho dos alunos e em práticas dos professores alfabetizadores da 1.ª Série do Ensino Fundamental, como, por exemplo, a porcentagem, por escola, de professores que completaram o Programa Letra e Vida e a participação do coordenador pedagógico da escola no Programa Letra e Vida para coordenadores¹³.

Deve-se destacar, contudo, que os indicadores estabelecidos são dependentes do contexto, ou seja, não teriam sentido fora da conjuntura do Programa; portanto,

¹³ A apresentação dos resultados obtidos na pesquisa extrapola os propósitos deste artigo; estes podem ser acessados no trabalho original, disponível em: <<http://www.teses.usp.br/teses/disponiveis/48/48134/tde-15072011-105940/pt-br.php>>. Acesso em: 01/12/2013.

não suportam generalização para análise de outros programas ou desse mesmo Programa em outros sistemas educativos.

5 Considerações finais

Neste artigo buscou-se destacar as contribuições da literatura que evidenciam a importância da etapa de formulação de indicadores no processo de avaliação de políticas sociais e apresentar elementos que ilustrassem os procedimentos adotados para sua construção.

Como uma das etapas da avaliação de programas, a delimitação de indicadores supõe assumir uma dada noção de qualidade como referência. No caso do programa de formação de professores aqui tratado, o Programa Letra e Vida, essa definição guardou estreita relação com os seus objetivos.

A escolha da metodologia mista, que integrou indicadores quantitativos e qualitativos, buscou enfrentar o desafio de apreender um fenômeno complexo, recorrendo-se a procedimentos e informações de natureza diversa, articulando-as, numa tentativa de superar os limites que métodos e dados apenas qualitativos ou quantitativos, tomados isoladamente, imporiam ao propósito de avaliar eventuais impactos do Programa no desempenho dos alunos e em práticas de professores alfabetizadores.

Diante desse desafio, procedeu-se à tradução dos propósitos anunciados pelo Programa em evidências observáveis, recorrendo-se a informações disponibilizadas em bancos de dados da Secretaria da Estadual de Educação de São Paulo, complementadas com dados obtidos diretamente em contextos escolares que contavam com a presença de professores egressos do Programa Letra e Vida.

Nos limites deste artigo, embora não se tenha tratado dos resultados obtidos com a avaliação do Programa, é oportuno registrar que os indicadores delineados no estudo em pauta trouxeram evidências para a compreensão das influências das atividades de formação continuada docente nos resultados de desempenho dos alunos, revelando a pertinência de esforços para avaliar impactos de programas de formação de professores, ainda escassos no Brasil.

Espera-se que a socialização dos procedimentos adotados no estudo aqui relatados possa alentar debates e iniciativas relativos à construção de indicadores, atividade inerente à avaliação de programas educacionais.

Referências

BAUER, A. *Avaliação de impacto de formação docente em serviço: o Programa Letra e Vida*. 2011. 250 f. Tese (Doutorado em Educação) - Faculdade de Educação, Universidade de São Paulo, São Paulo, 2011.

_____. Avaliação de impacto no Brasil: é possível mensurar impactos de programas de formação docente? *Estudos em Avaliação Educacional*, São Paulo, v. 21, n. 46, p. 229-252, maio/ago. 2010.

BOTTANI, N. Os indicadores da OCDE: propósitos, limites e processo de produção. In: SEMINÁRIO INTERNACIONAL DE AVALIAÇÃO EDUCACIONAL, 1997, Rio de Janeiro. *Anais...* Brasília: INEP, 1998. p. 53-64.

BROOKE, N.; SOARES, J. F. *Pesquisa em eficácia escolar: origem e trajetórias*. Belo Horizonte: Ed. UFMG, 2008.

CARO, F. Evaluation research: an overview. In: _____. *Readings in evaluation research*. New York: Russell Sage Foundation, 1971. p. 1-34.

CRESWELL, J. W. *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks: Sage Publications, 2007.

CRESWELL, J. W.; PLANO CLARK, V. Understanding mixed methods research. In: _____. *Designing and conducting mixed methods research*. Thousand Oaks: Sage Publications, 2007. p. 1-19.

DEVELLIS, R. F. Guidelines in scale development. In: _____. *Scale development: theory and applications*. Newbury Park: Sage Publications, 1991. p. 51-90.

FERREIRA, R. A.; TENÓRIO, R. M. A construção de indicadores de qualidade no campo da avaliação educacional: um enfoque epistemológico. *Revista Lusófona de Educação*, Lisboa, v. 15, p. 71-97, 2010. Disponível em: <<http://www.scielo.oces.mctes.pt/pdf/rle/n15/n15a06>>. Acesso em: 09 set. 2013.

FITZ-GIBBON, C. T. Evaluation in an age of indicators: challenges for public sector management. *Evaluation*, London, v. 8, n. 1, p. 140-148, 2002. <http://dx.doi.org/10.1177/1358902002008001741>.

GATTI, B. A. Estudos quantitativos em educação. *Educação e Pesquisa*, São Paulo, v. 30, n. 1, p. 11-30, jan./abr. 2004. Disponível em: <<http://www.scielo.br/pdf/ep/v30n1/a02v30n1.pdf>>. Acesso em: 30 out. 2008.

JANNUZZI, P. M. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil. *Revista do Serviço Público*, Brasília, v. 2, n. 56, p. 137-160, abr./jun. 2005.

_____. Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais. *Revista Brasileira de Administração Pública*, Rio de Janeiro, v. 36, n. 1, p. 51-72, jan/fev 2002.

MAC DONALD, B. Uma classificação política dos estudos avaliativos. In: GOLDBERG, M. A.; SOUSA, C. P. (Org.). *Avaliação de programas educacionais: vicissitudes, controvérsias, desafios*. São Paulo: EPU, 1982. p. 16-17.

MADAUS, G.; STUFFLEBEAM, D. Program evaluation: an historical overview. In: STUFFLEBEAM, D. L.; MADAUS, G. F.; KELLAGHAN, T. (Ed.). *Evaluation models: viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer-Nijhoff, 1983. p. 3-22.

MAROCO, J.; GARCIA-MARQUES, T. Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? *Laboratório de Psicologia*, Lisboa, v. 4, n. 1, p. 65-90, 2006.

MOHR, L. *Impact analysis for program evaluation*. Newbury Park: Sage Publications, 1992.

MORDUCHOWICZ, A. *Los indicadores educativos y las dimensiones que los integran*. Buenos Aires: UNESCO:IPE, 2006. Disponível em: <<http://www.iipe-buenosaires.org.ar/documentos/los-indicadores-educativos-y-las-dimensiones-que-los-integran>>. Acesso em: 17 jan. 2011.

SAMMONS, P.; HILLMAN, J.; MORTIMORE, P. *Key characteristics of effective schools: a review of school effectiveness research*. London: Institute of Education, 1995. Disponível em: <<http://www.eric.ed.gov:80/PDFS/ED389826.pdf>>. Acesso em: 23 fev. 2011.

SÃO PAULO (Estado). Secretaria de Estado da Educação. Coordenadoria de Estudos e Normas Pedagógicas. *Letra e Vida*. Disponível em: <<http://cenp.edunet.sp.gov.br/letravida/#>>. Acesso em: 31 jan. 2011.

_____. Secretaria da Educação do Estado de São Paulo. Coordenadoria de Estudos e Normas Pedagógicas. Portal da Biblioteca Virtual de São Paulo. *Programa Letra e Vida*, 2009. Disponível em: <<http://www.bibliotecavirtual.sp.gov.br/cgi-bin/wxis.exe?IsisScript=/projetos/bv/script2.xis&base=cds&fro m=00086&to=00086>>. Acesso em: 27 set. 2010.

SOARES, J. F.; ALVES, M. T. G. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a Sociologia da Educação. *Sociedade e Estado*, Brasília, v. 22, n. 2, p. 435-473, maio/ago. 2007.

SOUSA, S. Z. PROJÓVEM: elementos para subsidiar a avaliação do programa. In: SEMINÁRIO NACIONAL DO PROJÓVEM: “JUVENTUDE E AÇÕES DO ESTADO: PARÂMETROS DE AVALIAÇÃO DE PROGRAMAS GOVERNAMENTAIS”, 2006, Belo Horizonte. Texto apresentado por ocasião do evento.

STUFFLEBEAM, D. L.; CORYN, C. L. *Evaluation theory, models, and applications*. San Francisco: Jossey-Bass, 2012.

TRZESNIAK, P. Indicadores quantitativos: reflexões que antecedem seu estabelecimento. *Ciência da Informação*, Brasília, v. 27, n. 2, p. 159-164, maio/ago. 1998.

VIANNA, H. M. *Fundamentos de um programa de avaliação educacional*. Brasília: Liber Livro, 2005.

WORTHEN, B.; SANDERS, J.; FITZPATRICK, J. *Avaliação de programas: concepções e práticas*. São Paulo: EDUSP:Gente, 2004.

Indicators for the evaluation of educational programs: methodological challenges

Abstract

This paper discusses the methodological strategies used to meet the challenge of establishing indicators to compose the impact evaluation of an in-service teacher training program called “Letra e Vida”. This program was implemented in the

school system of the state of São Paulo, from 2003 to 2006, in order to improve the practice of literacy teachers working in the early years of elementary school. Preceding the methodological discussion, some concepts are discussed based on contributions from the academic literature, such as the notion of indicators, types of indicators, and their properties, as well as the political and technical aspects underlying the definition of indicators, which might support the evaluation of educational programs.

Keywords: Educational indicators. Program evaluation. Teacher training.

Indicadores para la evaluación de programas educativos: desafíos metodológicos

Resumen

Este artículo analiza las estrategias metodológicas utilizadas para cumplir con el reto de establecer indicadores con los que componer la evaluación del impacto de un programa de capacitación para profesores en servicio: el Programa Letra y Vida. Este programa se implementó en la red de enseñanza del estado de São Paulo, durante el período 2003-2006, con objetivo de mejorar la práctica de aquellos docentes alfabetizadores que trabajan en los primeros años de Enseñanza Primaria. Precediendo la discusión metodológica, se discuten, basándose en las contribuciones de la literatura, la noción de indicadores, de sus tipos y sus propiedades, así como los aspectos técnicos y políticos que subyacen en la definición de dichos indicadores y que pueden apoyar la evaluación de los programas educativos.

Palabras clave: Indicadores educativos. Evaluación de programas. Formación docente.

Informações dos autores

Adriana Bauer: adbauer@fcc.org.br

Sandra Zákia Sousa: sanzakia@usp.br