

Boxe 10 – O impacto de curto prazo do programa Aprender Valor no letramento financeiro de crianças e jovens

Introdução

O programa Aprender Valor é uma iniciativa do Banco Central do Brasil (BC) com o objetivo de apoiar escolas públicas e redes de ensino no desenvolvimento da Educação Financeira como tema transversal no ensino fundamental (EF) no país. Desde 2020, o programa oferece às escolas participantes: curso de formação para professores e gestores escolares; projetos escolares que abordam temas de educação financeira junto com temas das disciplinas obrigatórias; e avaliações (testes) de letramento financeiro para aplicação em sala de aula.¹ A escolha de quais recursos utilizar é uma opção da escola, e, assim, o nível de engajamento pode variar entre uma escola e outra.

São muitos os desafios de implementação de um programa educacional de abrangência nacional, como é o caso do Aprender Valor, especialmente em função das características do nosso país, de dimensões continentais. Nesse contexto, um primeiro desafio é fazer com que escolas de todo o país saibam que o Aprender Valor existe e se interessem em fazer adesão voluntária ao programa. O segundo desafio é estimular que gestores escolares e professores se engajem e utilizem todos os recursos educacionais oferecidos. Um terceiro desafio é medir a efetividade do programa, para identificar se está provocando os resultados planejados.

Em 2022, o BC se propôs a endereçar esse último desafio. Com o objetivo de mensurar o impacto de curto prazo dos cursos de formação e dos projetos escolares na aprendizagem dos estudantes, foi realizada uma avaliação de impacto, por meio da aplicação de testes de letramento financeiro para os alunos de uma amostra de escolas recém-aderidas ao programa.

A amostra foi composta por estudantes do 3º, 5º, 7º e 9º anos de 783 escolas públicas do ensino fundamental, das redes estaduais e municipais,² sorteadas dentre as escolas que haviam feito a adesão voluntária ao programa entre novembro de 2021 e fevereiro de 2022. Os 52.078 participantes da amostra estavam divididos em 2.064 turmas.³

1 Para mais detalhes sobre o programa Aprender Valor, veja o [link: Aprender Valor \(bc.gov.br\)](http://bc.gov.br).

2 Foram desconsiderados os estudantes de turmas multietapa, isto é, turmas que reúnem estudantes de mais de um ano escolar, de atendimento especial exclusivo (estudantes com necessidades educacionais específicas, a exemplo dos estudantes com deficiência), escolas indígenas e de local de funcionamento diferenciado. Essa exclusão se deveu ao pequeno número de escolas com esses perfis na população e às necessidades educacionais específicas desses estudantes, que provavelmente demandariam uma avaliação à parte. Por questões operacionais, foram consideradas elegíveis apenas escolas com dez ou mais alunos matriculados, segundo Censo Escolar 2020, o mais recente disponível à época.

3 O tamanho da amostra foi calculado visando possibilitar a identificação de uma diferença de 20 pontos das variações médias da nota de proficiência em letramento financeiro, numa escala única proposta para todo o ensino fundamental com variação de 0 a 1.000 pontos, a um nível de significância de 5% e com poder de 80%. Considerou-se um desvio padrão de 100 pontos, tamanho médio por turma de 25 estudantes e correlação intraclasses de 0,30. A partir de informações de execução do programa e de participação nas avaliações formativas no ano de 2021, um aumento de 50% foi aplicado no intuito de absorver eventuais perdas. O plano amostral aplicado é de uma amostra estratificada, cujos estratos são as combinações das cinco regiões do Brasil (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) e área (Região Metropolitana e Interior). Em cada um dos dez estratos, as escolas foram separadas em grupos: um das que ofertaram o programa nas etapas iniciais do ensino fundamental (1º a 5º anos) e outro das que ofertaram nos anos finais (6º a 9º anos). Em cada grupo foram sorteadas as escolas por meio de uma amostra aleatória simples; e, em cada escola sorteada, foi selecionada uma turma de 3º e de 5º anos, para as de anos iniciais, e uma de 7º e de 9º anos, para as de anos finais. Foram observados todos os alunos dessas turmas selecionadas. Após o sorteio da amostra, foi realizado o sorteio dos grupos de tratamento e controle, dividindo os quantitativos em cada estrato de forma uniforme nos grupos.

As escolas da amostra foram separadas em dois grupos: o grupo controle (n = 386 escolas) não teve acesso, ao longo de 2022, aos recursos educacionais oferecidos na plataforma do programa – formações para professores e gestores e projetos escolares com sugestões de aulas que integravam a educação financeira às disciplinas obrigatórias; o grupo tratamento (n = 397 escolas) teve acesso liberado na plataforma a esses recursos. Todas as escolas passaram a ter acesso aos recursos após a conclusão da avaliação de impacto (novembro de 2022).

Os testes de letramento financeiro foram elaborados especificamente para os participantes do programa com objetivo de mensurar a assimilação dos conteúdos abordados nos projetos escolares. Diferentes testes foram aplicados para as turmas do 3º, 5º, 7º e 9º anos, com 24, 27, 30 e 33 itens de prova, respectivamente. Alguns itens comuns e outros diferentes entre os testes garantem a progressão do nível de dificuldade e sustentam a proposta de escala única de proficiência de letramento financeiro para todo o ensino fundamental.

A primeira avaliação (teste de entrada) foi feita entre abril e maio de 2022; a segunda (teste de saída), entre outubro e novembro do mesmo ano.

Resultados da avaliação

No total, foram obtidas informações de 48.613 estudantes (93,3% da amostra).⁴ Entretanto, devido à mobilidade dos estudantes entre escolas e a dificuldades logísticas do campo, apenas 25.420 estudantes (52,3%) tiveram notas do teste de proficiências identificadas nos dois momentos de avaliação. Não se verificaram distribuições distintas de região-área de localização ($p=0,727$) e série ($p=0,810$) entre os grupos.

A pontuação no teste de proficiência em letramento financeiro foi obtida via Teoria de Resposta ao Item.⁵

Inicialmente, para se avaliar o efeito do tempo (diferença entre os testes de entrada e saída) e tratamento (acesso ou não acesso a recursos didáticos do Aprender Valor), foram utilizados modelos longitudinais multinível⁶ que avaliam a existência de três componentes: tempo, grupo e interação entre grupo e tempo. Os resultados apontam para uma evolução média similar da proficiência em letramento financeiro para a amostra como um todo ($p=0,329$) e por ano escolar. Observaram-se efeitos de tempo para a amostra global e por ano escolar: para a amostra global, 3º, 5º e 7º ano nota-se um aumento médio da primeira para a segunda avaliação. Padrão inverso foi observado para o 9º ano, provavelmente associado ao desenho do teste.⁷ Considerando a amostra global, não se verificaram efeitos de grupo, quando comparados controle e tratamento, isto é, não foi possível captar o efeito do tratamento aplicado.

Para compreender melhor os motivos que poderiam ter levado a esse resultado, foram realizadas análises complementares dos dados. Uma primeira conclusão a que se chegou nessas análises foi a de que o tratamento foi bastante heterogêneo, isto é, o nível de engajamento, ou o efetivo uso dos recursos disponibilizados, pelas escolas do grupo tratamento, variou bastante. Por isso, decidiu-se investigar se o nível de engajamento das escolas

4 A diferença entre a amostra planejada e a efetiva deve-se às perdas de campo, como dificuldade de acesso à escola para aplicação dos testes, não localização da turma devido à desatualização do censo escolar ou não realização do teste de letramento financeiro pela escola/turma ou estudante selecionado. Alguns diretores consideraram seus alunos despreparados para realizar as avaliações, outros estudantes não assinaram o cartão-resposta ou não sabiam ler e escrever, ou ainda a frequência escolar nos dias das provas foi prejudicada pelas chuvas, entre outros motivos.

5 Essa metodologia tem o objetivo de produzir uma escala única para os estudantes avaliados. O escore bruto gerado pela TRI não apresenta um limite inferior ou superior padrão. Entretanto, para facilitar a interpretação, o escore foi reescalado de tal forma que o valor mínimo e máximo observados na amostra desse estudo correspondam aos valores zero e mil, respectivamente.

6 Nesse modelo, os alunos correspondem ao primeiro nível, as turmas ao segundo, e as escolas ao terceiro nível, incorporando o efeito dos alunos, turma e escola na forma de efeitos aleatórios, acomodando uma possível dependência entre observações de um mesmo aluno (em momentos distintos), bem como de alunos de uma mesma turma, ou de turmas em uma mesma escola.

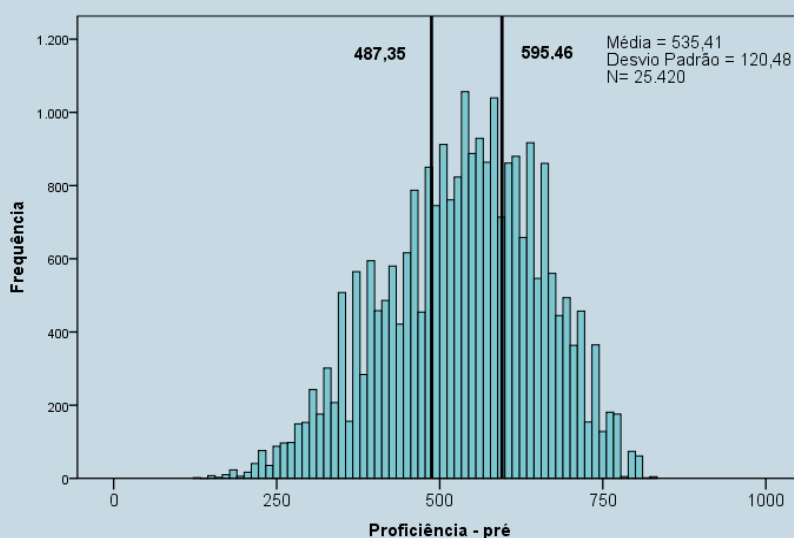
7 Notou-se a inclusão de dois descritores com baixo percentual de acerto apenas no teste de saída do 9º ano, ou seja, o teste de saída apresentou-se mais difícil que o teste de entrada.

no programa interferiu no desempenho dos estudantes. Para isso, foram levantadas as seguintes informações na plataforma do Aprender Valor:

- Realização de cursos de formação por professores ou gestores escolares na escola, havendo três níveis possíveis de engajamento:
 - sem formação iniciada (ou seja, não houve engajamento dos profissionais da escola nos cursos);
 - com formação iniciada (curso em andamento, mas não concluído);
 - com formação concluída (curso finalizado).
- Aplicação de projetos escolares, em sala de aula, pelos professores da escola, havendo três níveis possíveis de engajamento:
 - sem projeto selecionado (ou seja, não houve engajamento em projetos escolares);
 - com projeto selecionado (professor baixou o projeto escolar na plataforma do programa);
 - com projeto concluído (professor baixou o projeto, ministrou a aula e registrou a aula na plataforma).

Ainda, a análise complementar permitiu concluir que também poderia haver heterogeneidade no grau de proficiência em letramento financeiro inicial, isto é, no nível de conhecimento dos estudantes medido pela avaliação inicial (teste de entrada). Assim, com o objetivo de controlar o efeito dessa possível heterogeneidade, a amostra foi segmentada em três grupos de tamanhos similares (tercis) segundo essa variável (Gráfico 1). A variação das pontuações no teste de proficiência em letramento financeiro entre os dois momentos foi avaliada para cada um dos três grupos.

Gráfico 1 – Histograma do escore de proficiência em letramento financeiro na avaliação inicial



A partir dessas informações, foi proposta outra análise dos dados, empregando as árvores de decisão, tendo como variável dependente a **variação da pontuação nos testes de proficiência** entre a avaliação inicial (teste de entrada) e a avaliação final (teste de saída). Como variáveis predictoras, foram empregadas as variáveis de

status de engajamento do professor e do gestor e a aplicação de projetos escolares.⁸ A árvore de decisão permite avaliar as interrelações entre as variáveis preditoras de forma a identificar grupos com maiores e menores variações médias de proficiência.

As árvores de decisão identificaram que a variação entre as pontuações da avaliação inicial e da avaliação final tendeu a ser maior em escolas mais engajadas, ou seja, em que professores e gestores haviam iniciado ou concluído as formações e em que a aplicação dos projetos foi efetivamente concluída. Em relação aos tercis de proficiência na avaliação inicial, o grupo do primeiro tercil apresentou a variação maior entre a avaliação inicial e a final, e o grupo do terceiro tercil apresentou a menor variação.

Para se comparar os grupos formados nas árvores de decisão, ajustadas por ano escolar, região e área de localização, foi utilizado o modelo de regressão linear multinível com três componentes: tempo, grupo e interação entre grupo e tempo.⁹ A partir desse modelo, foram estimadas as variações médias por grupo (Tabela 1).

Tabela 1 – Estimativa das médias e respectivo intervalo de confiança de 95% estimadas via modelo linear multinível

Grupos	Descrição	Estimativa via modelo linear multinível	
		Média (IC95%)	p
T1: P-SI e G-SI ^B	1º Tercil – Professor e gestor sem formação iniciada	47,61 (45,10 a 50,13)	<0,001
T1: P-SI e G-I+A	1º Tercil – Professor sem formação iniciada e gestor com formação iniciado ou aprovada	55,13 (45,49 a 64,77)	<0,001
T1: P-I+A e Proj-NS+S	1º Tercil – Professor com formação iniciada ou formado e sem projeto selecionado ou selecionado mas não concluído	52,78 (46,23 a 59,33)	<0,001
T1: P-I+A e Proj-C ^A	1º Tercil – Professor com formação iniciada ou formado e projeto concluído	67,49 (53,47 a 81,50)	<0,001
T2: G-SI +I ^C	2º Tercil – Gestor sem formação iniciada ou iniciada	5,59 (3,40 a 7,78)	<0,001
T2: G-A + P- SI+I ^C	2º Tercil – Gestor aprovado e professor sem formação iniciada ou iniciada	16,12 (6,18 a 26,07)	0,002
T2: G-A + P- A ^C	2º Tercil – Gestor aprovado e professor formado	8,97 (2,48 a 15,47)	0,007
T3: P-SI +I ^D	3º Tercil – Professor sem formação iniciada ou iniciada	-28,73 (-31,03 a -26,43)	<0,001
T3: P-A ^D	3º Tercil – Professor formado	-25,60 (-34,64 a -16,57)	<0,001

N=25.420 alunos.

(A), (B), (C) e (D) e cores diferentes representam médias distintas segundo comparações múltiplas com correção de Bonferroni.

T1, T2 e T3 – Primeiro tercil, segundo tercil e terceiro tercil, respectivamente.

P – Professor, G – Gestor e Proj – Projeto.

SI-Sem formação iniciada; I – Formação iniciada, A – Aprovado/Formado; NS+S – Não seleção ou seleção sem conclusão; C: Concluído

Ajustados por região, área de localização e ano escolar, na amostra considerando todos as séries escolares (coluna 2), os grupos apresentaram variações médias distintas. Novamente, os grupos do primeiro tercil apresentaram variações médias superiores às do segundo tercil que, por sua vez, apresentaram médias superiores às do terceiro tercil. Entre os grupos do primeiro tercil, o subgrupo com maior engajamento no programa (T1:P-I+A e Proj-C) apresentou variação média superior à do grupo sem nenhum engajamento (T1: P-SI e G-SI). Já as médias entre grupos do segundo tercil não foram distintas. Padrão similar foi observado para os grupos do terceiro tercil.

⁸ As análises a partir desse ponto consideram as informações dos 25.420 estudantes que participaram dos dois momentos de avaliação.

⁹ Neste modelo, os alunos correspondem ao primeiro nível, a turma ao segundo e as escolas ao terceiro nível, incorporando o efeito dos alunos, turma e escola na forma de efeitos aleatórios, acomodando uma possível dependência entre observações de um mesmo aluno (em momentos distintos), bem como de alunos de uma mesma turma, ou de turmas em uma mesma escola.

Conclusões

A avaliação de impacto do Aprender Valor permitiu o levantamento de dados que auxiliarão não só na análise da efetividade do programa, como também no delineamento de estratégias para seu fortalecimento e monitoramento.

De forma geral, foi possível concluir que **os impactos do programa em relação ao letramento financeiro dos estudantes variam conforme o desempenho do estudante em letramento financeiro no início do ano e o grau de engajamento das escolas no programa.**

As maiores variações no desempenho em letramento financeiro entre o teste de entrada e o de saída foram entre os estudantes que sabiam menos no início do ano. Esses estudantes apresentaram um aumento médio em letramento financeiro, entre o início e o final do ano, superior aos demais.

Outro fator que influenciou na variação da proficiência em letramento financeiro entre a avaliação inicial e a final foi o **efetivo engajamento das escolas**. A diferença média entre os testes de entrada e saída foi maior para estudantes de escolas em que gestores e professores participaram dos cursos de formação e em que os projetos escolares foram aplicados em sala de aula, isto é, o engajamento da escola no programa contribuiu para uma maior evolução do letramento financeiro dos estudantes ao longo do ano.

Para além das conclusões sobre a efetividade do programa, é válido ressaltar que todos os desafios mencionados anteriormente para a implementação de uma iniciativa educacional de abrangência nacional ficaram bastante claros durante a aplicação da avaliação de impacto aqui relatada. Na prática, embora a amostra de escolas tenha sido ampla e representativa, e um número adequado de testes tenha sido aplicado na entrada e na saída, o critério de distinção entre controle e tratamento não foi plenamente atendido.¹⁰

Assim, apesar dos desafios de implementação do programa e da própria avaliação, podemos concluir que os **dados indicam que os recursos educacionais do Aprender Valor parecem de fato contribuir para a evolução do letramento financeiro dos estudantes alcançados pelo programa.**

Como desafios para o futuro, podemos mencionar, em primeiro lugar, o efetivo engajamento dos profissionais das escolas já participantes, com a realização das formações, a aplicação dos projetos escolares e das avaliações de letramento financeiro. Em segundo, a continuidade do monitoramento de resultados do programa, a fim de subsidiar a constante evolução do Aprender Valor e do letramento financeiro das crianças e dos jovens por ele alcançados.

10 Avaliando-se as informações provenientes dos professores, gestores e alunos, verificou-se, em virtude da mobilidade dos professores e gestores entre as escolas, uma contaminação parcial no grupo controle, bem como um nível de engajamento das escolas do grupo tratamento inferior ao esperado. A contaminação do grupo controle foi moderada: as porcentagens de estudantes com professores com formação concluída (17,1% versus 0,6%), gestores com formação concluída (12,5% versus 0,2%) e projetos escolares concluídos (21,5% versus 0,9%) são maiores no grupo tratamento, quando comparados com grupo controle. Em relação às escolas do grupo tratamento, observamos que parte significativa não se engajou plenamente no programa. Uma explicação pode estar no fato de que, em 2022, a maioria das escolas públicas ainda vivia um ano atípico, de retorno à normalidade após a pandemia da covid-19. Gestores e professores enfrentavam uma sobrecarga em função de dificuldades para a readaptação de estudantes às aulas presenciais e para a recomposição de aprendizagem, em virtude dos déficits deixados pelo período pandêmico. É preciso ainda considerar que o expressivo número de escolas envolvidas na avaliação e a sua elevada dispersão geográfica dificultaram um monitoramento mais próximo da equipe do BC em relação à aplicação do programa nas escolas do grupo tratamento.