

*CAMINHOS  
PARA A  
QUALIDADE  
DA EDUCAÇÃO  
PÚBLICA:*

***IMPACTOS E  
EVIDÊNCIAS***

*CAMINHOS  
PARA A  
QUALIDADE  
DA EDUCAÇÃO  
PÚBLICA:*

***IMPACTOS E  
EVIDÊNCIAS***



*CAMINHOS  
PARA A  
QUALIDADE  
DA EDUCAÇÃO  
PÚBLICA:*

***IMPACTOS E  
EVIDÊNCIAS***

APRESENTAÇÃO

- 8 *PROJETO JOVEM DE FUTURO:  
IMPACTOS, EVIDÊNCIAS E CONQUISTAS*  
**PEDRO MOREIRA SALLES**

INTRODUÇÃO

- 16 *EVIDÊNCIA PARA TRANSFORMAR*  
**MIRELA DE CARVALHO**

PARTE I

- 22 ***GESTOR E AVALIADOR:  
CONHECIMENTOS EM DIÁLOGO PARA  
TRANSFORMAR A EDUCAÇÃO PÚBLICA***

- 24 *PRODUÇÃO E USO DE EVIDÊNCIAS  
PARA GARANTIR O DIREITO À EDUCAÇÃO*  
**RICARDO HENRIQUES**

- 36 *JOVEM DE FUTURO: O IMPACTO DA GESTÃO  
SOBRE O APRENDIZADO DOS ESTUDANTES*  
**RICARDO PAES DE BARROS**

PARTE II

- 46 ***PRODUÇÃO DE CONHECIMENTO PARA  
APRIMORAR PROJETOS EDUCACIONAIS:  
O USO DA AVALIAÇÃO DE IMPACTO  
COM DESENHO EXPERIMENTAL***

- 48 *DESAFIOS PARA A PRODUÇÃO DE CONHECIMENTO  
CIENTÍFICO EM EDUCAÇÃO A SERVIÇO  
DAS POLÍTICAS PÚBLICAS*  
**RICARDO MADEIRA**

- 60 *USO DE AVALIAÇÃO EXPERIMENTAL  
PARA CÁLCULO DE EFEITOS DE PROGRAMAS  
EDUCACIONAIS SOBRE A DESIGUALDADE  
DE DESEMPENHO*  
SERGIO FIRPO
- 72 *COMO ABRIR A CAIXA-PRETA DO IMPACTO*  
PEDRO CARNEIRO
- 82 *CIÊNCIA PARA A EDUCAÇÃO:  
UMA JANELA DE OPORTUNIDADE PARA O BRASIL*  
ROBERTO LENT

PARTE III

- 98 ***META-ANÁLISES:  
OS IMPACTOS DAS POLÍTICAS EDUCACIONAIS  
E SUAS REVISÕES SISTEMÁTICAS***
- 100 *META-ANÁLISE DE IMPACTOS: USO DE EVIDÊNCIAS  
PARA PRIORIZAR INVESTIMENTOS EDUCACIONAIS*  
ANDRÉ PORTELA
- 114 *POLÍTICA E PRÁTICAS BASEADAS EM EVIDÊNCIAS:  
DE ONDE VIERAM, PARA ONDE VÃO E  
POR QUE SÃO IMPORTANTES*  
HERB TURNER
- 128 *LIÇÕES EXTRAÍDAS DE UMA REVISÃO SISTEMÁTICA  
DO QUE FUNCIONA EM EDUCAÇÃO*  
EDOARDO MASSET

PARTE IV

- 144 **PRODUÇÃO E USO DE EVIDÊNCIAS  
EM EDUCAÇÃO NA VISÃO DOS  
PESQUISADORES BRASILEIROS**
- 146 *PRODUÇÃO E USO DE EVIDÊNCIAS  
NA EDUCAÇÃO BÁSICA*  
JOSÉ FRANCISCO SOARES
- 156 *O PAPEL DOS PESQUISADORES DA EDUCAÇÃO  
NA PRODUÇÃO E DIFUSÃO DE PESQUISA EDUCACIONAL*  
PAULA LOUZANO
- 172 *EM BUSCA DA MELHORIA DA QUALIDADE  
DO CLIMA ESCOLAR*  
TELMA VINHA
- 192 *PESQUISA SERVE PARA ALGUMA COISA?*  
CLÁUDIO DE MOURA CASTRO

PARTE V

- 206 **GESTORES EDUCACIONAIS  
E O USO DE EVIDÊNCIAS:  
EXPERIÊNCIAS E REFLEXÕES**
- 208 *FINLÂNDIA: EVIDÊNCIAS PARA CONSTRUIR UM DOS  
MELHORES SISTEMAS EDUCACIONAIS DO MUNDO*  
ILKKA TURUNEN
- 226 *A COMPLEXIDADE DA AVALIAÇÃO EM EDUCAÇÃO:  
A EXPERIÊNCIA MEXICANA*  
MIGUEL SZÉKELY
- 246 *SENSO COMUM VERSUS EVIDÊNCIAS:  
RISCOS DA CULTURA DO "ACHISMO"*  
EDUARDO DESCHAMPS

- 256 *PERNAMBUCO: GESTÃO E EVIDÊNCIAS*  
**FREDERICO AMANCIO**
- 272 *A IMPORTÂNCIA DAS RELAÇÕES  
COM A ACADEMIA PARA A CONSTRUÇÃO  
DE EVIDÊNCIAS NA REDE DO ESPÍRITO SANTO*  
**HAROLDO CORRÊA ROCHA**
- 280 *EXPERIÊNCIA DA REDE DE ENSINO DE GOIÁS  
COM O USO DE EVIDÊNCIAS:  
IMPORTÂNCIA DA GESTÃO ESTRATÉGICA*  
**RAQUEL TEIXEIRA**

ANEXO I

- 290 *O PROJETO JOVEM DE FUTURO*

ANEXO II

- 312 *O IMPACTO DO PROJETO JOVEM DE FUTURO  
SOBRE A APRENDIZAGEM EM ESCOLAS PÚBLICAS*



# ***PROGRAMA JOVEM DE FUTURO: IMPACTOS, EVIDÊNCIAS E CONQUISTAS***

**PEDRO MOREIRA SALLES** é presidente do Instituto Unibanco e do Conselho de Administração do Itaú Unibanco Holding, sócio e copresidente da Cambuhy Investimentos, conselheiro da TOTVS e membro do Conselho de Administração da Falconi Consultores Associados. Membro do Conselho Orientador da Fundação Osesp e do Conselho Deliberativo do Insper, é graduado *magna cum laude* em economia e história pela Universidade da Califórnia em Los Angeles, cursou o programa de mestrado em relações internacionais da Universidade de Yale e o Owners/President Management Program da Universidade de Harvard.

O seminário “Caminhos para a qualidade da Educação pública: impactos e evidências” nos reuniu em setembro de 2016 em torno de uma crença fundamental que compartilhamos: o poder transformador que a Educação pode exercer na sociedade. Acreditamos que somente juntos – educadores, gestores públicos, representantes do terceiro setor e da academia – seremos capazes de atuar de maneira mais consequente e transformadora na Educação. O que buscamos naquele encontro de ideias que resultou neste livro foi mostrar aos jovens algo que, por incrível que pareça, esteve desacreditado em nosso país. Queremos provar um conceito que parece óbvio em muitas nações, mas que aqui ainda guarda contornos quase revolucionários: estudar vale a pena. Sabemos que muito vem sendo feito, mas não há a menor dúvida de que é preciso avançar muito mais.

Nossos resultados educacionais estão longe do desejado, especialmente no Ensino Médio. Alguns números provam esse ponto com contundência. Quem frequenta o Ensino Médio sem estar atrasado deveria ter entre 15 e 17 anos, mas apenas 61% dos nossos jovens nessa faixa etária estão no nível esperado. Todos

*Queremos provar um conceito que parece óbvio em muitas nações, mas que aqui ainda guarda contornos quase revolucionários: estudar vale a pena. Sabemos que muito vem sendo feito, mas não há a menor dúvida de que é preciso avançar muito mais*

os demais estão atrasados porque repetiram séries anteriores ou porque simplesmente estão fora da escola – a dura realidade é que 16% estão longe das salas de aula. De acordo com o Plano Nacional de Educação (PNE) aprovado em 2014, precisaríamos terminar o ano de 2016 com 100% dos jovens entre 15 e 17 anos na escola. Mas estamos muito distantes desse objetivo. O mesmo PNE estabelece que devemos chegar a 2024 com 85% dos jovens entre 15 e 17 anos frequentando o Ensino Médio. Mantido o ritmo histórico, o país não alcançará essa meta.

Esses números dizem respeito à participação dos jovens brasileiros nos programas educacionais do país. No entanto, tão importante quanto o universo de alunos que vão à escola é a qualidade da aprendizagem oferecida. Nesse sentido, o Índice de Desenvolvimento da Educação Básica (Ideb), que combina proficiência e aprovação nas escolas das redes públicas de ensino e foi divulgado em setembro de 2016, ratifica que o Ensino Médio se mantém em situação de extrema fragilidade. O Ideb desse segmento, que vinha desde 2009 em 3,4, apresentou uma pequena oscilação positiva e chegou a 3,5. Ainda assim, encontra-se muito abaixo da meta para 2015, que é de 3,9. O país tem o compromisso de chegar a um Ideb de 6,0 no Ensino Médio até 2028, que equivaleria ao nível registrado nos países que fazem parte da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) no já distante ano de 2003.

## **A IMPORTÂNCIA DA BOA GESTÃO**

O Instituto Unibanco aposta na gestão – acima de tudo – para dar um passo decisivo em direção às transformações profundas e necessárias para atingir os objetivos que acabo de mencionar. Se há algo em que apostei ao longo de toda a minha trajetória empresarial – e que jamais me trouxe qualquer tipo de desapontamento – é a boa gestão. Acredito que as boas práticas de gestão são princípios universais que podem – e devem – ser aplicados às escolas na construção de um projeto de Educação para o país.

Investir em gestão significa criar condições para o surgimento de boas ideias e iniciativas nas escolas e nas secretarias de Educação. Ela lança luz sobre os processos de aprendizagem e ajuda a produzir evidências sobre avanços e estagnações nesses processos. Além disso, o exercício da boa gestão nos condiciona a buscar sempre respostas para questões decisivas. Nosso esforço está sendo suficiente? Se não está, por quê? Se está, por quê? O que deveríamos mudar para obter melhores resultados?

Em ambientes onde existe uma cultura de boa gestão, aprende-se sistematicamente sobre quais práticas funcionam melhor. Esses aprendizados são colocados à disposição dos processos de trabalho e da tomada de decisão, que, por sua vez, tornam-se mais eficazes e alimentam um novo ciclo virtuoso. A boa gestão é um instrumento decisivo para ultrapassar o muro que separa a sociedade de hoje daquela que queremos construir. Esse muro obstrui nossa visão e nos impede de ir além. Para ultrapassá-lo, precisamos focar em ter os alunos na escola e em sua série; priorizar desafios e formular boas soluções; e persistir e reformar as soluções sempre que necessário.

Para isso, é importante fortalecer o conhecimento do que funciona – de tudo o que aprendemos com tantas pesquisas, avaliações, dados de monitoramento e boas práticas. Esta publicação sintetiza o conhecimento de cientistas de diversas áreas, avaliadores e gestores reunidos no seminário “Caminhos para a qualidade da Educação pública: impactos e evidências” para discutir como as evidências apuradas até aqui podem ajudar na tomada de decisão na Educação.

Por um lado, perguntamos a cientistas e avaliadores: “O que se sabe hoje, com base em evidências, sobre o que precisa ser feito em Educação para alcançar melhores resultados de aprendizagem?”. Por outro, solicitamos aos gestores que expusessem quais perguntas cientistas e avaliadores precisam priorizar. Muitos projetos são excelentes em termos de potencial, mas dependem da implantação e do contexto para gerar resultados efetivos. Apenas a prática mostrará se funcionarão ou não.

O projeto Jovem de Futuro do Instituto Unibanco, iniciado em 2007, foi criado, implementado e é permanentemente aprimorado seguindo exatamente esse princípio. Ele prevê formação em gestão, assessoria técnica e acompanhamento para gestores, orientando-os para perseguir metas de aprendizagem. Isso se dá sem qualquer transferência de recursos financeiros entre as secretarias estaduais de Educação e o Instituto Unibanco. O projeto encontra-se na terceira versão, pois vem sendo ajustado em função de evidências, exatamente como defendemos. Alguns exemplos de ajustes realizados a partir das evidências merecem ser destacados.

## ***JOVEM DE FUTURO: EVIDÊNCIAS E AJUSTES***

Sobre a primeira versão – ainda na fase piloto –, destacamos o aprendizado associado à importância de desenvolver uma avaliação estritamente externa dos estudantes, sem qualquer participação dos gestores, mesmo quando os conteúdos das avaliações eram tecnicamente consistentes. As evidências mostravam que havia elevado risco de incentivos negativos para treinar para os testes.

Em sua segunda versão, o projeto estava atrelado ao Programa Ensino Médio Inovador (ProEMI) do Ministério da Educação (MEC) e fazia transferência de recursos financeiros para as escolas, o que deveria incentivá-las a transformar sua gestão. Nossas avaliações demonstraram que o envio de recursos não necessariamente as incentivava no caminho desejado. Pelo contrário, recolhemos evidências de que os gestores permaneciam presos a um plano de gastos e não modificavam sua maneira de plane-

## *Investir em gestão significa criar condições para o surgimento de boas ideias e iniciativas nas escolas e nas secretarias de Educação*

jar. Para gerar os incentivos necessários, avançamos na direção de construir um compromisso de metas envolvendo secretarias, regionais de ensino e escolas.

Outro exemplo de mudança de desenho está em nosso conteúdo de formação para gestores. Em versões anteriores, embora bastante aplicado, o material não era suficiente para que eles conseguissem instaurar um processo de gestão realmente novo. Agora, criamos protocolos capazes de estabelecer esses novos processos transformadores, fortalecendo e sendo fortalecidos pelas formações oferecidas.

Desde o início, o Jovem de Futuro foi submetido a uma rigorosa avaliação de impacto com desenho experimental. Veremos nesta publicação mais detalhes sobre esse tipo de avaliação, comumente usado na área biomédica e segundo o qual implementa-se a intervenção em um grupo chamado “tratado” e mantém-se um grupo de controle para poder comparar os efeitos gerados.

Destaco algumas datas-chave em uma breve cronologia do projeto Jovem de Futuro, para exemplificar o poder transformador das evidências. Embora tenhamos iniciado o diálogo com as primeiras redes parceiras em 2007, o atendimento na primeira geração do programa só começou em 2008, em escolas da região metropolitana de Belo Horizonte (MG) e Porto Alegre (RS). Naquela época, havia cerca de 50 escolas participantes em cada localidade. Metade era atendida nos primeiros três anos, e a outra metade, nos três anos seguintes. A melhora observada no grupo inicial – quando comparada com os resultados obtidos pelo grupo que só entrou na segunda etapa, o chamado grupo de controle – serviu para avaliar os avanços conquistados. Assim, foi possível isolar exatamente o quanto da melhora nas escolas estava relacionada ao Jovem de Futuro.

Em 2012, na segunda geração do programa, nos tornamos parceiros dos estados de Ceará, Goiás, Mato Grosso, Pará e Piauí, e nos associamos ao Programa Ensino Médio Inovador (ProEMI) do Ministério da Educação (MEC). As escolas também entraram no projeto em ciclos, o que permitiu continuar com a avaliação de impacto. Nesses estados, os gestores se comprometeram com a produção de evidências sobre a relevância do projeto, antes de universalizá-lo. Até 2014, 2.100 escolas foram atendidas nesses estados. Em 2015, o projeto se desvincula do ProEMI e ganha mais força na atuação junto às regionais de ensino e ao órgão central da Secretaria de Educação.

Na terceira geração do Jovem de Futuro, o Espírito Santo se torna parceiro. Pará e Piauí migram para esse novo desenho, sempre com o compromisso de gerar conhecimento sobre o impacto da intervenção. Em 2016, Ceará e Goiás também se juntam à nova geração do programa. Nesses estados, o projeto já está sendo implantado com transferência de ações – que tradicionalmente cabiam ao Instituto Unibanco – para as secretarias de Educação. Estamos aprendendo juntos sobre como tornar o projeto cada vez mais duradouro e sustentável, para que todas as escolas dessas redes sejam atendidas.

Nossa capacidade efetiva de transferir para os estados a tecnologia do programa é o teste definitivo da qualidade e robustez do Jovem de Futuro. A dura realidade é que, se esse projeto, construído com a contribuição de tantas pessoas, não se transformar em política pública, terão sido desperdiçados oito anos de vultuosos investimentos e aprendizados muito relevantes. Neste momento, nosso ânimo é enorme. Esperamos que até 2018, nos estados onde estabelecemos parcerias, o projeto Jovem de Futuro atenda 2.500 escolas públicas, beneficiando 2 milhões de alunos.

*Nossa capacidade efetiva de transferir para os estados a tecnologia do programa é o teste definitivo do Jovem de Futuro*

Acredito que cabe uma reflexão sobre o impacto do Jovem de Futuro, que avalio como bastante significativa. O projeto fez com que a proficiência em língua portuguesa e matemática, medida pela escala do Sistema de Avaliação da Educação Básica (Saeb), ficasse cinco pontos acima do que as escolas teriam conseguido sem esse apoio. Vale ressaltar que a operação em escala do programa reduziu – como seria esperado – os primeiros valores de impacto observados na fase piloto, quando a abordagem era quase laboratorial. Contudo, o patamar de cinco pontos dos dados em escala é robusto e mostra um resultado indiscutivelmente relevante. Isso é muito? Em um país com o gigantesco desafio educacional do Brasil não é suficiente. Por outro lado, frente à dificuldade de melhora que enfrentamos, certamente não é pouco.

## **CONCLUSÃO: PRECISAMOS ACELERAR**

A lição que fica é de que precisamos aprender mais rápido sobre o que funciona e sobre os caminhos que nos farão acelerar. Precisamos de todo o conhecimento disponível para enfrentar esse desafio crucial. Ele não se encontra apenas nas publicações científicas: pode e deve ser buscado na experiência dos gestores do setor público e do terceiro setor. Temos plena consciência de que a gestão é apenas uma parte da solução.

Por outro lado, acreditamos que, se a boa gestão não é suficiente para mudar por si só o padrão de aprendizagem, ela é absolutamente necessária. O Jovem de Futuro apresenta fortes evidências nessa direção, apesar das dificuldades inerentes à implantação de um programa padronizado em um ambiente heterogêneo e volátil. O histórico apresentado aqui mostra como um projeto construído em conjunto com as secretarias é robusto e bem-sucedido, demonstrando que é possível transformar o panorama da Educação no Brasil – algo pelo qual lutamos há tantos anos.

O objetivo desta publicação é compartilhar o conhecimento acerca de impactos e evidências para que consigamos ultrapassar o muro que nos impede de transformar o Brasil em um país onde o direito à aprendizagem seja realmente respeitado e inquestionavelmente valorizado.



# ***EVIDÊNCIA PARA TRANSFORMAR***

**MIRELA DE CARVALHO** é gerente de gestão do conhecimento do Instituto Unibanco. Foi gerente de projetos no Instituto Ayrton Senna, assessora na Secretaria de Educação do Estado do Rio de Janeiro, pesquisadora no Instituto de Pesquisa Econômica Aplicada (Ipea) e no Instituto de Estudos do Trabalho e Sociedade (IETS). Economista, é doutora em sociologia pelo Instituto Universitário de Pesquisas do Rio de Janeiro (IUPERJ).

Os desafios do Brasil para garantir a todos o direito à aprendizagem precisam de enfrentamento urgente. São milhares de jovens com esse direito negado ano após ano, nos afastando de uma sociedade mais justa e desenvolvida. Como fazer melhor? Quais práticas são mais efetivas e em que contextos? Para trilhar um caminho de transformações reais, há muito a ser feito e, por isso, será fundamental aprender permanentemente com acertos e erros. A ciência aplicada pode nos ajudar muito nesse percurso, pois aporta conhecimento consolidado e rigoroso, gera novas evidências e adensa reflexões. Entretanto, sem gestores atentos e envolvidos na geração de conhecimento, a ciência pouco pode contribuir para a transformação educacional. Afinal, são eles, com suas decisões do dia a dia, que vão moldando, em conjunto com os estudantes, professores e demais membros da comunidade escolar, um novo ambiente, novas políticas e maiores oportunidades de desenvolvimento.

O Brasil é extremamente rico em práticas pedagógicas e de gestão educacional, com resultados bastante promissores. De

Sobral (CE) a Cacau dos Alves (PI), do Espírito Santo a Pernambuco, há várias escolas fazendo a diferença, acertando em ações que dão resultado e melhoram o desenvolvimento dos estudantes, no funcionamento e na qualidade das relações humanas estabelecidas. Por essa razão, se é dito que os gestores têm a aprender com a ciência, no Brasil, a ciência tem muito o que aprender com a prática.

O diálogo virtuoso entre os gestores educacionais e a ciência é o objeto tratado neste livro, *Caminhos para a qualidade da Educação: impactos e evidências*, resultante do seminário de mesmo nome ocorrido em setembro de 2016 em São Paulo. É o segundo volume da série promovida pelo Instituto Unibanco e seus parceiros<sup>1</sup>.

A primeira parte do livro aborda a avaliação de impacto do projeto Jovem de Futuro, implementado pelo Instituto Unibanco em parceria com redes públicas estaduais de ensino. O impacto sobre a proficiência dos alunos foi aferido entre os anos de 2008 e 2014 em oito estados brasileiros e foi de 5 pontos na escala Saeb. Para facilitar a comparação, tomemos como parâmetro o desvio padrão em proficiência tanto em matemática como em língua portuguesa no Brasil, que é da ordem de 33 pontos na escala Saeb. Assim, um impacto de 5 pontos alcançado pelo Jovem de Futuro corresponde a 15% de um desvio padrão. Com desenho experimental e resultados relevantes, o projeto indica possibilidades de uso efetivo de evidências para aprimorar o desenho da intervenção.

Ricardo Henriques, superintendente executivo do Instituto Unibanco, e Ricardo Paes de Barros, o principal nome em avaliação de impacto no país, ilustram, na prática, como travar uma conversa bem-sucedida entre gestor e avaliador. Após definir o que se entende no meio científico por avaliação de impacto com desenho experimental e defender que apenas ela permite isolar a contribuição do projeto para o resultado final alcançado, o arti-

---

1. O primeiro volume, *Caminhos para a qualidade da gestão pública: gestão escolar* encontra-se disponível em <[https://observatoriodeeducacao.org.br/wp-content/uploads/2017/08/caminhos\\_seminario\\_completo.pdf](https://observatoriodeeducacao.org.br/wp-content/uploads/2017/08/caminhos_seminario_completo.pdf)> e em <[http://www.fundacaosantillana.com.br/pdfs/caminhos\\_educacao.pdf](http://www.fundacaosantillana.com.br/pdfs/caminhos_educacao.pdf)>. Últimos acessos em setembro de 2017.

go sugere “um lugar” para a avaliação de impacto no processo de geração de conhecimento sobre um projeto. É um lugar importante, sem dúvida, mas não é o único – afinal, a avaliação experimental não é suficiente para orientar o redesenho de uma ação. Avaliar a percepção e o engajamento dos atores, assim como o processo de implementação, e adotar métodos qualitativos de observação ou métodos etnográficos para compreender como se dão as relações entre o público afetado pela intervenção também integram o conjunto de saberes científicos de que o gestor precisa dispor para decidir com base em evidências e ter maior chance de sucesso. O artigo defende ainda que, para acumular conhecimento a partir das avaliações, é imprescindível contar com um modelo lógico que descreva a intervenção e uma sólida teoria sobre a mudança desencadeada. Afinal, a cada etapa concluída nos processos de avaliação, a teoria da mudança é questionada, e o modelo lógico, revisado.

A segunda parte do livro traz artigos de cientistas brasileiros – Ricardo Madeira, Sergio Firpo e Roberto Lent – e de um estrangeiro – Pedro Carneiro – que se dedicam à avaliação de impacto com desenho experimental de projetos educacionais. São apresentados modelos de geração de conhecimento voltados para alimentar o redesenho de projetos, além de situações e dilemas práticos enfrentados pelos avaliadores e gestores durante a condução das avaliações.

*Os desafios do Brasil para garantir a todos o direito à aprendizagem precisam de enfrentamento urgente. [...] A ciência aplicada pode nos ajudar muito nesse percurso, pois aporta conhecimento consolidado e rigoroso, gera novas evidências e adensa reflexões*

## *Sem boas teorias da mudança para a Educação tampouco se conseguirá fazer a diferença, pois as evidências deixam de ecoar e não podem gerar interpretações sobre o que mudar nas intervenções*

Na terceira parte do livro, André Portela, Herb Turner e Edoardo Masset nos explicam o que são meta-análises e como elas podem ser consultadas e utilizadas pelos gestores. De maneira geral, as meta-análises sistematizam o impacto de diferentes tipos de programas em todo o mundo, e a importação de projetos de um contexto para outro exige compreender, em detalhes, não só o desenho do mesmo, mas as condições de sua implementação. De qualquer forma, a existência de regularidade de impacto para certos tipos de programas em diferentes contextos indica robustez e dá alguma confiança para que se implementem ações da mesma natureza.

A quarta parte do livro amplia o olhar sobre a produção científica e o uso de evidências abrangendo a produção de ponta das pesquisas aplicadas no campo epistemológico da Educação brasileira. Francisco Soares defende que a produção científica precisa avançar no sentido de formular uma teoria pedagógica, pois hoje está na etapa de síntese de evidências e registro de associação entre elas. Falta explicação. Já Paula Louzano debate a importância da participação de pesquisadores com formação em Educação para o desenho e a produção de pesquisas educacionais e o rigor para desenvolver esses estudos, usando como exemplo currículos na área de matemática e sua relação com a pesquisa acadêmica. Telma Vinha mostra o esforço conjunto de pesquisadores brasileiros para construir um instrumento nacional de avaliação do clima escolar e defende que os resultados precisam retornar às escolas. As avaliações não podem ser um fim em si mesmas, precisam ser o meio a serviço de um projeto de transformação. Por fim, Cláudio Moura Castro levanta os

principais problemas que identifica na produção científica em Educação e os obstáculos que limitam a divulgação e a capacidade de influenciar.

Na quinta e última parte do livro, a palavra é dos gestores públicos nacionais e internacionais. O conhecimento empírico que acumulam a partir da vivência e proximidade com o processo educativo tem tanto valor quanto o conhecimento científico. Ilkka Turunen, da Finlândia, apresenta a relevância, a estratégia e os novos desafios da produção de conhecimento no país nórdico. Miguel Székely, do México, ressalta a complexidade de produzir evidências úteis no campo da Educação e aponta falhas vividas em seu país. Por fim, os secretários de Educação de estados brasileiros – Eduardo Deschamps, Frederico Amancio, Haroldo Rocha e Raquel Teixeira – tratam dos desafios de aproximação com a academia, a importância de os gestores fazerem as perguntas e de que maneira as evidências têm sido utilizadas na tomada de decisão.

Por fim, o seminário e a publicação deixam o legado de que é preciso mudar nossa atitude de gestão frente à rotina. Copiando o pensamento científico, que indaga incessantemente se as conclusões estão corretas e quais seriam as condições empíricas a indicar os erros, a gestão tem muito a realizar! Um gestor atento aos resultados aprende permanentemente e corrige rota. Nesse processo, qualquer evidência científica é útil, sejam elas oriundas de avaliações experimentais, não experimentais ou qualitativas. Sejam elas da Educação, da psicologia, da economia, das políticas públicas ou de outros campos da ciência e epistemologias. São fundamentais as evidências que nascem do monitoramento dos projetos e também aquelas fundamentadas na percepção apurada pela experiência do gestor. O sucesso depende do uso de boas evidências, de natureza variada! Sem boas teorias da mudança para a Educação tampouco se conseguirá fazer a diferença, pois as evidências deixam de ecoar e não podem gerar interpretações sobre o que mudar nas intervenções. Como dizia Kant, evidência sem teoria é cega, teoria sem evidência é inócua.

**GESTOR E  
AVALIADOR:  
CONHECIMENTOS  
EM DIÁLOGO  
PARA  
TRANSFORMAR  
A EDUCAÇÃO  
PÚBLICA**





# *PRODUÇÃO E USO DE EVIDÊNCIAS PARA GARANTIR O DIREITO À EDUCAÇÃO*

**RICARDO HENRIQUES** é superintendente executivo do Instituto Unibanco. Economista, foi secretário nacional da Secretaria de Educação Continuada, Alfabetização e Diversidade (Secad) do Ministério da Educação (MEC) e secretário executivo do Ministério de Desenvolvimento Social. No Rio de Janeiro, foi secretário estadual de Assistência Social e Direitos Humanos, além de presidente do Instituto Municipal de Urbanismo Pereira Passos.

A Educação é um direito e, por isso, é para todos. Mas do que estamos falando exatamente quando afirmamos que a Educação é um direito? O acesso à escola é certamente um bom começo. Contudo, uma vez na escola, é preciso que crianças e jovens possam lá permanecer até concluir todas as etapas da Educação Básica. Sem acesso, permanência e conclusão, não haverá aprendizagem. Afinal, o que se pretende como sociedade é que cada criança e jovem possa desenvolver seu potencial a fim de participar da vida cidadã.

Diversos países, incluindo o Brasil, reconheceram durante a Conferência de Incheon, em 2015, que o direito à Educação deve ir além do acesso à escola e precisa chegar às aprendizagens. Foi preconizada uma Educação de qualidade inclusiva e equitativa. O documento do Fundo das Nações Unidas para a Educação, Ciência e Cultura (Unesco) intitulado “Educação de Qualidade para Todos” vai nessa direção. O Fórum Mundial da Educação, ocorrido em 2015, fez um balanço do período 2000-2015 em termos do alcance das metas de universalização e pactuou novas metas para o período de 2016-2030. As novas metas incorporam

## *Para atender às metas, as escolas e todo o sistema educacional precisam organizar seus processos de trabalho, levando em conta quatro dimensões da gestão*

essa concepção mais ampla do direito à Educação que inclui não só o acesso mas também a aprendizagem. No Brasil, as 21 metas do Plano Nacional de Educação (PNE) para o período de 2014-2024 e as metas do Índice de Desenvolvimento da Educação Básica (Ideb), um indicador de qualidade que mensura o alcance do direito à aprendizagem em toda a Educação Básica, são exemplos na mesma direção.

Portanto, no Brasil e no mundo, há metas pactuadas e resultados monitorados a serviço de uma concepção de direito à Educação que vai muito além do acesso à escola.

Para atender a essas metas, as escolas e todo o sistema educacional precisam organizar seus processos de trabalho, levando em conta quatro dimensões da gestão. Uma delas é a dimensão de pessoas, que engloba o corpo de professores, gestores e demais funcionários, que precisam se desenvolver profissionalmente e se manter engajados no trabalho. A participação de todos os atores na resolução dos problemas da escola e seu compromisso com a aprendizagem dos alunos é crucial. Uma segunda dimensão é a físico-financeira, marcada por muitos desafios impostos pela enorme desigualdade no Brasil em oferta física, permitindo que escolas em condições precárias e de alta vulnerabilidade convivam com outras de melhores condições. A terceira dimensão é a relacional, que contempla tanto as relações no interior da escola (entre estudantes, entre eles e os profissionais e toda a comunidade escolar) quanto as relações da escola com as demais instâncias dos setores público ou privado. E, por fim, a quarta dimensão é a pedagógica, a que cuida do processo de ensino-aprendizagem, principal experiência no interior de uma escola.

## GESTÃO NA ESCOLA E SEUS PROCESSOS

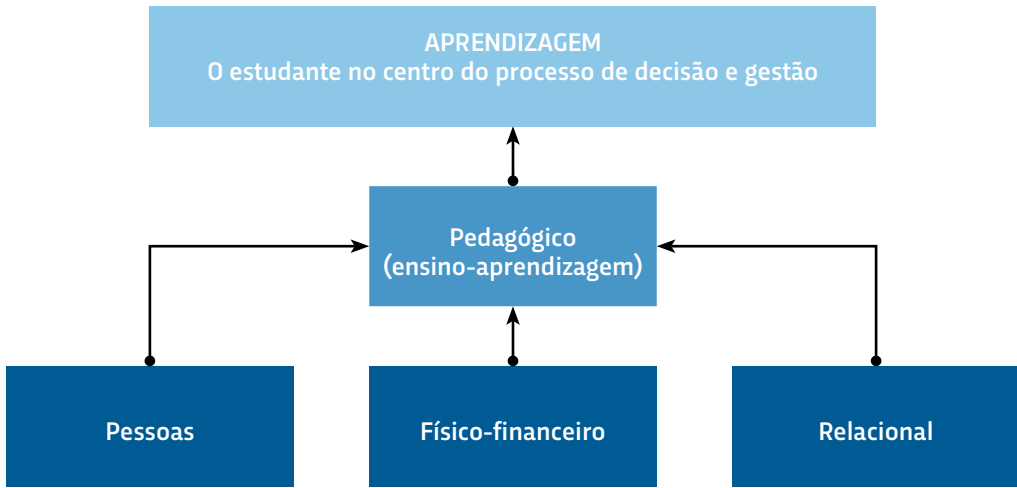


Fonte: elaboração do autor.

A escola que cumpre seu papel e consegue garantir o direito à Educação faz um bom trabalho em todas essas dimensões de gestão, mas não perde de vista a hierarquia existente entre elas. A gestão de pessoas, a físico-financeira e a relacional devem estar a serviço da gestão pedagógica.

*O gestor educacional é um educador, que compreende os processos no interior da escola e os integra em prol da aprendizagem dos alunos*

## HIERARQUIA DE PROCESSOS



Fonte: elaboração do autor.

O gestor educacional é um educador, que compreende os processos no interior da escola e os integra em prol da aprendizagem dos alunos. O corolário do gestor educacional é uma frase dita com muito consenso, mas difícil de ser efetivamente garantida: deve-se colocar o estudante no centro de decisão da escola.

## EM BUSCA DE TRANSFORMAÇÕES

Pautar a gestão educacional por evidências permitirá ao Brasil avançar mais e reduzir as enormes desigualdades mais rapidamente. Precisamos utilizar e gerar conhecimento. O Ideb, as taxas bruta e líquida de matrícula, o abandono escolar e outras métricas são evidências que nos permitem monitorar o direito de aprender. Um sistema completo de monitoramento inclui também resultados intermediários, os quais acredita-se estarem associados à garantia desse direito. São, por exemplo, medidas de qualidade da escola e seu bom funcionamento. Da mesma forma, as ações e políticas voltadas para transformar os resultados também precisam ser monitoradas. Em suma, tudo aquilo que queremos que mude e tudo o que estamos fazendo para mudar deve ser acompanhado e analisado.

Não se faz boa gestão sem monitoramento. É comum pensar no uso de evidências na esfera do monitoramento. Contudo, o campo das ciências também tem muito a dizer aos gestores. Estudos científicos sobre o impacto de programas, as teorias da mudança em Educação, as avaliações sobre a implementação de ações e projetos, bem como os estudos etnográficos são fontes importantíssimas para alimentar o conhecimento dos gestores e apoiar seu processo de tomada de decisão sobre o que priorizar e como fazê-lo.

Mas não só as evidências de monitoramento e as análises científicas indicam caminhos. A própria percepção dos gestores sobre o que pode funcionar e sobre como fazer, calibrada por larga experiência e fundamentada por evidências, é uma fonte igualmente valiosa de conhecimento sobre políticas públicas.

Com todas as possibilidades de geração e acúmulo de conhecimento para subsidiar o desenho de políticas educacionais, há de se ter nitidez sobre o que é o “impacto de uma ação” e qual a sua importância. Para definir impacto, é preciso distinguir entre o progresso de um determinado resultado e qual, efetivamente, foi o efeito gerado por uma ação. Qual a parcela do progresso que pode ser atribuída à ação? Afinal, os resultados podem ter melhorado por várias razões, sem que isso tenha relação direta com a ação avaliada.

Qualquer tentativa de se falar em causalidade – nas ciências em geral, em particular nas ciências humanas e, em especial, na Educação – pode ter armadilhas. É possível cogitar associações e relações. Eventualmente se poderá conhecer causas, o que nos convoca necessariamente a trabalhar com experimentos científicos. Aprender sobre o que causou a melhora de resultados exige caminhar na fronteira e fazer conjecturas precisas para tentar explicar o real e submetê-las a testes consistentes com os mais rigorosos métodos. Envolve também acumular o máximo de evidências a partir da combinação de métodos, todos voltados para testar uma bem construída teoria da mudança.

## CAMINHOS INOVADORES

A busca por estar nessa fronteira tem organizado a agenda do projeto Jovem de Futuro. Muitos dos desafios que se apresentam têm a mesma natureza daqueles enfrentados por inúmeros gestores que desenvolvem e implementam ações no Ensino Médio público brasileiro.

O projeto Jovem de Futuro teve como ponto de partida as evidências de pesquisas já acumuladas em Educação. As teorias da mudança indicam que são múltiplos os determinantes da qualidade da Educação. Sabe-se, pela evidência empírica e pela literatura, que dimensões como a formação dos professores, sobretudo a formação inicial, mas também a continuada e as práticas pedagógicas, têm enorme impacto sobre a probabilidade de sucesso da aprendizagem e, portanto, sobre a garantia desse direito dos estudantes. Essa poderia ter sido uma dimensão priorizada pelo novo programa que nascia. Contudo, a escolha não foi por buscar simplesmente aquilo que poderia gerar o maior impacto. Pretendia-se encontrar caminhos inovadores em meio a uma cultura escolar que não aproveita o poder de uma boa gestão. Todo o esforço realizado teve como objetivo incrementar a gestão, o que significa hierarquizar os processos de trabalho no interior da escola e organizá-los a serviço da aprendizagem. Sabemos que gestão não é uma panaceia, que não é ela, sozinha, que gerará impacto. No entanto, é indispensável melhorá-la para potencializar outras qualidades disponíveis. A gestão é o solo fértil que faz com que outras ações tenham ainda mais efeito, inclusive aquelas ligadas à formação dos professores e outras

*A escolha não foi por buscar simplesmente aquilo que poderia gerar o maior impacto. Pretendia-se encontrar caminhos inovadores em meio a uma cultura escolar que não aproveita o poder de uma boa gestão*

que atuem diretamente no processo de ensino-aprendizagem. Isso é muito relevante, sobretudo em um país continental tão desigual como o Brasil, marcado por uma cultura de rupturas e de fragmentação dos processos decisórios. Temos professoras e professores que investem até 25 anos de suas vidas na Educação, assim como diretores de escola, o que é absolutamente fantástico. Uma boa gestão amplia a capacidade de converter esse investimento em algo mais transformador.

## **APRENDER POR APROXIMAÇÕES SUCESSIVAS NAS REDES PÚBLICAS**

O Jovem de Futuro estimula o diálogo dessa fronteira das evidências com a ciência, buscando garantir o direito à aprendizagem. Os gestores em todas as instâncias (escolas, regionais e secretarias) instituem um ciclo que envolve meta (reflete um direito de aprendizagem), planejamento, execução, monitoramento, troca de boas práticas e correção de rotas. A cada rodada do Circuito de Gestão, sucessivamente, os gestores, atentos ao que praticaram e aos resultados obtidos, compreendem melhor os acontecimentos e tomam novas decisões sobre o que fazer e como fazer melhor. A cada rodada, fortalecidos pelas evidências da ciência, formulam novas perguntas para que as agendas de pesquisa os apoiem.

A atuação do programa em prol dessa construção de aprendizado sucessivo se dá por formação em serviço aliada à instituição de processos de gestão. É oferecida aos gestores das secretarias e das escolas uma quantidade relevante de horas de formação presencial e a distância, focadas em estruturar de maneira consistente e analítica os desafios. Os temas são gestão estratégica, gestão tática e um pouco de gestão de rotina. É fundamental articular o exercício proposto ao conhecimento prático de cada um dos atores envolvidos. Mas a formação é apenas um dos componentes. É necessário associá-la à institucionalização dos processos, o que é feito com o uso de protocolos que dão forma ao Circuito de Gestão. Os protocolos estabelecem uma base



*A cada rodada do Circuito de Gestão, os gestores, atentos ao que praticaram e aos resultados obtidos, compreendem melhor os acontecimentos e tomam novas decisões*

comum, desfrutada e compartilhada por todos os atores de cada escola e rede de ensino e que, por isso, constituem instrumentos redutores da enorme desigualdade de oferta de qualidade entre escolas. Protocolos definem processos que são comuns, e não mínimos. Aprender a partir deles estimula a autonomia, a criatividade e a inovação.

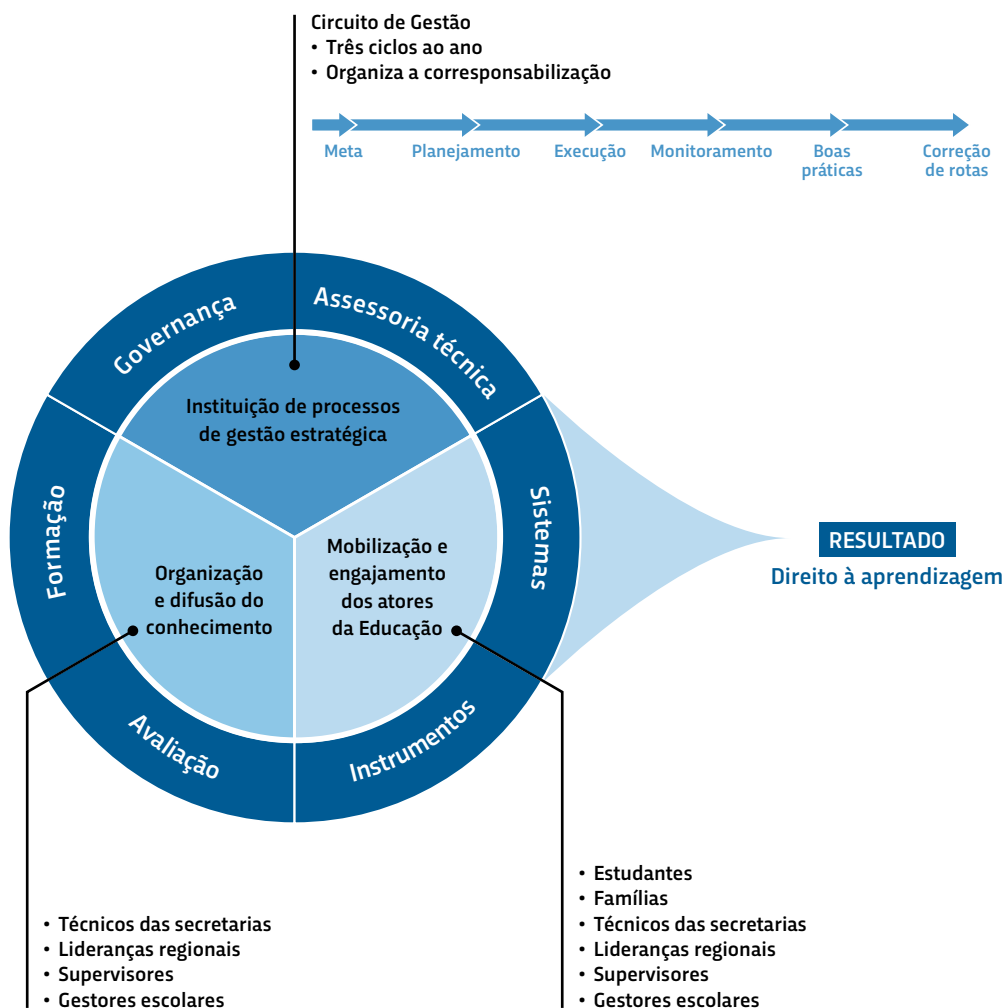
Assim, de braços dados com os gestores, o Circuito de Gestão roda três vezes por ano. É um processo difícil, que gera tensões e angústias, mas que permite, diante da clareza dos indicadores, construir uma visão de determinantes, acompanhar o que deu certo e o que deu errado, aprender com os erros e corrigi-los. Essas são competências individuais desenvolvidas e fortalecidas.

## **CORRESPONSABILIZAÇÃO**

É imprescindível também que sejam desenvolvidas competências coletivas. Conhecer, discutir e contribuir compõem o que chamamos de corresponsabilização. Além disso, a boa comunicação dos atores, geradora de engajamento na mudança. Esses são os outros pilares que sustentam o programa. A agenda precisa envolvê-los no campo dos sonhos e desejos, disseminando o valor da Educação e a ideia de que estudar vale a pena.

Os jovens precisam ser cada vez mais protagonistas de seu aprendizado. Os professores, mais atentos às aulas, que passam a ser dadas para “aquela” turma, com “aquele” desafio de aprendizagem. Da mesma forma, a governança do sistema deve se fortalecer. Precisam ser ágeis os apoios da regional e da secretaria para o sucesso dos planos de mudança das escolas.

## JOVEM DE FUTURO DESENVOLVIMENTO DE COMPETÊNCIAS INDIVIDUAIS E COLETIVAS E CAPACIDADE INSTITUCIONAL



Fonte: projeto Jovem do Futuro.

A fim de fortalecer a governança e melhor estruturar os apoios das regionais e secretarias às escolas, o Jovem de Futuro oferece assessoria que coloca à disposição das redes públicas de ensino métodos, técnicas e conhecimentos para desenhar e acompanhar processos. Também são disponibilizadas ferramentas que permitem acompanhar o que ocorre o tempo todo, com relatórios que agilizam o diagnóstico e a tomada de decisão.

## **APRENDER POR APROXIMAÇÕES SUCESSIVAS TAMBÉM NO INSTITUTO**

Nos últimos dez anos de existência, o Jovem de Futuro também vem aprendendo por aproximações sucessivas, com acertos e erros. Esse processo pressupõe uma postura generosa, baseada no fato de que não se sabe a verdade *a priori*. Isso faz com que o uso da ciência seja rigoroso e regular, o monitoramento seja uma busca incessante e que nossa percepção seja sempre considerada. Aprendemos que o Jovem de Futuro gera impacto, e esse é de magnitude relevante.

O ambiente de aprendizagem permanente para aprimorar políticas públicas pressupõe algumas diretrizes que estão sendo construídas. A primeira é valorizar igualmente o conhecimento prático e o científico. Esse exercício é difícil e corre-se, às vezes, o risco de cair no senso comum. Requer, portanto, questionamento contínuo. Implica atenção para não acreditar em soluções fáceis e, ao mesmo tempo, uma interação com a pesquisa aplicada que busca produzir soluções, o que é, com frequência, de natureza diferente da pesquisa aplicada voltada à publicação em revistas acadêmicas. A segunda diretriz é: como capturar de maneira rigorosa esse conhecimento prático dos gestores? Como se apreende isso? Novamente, é um jogo complexo de ir e vir, em que há questionamentos o tempo todo – é preciso estar aberto para encontrar soluções reais. É comum aparecer uma pergunta vinda da academia, que parece inteligente, mas simplesmente não encontra aderência com o que está acontecendo na escola. Quais são as perguntas da escola?

A terceira diretriz diz respeito à necessidade de se recorrer a diferentes epistemologias e métodos para de fato compreender a realidade. Como se ancora isso, aproximando mundos do conhecimento? É possível olhar para o mesmo fenômeno e buscar explicações em economia, antropologia, psicologia, pedagogia, sociologia e ciência política? Como combinar diferentes métodos? Conhecer o impacto das ações, tal como defini anteriormente – a tentativa de isolar um efeito –, é um passo muito importante, tanto para a prestação de contas à sociedade como para o acú-

mulo de conhecimento, e só pode ser feito por meio de avaliações experimentais. Isso, porém, está longe de ser suficiente para aprimorar as políticas públicas. Outros tipos de avaliação e pesquisa, que respondam por que determinada ação funciona ou não, dados seus objetivos, virão de outras análises quantitativas e qualitativas que adensam o processo de acúmulo de conhecimento.

O que não pode acontecer é cairmos em armadilhas de imaginar que não importa conhecer o impacto de uma ação ou de que existe uma única visão de mundo suficiente para explicar a realidade. Nada disso é verdade. O desafio é construir conhecimentos, elaborar soluções aplicadas e aproximar mundos, sem abraçar uma razão pedagógica descompromissada com resultados ou uma razão tecnocrática que traga apenas uma visão formal sobre os processos.

Vamos construir um caminho que, uma vez trilhado, provavelmente ajudará a fazer com que a maioria das crianças e dos jovens aprenda. Talvez exista espaço para criar uma terceira margem nessa agenda, juntando secretarias de Educação, pesquisadores de diversas academias, fundações privadas e movimentos sociais. Uma via que dê vazão à força transformadora da sociedade e que seja plural. A construção dessa terceira margem poderia ser um objetivo de todos nós.

# ***JOVEM DE FUTURO: O IMPACTO DA GESTÃO SOBRE O APRENDIZADO DOS ESTUDANTES***

**RICARDO PAES DE BARROS** é economista-chefe do Instituto Ayrton Senna e professor no Insper. É também coordenador do Núcleo de Ciência pela Educação no Centro de Políticas Públicas. Atuou no Instituto de Pesquisa Econômica Aplicada (Ipea) por mais de 30 anos. Foi subsecretário de Ações Estratégicas da Secretaria de Assuntos Estratégicos da Presidência da República. Possui pós-doutorado em economia pela Universidade de Chicago e pela Universidade de Yale.

O projeto Jovem de Futuro foi criado pelo Instituto Unibanco em 2007 partindo do princípio de que a gestão escolar voltada para resultados pode proporcionar um impacto significativo no aprendizado dos estudantes do Ensino Médio. O esforço para avaliar esse impacto, no entanto, não é trivial. Em primeiro lugar, é preciso que algumas premissas sejam definidas. Trata-se de impacto sobre o quê? Sobre quem? Quando? (ver Anexo 1, p. 290.)

As premissas de avaliação do Jovem de Futuro são três. O impacto relevante refere-se às proficiências em matemática e língua portuguesa, ou seja, o programa pretende afetar, em última instância, a aprendizagem dos alunos que estão terminando o Ensino Médio, em escolas participantes (ver Anexo 2, p. 312). A proficiência é mensurada a partir de avaliações externas conduzidas pelos próprios estados parceiros. Esse efeito deve ser mensurado após três anos. A explicação para a janela de avaliação é que, ao final desse período, os alunos que estavam na primeira série do Ensino Médio no primeiro ano de implementação do programa, se aprovados, chegariam à terceira série. Assim, poderia-se capturar o impacto máximo, uma vez que esses alunos

teriam frequentado todo o Ensino Médio em uma escola que mudou sua gestão.

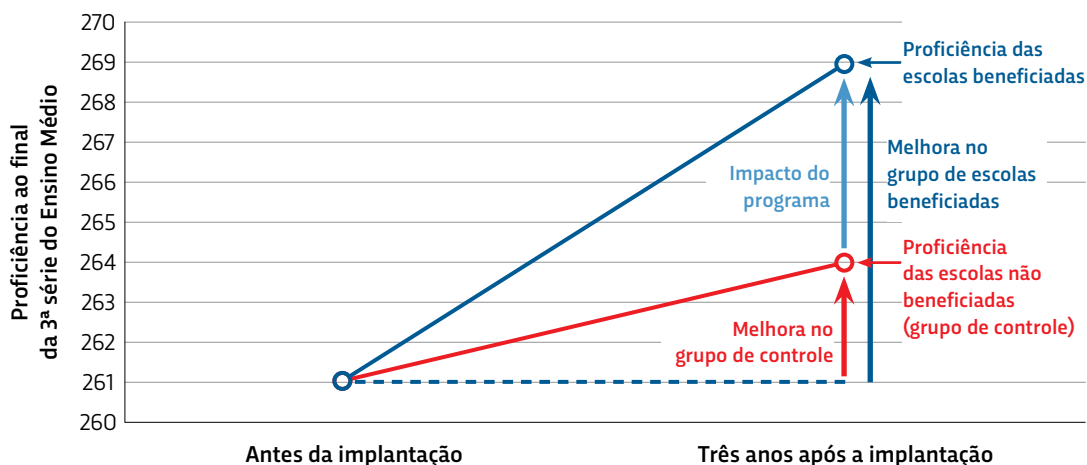
A linha de base da avaliação é dada pelos níveis de proficiência em matemática e língua portuguesa dos alunos da 3ª série das escolas no ano anterior à chegada do programa. Três anos depois, mede-se novamente o aprendizado nessas disciplinas dos alunos da 3ª série. O objetivo é verificar o quanto o aprendizado após três anos foi maior do que no momento anterior ao início da implementação. Isso é fácil de fazer, e os resultados mostram que o desempenho dos alunos tem crescido (ver Anexo 2, p. 312).

Mas o impacto se resume à melhoria no aprendizado dos alunos na 3ª série? É preciso lembrar que houve uma ação conjunta do Instituto Unibanco e das secretarias estaduais de Educação sobre o aprendizado, porém o que se deseja é isolar e identificar a contribuição específica dessa intervenção. Esse é o real significado de impacto. Não basta medir a evolução das notas: é preciso saber o que teria acontecido com essas escolas se a secretaria e o Instituto Unibanco não tivessem implantado o Jovem de Futuro (situação contrafactual). Esse impacto não é uma simples comparação entre antes e depois. Tampouco interessa saber tudo o que mudou, mas sim aquilo que foi modificado por conta da intervenção. Como é possível calcular isso?

É impossível mensurar com exatidão, a menos que houvesse uma máquina do tempo e fosse possível aplicar o programa e verificar os resultados; então, voltar no tempo, não oferecer o programa e verificar novamente os resultados. Se o projeto foi aplicado em uma determinada escola, não é possível saber o que teria ocorrido com ela caso a intervenção não tivesse existido.

No início do século 20, Ronald Fischer (1890-1962), um célebre estatístico inglês, encontrou uma maneira de resolver esse problema. Ele descobriu o seguinte: caso os participantes de um programa sejam escolhidos aleatoriamente, por meio de sorteio, aqueles que não foram sorteados apontarão exatamente o que aconteceria com os participantes se eles não tivessem acessado o programa. A diferença entre a melhora nos chamados grupo de tratamento (beneficiado pelo programa) e grupo de controle (não beneficiado) corresponde ao impacto do programa.

## EVOLUÇÃO DA PROFICIÊNCIA NO GRUPO DE ESCOLAS BENEFICIADAS E NO GRUPO DE CONTROLE



Fonte: elaboração do autor.

## O PROCESSO DE ALEATORIZAÇÃO – SORTEIOS

São diversas ações previstas no programa e, em geral, os estados viam-se com dificuldades de atender a todas as escolas logo no primeiro ano de implementação. Uma curva de aprendizagem em implementação era gerada, e o planejamento da entrada das escolas no programa em ciclos fez-se necessário. A entrada em ciclos acabou sendo fundamental para viabilizar a avaliação de impacto. Para os resultados serem confiáveis, esses sorteios precisam ser feitos de forma adequada.

As escolas que aderiram à iniciativa foram agrupadas por semelhança, às vezes em duplas, outras em trios ou grupos um pouco maiores. Dentro desses grupos, foram sorteadas as escolas para participar do programa logo no primeiro ano. Também por sorteio, foram definidas escolas que ingressariam um ano depois, dois ou três anos depois. As que ingressam por último, ou seja, três anos após o início da implementação na primeira escola, permanecem sem o programa durante todo o período compreendido pela janela de avaliação. Elas serão as escolas dos grupos de controle.



No Jovem de Futuro, ao longo de oito anos foram realizados 400 experimentos em dez estados brasileiros: Ceará, Espírito Santo, Goiás, Mato Grosso do Sul, Minas Gerais, Pará, Piauí, Rio de Janeiro, Rio Grande do Sul e São Paulo.

Dos 400 grupos dos quais foram sorteadas escolas de controle e de tratamento, apenas 141 alcançaram três anos de maturidade até agora: 15 no Pará, 25 no Ceará, 12 no Rio de Janeiro e 20 no Vale do Paraíba (interior de São Paulo), além de outros 17 na região metropolitana de São Paulo, quatro em Minas Gerais, 25 em Goiás e 23 no Mato Grosso do Sul. São 141 experimentos, com 141 grupos de tratamento e 141 grupos de controle.

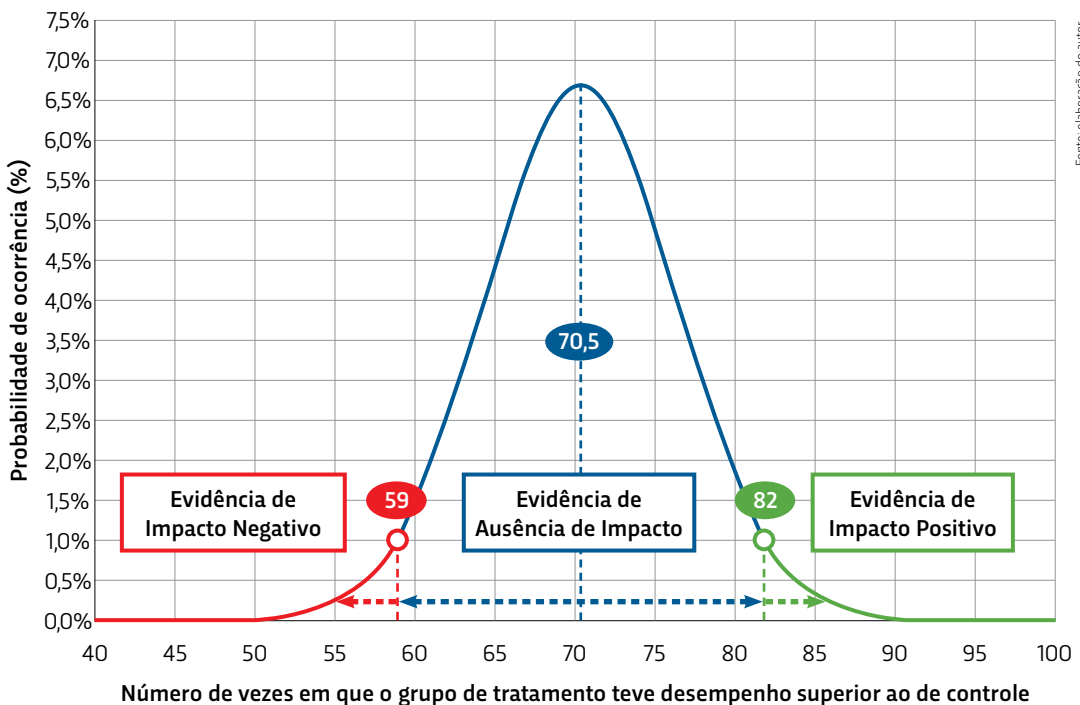
## **COMPARAÇÃO DOS RESULTADOS DOS GRUPOS DE TRATAMENTO COM OS DE CONTROLE**

A primeira tarefa é verificar se o programa foi bem-sucedido. Todo estatístico tem uma forma curiosa de pensar: considera sempre a hipótese nula, isto é, que o programa não teve impacto. Seu objetivo é se perguntar se há evidências que afastariam essa hipótese. Se o Jovem de Futuro não servisse para nada, o aprendizado melhoraria mais entre as escolas do grupo de tratamento ou de controle? O resultado, nesse caso, seria aleatório. Isso significa que, se o programa não gerou nenhum impacto, verificaríamos que, dos 141 grupos considerados, em metade as escolas de tratamento melhorariam mais e, na outra metade, as escolas de controle melhorariam mais. É como jogar uma moeda para cima e constatar se deu cara ou coroa.

Buscar evidências que afastem a hipótese nula equivale a descobrir em quantos casos, dos 141 experimentos, a melhora no grupo de tratamento foi maior do que a melhora no grupo de controle. Se o impacto foi positivo, deveria haver maior frequência de escolas de tratamento melhorando mais do que as de controle. Se o impacto foi negativo, deveria haver maior frequência de escolas de controle melhorando mais do que as de tratamento.

Quando atiramos 141 moedas para o alto, não é sempre que o resultado se equilibra com 71 caras e 70 coroas (ou vice-versa). Qualquer pessoa que fizer essa experiência chegará à conclusão de que só em 13% das tentativas o resultado é exatamente 71 caras e 70 coroas. Na maioria das vezes, se observará um número próximo a 70, podendo ser um pouco mais ou um pouco menos. Um estatístico pode calcular isso: em 95% das vezes, esse resultado estará entre 59 e 82. Portanto, pode-se dizer que, quando o número de casos em que o tratamento melhora mais do que o controle estiver entre 59 e 82, isso é perfeitamente compatível com um programa sem impacto nenhum. Dito de outra forma: se o experimento foi realizado 141 vezes e em 80 casos o tratamento melhorou mais do que o controle, tal incidência está dentro da margem de impacto nulo. Mas se esse número for superior a 82 vezes, o impacto do programa é positivo. Se for inferior a 59, o impacto do programa é negativo (ver Anexo 2, p. 312).

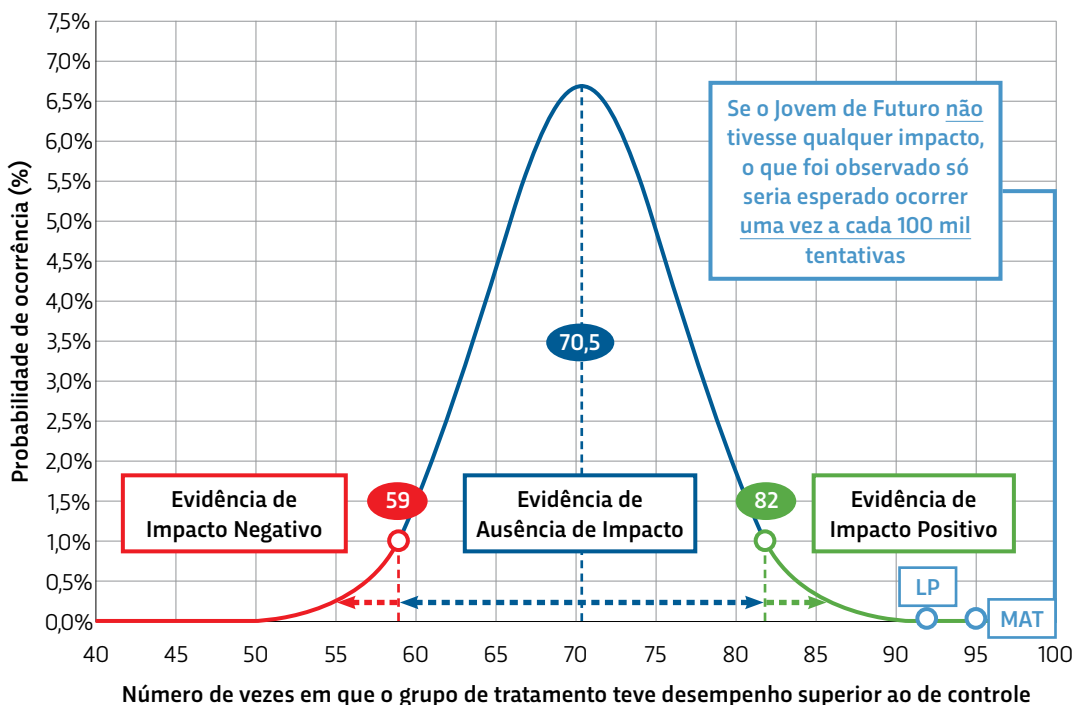
**DISTRIBUIÇÃO DAS VEZES, EM 141 EXPERIMENTOS, EM QUE O GRUPO DE TRATAMENTO TEM DESEMPENHO SUPERIOR AO DO GRUPO DE CONTROLE, CASO O PROGRAMA NÃO TENHA IMPACTO**



Fonte: elaboração do autor.

Na avaliação do Jovem de Futuro, constatou-se que, dos 141 experimentos, houve melhora no grupo de tratamento maior do que no grupo de controle em 95 casos, considerando a proficiência em matemática, e em 92 casos, considerando a proficiência em língua portuguesa.

**DISTRIBUIÇÃO DAS VEZES, EM 141 EXPERIMENTOS, EM QUE O GRUPO DE TRATAMENTO TEM DESEMPENHO SUPERIOR AO DO GRUPO DE CONTROLE, CASO O PROGRAMA NÃO TENHA IMPACTO**



Fonte: elaboração do autor.

Se atirmos uma moeda 141 vezes para o alto, é possível dar cara 95 vezes? Pode acontecer. É provável que aconteça? Não é. O quanto isso é improvável? Também é possível calcular, e a resposta é: uma vez a cada 100 mil. Logo, é possível acreditar que esse programa realmente teve impacto positivo porque, com esse desempenho, é muito pouco provável que o resultado tenha vindo por pura sorte. Avaliações experimentais não conseguem afirmar com certeza absoluta que houve impacto. Contudo, permitem afirmar que, segundo as evidências, o projeto muito pro-

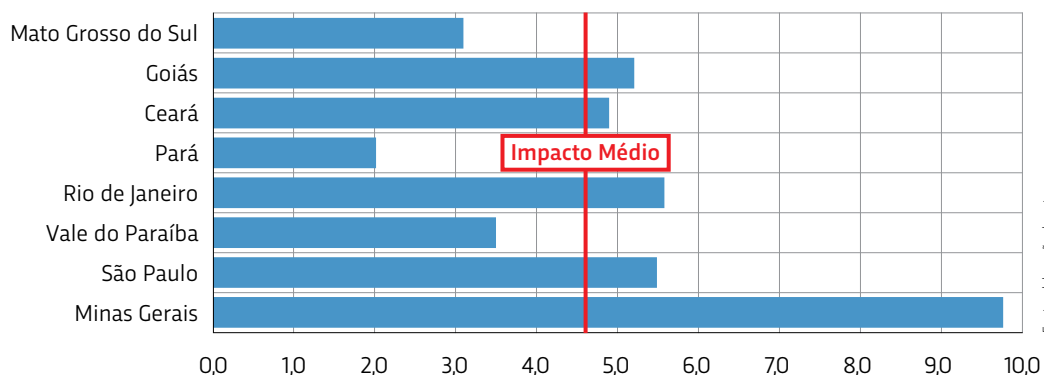
vavelmente gerou impacto – no caso do Jovem de Futuro, há apenas uma chance em 100 mil de que isso não seja verdade.

## MAGNITUDE DO IMPACTO

E qual seria a magnitude desse impacto? É próximo de 5 pontos na escala de proficiência do Sistema de Avaliação da Educação Básica (Saeb) para o Ensino Médio. Trata-se de um número relativamente alto. Para se ter uma ideia, os alunos brasileiros avançam 16 pontos na escala Saeb ao longo dos três anos dessa etapa. O ideal seria que a proficiência dos alunos crescesse 40 pontos nessa escala, mas o que se aprende de fato equivale a uma elevação de apenas 16 pontos. Fazendo uma conta simplificada, conclui-se que, a cada ano do Ensino Médio, os alunos aprendem o equivalente a 5 pontos na escala Saeb.

Se o impacto do projeto Jovem de Futuro é de 5 pontos, pode-se afirmar que, nas escolas participantes, é como se os alunos tivessem sido expostos a um ano a mais de aprendizado. Em três anos, eles aprendem o que demorariam, em circunstâncias naturais, quatro anos para aprender.

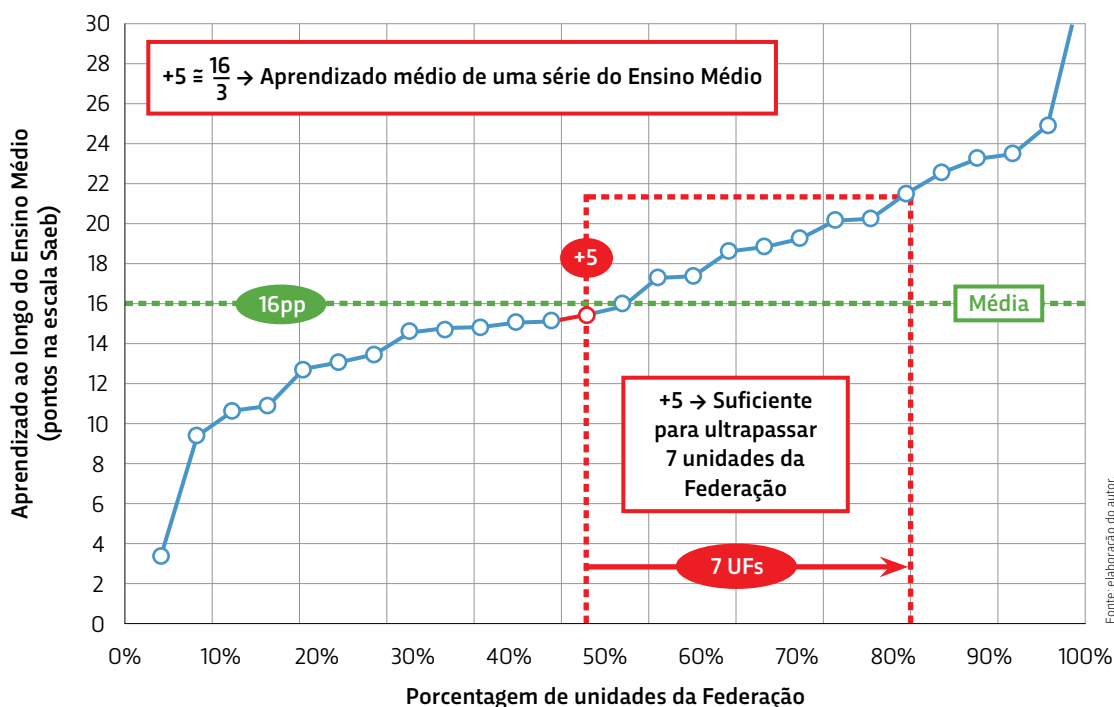
### MAGNITUDE DO IMPACTO DO JOVEM DE FUTURO SOBRE O APRENDIZADO DE LÍNGUA PORTUGUESA POR ÁREA DE ATUAÇÃO



Fonte: elaboração do autor.

O quanto os alunos normalmente aprendem durante o Ensino Médio é desigual entre os estados brasileiros. Há unidades da Federação em que os estudantes avançam 26 pontos na escala Saeb e outras em que o crescimento é de apenas 4 pontos. Se um estado com aprendizagem mediana dos alunos durante esse período (a mediana é aquele valor que separa a metade maior e a metade menor de uma amostra) conseguisse aumentar o desempenho em 5 pontos, ultrapassaria nada menos do que sete outros estados em termos de resultados no Ensino Médio.

### APRENDIZADO EM MATEMÁTICA AO LONGO DO ENSINO MÉDIO POR UNIDADE DA FEDERAÇÃO



### ABRINDO A CAIXA-PRETA DA AVALIAÇÃO

Dado que o Jovem de Futuro reúne um conjunto variado de ações, qual delas será a mais importante? Qual é o impacto de

cada componente do programa? Do ponto de vista teórico, avaliar o impacto de um componente é perfeitamente possível. Bastaria extrair aquele componente do cálculo, verificar o impacto do projeto sem ele e comparar com os resultados do projeto quando se inclui o componente.

Evidentemente, não faz sentido tentar descobrir essa resposta quando os componentes dos quais se deseja isolar o impacto forem igualmente importantes. Se a hipótese é de que retirando um dos componentes o programa não consegue funcionar, não é razoável entregar o programa sem ele.

O Jovem de Futuro vem buscando ampliar o conhecimento proveniente de avaliações, abrindo a caixa-preta da avaliação experimental. Com isso, pretende descobrir como o impacto é transmitido em todas as instâncias do sistema e dentro da escola até chegar à aprendizagem dos alunos.

É fácil calcular o impacto do projeto sobre o clima escolar. Também é fácil calcular o impacto sobre o funcionamento da escola. O difícil é saber quanto do impacto sobre a aprendizagem advém do impacto que o próprio programa gerou sobre o clima escolar. Calcular o impacto do clima escolar sobre a aprendizagem dos alunos é uma tarefa muito difícil, porque se trabalha com uma variável que não é controlável de maneira direta.

Decifrar esse tipo de detalhe depende da criação de um modelo de mudança extremamente completo. É preciso muita teoria, talvez mais do que se tem hoje. Enquanto o programa afetar o aprendizado de maneira direta e também indireta, por exemplo por interferência do clima escolar, não se conseguirá isolar o impacto do clima escolar sobre a aprendizagem dos alunos.

Há desafios cuja solução é bastante simples e outros que envolvem a combinação de um aprofundamento teórico sobre o modelo de mudança do projeto e, evidentemente, mais evidências empíricas.

**PRODUÇÃO DE  
CONHECIMENTO  
PARA  
APRIMORAR  
PROJETOS  
EDUCACIONAIS:  
O USO DA  
AVALIAÇÃO  
DE IMPACTO  
COM DESENHO  
EXPERIMENTAL**





# *DESAFIOS PARA A PRODUÇÃO DE CONHECIMENTO CIENTÍFICO EM EDUCAÇÃO A SERVIÇO DAS POLÍTICAS PÚBLICAS*

**RICARDO MADEIRA** é professor de economia da Universidade de São Paulo (USP) e pesquisador associado da Fundação Instituto de Pesquisas Econômicas (Fipe). Ph.D. em economia pela Universidade de Boston (EUA) e mestre em economia pela Fundação Getúlio Vargas, suas principais áreas de pesquisa são microeconomia aplicada, economia da Educação e avaliação de políticas educacionais.

Este artigo traz algumas reflexões sobre a dificuldade de colocar a produção de conhecimento científico a serviço das políticas públicas. Minha linha de argumentação é que a academia tem muito a contribuir para a definição de políticas públicas, porém existem desafios conceituais e práticos que precisam ser enfrentados. Nas considerações finais, reflito sobre um modelo de atuação conjunta da comunidade científica com o setor público.

## **MOTIVAÇÃO**

Sabe-se que se aplica uma fatia significativa do orçamento público em políticas educacionais – em torno de 7% do Produto Interno Bruto (PIB) é gasto com Educação e políticas que procuram melhorá-la. A Educação é algo da máxima importância, e ninguém discorda de que se deve investir muito nela. No entanto, um custo de oportunidade social gigantesco está associado a esse recurso, ou seja, ele poderia ser investido em políticas alternativas ou em outros empreendimentos sociais, inclusive

em outras áreas de alta necessidade, a exemplo da saúde pública. Obviamente, não se pode entender o gasto como uma medida de sucesso. É preciso levar em conta a relação custo-benefício.

Há dois elementos centrais para avaliar o desempenho de políticas educacionais. O primeiro é sua eficiência, ou seja, quais iniciativas, de um menu de possibilidades, são mais efetivas do ponto de vista de real gasto para benefício gerado; o segundo, a possibilidade de incrementar o “cardápio” de políticas com inovação. A ciência pode ajudar nessas tarefas, provendo subsídios para o desenho de políticas novas e, também, para o aprimoramento de políticas existentes. Como o conhecimento científico pode colaborar nessas duas dimensões?

## **CONTRIBUIÇÕES DA CIÊNCIA PARA AVALIAÇÃO E DESENHO DAS POLÍTICAS EDUCACIONAIS**

Vale destacar ao menos quatro pontos críticos de contribuição da ciência. Primeiro, é preciso propor teorias de impacto empiricamente refutáveis e falseáveis. A proposta de um programa ou intervenção deve ser subsidiada por uma teoria, um conjunto de hipóteses que justificam a crença de impacto. Por que a iniciativa funcionará? Que alavancas se moverão para gerar o efeito esperado? É muito importante que haja uma teoria falseável para isso, ou seja, um método pelo qual se possa testar empiricamente se as hipóteses que justificam a adoção daquela política são verdadeiras ou não.

Segundo, a ciência também pode contribuir com a mensuração. Quando se propõe uma política, existe um conjunto de dimensões em que se imagina produzir um impacto, por exemplo: “A política vai melhorar o desempenho em língua portuguesa ou em matemática? Vai influenciar determinadas habilidades socioemocionais ou atributos atitudinais? Em qual direção?”.

O terceiro ponto refere-se à importância de realizar avaliações de impacto de políticas educacionais que tenham validade causal, isto é, que permitam atribuir às políticas a causa de determinados fenômenos, como: “As secretarias que receberam o

projeto Jovem de Futuro melhoraram seu desempenho na escala do Sistema de Avaliação da Educação Básica (Saeb) em 5 pontos por causa do programa”. Trata-se de fazer a análise contrafactual sobre a qual Ricardo Paes de Barros discorre em seu artigo neste livro (p. 36), por exemplo: “O que teria acontecido com as unidades que receberam a intervenção se esta não tivesse ocorrido?”. Como ele explica, existem recursos estatísticos para responder a isso. O padrão-ouro é a utilização do método de avaliação experimental (randomized control trial, em inglês, ou RCT), que separa a unidade de intervenção e a de controle por sorteio e acompanha as duas para aferir o impacto da intervenção.

O quarto aspecto importante da contribuição da ciência é produzir evidências empíricas sobre as hipóteses que sustentam a escolha de determinadas políticas educacionais, o que é particularmente relevante para o desenho de novas políticas. Uma série de hipóteses subsidiam a adoção de uma política – sobre como alunos e professores se comportam, como reagem a certa intervenção etc. –, e é muito importante validá-las. Acontece que esse tipo de investigação, em geral, não é de interesse direto do avaliador; trata-se, tipicamente, de uma agenda de pesquisa conduzida dentro da academia. No entanto, o conhecimento produzido por essas investigações é extremamente importante para o desenho de novas políticas. Quando se documentam fenômenos educacionais e se entende por que eles ocorrem, ou seja, as razões por trás dos fenômenos, consegue-se pensar em políticas que possam atuar sobre eles.

*Quando se documentam fenômenos educacionais e se entende por que eles ocorrem, ou seja, as razões por trás dos fenômenos, consegue-se pensar em políticas que possam atuar sobre eles*

## **DESAFIOS CONCEITUAIS**

Do ponto de vista conceitual, quais são os principais desafios para que a ciência possa contribuir nas quatro dimensões mencionadas?

1. Extrapolar os resultados das avaliações de impacto para outros contextos.
2. Diferenciar a avaliação de impacto de intervenções/políticas de testes empíricos das hipóteses que as justificam.
3. Identificar os canais/mecanismos pelos quais as políticas atuam (hipóteses que as justificam).
4. Reconhecer a natureza interdisciplinar dos fenômenos educacionais.
5. Delimitar a fronteira entre aspectos positivos e normativos.

## **DIFICULDADE DE EXTRAPOLAR APRENDIZADOS OBTIDOS POR MEIO DE MÉTODOS EXPERIMENTAIS**

O problema da extrapolação de resultados é bastante agudo, inclusive para avaliações experimentais. Por exemplo: determinado estado decide adotar uma política de incentivo a professores, tal iniciativa é avaliada pelo método experimental (separando as escolas de controle e as de tratamento por sorteio) e, ao final, os pesquisadores mensuram o impacto da política no desempenho dos alunos em matemática ou língua portuguesa e não encontram efeito algum. A conclusão é: ou essa política não teve impacto por azar, ou, com muita probabilidade, ela realmente não funciona. Ocorre que uma intervenção pode não funcionar devido a elementos relevantes do contexto. No caso apresentado, poderiam ser condições externas à intervenção, como as opções profissionais para professores naquele mercado de trabalho ou o próprio arcabouço institucional no qual as escolas estão envolvidas. Será, portanto, que a conclusão pode ser extrapolada para outros contextos? Para que isso seja possível, é preciso saber quais elementos do contexto são relevantes para o impacto da intervenção.

Essa questão, conhecida como problema do enquadramento, é bastante comum nas ciências sociais. Para identificar quais elementos do contexto são relevantes, deve-se lançar mão de uma avaliação de impacto. No entanto, as conclusões de uma avaliação de impacto só podem ser extrapoladas se for possível identificar o que é relevante. Isso gera uma circularidade difícil de equacionar.

No caso da política de incentivo aqui mencionada observa-se, ao compilar diversas experiências internacionais, que nos últimos 15 anos foram realizadas várias, com desenhos diferentes. De acordo com avaliações rigorosas, algumas foram bem-sucedidas e tiveram efeitos positivos; outras não. Como compará-las? É preciso identificar o que é comparável e o que não; o que está no mesmo contexto e o que não; o que é relevante para o contexto e o que não. Assim, para garantir a comparabilidade dos estudos, é importante que tenham sido implementados em contextos parecidos.

Portanto, devemos ser muito cuidadosos ao interpretar os resultados da avaliação de impacto de uma intervenção. Rejeitar que determinada política funciona é, na verdade, rejeitar que tal política funciona no contexto específico em que foi implementada; a mesma política poderia funcionar em um contexto distinto.

## **DIFERENCIAÇÃO ENTRE A AVALIAÇÃO DE IMPACTO DE INTERVENÇÕES/POLÍTICAS DE TESTES EMPÍRICOS E AS HIPÓTESES QUE AS JUSTIFICAM**

Como mencionado, existem hipóteses que justificam a expectativa de impacto acerca de uma intervenção. Quando se rejeita uma política implementada, é *aquela política* que está sendo rejeitada, e não as hipóteses associadas a ela. Talvez a política não tenha funcionado porque as hipóteses que a justificavam não eram válidas, mas outras razões podem explicar o resultado.

Tomando novamente o exemplo da políticas de incentivo, alguém poderia reagir ao não resultado afirmando: “Professores não são motivados por incentivo extrínseco. Basicamente,

o exercício da docência tem motivação intrínseca, de maneira que qualquer política de incentivo extrínseco não será eficaz”. Assim, não é porque se rejeitou determinado programa de incentivo implementado com um desenho particular que os professores não respondem a incentivos extrínsecos. É importante fugir dessa armadilha – de que a hipótese X não é verdadeira, pois uma política baseada em X não se mostrou eficaz.

### *IDENTIFICAÇÃO DOS CANAIS/MECANISMOS PELOS QUAIS AS POLÍTICAS ATUAM (HIPÓTESES QUE AS JUSTIFICAM)*

Já se disse antes que se devem saber quais são as hipóteses associadas a uma política, pois isso é fundamental para a generalização, ou seja, para a extrapolação dos resultados. O ideal seria identificá-las por meio de experimentos controlados, de análise experimental, mas nem sempre é fácil fazê-lo, simplesmente porque as hipóteses e o contexto podem se relacionar com a política de maneira muito complexa, sem que haja controle sobre todas as dimensões em um formato experimental (ou que emule tal método). Assim, muitas vezes é necessário recorrer a outras fontes de conhecimento, como o conhecimento prático, que pode nos guiar em apontar, com base na experiência prática, quais são os mecanismos de efeito mais prováveis.

### *INTERDISCIPLINARIDADE DA NATUREZA DOS FENÔMENOS EDUCACIONAIS*

É inquestionável que a tarefa de estudar os fenômenos educacionais é interdisciplinar. Os fenômenos educacionais estão relacionados a objetos de estudos de diversas áreas do conhecimento, entre elas a pedagogia, a psicologia, a sociologia, a medicina e a economia. Essas áreas dispõem de métodos de investigação distintos. Algumas se debruçam mais sobre evidência qualitativa; outras, sobre quantitativa. Articular tais saberes não é fácil.

*[...] é difícil chegar a um consenso entre pesquisadores das diferentes áreas sobre o que é uma evidência cientificamente válida*

Muitas vezes, é difícil chegar a um consenso entre pesquisadores das diferentes áreas sobre o que é uma evidência cientificamente válida.

### **DELIMITAÇÃO DA FRONTEIRA ENTRE ASPECTOS POSITIVOS E NORMATIVOS**

Não raro, ao debaterem sobre resultados de avaliações de impacto de políticas educacionais, pesquisadores das diferentes áreas do conhecimento parecem discordar sobre os aspectos de caráter positivo e os de caráter normativo. Por vezes, uma política se mostra eficaz em determinada dimensão, mas, quando seus resultados são apresentados à comunidade interessada, muitos alegam que ela não é adequada ou desejável. Para que tais dilemas sejam superados, as razões de descontentamento precisam ser explicitadas. Por que a política não é desejada? É porque existe algum objetivo que ela compromete e que deveria ter sido enunciado na partida? Por que membros da comunidade rejeitam essa política?

Ao refletir sobre os resultados que se deseja alcançar e as políticas que entregam esses resultados, quais são as dimensões de impacto relevantes? Busca-se proficiência em língua portuguesa, em matemática, mas também se quer levar em conta atributos atitudinais, para habilidades socioemocionais? É necessário haver concordância acerca de quais são as dimensões de impacto relevantes para então pensar em políticas eficazes para atingir essas dimensões definidas *a priori* como importantes. Tal fronteira deve ser bem delimitada, o que não é fácil em um campo tão interdisciplinar como o da Educação.



## DESAFIOS PRÁTICOS

Existem ainda os desafios de ordem prática, como discutido em outros artigos desta publicação. Aqui se destacam três:

1. Resistências aos métodos experimentais.
2. Desinteresse pela produção de evidências para além da avaliação de impacto de intervenções (identificação de canais).
3. Garantias de implementação dos protocolos de avaliação.

### RESISTÊNCIAS AOS MÉTODOS EXPERIMENTAIS

A alocação de unidades de tratamento e unidades de controle por sorteio nem sempre é bem-aceita, especialmente no setor público, com frequência sob a alegação de que tal método fere o princípio da isonomia. Outro ponto de resistência é o custo da implementação de avaliações experimentais. Existem cada vez mais dados administrativos à disposição do pesquisador, de modo que, muitas vezes, não há necessidade de uma coleta primária de dados extensa (que costuma ser o componente mais caro dos estudos experimentais)

*É necessário haver concordância acerca de quais são as dimensões de impacto relevantes para então pensar em políticas eficazes para atingir essas dimensões definidas a priori como importantes. Tal fronteira deve ser bem delimitada, o que não é fácil em um campo tão interdisciplinar como o da Educação*

Mesmo assim, o custo dessas investigações é alto, e o tempo de produção de resultados, muito longo. Por exemplo: o prazo de maturação dos resultados da avaliação de impacto do Jovem de Futuro foi de três anos, ou seja, até que uma geração de alunos da 1ª série do Ensino Médio chegasse à 3ª. O esforço de mobilização de equipes é enorme, e esse tipo de comprometimento gera resistências.

Há também uma resistência psicológica natural à avaliação – o temor de obter resultados negativos. Contribui para isso o fato de que o reconhecimento público está na execução, e não na avaliação.

## **DESINTERESSE PELA PRODUÇÃO DE EVIDÊNCIAS PARA ALÉM DA AVALIAÇÃO DE IMPACTO DE INTERVENÇÕES (IDENTIFICAÇÃO DE CANAIS)**

Validar as hipóteses que justificam as políticas é muito importante, sobretudo para o desenho de novas intervenções. O problema é que, tipicamente, o responsável pela implementação está mais preocupado em conhecer o resultado imediato, prático, do que em saber quais são as hipóteses que justificam a política, ou seja, em validar tais hipóteses.

Por exemplo: determinada política de incentivo não funcionou porque uma dinâmica de comportamento dos professores não foi mapeada. Identificar se essa dinâmica está presente ou não é algo de menor interesse para quem propõe a intervenção. O raciocínio é: “Sempre haverá algum pesquisador que vai responder a essa questão; não é preciso investir tempo nisso”. Ou seja, existe a expectativa de “pegar carona” em outra investigação que será feita em algum momento.

## **GARANTIAS DE IMPLEMENTAÇÃO DOS PROTOCOLOS DE AVALIAÇÃO**

Os protocolos de avaliação de impacto exigem um esforço operacional muito grande, ou seja, mobilização de equipes para exe-

*É fundamental fomentar e desenvolver um ecossistema de investigação científica interdisciplinar e contínuo, com a participação ativa da comunidade interessada: o gestor público, o terceiro setor e a comunidade científica*

cutá-los. Sempre há também o risco de contaminação do grupo de controle, seja por ação deliberada (o grupo de controle tem a expectativa de impacto da intervenção, então, de algum modo, seus integrantes tentam se beneficiar da intervenção burlando os protocolos), seja por uma razão não intencional (alguma rotina administrativa de alocação de professores entre escolas gera um transbordamento do grupo de tratamento para o de controle, de maneira que este é indiretamente afetado pela política).

## **CONSIDERAÇÕES FINAIS**

Para que a ciência possa ser mais efetiva em sua tarefa de colaborar com o desenho das políticas educacionais, precisamos achar meios de enfrentar os desafios aqui apresentados. Para isso, é fundamental fomentar e desenvolver um ecossistema de investigação científica interdisciplinar e contínuo, com a participação ativa da comunidade interessada: o gestor público, o terceiro setor e a comunidade científica.

O gestor público é quem apresenta as perguntas relevantes para o desenho das políticas, devendo contribuir com seu conhecimento prático e intuitivo, muito relevante para a reflexão a respeito delas. O terceiro setor tem a flexibilidade e a agilidade necessárias para inovar e executar a operação, além de uma

agenda de longo prazo ligada ao interesse pela disseminação dos resultados. Vale ressaltar que um dos desafios relacionados ao tempo da avaliação é a questão do ciclo eleitoral, que faz parte da vida política. Normalmente, o resultado de uma intervenção fica disponível depois de encerrado o ciclo eleitoral. Um secretário começa uma gestão e tem quatro anos, se tanto, para implementar uma política. No entanto, esses quatro anos podem não ser suficientes para realizar uma avaliação rigorosa do que foi feito. A existência de um agente externo como o terceiro setor, com outro horizonte de tempo, pode ajudar na continuidade e na manutenção do interesse pela pesquisa.

Também é preciso envolver a comunidade científica relacionada, articular as diferentes áreas do conhecimento e gerar concordância sobre a evidência considerada cientificamente válida, necessária para pautar de maneira efetiva o desenho de políticas públicas. Outras contribuições importantes da comunidade científica nesse processo são a garantia do comprometimento com métodos científicos reconhecidos e a postura isenta e independente dos pesquisadores responsáveis pela avaliação.

Por fim, destaca-se um aspecto muito importante: a documentação rigorosa das avaliações de impacto e dos resultados produzidos. Só é possível disseminar boas práticas por meio de documentação rigorosa, explicitando quais foram as perguntas a que a avaliação de impacto buscou responder e as hipóteses que ela se propôs a testar. Desse modo, saberemos o que foi respondido pela avaliação e quais são as limitações dos resultados produzidos. Somente assim poderemos dizer o que funciona e para quem funciona e então pensar na extrapolação, ou seja, em levar os resultados obtidos a outros contextos.

# USO DE AVALIAÇÃO EXPERIMENTAL PARA CÁLCULO DE EFEITOS DE PROGRAMAS EDUCACIONAIS SOBRE A DESIGUALDADE DE DESEMPENHO

**SERGIO FIRPO** é professor titular da Cátedra Instituto Unibanco no Insper, em São Paulo. Possui mestrado em economia pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ) e mestrado em estatística e Ph.D. em economia pela Universidade da Califórnia em Berkeley. É membro da Econometric Society, *fellow* do Institute for the Study of Labor e membro fundador do Instituto Nacional de Ciência e Tecnologia de Educação, Desenvolvimento Econômico e Inserção Social.

O Insper é um instituto de pesquisa e ensino voltado para a produção de conhecimento prático e do qual a sociedade se apropria em busca de soluções para seus problemas. A qualidade da Educação é um deles, e talvez dos mais importantes. O Instituto Unibanco, ao promover e financiar programas que tentam melhorar a qualidade da Educação, sobretudo por meio da gestão escolar, mostra, em sua parceria com o Insper – da qual resulta essa cátedra –, que se preocupa com o rigor acadêmico, em especial com a maneira como ele pode ajudar na formulação de políticas públicas.

Em sintonia com tal proposta, este artigo visa apresentar como é feita a medição do impacto de políticas públicas no desempenho médio dos alunos, focando, principalmente, a eventual desigualdade que programas educacionais venham a gerar entre alunos e entre escolas. Fazendo uso de experimentos aleatórios, é possível aprender sobre desigualdade de rendimento, desempenho ou proficiência.

## **EXPERIMENTOS ALEATÓRIOS: POR QUE USAR?**

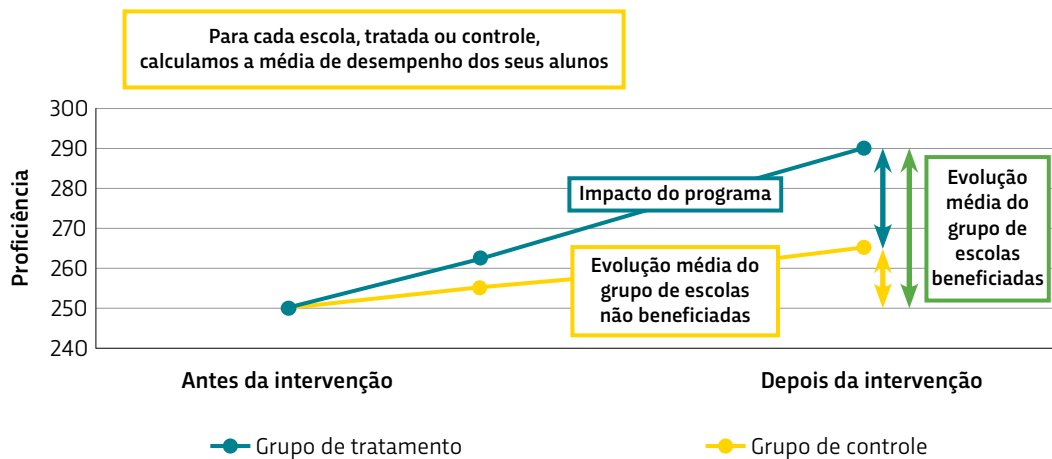
Embora sejam objeto de algumas críticas, os experimentos aleatórios permitem que características relevantes para explicar o desempenho dos alunos sejam as mesmas entre escolas que recebem o programa, aqui chamadas de “escolas tratadas”, e aquelas que não o recebem, designadas “escolas controle”. Se fossem empregados outros métodos que não garantissem esse equilíbrio desde o início, decerto seria necessário utilizar alguma premissa para assegurar que os resultados encontrados não decorressem desse desbalanceamento entre características previamente observadas. Em programas educacionais, tais características são relevantes; proficiência anterior é importante para explicar o desempenho futuro, bem como evasão anterior e eficiência da gestão.

Em um experimento aleatório, é possível controlar características que não são necessariamente observáveis, como a eficiência da gestão. Será que essa eficiência não está sendo medida com certo erro? Em relação às práticas pedagógicas, pode-se saber com exatidão todo o rol que os professores empregam em sala de aula? Provavelmente não, mas, se o experimento é aleatório, há como garantir que escolas tratadas e controle tenham a mesma distribuição ou as mesmas características prévias em relação a fatores que possam afetar o desempenho. Afinal, nesse método, atira-se uma moeda para o alto e o acaso é que determinará quem receberá o tratamento.

Então, sendo factível, esse desenho experimental permite computar o efeito de uma intervenção em Educação da maneira mais simples possível. Isso se dá porque, na verdade, quando se está interessado em um efeito médio da política, basta calcular duas médias, entre tratadas e entre controles, e obter uma diferença entre elas. É necessário apenas comparar, após a intervenção, as médias dos grupos de tratamento e os grupos de controle – algo relativamente simples.

Sem dúvida, nem sempre se consegue usar experimentos aleatórios. Sabe-se que certas barreiras éticas e questões institucionais inviabilizam seu uso. Porém, quando se tem a oportunidade, devem ser utilizados.

## EVOLUÇÃO DA PROFICIÊNCIA DURANTE A INTERVENÇÃO



O gráfico acima mostra por que é necessário existir um grupo de controle. Ao longo do tempo, percebe-se que houve ganho no grupo que recebeu a intervenção, indicado pela linha verde, embora não seja claro que isso se deveu apenas à intervenção. Da mesma forma, ocorreu ganho no grupo de controle, dado pela linha amarela. O impacto da intervenção é a diferença entre a seta dupla verde e a marrom, representada pela seta dupla azul em impacto do programa.

## MEDINDO A DESIGUALDADE

Muitas vezes, o pesquisador está interessado não apenas nos efeitos médios, mas também nos efeitos distributivos da política educacional. Ele quer saber, por exemplo, se o programa educacional aplicado reduziu disparidades ou se, diferentemente, ampliou-as. O primeiro passo para isso é definir uma medida de desigualdade, pois há diversas. Este artigo se restringirá a uma delas: vai comparar o desempenho dos melhores alunos com o dos piores e verificar se a política educacional conseguiu, além de elevar a média dos alunos ao longo do tempo, reduzir o hiato entre os melhores e os piores.

Não é certo que políticas educacionais, mesmo que muito bem pensadas e desenhadas, alcancem tal resultado. Não raro,

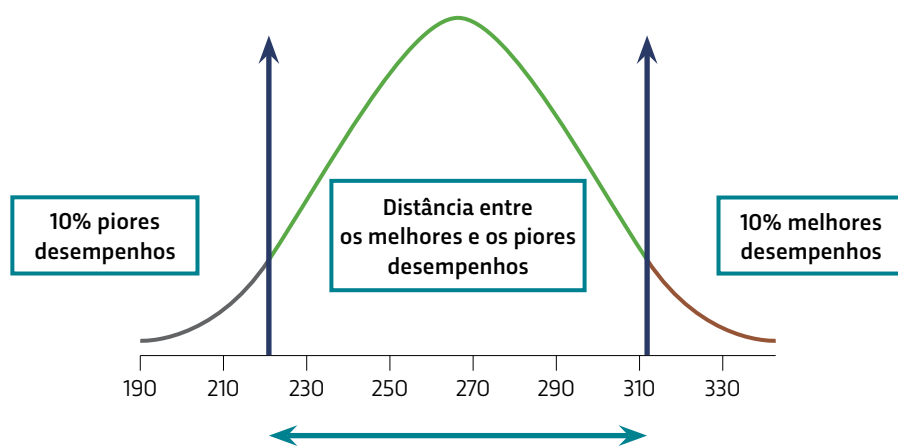


esse cuidado nem sequer está presente. Porém, para que haja redução da desigualdade, são necessárias políticas educacionais que tenham esse foco, além de aumentar a média.

Para exemplificar tal desafio, vou basear minha explicação no histograma a seguir.

Primeiro, é importante perceber que as notas dos alunos se agrupam por conveniência; na verdade, são uma aproximação do que seria a verdadeira distribuição de notas, dada, como se observa no gráfico a seguir, pela linha verde. Essa linha verde, que os estatísticos chamam de densidade, é usada para calcular a desigualdade de desempenho. Linhas verdes mais achatadas representam distribuições com dispersão maior, ou seja, situações em que a desigualdade é maior; linhas verdes mais altas e mais estreitas são densidades indicativas de menor desigualdade. Por quê? Porque a diferença entre o melhor e o pior desempenho diminui.

#### EFEITO DO PROGRAMA NA DISTRIBUIÇÃO DO DESEMPENHO



Fonte: elaboração do autor.

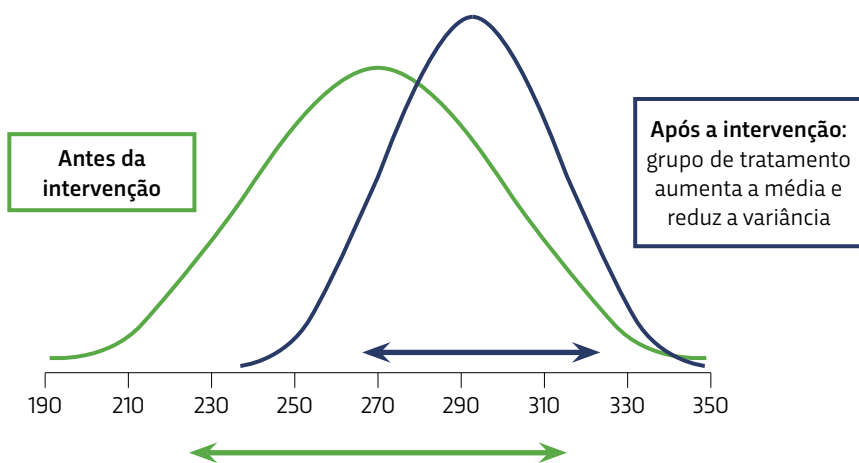
É possível utilizar a linha verde para construir medidas de desigualdade. As áreas abaixo das linhas mais escuras contêm, do lado esquerdo, os 10% de alunos com notas em torno de 200 e, do lado direito, os 10% com notas superiores a 310; são, respectivamente, os grupos com pior e com melhor desempenho. Por

meio desse gráfico, pode-se, ainda, ter noção da distância entre os 10% piores e os 10% melhores ou entre o melhor dos 10% piores e o pior dos 10% melhores, como indicado pela seta dupla azul clara.

Essa é uma medida de distância entre pontos da distribuição utilizada com frequência para o cálculo de medidas de dispersão ou desigualdade. Existem diversas, mas nesse gráfico apenas se observará a distância dada pela linha azul.

Uma vez decidida a medida, o segundo passo é definir: desigualdade entre quem? Entre alunos de uma mesma sala? De um mesmo ano? De uma mesma série? De um mesmo ciclo? De uma mesma escola? Ou é desigualdade da média das escolas entre si? É possível computar a média das escolas e observar como o desempenho médio ficou mais disperso, ou seja, a melhor escola agora talvez tenha se distanciado da pior, ou, se a política educacional promoveu redução da desigualdade, a distância pode ter encurtado. Outra possibilidade é avaliar o desempenho médio de cada uma das escolas e reportar a distância entre as 10% melhores e as 10% piores. A definição desse ponto é fundamental, porque, dependendo dela, as políticas públicas talvez tenham efeitos diferentes. É possível que uma política pública não afete em nada a dispersão de desempenho médio entre as escolas, mas influencie bastante a desigualdade dentro das escolas.

### EFEITO DO PROGRAMA NA DISTRIBUIÇÃO DO DESEMPENHO

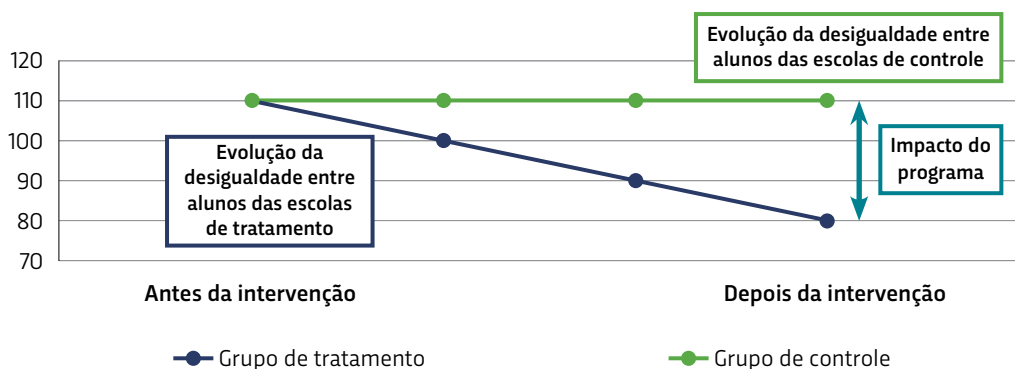


Fonte: elaboração do autor.

O gráfico da página anterior mostra o grupo de controle, exatamente igual ao grupo de tratamento antes da intervenção ou da implantação da política em avaliação. A linha azul representa uma política educacional que aumenta o desempenho médio e, ao mesmo tempo, reduz a dispersão entre alunos ou entre escolas, e a seta dupla indica exatamente a nova distância entre alunos de uma mesma turma.

O que se tem nesse gráfico é um exemplo de política educacional que levou à redução da desigualdade entre alunos. Tal medida permite calcular o que havia antes da intervenção e o que ocorreu depois. Também é possível representar o grupo de controle e o de tratamento; este, ao longo do tempo, teve sua dispersão ou desigualdade reduzidas. O ganho da política sobre a desigualdade é dado pela linha azul.

### EVOLUÇÃO DA DISTÂNCIA ENTRE AS MÉDIAS DOS 10% MELHORES E DOS 10% PIORES DESEMPENHOS



Fonte: elaboração do autor.

Por que a linha verde do gráfico foi mantida sem inclinação? Porque se pressupõe que, na ausência de intervenção, não houve interferência na desigualdade de desempenho. Isso pode ser verificado nos dados; talvez o tempo tenha feito a desigualdade aumentar ou diminuir. Aqui, porém, parte-se do princípio de que nada aconteceu. Sendo assim, o ganho do programa sobre a redução da desigualdade é, agora, diferente do gráfico anterior, que representava um ganho de médias; trata-se de uma redução da desigualdade.

## *Em certas situações, o programa gera aumento da desigualdade [...] Os experimentos aleatórios permitem fazer essa conta*

No entanto, em certas situações, o programa gera aumento da desigualdade, e é importante que isso possa ser medido. Os experimentos aleatórios permitem fazer essa conta, porque, adotando-os, tem-se a distribuição do grupo de tratamento e a do grupo de controle. Logo, é viável calcular tais medidas de desigualdade: testar a hipótese de que o programa reduziu a desigualdade ou checar e concluir que, na realidade, o programa contribuiu para aumentá-la.

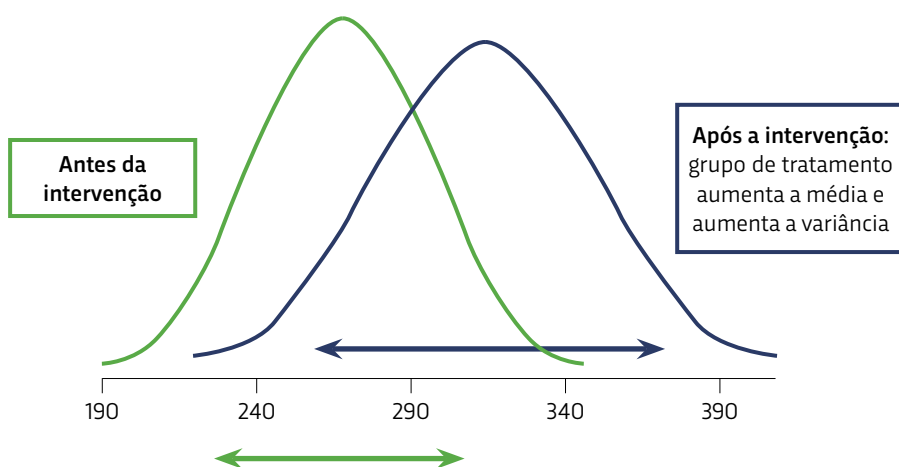
Um exemplo em que isso ocorre é quando se instauram metas, porém professores e gestores não recebem instrumentos pedagógicos adequados para atingi-las. Imagine-se uma situação em que o professor é desafiado a aumentar a nota média de seus alunos na escala do Sistema de Avaliação da Educação Básica (Saeb). Sem as ferramentas apropriadas, corre-se o risco de que ele concentre seus esforços em um grupo de estudantes que responderá mais rapidamente. Isso pode gerar desigualdade em sala de aula; afinal, talvez esses alunos já fossem os melhores antes da intervenção.

Também é possível que haja aumento da desigualdade entre escolas. Uma política pode gerar efeitos perversos, causando impactos diversos entre escolas diferentes. Mesmo que tenha sido muito bem desenhada, ela só funcionará perfeitamente, na maior parte das vezes, em escolas cuja complexidade de gestão é pequena. Em escolas de gestão altamente complexa – e aqui parte-se do pressuposto de que, quanto maior a complexidade da gestão, menor o desempenho prévio –, é de esperar que o resultado final seja um aumento da dispersão educacional entre escolas; afinal, essa política terá maiores impactos positivos en-

tre aquelas em que a complexidade de gestão é menor. As políticas públicas educacionais precisam ser pensadas levando em conta esses eventuais subprodutos de sua intervenção. Contudo, é possível medir seu impacto.

O foco recente dos experimentos aleatórios tem sido, sobretudo, mostrar como os desempenhos, em média, melhoraram ou não, dependendo da política. A princípio, é o que se considera quando se realiza um experimento aleatório. Isso, porém, não deve tolher o trabalho do pesquisador: é perfeitamente possível usar números e estatísticas para calcular aumentos e reduções da desigualdade de desempenho entre alunos e entre escolas. O gráfico a seguir traz uma situação esquemática em que houve aumento da dispersão.

#### EFEITO DO PROGRAMA NA DISTRIBUIÇÃO DO DESEMPENHO



Fonte: elaboração do autor.

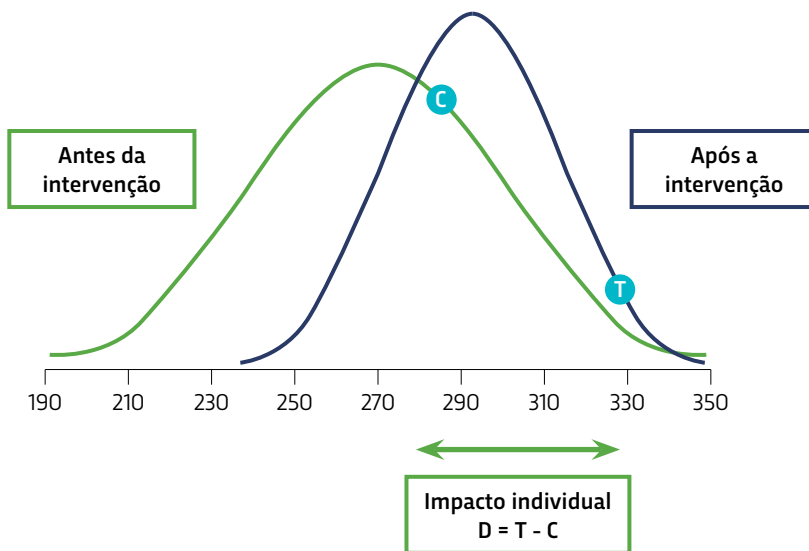
Nota importante aqui é que a diferença entre os piores alunos não cresce tanto quanto a diferença entre os melhores alunos, e é por isso que a desigualdade aumenta. A política educacional adotada amplificou o potencial dos melhores alunos. Cabe checar se isso fazia parte ou não da carta de intenções e, se for o caso, repensar se esse programa, apesar de aumentar o desempenho médio, não precisa ser readequado para que, no mínimo, não se aumente a dispersão.

## EFEITOS INDIVIDUAIS

Por fim, é preciso se perguntar: os dados experimentais oferecerão todas as respostas sobre desigualdades? Não; mesmo que se tenha o experimento mais controlado do mundo, não há como falar tudo sobre distribuição do desempenho. Talvez o interesse do pesquisador sejam os efeitos para um indivíduo. O problema, quando se pensa em um indivíduo, é que não será possível observá-lo nas duas situações: sem e com tratamento. Portanto, é algo que não se consegue fazer apenas com experimento. Seria necessário, por exemplo, saber como escolas ou alunos mudam de posição relativa diante das demais escolas ou de alunos que receberam a intervenção.

### EFEITO DO PROGRAMA NA DISTRIBUIÇÃO DO DESEMPENHO

#### EFEITOS INDIVIDUAIS DA INTERVENÇÃO



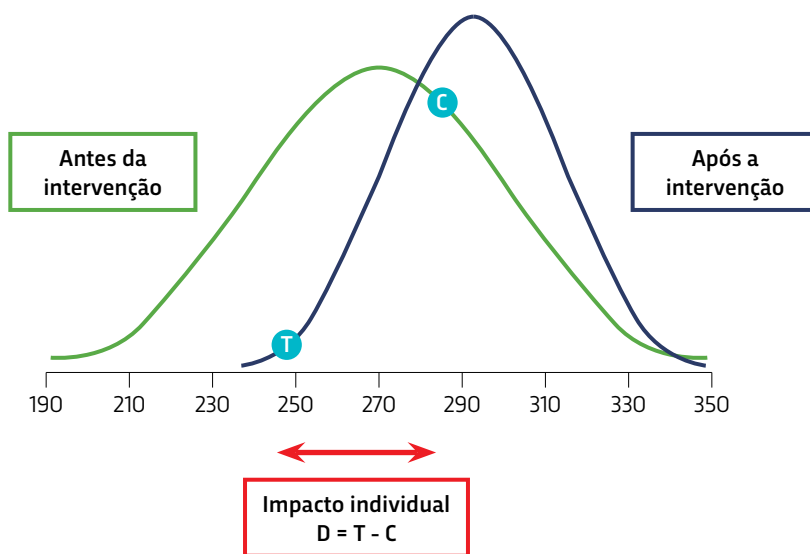
Fonte: elaboração do autor.

Imagine-se, no gráfico acima, que, na situação sem intervenção, uma escola estivesse no ponto C. Com a intervenção, a posição relativa melhora e chega-se ao ponto T. O ganho é dado pela seta dupla verde. O problema é que não se pode medir exatamente a distância em verde; essa informação não está dispo-

nível nos dados colhidos, porque, para isso, seria necessário que os pontos T e C representassem a mesma escola. Ocorre, porém, que não é possível saber onde a escola está em uma das situações: se ela não tiver recebido a intervenção, fará parte do grupo de controle e lá ficará; se tiver recebido a intervenção, não há como ter certeza do que teria acontecido com ela caso pertencesse ao outro grupo. Consegue-se discorrer sobre médias e sobre medidas de desigualdade, mas não sobre o que teria acontecido com essa escola em particular.

### EFEITO DO PROGRAMA NA DISTRIBUIÇÃO DO DESEMPENHO

#### EFEITOS INDIVIDUAIS DA INTERVENÇÃO



Fonte: elaboração do autor.

O gráfico acima ilustra bem essa incerteza. É possível que, em vez de a escola ter ganhado individualmente com a intervenção, ela tenha perdido. A distância representada pela seta dupla vermelha mede essa perda.

O que se vê é uma situação em que houve aumento, em média, bem como redução da dispersão entre escolas; no entanto, não é possível saber se essa escola sofreu com a intervenção ou se teria sofrido caso a tivesse recebido. Ela pode ter tido um ganho ou uma perda, mas não há como dizer, pois não se tem in-

formação sobre a distribuição dos efeitos do programa do ponto de vista da escola individualmente. Por uma questão de mensuração de ganhos de bem-estar, esse ponto é relevante.

Porém, ainda que não se consiga explicitar precisamente o que teria acontecido com essa escola em particular, é possível construir alguns intervalos em que tais efeitos estariam. É só pensar que, para cada escola, haveria um efeito máximo e um mínimo; então, talvez a resposta não se apresentasse como um ponto, e sim como um intervalo. Diante desse raciocínio, pode-se tomar uma das seguintes atitudes: admitir que, mesmo que haja pouco a fazer em relação ao indivíduo com os resultados de experimentos aleatórios, talvez se possa construir intervalos de efeitos; ou impor uma premissa simplesmente para facilitar a avaliação de que as pessoas e as escolas permanecem nos mesmos *rankings* ou posições relativas, tendo sido tratadas ou não.

## CONCLUSÃO

Este artigo empenhou-se em mostrar que um experimento aleatório pode, sim, responder a perguntas que os pesquisadores e os gestores se fazem. É possível calcular efeitos médios, efeitos de distribuição, isolar o efeito causal de um programa e, portanto, definir o que é o efeito causal na média e sobre distribuições. Para o dia a dia do gestor, pode ser bastante relevante saber se o programa beneficia mais os alunos que já apresentavam melhor desempenho ou se ajuda aqueles de pior desempenho e, conseqüentemente, reduz a desigualdade. Do ponto de vista das escolas, trata-se de um programa que aumenta a diferença entre elas ou a reduz? Isso é relevante. No entanto, o experimento aleatório não atende à necessidade de tomar atitudes que levem em conta o ponto de vista do indivíduo. Para isso, seria necessário haver mais restrições acerca do comportamento ou alguma teoria subjacente. Só assim se poderia afirmar algo sobre a distribuição dos efeitos individuais.



# *COMO ABRIR A CAIXA-PRETA DO IMPACTO*

**PEDRO CARNEIRO** é professor associado da University College London e economista e pesquisador do Centre for Microdata Methods and Practice, da mesma instituição. Foi professor associado visitante e professor associado da Universidade de Georgetown e atuou como economista visitante do Banco Mundial. Possui Ph.D. em economia pela Universidade de Chicago, sob a orientação de James Heckman.

Este artigo trata da utilização das estimativas de impacto.

Em geral, dá-se grande ênfase ao rigor, a demonstrar o resultado causal de uma intervenção ou programa. No entanto, muitas vezes, quando se consegue um resultado, é difícil usá-lo, por várias razões: o programa é complexo demais; a população, bastante específica ao contexto; ou, talvez, a implementação da intervenção foi muito cuidada, porque o programa, para ser tão bem avaliado, pode ter sido menor, com os envolvidos muito motivados. Enfim, um conjunto de circunstâncias especiais muitas vezes dificulta o uso desse resultado. Então, o que é a caixa-preta do impacto?

Vejamos, por exemplo, um programa muito complexo, como a construção de uma escola com habilitação do gestor e treinamento do professor. Aí obtém-se um impacto, que é o que chamo de caixa-preta. O que ocorreu dentro dela?

É fundamental entender o que aconteceu para que se possa usar essa avaliação ou esse resultado. Caso contrário, torna-se difícil responder às várias perguntas que surgem:

- Que aspectos do programa funcionaram?
- Por que alguns funcionaram e outros não?
- Que aspectos necessitam de reforço?
- Como melhorar resultados?
- Como replicar resultados?
- Que populações são mais afetadas?
- Onde a implementação foi mais bem-sucedida?
- Que mecanismos estão em ação?
- Qual seria o impacto de um programa ligeiramente distinto?

Todos esses aspectos estão contidos na caixa-preta. No entanto, se o programa é muito complexo, compreender que aspectos foram bem-sucedidos pode ser mais interessante do que saber se ele funcionou como um todo ou não. Assim, cabem também outras questões, a saber:

- Se não há dinheiro para realizar todo o programa, opta-se por fazer uma parte, mas qual parte escolher?
- Será que se deve eliminar parte do programa?
- Que aspectos necessitam de reforço ao redesenhar o programa e melhorar um componente?

## **ABRIR A CAIXA-PRETA**

Durante bastante tempo, muitos na área de Educação eram totalmente avessos a economistas trabalhando nesse campo. Era como se dissessem: “Vocês não entendem nada de Educação, mas vêm fazer essas avaliações e dizer que o programa A causa o resultado B”.

É claro que os economistas não sabem o que há dentro da caixa-preta e, portanto, têm dificuldade em usar esses resultados ou até em interpretá-los. Para abri-la, para ir além de “O programa A causa o resultado B”, é preciso dispor de algum conhecimento prévio na partida.

Simplificando a discussão, há duas maneiras principais de tentar entender o que ocorreu com uma avaliação, sobretudo quando os programas são complexos ou têm resultados inesperados.

*Se o programa é muito complexo, compreender que aspectos foram bem-sucedidos pode ser mais interessante do que saber se ele funcionou como um todo ou não*

Uma delas consiste em avaliar modalidades adicionais dos programas, algo bastante difícil, mas que, basicamente, envolve construir avaliações mais complexas. Por exemplo: um programa abarca, ao mesmo tempo, oferecer livros novos para os alunos, mudar o currículo e treinar os professores. Talvez não sejam necessários três programas – em uma região de São Paulo, só dar livros novos; em outra, apenas treinar os professores; em uma terceira, somente treinar os gestores. Então, uma primeira maneira é fazer uma avaliação mais complexa, com mais ramificações. Embora muitas vezes seja bastante difícil, esse é um modo de procurar compreender qual componente, ou qual combinação de componentes, está funcionando. Outro exemplo: realizar um programa com maior intensidade para uma população mais pobre ou que tenha menos conhecimentos de base.

A segunda maneira de abrir a caixa-preta consiste em usar as avaliações já disponíveis de modo mais rico. Como fazer isso? De um lado, coletando mais informações em cada avaliação, o que também envolve maior custo. Por exemplo: em vez de só captar dados dos insumos e resultados, ou só captar dados sobre o programa que entra e o resultado que sai, tentar criar alguma informação sobre o processo e sobre o contexto. Isso envolve coleta de mais informações e análise de mais dados. Basicamente, trata-se de buscar encontrar correlações em conjunção com o experimento, com a avaliação de impacto experimental que se está realizando, de contar uma história mais rica sobre o que acontece.

Vejamos alguns exemplos. Em geral, tentamos usar essas duas maneiras de abrir a caixa-preta em conjunção, pois são complementares. Primeiro, discorrerei sobre a ideia de como pode ser útil desenhar programas mais complexos. É o caso do Bolsa Família da Macedônia, um país muito pequeno, com pouco mais de 2 milhões de habitantes, que era parte da antiga Iugoslávia. Trata-se de um programa de transferência de renda com várias condicionalidades destinado a pessoas mais pobres. Para receber a transferência de renda, a família precisa assegurar que a criança frequente a escola em pelo menos 95% dos dias letivos.

Há muitas perguntas sobre como esse programa de transferência de renda condicional funciona exatamente, uma vez que é uma intervenção complexa. De um lado, podemos debater questões como: “O que acontece se aumentarmos ou reduzirmos a transferência?”; de outro, levar em conta o componente do *target*: “Estamos atingindo essa população?”; “E se mudarmos o limiar de pobreza abaixo do qual a pessoa tem acesso a essa renda?”.

Muitas outras condicionalidades podem estar envolvidas nesse tipo de programa. Por exemplo: “O que causa impacto nas crianças é renda ou suas condicionalidades?”. Outro exemplo: na maior parte das transferências condicionadas de renda, os pagamentos não são feitos a qualquer pessoa do domicílio, mas à mulher. Entender como os diversos componentes desse programa impactam as famílias e as crianças é importante para o redesenho dele.

*Muitas condicionalidades podem estar envolvidas em um programa de transferência condicionada de renda. Entender como os diversos componentes dele impactam as famílias e as crianças é importante para seu redesenho*

Contudo, talvez a condicionalidade não seja tão importante e possa ser mudada; talvez a transferência seja ineficaz ou grande demais; ou talvez, ainda, não seja necessário pagar à mulher para atingir o impacto que se observa nas crianças ou na diminuição da pobreza.

Uma questão muito simples, não tanto de Educação, mas de economia, ajuda a explicar o uso do desenho de avaliações mais complexas no abrir dessa caixa-preta. Imaginemos que foi implementado o Bolsa Família da Macedônia e obteve-se um impacto: houve aumento da escolaridade das crianças, da despesa das mães com a educação dos filhos, do gasto das famílias com o consumo de alimentos etc.

Existe, porém, uma questão que é um quebra-cabeça: em vários programas de transferência de renda condicional, com as condicionalidades aqui mencionadas, cresce a despesa dos domicílios com alimentação, porque as famílias são muito pobres e sua renda se eleva, então é normal que gastem mais com comida. No entanto, na Macedônia, não acontecia apenas isso: verificamos o aumento do peso da despesa com alimentação no orçamento total desses domicílios. Era um mistério. Em quase todo o mundo, quando a renda da família cresce, a despesa com alimentação como porcentagem do orçamento total normalmente diminui, porém, no caso dessas transferências condicionais, houve incremento, algo, portanto, que não se esperava.

Talvez possa parecer um pouco sem importância, mas a hipótese mais provável para explicar por que isso acontece é a seguinte: tais programas levam ao aumento do peso da alimentação na despesa total do domicílio porque os pagamentos são feitos à mulher e não ao homem. Dessa maneira, além de a renda do domicílio crescer, dá-se mais poder de decisão à mulher e talvez, por essa via, a alocação do recurso se reencaminhe mais de acordo com as preferências femininas.

Na Macedônia, transferir renda para a mulher não é natural, de acordo com a cultura local, e reivindicou-se que o recurso fosse entregue ao chefe do domicílio, em geral homem. Foram feitos vários desenhos desse programa para tentar descobrir que componentes funcionavam e a causa de seu impacto. Um deles variou a modalidade de pagamento conforme o gênero do receptor da família.

Assim, em metade dos municípios da Macedônia, quem recebia a renda da transferência era a mulher; na outra metade, o chefe do domicílio, quase sempre homem. Em resumo: fizemos um desenho mais complexo de um programa típico de transferência de renda condicional para entender um pouco mais o resultado de desenhos diversos. De fato, encontramos que, quando o pagamento era feito ao homem, a porcentagem da alimentação no orçamento era de 50%; quando à mulher, de 55% – ou seja, pagar à mulher fazia diferença. Esse componente era importante em nossa avaliação de impacto.

## MODELOS

Muitas vezes, é difícil pensar em todos os desenhos alternativos de um programa; na verdade, é impossível. Existe a seguinte teoria subjacente: “Acho que essa mudança está ocorrendo por tais razões. Vou testar talvez uma, no máximo duas delas. Ao longo do tempo, talvez possa testar mais, porém em número limitado, informado por minha teoria, por minha ideia do que está acontecendo com os mecanismos desse programa”.

Normalmente, não se consegue avaliar todos os possíveis desenhos de um programa. Então, a alternativa, como mencionado, envolve coletar e processar mais informações em cada avaliação.

Um exemplo específico da área de Educação da 3ie<sup>1</sup> foi um estudo sobre bolsas para escolas no Senegal, realizado em 2008-2009. O programa existia no país desde o início dos anos 1990 e nunca tinha sido avaliado, ou seja, não se sabia se era um bom uso dos recursos ou não.

As escolas, além do financiamento centralizado normal que recebiam do governo, podiam criar variados projetos pedagógicos ou de infraestrutura, como comprar livros, treinar professores ou construir bibliotecas. Elas formulavam uma can-

---

1. 3ie é uma organização não governamental internacional que financia avaliações de impacto, divulga resumos de evidências e fornece recomendações para subsidiar políticas públicas.

*É difícil pensar em todos os desenhos alternativos de um programa [...] Normalmente, não se consegue avaliar todos os possíveis desenhos. Então, a alternativa envolve coletar e processar mais informações em cada avaliação*

didatura a financiamento dessas iniciativas e, se aprovadas, recebiam o dinheiro do Ministério da Educação para investir nelas. No entanto, o órgão sempre se perguntava: “Será que esse financiamento está fazendo alguma diferença na aprendizagem das crianças ou não?”. Então, basicamente, o programa era uma grande caixa-preta.

Fizemos a avaliação. O que ocorria? O Ministério da Educação repassava dinheiro para as escolas e verificava-se 0,12 de desvio padrão (DP). Analisamos escolas do país inteiro, e só algumas concorriam todos os anos; umas recebiam dinheiro e outras não, pois não havia verba para financiar todos os projetos bons. A conclusão da avaliação de impacto experimental foi a de que, quando as escolas tinham acesso a esse dinheiro adicional, em média, no Senegal, a aprendizagem dos alunos em matemática e em francês aumentava 0,12, ou seja, 12% de desvio padrão. Apesar de não ser um efeito muito alto, era positivo.

Tínhamos coletado muitas informações, pois queríamos abrir um pouco essa caixa-preta. Uma divisão muito simples que podia ser feita era entre o Norte e o Sul do país. O Norte era mais rico e pacífico do que o Sul, região com muito terrorismo, tanto que em várias escolas foi difícil realizar a pesquisa, porque nelas havia terroristas com metralhadoras. Queríamos entender as diferenças regionais, uma vez que o contexto devia importar; achávamos que no Sul seria muito difícil a qualquer programa produzir impacto, pois as condições eram realmente desfavoráveis.



Entretanto, ocorreu exatamente o contrário. No Sul, o impacto foi de 39% de desvio padrão, e no Norte, 7% – nesse caso, não podemos rejeitar que não houve impacto algum, estatisticamente. Então, todo o impacto ocorre nas escolas do Sul. Por que existem essas diferenças regionais? O resultado foi contra o que se imaginava que aconteceria. A questão foi levada ao Ministério da Educação para que se levantassem hipóteses.

## **HIPÓTESES**

É possível ter um modelo bem simples: quando se entrega dinheiro às escolas, elas podem gastar em equipamentos ou pessoas, duas categorias bem definidas do uso dos recursos; isso será útil para o que vou mostrar a seguir. O pensamento, na realidade, foi um pouco mais complexo do que exposto aqui. E, claro, havia variáveis de contexto e ambiente familiar, fatores que interferem no resultado. Também foram observados diversos aspectos do ambiente familiar e, com bastante cuidado, as categorias em que as escolas estavam utilizando os recursos.

Havia diferenças drásticas entre o uso dos recursos no Sul e no Norte. Essas eram as variações mais importantes nos projetos – o montante de dinheiro era parecido, mas a utilização era bastante diversa. Quatro categorias de gastos podiam estar incluídas: livros, computadores, formação em gestão escolar e formação de professores.

Em algumas escolas, gastava-se sobretudo com livros e computadores; em outras, com recursos humanos, ou seja, na formação de professores ou em gestão escolar.

No Sul, a porcentagem de projetos com um componente importante de livros era 80%; no Norte, 90%. No Sul, a porcentagem de projetos com um componente importante de compra de computadores era 3%; no Norte, bastante maior, 12%. No entanto, quando observamos os gastos com recursos humanos, ocorria o contrário. No Sul, 91% dos projetos tinham um componente grande de gastos com formação em gestão escolar; no Norte, só 75%. No Sul, 63% dos projetos tinham uma componente grande de gastos com formação de professores; no Norte, apenas 37%.

O que isso sugere, em conjunção com o experimento? Conseguimos medir um impacto causal rigoroso: dar mais dinheiro às escolas produz resultado nos alunos, e esse impacto é importante sobretudo no Sul, onde os recursos são mais usados para apoiar os professores e menos para adquirir livros e computadores. O Ministério da Educação diz: “Isso faz sentido”.

Visitei quatro escolas no Senegal, todas na capital, Dacar. Na primeira, o diretor tinha comprado um computador. Fomos até o laboratório de informática e a primeira coisa que vi foi um cartaz na parede: “Meia hora: 5 francos. Uma hora: 10 francos”. O diretor havia perguntado nas salas: “Vocês gostam de computador?”. “Sim”, responderam os alunos. “Vocês usam computador?”. Silêncio total. “Por que não usam computador?”. “Porque é muito caro”, foi a explicação. O diretor tinha basicamente uma lan house na escola. Esse resultado correspondia à evidência causal do que estava acontecendo no programa.

## COMENTÁRIOS FINAIS

Obviamente, é fundamental saber se o projeto funciona ou não e ter uma avaliação rigorosa de seu impacto, com grupo de controle definido, experimentos etc. Contudo, do ponto de vista do gestor público e mesmo da perspectiva acadêmica, é muito mais útil entender por que o programa teve sucesso: quais componentes funcionaram melhor, o que exatamente mudou no comportamento do professor ou do gestor, entre outros fatores. Também se deve questionar: “O que é preciso alterar no programa para o gestor dar mais atenção à aprendizagem dos alunos?”; “Em que populações o programa funcionou melhor?”; “Como posso fazer um *targeting* mais eficiente desse programa?”.

Essas perguntas são tão ou mais importantes do que saber se o programa funcionou ou não. Fica neste artigo, portanto, uma reflexão sobre o valor de tentar responder a perguntas adicionais e fazer um uso mais rico das avaliações, acrescentando um pouco de complexidade. É bastante promissor buscar coletar e analisar mais informações sobre a avaliação que temos a fazer, e essas duas condutas são complementares.

# *CIÊNCIA PARA A EDUCAÇÃO: UMA JANELA DE OPORTUNIDADE PARA O BRASIL*

**ROBERTO LENT** é graduado em medicina pela Universidade Federal do Rio de Janeiro (UFRJ), com mestrado e doutorado pelo Instituto de Biofísica da UFRJ e pós-doutorado pelo Massachusetts Institute of Technology (MIT). É professor titular da UFRJ, onde chefia o Laboratório de Neuroplasticidade do Instituto de Ciências Biomédicas. É também coordenador da Rede Nacional de Ciência para Educação e membro titular da Academia Brasileira de Ciências.

Ciência para a Educação é um conceito relativamente novo no âmbito internacional e em nosso país, com implicações cruciais para o futuro da Educação. Esse conceito deriva de um outro, mais antigo, o de pesquisa translacional, que surgiu de um famoso relatório produzido por um assessor importante do presidente Franklin Roosevelt, Vannevar Busch (1890-1974), logo após a Segunda Guerra Mundial. A pesquisa translacional significava então uma transição linear entre a pesquisa básica, pela qual o cientista trabalha desinteressadamente, movido pela sua curiosidade, e a pesquisa aplicada, ao cargo de profissionais capazes de gerar produtos e processos derivados do que se havia descoberto antes.

Uma nova interpretação surgiu nos anos 1990, também nos Estados Unidos, de autoria de um cientista político da Universidade de Princeton, chamado Donald Stokes (1914-1997). Ele dividiu a pesquisa científica em quadrantes, conforme a figura na próxima página, a seguir. Na visão de Stokes, a pesquisa científica deveria atender a dois objetivos essenciais: gerar conhecimentos fundamentais sobre a natureza e a sociedade e/ou gerar

resultados que pudessem ter utilização social. Organizando esses dois objetivos em eixos ortogonais, Stokes chegou a quatro quadrantes, deu-lhes nomes ilustrativos em homenagem a cientistas famosos e identificou um deles como o mais importante do ponto de vista das políticas públicas de fomento à Ciência – o quadrante de Louis Pasteur (1822-1895). Outros dois quadrantes homenageavam Niels Bohr (1885-1962) e Thomas Edison (1847-1931), e Stokes deixou em branco o quarto quadrante, que bem poderia, no entanto, homenagear o zoólogo Carl Linnaeus (1707-1778), autor de um sistema de classificação dos seres vivos que não tem interesse social maior nem acrescenta algum novo conceito à ciência, mas é muito útil aos próprios pesquisadores.

### PESQUISA TRANSLACIONAL – CONCEPÇÃO 2D

		Utilidade prática?	
		NÃO	SIM
Conceitos fundamentais?	SIM	<b>PESQUISA BÁSICA PURA</b> (Niels Bohr)	<b>PESQUISA INSPIRADA PELO USO</b> (Louis Pasteur)
	NÃO	<b>PESQUISA BÁSICA APLICADA</b> (Carl Linnaeus)	<b>PESQUISA APLICADA PURA</b> (Thomas Edison)

Fonte: modelo Donald Stokes, 1997.

Niels Bohr foi um físico quântico do início do século passado. O trabalho que ele fazia, na época em que o fazia, não tinha nenhuma implicação prática – tampouco ele estava preocupado com isso. Posteriormente, viu-se que suas descobertas permitiram desenvolver conceitos de grande relevância para a física e importantes aplicações tecnológicas. Seu trabalho era movido pela curiosidade de conhecer a natureza.

Thomas Edison, por outro lado, estava interessado em inventar um dispositivo que iluminasse ambientes. Ele não queria saber qual o papel do vácuo dentro da lâmpada, qual a velocidade de propagação da luz dentro do bulbo e assim por diante. Não estava preocupado com conceitos fundamentais, mas apenas com a utilidade prática das descobertas e invenções.

Donald Stokes privilegiou o quadrante de Louis Pasteur (1822-1895). Pasteur revolucionou as ciências da vida ao derrotar a ideia aristotélica de que a vida podia ser gerada espontaneamente. Essa ideia não era verdadeira, e o cientista descobriu isso inspirando-se nos vinicultores e cervejeiros da França, que queriam saber por que seus vinhos e cervejas azedavam. Ninguém tinha a resposta para aquele fenômeno, e Pasteur foi estudá-lo, chegando a conceitos fundamentais sobre a existência de microrganismos e sua ação. Ao mesmo tempo, resolveu problemas práticos e desenvolveu processos tecnológicos, como foi o caso da técnica que levou seu nome, a pasteurização. Portanto, seu trabalho respondia “sim” à pesquisa criadora de conceitos fundamentais e “sim” às questões de utilidade prática. Essa estratégia ilustra o chamado quadrante de Pasteur (Stokes publicou um livro muito interessante com esse título<sup>1</sup>), e é considerada a abordagem mais interessante para as políticas públicas, pois não abandona nem a ciência básica, nem a inovação tecnológica, incorporando-as no mesmo conjunto. Este artigo discutirá brevemente a pesquisa translacional com esse viés, que é muito prevalente na área de saúde. Depois, fará a relação com a Educação.

## **PESQUISA TRANSLACIONAL EM SAÚDE**

Na saúde, a pesquisa translacional funciona muito bem há décadas – usa-se até a expressão “da bancada à beira do leito”, refletindo a concepção linear descrita antes, ou adicionando um “vice-versa” ao final, para ilustrar algo mais próximo do quadrante de Pasteur. Do lado da bancada estão as universidades e os ins-

---

1. *O quadrante de Pasteur – a ciência básica e a inovação tecnológica* (trad. J. E. Maiorino), Editora Unicamp, 2005, pp. 247.

## *A abordagem mais interessante para as políticas públicas não abandona nem a ciência básica, nem a inovação tecnológica, incorporando-as no mesmo conjunto*

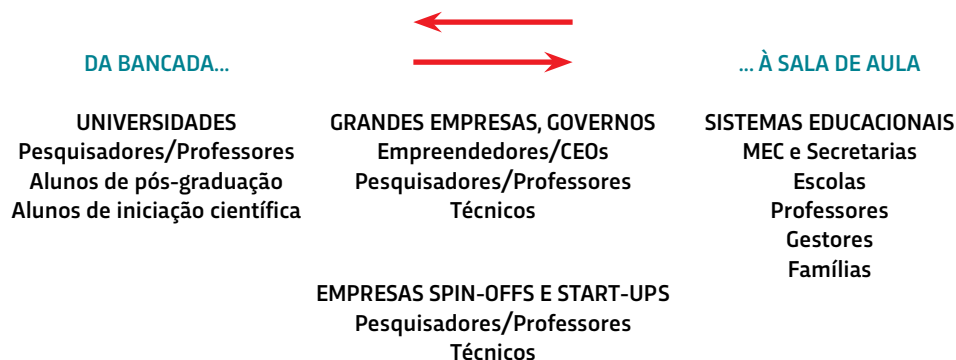
titutos de pesquisa; depois, as empresas pequenas e as grandes, que dão escala às invenções. Nesse processo, elas eventualmente chegam à beira do leito por meio dos sistemas de saúde, dos quais fazem parte hospitais, clínicas, médicos, enfermeiros e assim por diante. Trata-se de um processo aceito socialmente no mundo inteiro e que levou a progressos inegáveis: a despeito da diversidade social e econômica entre os países, houve redução da mortalidade infantil e geral, aumento da longevidade e cura de várias doenças.

Há diversos exemplos do sucesso da pesquisa translacional em saúde. Um deles esteve em foco recentemente: a epidemia de infecção pelo vírus Zika que o Brasil enfrentou recentemente. Os governos federal e estaduais criaram editais de fomento que reuniram pesquisadores básicos, médicos e tecnólogos em universidades e instituições de saúde, que rapidamente conseguiram estabelecer a relação de causa e efeito entre o vírus Zika e a microcefalia, conhecer as características do desenvolvimento embrionário das crianças infectadas e começar a desenvolver uma vacina para evitar que a epidemia se espalhasse.

## **PESQUISA TRANSLACIONAL EM EDUCAÇÃO**

Pouca gente se dá conta, mas é possível pensar dessa maneira para a Educação também, ou seja, pode-se conceber uma pesquisa translacional para a área, seja utilizando o conceito 1D ou o conceito 2D do quadrante de Pasteur. Como funcionaria isso?

## PESQUISA TRANSLACIONAL EM EDUCAÇÃO



Fonte: elaboração do autor.

O processo iria da bancada à sala de aula e vice-versa. Os atores seriam muito semelhantes: de um lado, universidades, pesquisadores, professores, alunos de pós-graduação etc. No meio, a possibilidade de criação de *start-ups* e *spin-offs*, como mostra a figura acima. Não existe aqui ainda o equivalente à *big pharma*<sup>2</sup> no caso da saúde, mas há grandes empresas, e, fundamentalmente, existem os governos. Pode-se pensar, por exemplo, em grandes empresas – como o Google – que investem muito em dispositivos que influenciam a Educação, ainda que não tanto a Educação formal, mas muito a informal. Na outra ponta estão os sistemas educacionais, que são o Ministério da Educação (MEC), as secretarias, as escolas, os professores e assim por diante. Esses atores compõem uma cadeia de pesquisa translacional voltada para a Educação: a ciência para a Educação.

Será que isso é viável? Será que é interessante? Alguns exemplos podem ser citados para ilustrar essa possibilidade, tanto no Brasil como em outros países. Menciono apenas um deles.

Há algum tempo se conhece quais são as regiões cerebrais envolvidas no processamento da leitura pelas crianças, e um grupo israelense<sup>3</sup> descobriu que essas regiões estão menos co-

2. Expressão em inglês pela qual são conhecidas as grandes multinacionais farmacêuticas.

3. Tzipi Horowitz-Kraus e colaboradores (2016). *Annals of Dyslexia*, vol. 66, p. 301-318.



nectadas nas crianças disléxicas. Pois bem: o grupo desenvolveu também um dispositivo capaz de acelerar a leitura tanto de disléxicos como de pessoas normais<sup>4</sup> chamado RAP, usando a abreviatura em inglês para Reading Acceleration Program (Programa de Aceleração de Leitura). O RAP apresenta um texto curto que aparece em um monitor e é depois apagado no mesmo sentido da leitura. Só que, durante o treinamento, a velocidade com que o texto é apagado é aumentada gradualmente, como que “empurrando” a leitura para velocidades crescentes. É claro que, ao final do treinamento, os sujeitos tinham que responder a um pequeno questionário de conteúdo para verificar que não só leram, mas também compreenderam o sentido do que leram. O que se constatou é que o RAP constitui um treinamento capaz de melhorar o desempenho dos disléxicos (pesquisa translacional em saúde), mas também impacta positivamente o desempenho das pessoas normotípicas (pesquisa translacional em Educação).

Não seria o caso de testar a utilização em escala dessa ferramenta, para verificar a sua eficácia terapêutica (no caso dos disléxicos) e educacional (no caso dos normotípicos)? O quadrante de Pasteur fica plenamente preenchido nesse exemplo, pois a conectividade cerebral da leitura é elucidada, e ao mesmo tempo um instrumento com utilidade potencial tanto em saúde como em Educação pode ser aproveitado socialmente.

## **CIÊNCIA PARA A EDUCAÇÃO: POR QUE É NECESSÁRIA?**

O conceito de ciência para a Educação é novo no cenário internacional. Algumas iniciativas têm sido criadas sobre ciência da aprendizagem (do inglês *science of learning*), abordando quase exclusivamente os aspectos neuropsicológicos da parte mais visível da Educação, que é a aprendizagem. No Brasil, preferimos usar o termo mais amplo, porque consideramos que existe *learning* (aprendizagem), mas existe também *teaching* (ensino), e é preciso ser mais abrangente do que apenas considerar o lado

---

4. Zvia Breznitz e colaboradores (2013). *Nature Communications*, vol. 4, p. 1-6.

## *Ciência para a Educação é uma janela de oportunidade que está se abrindo no mundo inteiro*

do aprendiz; é necessário levar em conta também o do educador. Um primeiro encontro internacional foi realizado no Rio de Janeiro como simpósio satélite do Congresso Mundial sobre o Cérebro, em 2015, promovido pela então recém-criada Rede Nacional de Ciência para Educação (ver mais detalhes abaixo). Um segundo encontro está sendo organizado para setembro de 2017 em Brisbane, Austrália. A Fundação Nacional de Ciência dos Estados Unidos tornou permanente um programa de fomento à ciência da aprendizagem, envolvendo alguns centros universitários onde se pesquisa o tema. A Universidade de Queensland, na Austrália, e a Universidade de Hong Kong têm centros semelhantes. Nenhuma iniciativa mais orgânica foi criada nos países europeus, além da atuação esparsa de grupos de pesquisa em universidades de inúmeros países.

Ciência para a Educação, então, é uma janela de oportunidade que está se abrindo no mundo inteiro. Se o Brasil entrar nesse campo agora, poderemos estar à frente do movimento em curso no restante do mundo; se esperarmos dez anos, perderemos a oportunidade. A ampla visão brasileira proposta pela Ciência para Educação é mais abrangente e estratégica, porque aproxima as mais variadas disciplinas científicas – e não apenas as neurociências – do quadrante de Pasteur em benefício da Educação. Assim, a rede engloba pesquisadores em psicologia, pedagogia, ciência da computação, matemática, estatística, sociologia, economia, biologia e tantas outras disciplinas sem limitações em um movimento que se pretende forte para acelerar a aquisição de conhecimento científico sobre a Educação – tanto o ensino como a aprendizagem; tanto o desempenho individual como a performance sistêmica de toda a sociedade; tanto as redes neuronais como os agrupamentos sociobiológicos; tanto as crianças com transtornos de aprendizagem como as normotípicas.

Por que razão é tão importante que o Brasil não perca essa janela de oportunidade? Porque a posição do país no Pisa (Programme for International Student Assessment ou, em português, Programa Internacional de Avaliação de Estudantes) da Organização para Cooperação e Desenvolvimento Econômico (OCDE) é deplorável. Os dados são conhecidos de todos que atuam em Educação: nas três avaliações feitas pelo Pisa em matemática, leitura e ciência, bem como na comparação com as médias de todos os países avaliados pela OCDE, o Brasil se encontra entre as posições mais baixas. Se comparados os índices médios ao longo dos últimos 15 anos, o Brasil patina – talvez esteja até mesmo em declínio.

Os dados, portanto, indicam que é preciso tomar medidas que não apenas garantam um crescimento modesto do Brasil, como o que marcou a década de 2000, mas que acelerem decisivamente esse crescimento. Trata-se não apenas de infletir o crescimento para cima, mas de torná-lo mais vigoroso.

Façamos um exercício de futurologia (figura do alto na próxima página). Se considerarmos uma curva hipotética que descreva a taxa de crescimento do Brasil na última década (em verde); se a extrapolarmos no tempo; e se considerarmos que a média da OCDE se manterá constante, só atingiremos essa média por volta do ano 2060. Observar que nessa figura consideramos apenas como uma oscilação o declínio medido em 2015 (figura do meio na próxima página), e, além disso, imaginamos que a média dos demais países permanecerá constante, até mesmo a dos dez que se encontram no topo.

*É preciso tomar medidas que não apenas garantam um crescimento modesto dos resultados educacionais, mas que acelerem decisivamente esse crescimento. Trata-se não apenas de infletir o crescimento para cima, mas de torná-lo mais vigoroso*

## HIPÓTESE DE INFLEXÃO DA CURVA O EFEITO DE DIFERENTES POLÍTICAS



Fonte: elaboração do autor e Pisa.

Uma política educacional conservadora, isto é, que mantenha o nível vigente dos investimentos em Educação, sem acréscimos, e deixe de tomar quaisquer medidas de reforma, mantendo o *status quo*, conseguirá no máximo atingir a média da OCDE daqui a 40 anos. Uma política educacional acumuladora seria aquela que, sem maiores reformas, mas com algum grau de vontade política, acrescente recursos ainda que modestos às medidas já conhecidas – e justíssimas – de melhoria salarial dos professores, implementação do período integral, viabilização da dedicação docente integral à escola etc. O resultado seria o alcance da média da OCDE em 20 anos (última figura, acima). Em nossa hipótese, é preciso infletir a curva de modo decisivo e acentuado, o que só seria possível com uma política educacional multiplicadora, ou seja, capaz de realizar com vigor as medidas acima e, além disso,

promover um forte movimento de adesão da comunidade científica e acadêmica brasileira ao esforço de criar conhecimento sobre Educação e propor medidas verdadeiramente revolucionárias, posto que desconhecidas do restante do mundo. Nesse caso, talvez seja possível chegar à média da OCDE em menos de 10 anos (página anterior, terceira figura de cima para baixo), e quem sabe atingir os dez países líderes em outro tanto.

Sonho? Talvez. Mas não há outra alternativa, e o momento é agora, em que os países da OCDE ainda não “descobriram” a chave da ciência para a Educação.

O que saliento neste artigo, e que tenho exposto a vários públicos, é que o investimento em ciência para a Educação pode contribuir de modo fundamental com políticas públicas que trabalham outros aspectos indispensáveis e cruciais no sentido de aumentar a derivada de crescimento de nossos indicadores educacionais. Foi com esse objetivo que, em 2014, um conjunto então pequeno de pesquisadores de várias universidades do Brasil, com o apoio do Instituto de Ciências Biomédicas da Universidade Federal do Rio de Janeiro (UFRJ), do Instituto Ayrton Senna, do Instituto D’Or de Pesquisa e Ensino e, posteriormente, da Empresa Brasileira de Pesquisa e Inovação Industrial (Embrapii), criou uma rede nacional de ciência para a Educação chamada Rede CpE<sup>5</sup>.

## **A REDE NACIONAL DE CIÊNCIA PARA EDUCAÇÃO**

Em 2015, descobrimos que não sabíamos quem éramos, nem onde estávamos, nem o que fazíamos, porque essa abordagem pasteuriana da ciência para a Educação não é adotada na maioria das instituições de pesquisa do Brasil. Decidimos então fazer uma prospecção, uma espécie de censo de ciência para a Educação no país. A tarefa coube à bióloga Daniele Botaro, bolsista de pós-doutorado do Instituto Ayrton Senna na UFRJ, e a Jesus Mena-Chalco, professor de ciência da computação da Universidade Federal do ABC (SP). O trabalho consistiu em uma prospecção na platafor-

---

5. Disponível em: <<http://cienciaparaeducacao.org>>. Último acesso em setembro de 2017.

## *Palavras-chaves identificam os professores e pesquisadores que trabalham com alguma linha de pesquisa ou disciplina científica inspiradas pela aplicação na Educação*

ma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)<sup>6</sup> e no banco de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)<sup>7</sup>.

Usando palavras-chaves, foram levantados vários indicadores que permitiram identificar, na comunidade científica brasileira, quem são os professores e pesquisadores que trabalham com alguma linha de pesquisa ou disciplina científica inspiradas pela aplicação na Educação. Resultou disso uma ferramenta aberta intitulada Plataforma CpE, disponível no site da Rede<sup>8</sup>. Qualquer pessoa pode procurar, pelo nome do pesquisador, assunto ou local de trabalho o que ele faz e seu grau de senioridade e produtividade na carreira.

A plataforma permitirá utilizar técnicas de grafos, como mostra a figura na próxima página. Esse estudo-piloto foi realizado da seguinte maneira: para o lado esquerdo, em azul, estão os pesquisadores que trabalham na área de neurociência; do lado vermelho encontram-se os que trabalham em Educação; no meio, em amarelo, localizam-se os que combinam os dois campos. Há três nomes de excelentes pesquisadores nos três perfis, que podem servir de exemplo: em neurociência, Jorge Moll Neto, do Instituto D'Or, Rio de Janeiro; em Educação, Patrícia Alejandra Behar, da Universidade Federal do Rio Grande do Sul (UFRGS); e na intersecção, Sylvia Maria Ciasca, da Universidade Estadual de Campinas (Unicamp).

---

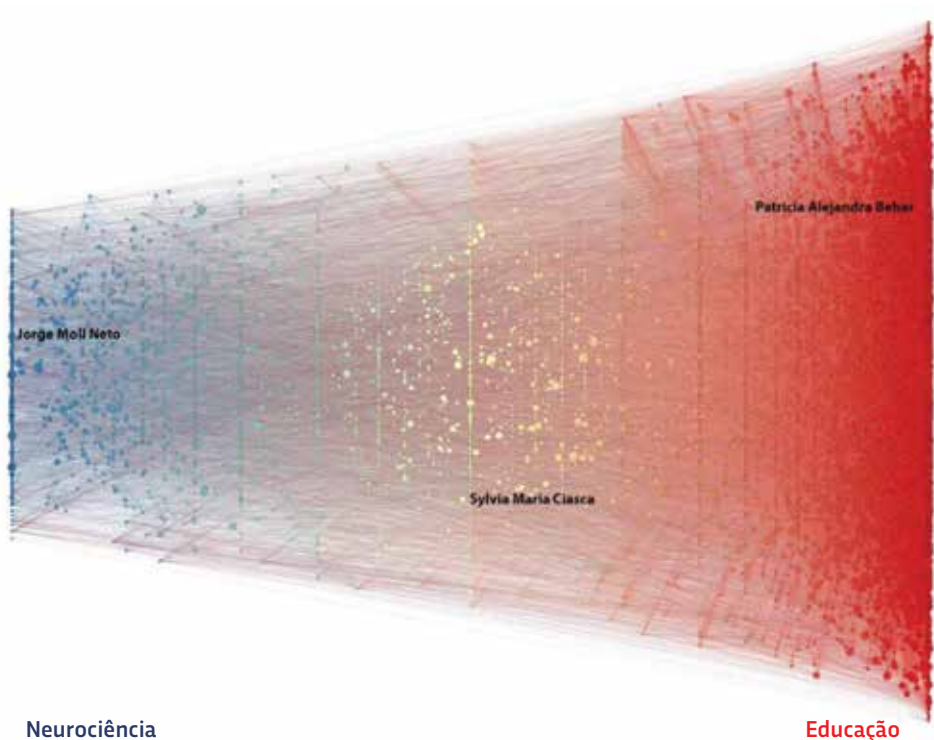
6. Disponível em: <[www.cnpq.br](http://www.cnpq.br)>. Último acesso em setembro de 2017.

7. Disponível em: <[www.capes.gov.br](http://www.capes.gov.br)>. Último acesso em setembro de 2017.

8. Disponível em: <<http://cienciaparaeducacao.org/plataforma-cpe>>. Último acesso em setembro de 2017.

## A REDE NACIONAL DE CIÊNCIA PARA A EDUCAÇÃO

CENSO NACIONAL DE CPE



Neurociência

Educação

Fonte: Rede Nacional Ciência para Educação.

Na plataforma, colocando-se o cursor sobre cada ponto, aparece o nome do pesquisador e toda sua rede de colaborações, ou seja, os outros pesquisadores que formam, como ele, uma minirrede de pesquisa. Para o gestor público formulador de políticas, a ferramenta facilitará a prospecção da massa crítica científica já qualificada para atuar em temas que possam impactar a Educação. Por exemplo, suponha-se que um gestor do MEC tenha como prioridade a alfabetização. Ele pode descobrir quais são as pessoas que trabalham com neurociência da leitura, processos de alfabetização em sala de aula, alfabetização de deficientes auditivos, bilinguismo etc. É possível, assim, dimensionar um edital de fomento, por exemplo, para mobilizar a comunidade científica brasileira especializada a apresentar projetos que acelerem o desenvolvimento nessa área específica. O mesmo pode ser feito com vários outros temas.

A Rede CpE tem uma estrutura muito simples, no formato jurídico de uma associação privada sem fins lucrativos. A governança envolve um Conselho de Administração, um Conselho Técnico-Científico e uma Coordenação Executiva. Os associados são agrupados em duas categorias principais: pesquisadores (quase uma centena, atualmente) e parceiros (instituições). Como um de seus objetivos é estabelecer uma ponte com educadores, gestores e inovadores em geral, está em curso uma chamada para atrair afiliados amigos da rede, profissionais de qualquer categoria cujo trabalho, embora não seja de pesquisa, tenha relação com a Educação.

Dentre as atividades já realizadas nos últimos anos estão alguns eventos que reuniram cientistas e educadores, como o I Simpósio Internacional de Ciência para Educação, que ocorreu em julho de 2015 como parte do Congresso Mundial sobre o Cérebro. Além disso, outros encontros foram promovidos, salientando o I Encontro Nacional entre Cientistas e Educadores (dezembro de 2016), que pretendemos repetir anualmente. Também um I Encontro Nacional entre Cientistas e Empreendedores pela Educação está no horizonte próximo, bem como um segundo simpósio internacional (2017, Austrália). A rede também tem reunido grupos de trabalho para a elaboração de documentos temáticos<sup>9</sup> com sugestões para pesquisa e política pública. Outras metas em estudo consistem no lançamento de editais sobre temas de relevância para a Educação, com o objetivo de fomentar o trabalho de grupos de pesquisa organizados de modo multidisciplinar e estimulá-los a trabalhar no espírito do quadrante de Pasteur.

Finalmente, a meta mais ambiciosa, programada para os próximos anos, consiste na criação do Centro Nacional de Ciência para Educação, um laboratório multiusuário e multipropósito que funcionará em um prédio novo em construção no campus Fundão da UFRJ.

Há negociação em curso com o Banco Nacional do Desenvolvimento Econômico e Social (BNDES) para captar recursos que permitam concretizar as seguintes instalações:

---

9. Disponível em: <<http://cienciaparaeducacao.org/publicacoes-da-rede-cpe>>. Último acesso em setembro de 2017.



- Laboratório de neuroimagem com diferentes tecnologias, que possa realizar estudos de neuropsicologia em um ambiente propício ao estudo de crianças.
- Laboratório de multirregistro fisiológico, que permita realizar, em ambientes semirrealistas, pesquisas para registrar telemetricamente (sem fios) o eletroencefalograma de crianças (que informa sobre atividade cerebral *on-line*); estudar movimentos oculares (que indica foco atencional); e outras técnicas aplicáveis à Educação.
- Laboratório de geração de modelos matemáticos e tecnologias digitais.
- Laboratório de modelos animais que possa realizar pesquisa translacional em modo reverso, isto é, trazendo para a bancada a tentativa de comprovação mais precisa do que se consegue demonstrar em seres humanos.

Por que isso é importante? Vejamos alguns exemplos. Já existem vários trabalhos que documentam o funcionamento em redes de diferentes áreas cerebrais durante a interação entre pessoas (hiperescaneamento), o que inclui a interação entre aprendizes e instrutores. Algumas das técnicas utilizadas são

*O momento é imperdível, já que essa é uma estratégia científica ainda não "descoberta" pela maioria dos países, que poderá colocar o Brasil com mais rapidez no topo do ranking dos países com melhor Educação em todo o mundo. Não é preciso salientar o que isso trará para a população brasileira*

de difícil aplicação, como é o caso do imageamento por ressonância magnética. Outras já se tornam viáveis para experimentos interativos, como a chamada espectroscopia de infravermelho. Além disso, a monitoração da movimentação ocular revela o foco atencional de uma pessoa (uma criança em sala de aula, por exemplo) – se estudamos os movimentos oculares, sabemos aonde a criança está dirigindo a atenção, porque na maioria das vezes o foco atencional coincide com o foco visual. Outros dispositivos podem monitorar a chamada condutância eletrodérmica, medição de microgotículas de suor nos dedos que emanam de situações emocionais. Dessa maneira, é possível, por exemplo, realizar um experimento de uma situação de sala de aula em que as crianças se emocionam de algum modo com alguma atividade proposta pelo professor.

O objetivo da Rede Nacional de Ciência para Educação é prover a comunidade científica de um conjunto de tecnologias sofisticadas que possam ser utilizadas para a realização de projetos multidisciplinares. Obviamente, os laboratórios planejados são os primeiros de uma série que, em futuro próximo, possa estar à disposição da comunidade científica que deseje se inspirar na Educação para realizar suas pesquisas.

## CONCLUSÃO

Ciência para a Educação é uma janela de oportunidade que o Brasil não deveria perder, com o objetivo de tornar multiplicadora a política educacional, fomentando um amplo movimento de cientistas em conjunto com educadores e gestores para prover evidências científicas – não apenas para validar as políticas em curso, mas para inspirar aquelas que serão planejadas a cada passo do futuro.

O momento é imperdível, já que essa é uma estratégia científica ainda não “descoberta” pela maioria dos países, que poderá colocar o Brasil com mais rapidez no topo do *ranking* dos países com melhor Educação em todo o mundo. Não é preciso salientar o que isso trará para a população brasileira. Essa parte já é de domínio público.

***META-ANÁLISES:  
OS IMPACTOS  
DAS POLÍTICAS  
EDUCACIONAIS E  
SUAS REVISÕES  
SISTEMÁTICAS***



# *META-ANÁLISE DE IMPACTOS: USO DE EVIDÊNCIAS PARA PRIORIZAR INVESTIMENTOS EDUCACIONAIS*

**ANDRÉ PORTELA** é professor da Escola de Economia de São Paulo da Fundação Getúlio Vargas (FGV/EESP), coordenador do Centro de Estudos em Microeconomia Aplicada e diretor do Centro de Aprendizagem em Avaliação e Resultados para o Brasil e África Lusófona (Clear), ambos da FGV/EESP.

Este artigo trata da importância das decisões baseadas em evidências, da revisão sistemática e meta-análise de impactos e de alguns exemplos de meta-análise na área de Educação, em particular um no Brasil.

## **A IMPORTÂNCIA DAS DECISÕES BASEADAS EM EVIDÊNCIAS**

O primeiro ponto a enfatizar é que, obviamente, tomadas de decisões relacionadas a políticas públicas baseadas em evidências são importantes sob muitos aspectos, a saber:

1. Identificar avanços e fracassos.
2. Identificar novas tendências.
3. Identificar os determinantes do progresso e de novas tendências.
4. Realizar avaliações de impacto *ex ante*.
5. Realizar avaliações de impacto *ex post*.
6. Aprender com os outros (meta-análise).

Em muitos casos, vê-se, não só em Educação, mas em saúde e em outros domínios da política pública, que os gestores tomam decisões somente com base em observações pessoais ou experiências de uma circunstância específica, utilizando critérios subjetivos.

No entanto, quando se quer fazer políticas públicas abrangentes, a sistematização das informações dá uma boa ideia do que ocorre com uma política destinada a uma gama de pessoas maior.

## **REVISÃO SISTEMÁTICA E META-ANÁLISE DE IMPACTOS**

Ainda há um aspecto pouco explorado no Brasil: o uso dessas sistematizações de informações para meta-análise. O que de fato se conhece sobre determinada política? Pode-se saber muito acerca de uma intervenção, mas, por mais cuidadosa que tenha sido, será que o resultado encontrado é o que se obteria ao generalizá-la para outros contextos? Como um resultado se compara a outros?

É essa a importância da meta-análise: aprender com os outros, para analisar como uma política ou avaliação em particular está em consonância ou não com experiências de políticas semelhantes realizadas tanto no Brasil como em outros lugares.

Apresento aqui, então, duas definições rápidas:

- Revisão sistemática: revisão da literatura científica com base em métodos sistemáticos que seleciona e avalia estudos relevantes sobre uma questão claramente formulada.
- Meta-análise: método estatístico utilizado na revisão sistemática para integrar os resultados de vários estudos, de modo a aumentar o poder estatístico da pesquisa primária.

No dia a dia, muitas vezes as expressões “meta-análise” e “revisão sistemática” são empregadas como sinônimos. Há casos em que a compreensão é essa e tem sentido. Contudo, a meta-análise faz um pouco mais. Não se trata apenas de uma revisão sistemática da literatura – é a revisão que vai verificar, no caso da meta-análise de impactos, os resultados de vários estudos e

tentar, de algum modo, responder: “Usando todos esses estudos, qual é o impacto dessa política?”.

Uma forma de organizar isso é pensar em algumas etapas, conforme segue.

#### META-ANÁLISE E REVISÃO SISTEMÁTICA DE IMPACTOS – ETAPAS

1. Definir objetivamente a questão ou problema.
2. Levantar os estudos que abordam a questão.
3. Selecionar os estudos de acordo com critérios claros de inclusão e exclusão.
4. Avaliar a qualidade dos estudos e a heterogeneidade entre eles.
5. Tornar os resultados dos estudos comparáveis e calcular os impactos médios.
6. Interpretar os resultados levando em conta sua generalização.

Primeiro, define-se objetivamente a questão ou problema que queremos tratar. Por exemplo: será que reduzir o tamanho da turma tem impacto sobre o aprendizado em matemática? Sim ou não? Qual é o impacto?

Depois, realiza-se o levantamento dos estudos que abordam essa questão e, então, sua seleção de acordo com critérios claros de inclusão e exclusão. Aqui pode existir algum elemento de subjetividade ou não, mas é preciso explicitar quais são os critérios.

Em seguida, avaliam-se a qualidade desses estudos e a heterogeneidade entre eles. Posteriormente, tenta-se, de algum modo, compatibilizar os vários resultados, ou seja, torná-los homogêneos para serem comparados e, assim, viabilizar a construção de alguma métrica acerca do que se sabe. Por fim, vem a interpretação dos resultados, levando em conta o grau de generalização deles.

Como comentado em outros artigos desta publicação, é importante separar o que é causalidade do que é associação. A meta-análise preocupa-se com estudos de causalidade, porque se quer saber sobre os impactos das intervenções – se é preciso expandir uma política ou não, levá-la para outros lugares ou não.

Para o desenho das políticas educacionais, o que se deseja conhecer é quanto o aprendizado mudará caso determinada ação

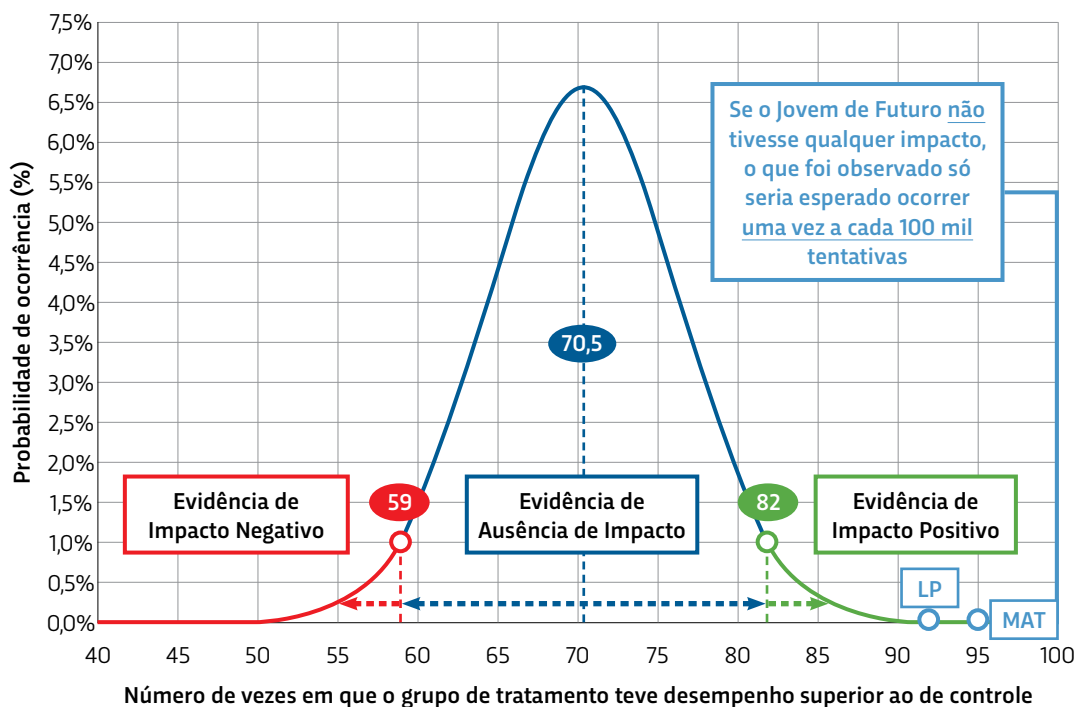


seja implementada (impacto causal da ação), e não simplesmente quanto maior será o aprendizado nas escolas que implantaram a ação quando comparadas com aquelas que não implantaram (associação). Busca-se identificar, isolar e quantificar a magnitude do impacto de diversos fatores determinantes do aprendizado. Assim, constrói-se um mapa dos impactos e não apenas um mapa de fatores associados ao aprendizado.

## EXEMPLOS

O primeiro exemplo de meta-análise é o projeto Jovem de Futuro, do Instituto Unibanco, descrito por Ricardo Paes de Barros em seu artigo “Jovem de Futuro: o impacto da gestão sobre o aprendizado dos estudantes”. A figura a seguir é um dos gráficos que ele apresenta.

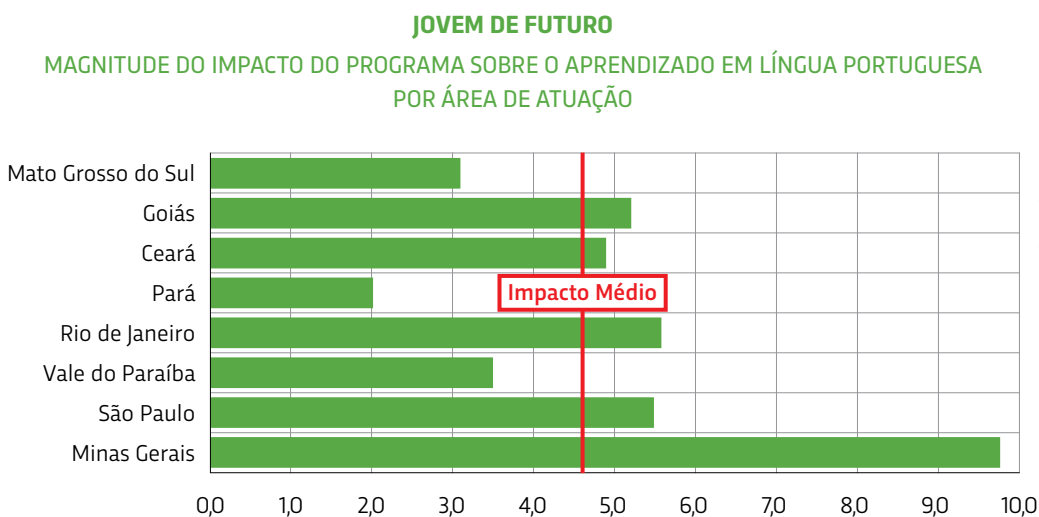
### DISTRIBUIÇÃO DAS VEZES, EM 141 EXPERIMENTOS, EM QUE O GRUPO DE TRATAMENTO TEM DESEMPENHO SUPERIOR AO DO GRUPO DE CONTROLE, CASO O PROGRAMA NÃO TENHA IMPACTO



Fonte: Ricardo Paes de Barros.

O gráfico mostra vários experimentos: 141 intervenções muito parecidas, porém com certas variações – certamente mudam o contexto, as pessoas que aplicam o programa, como foi feito o teste para saber se de fato os estudantes aprenderam matemática ou português etc. Ainda assim, são intervenções com alguma homogeneidade, todas experimentais: em um lugar, encontra-se um efeito; em outro, acha-se um diferente. Como consta do artigo de Paes de Barros, todo o programa tem um efeito, que ele modestamente chamou de 5 pontos na escala Saeb – isso é uma meta-análise: Mato Grosso teve um impacto, Goiás, outro, Minas Gerais, outro, e o impacto médio foi 5%, como se tivesse replicado os experimentos 141 vezes. Então, será possível dizer: “Depois de 141 replicações, vimos que a maioria desses experimentos teve resultado positivo, portanto achamos que a chance de o programa não ser positivo é muito baixa: de 1 a cada 100 mil”. Disso nasce uma confiança muito grande de que, de fato, o impacto é positivo.

Em grande parte das vezes, o que se tem é uma avaliação. No entanto, se queremos realmente ter confiança e saber se um programa funciona ou não, replicar os experimentos é importante, e aqui entra a meta-análise. No caso do Jovem de Futuro, essa meta-análise não foi tão complexa, por ser experimental: trata-se da mesma intervenção; só mudou o contexto.



## *O que se conhece sobre tal tipo de intervenção? É possível aproveitar as meta-análises de outros experimentos ou intervenções semelhantes*

Imagine-se agora um gestor que quer saber se vale a pena fazer determinada intervenção. A pergunta é: o que se conhece sobre tal tipo de intervenção? É possível aproveitar as meta-análises de outros experimentos ou intervenções semelhantes. O exemplo a seguir foi liderado por Paes de Barros e Mirela de Carvalho, responsável pela gestão de conhecimento do Instituto Unibanco. Trabalhamos nesse programa até 2011.

Foi minha primeira experiência em meta-análise em Educação, talvez uma das primeiras no Brasil, e o resultado intitula-se “Caminhos para melhorar o aprendizado”<sup>1</sup>. Os apoiadores foram o Instituto Ayrton Senna e o movimento Todos Pela Educação. Apresentarei o processo, as dúvidas, as decisões que tomamos ao longo do tempo e alguns resultados. É importante registrar que a meta-análise também traz dificuldades, que serão apontadas.

Em primeiro lugar, estávamos preocupados em saber quais intervenções teriam impacto sobre o aprendizado, com um foco muito claro: língua portuguesa e matemática.

Em seguida, restringimos a pesquisa a estudos que tivessem rigor, ou seja, estudos experimentais ou quasi-experimentais que contassem com grupo de comparação. A dificuldade foi que não havia muitos estudos experimentais no país – nem mesmo no mundo – para os fatores que queríamos analisar. Então, foi preciso fazer a meta-análise para todos os estudos que existiam no mundo na área de impacto sobre o aprendizado e nos quais se utilizava algum grupo de comparação.

Chegamos a 400 estudos, tanto nacionais como internacionais, levando em conta apenas os experimentais e quasi-experi-

---

1. Disponível em: <<http://www.paramelhoraroaprendizado.org.br>>. Último acesso em setembro de 2017.

mentais. A partir daí, definimos outros critérios – o principal foi que cada estudo tivesse uma amostra razoavelmente representativa, com pelo menos 1.500 alunos. Assim, ficamos com 165. Os resultados foram catalogados por temas, comparados por magnitude e avaliados segundo sua validade interna, sua validade externa, sua precisão e a robustez dessas estimativas. Validade interna era se o estudo de fato tinha o cuidado de estimar, com algum grau de confiança, o impacto da intervenção. Validade externa era quanto se podia generalizá-lo. Existe um dilema, um *trade off* entre validade interna e validade externa.

Às vezes, o experimento mais bem desenhado é muito específico para uma subpopulação de uma região, e a generalização torna-se mais difícil. Além de diferentes graus de validade interna, os estudos também apresentavam algumas diferenças. Por exemplo: qual a unidade de impacto utilizada? Para homogeneizar esse ponto, atribuímos a cada impacto um numerador e um denominador, como: aumento do aprendizado por ano de experiência do professor; aumento do aprendizado por aluno a menos em cada turma. Os impactos podiam ser por sistema educacional, professor, infraestrutura etc. Analisamos mais de 20 fatores, sempre sobre aprendizado.

Também padronizamos todos os resultados, ou seja, todos os impactos, em termos da escala do Sistema de Avaliação da Educação Básica (Saeb), alguns com mais facilidade, outros com certa generosidade nossa, sempre considerando seus desvios

*A preocupação de homogeneizar o resultado e traduzi-lo não só na escala Saeb, mas como “aceleração do aprendizado” para o público em geral e para o gestor em particular, visava dar uma noção mais clara do impacto daquele conjunto de informações*

padrão. Por exemplo: alunos tendo aula com um bom professor entre os 20% melhores aprendem por ano 29% de um desvio padrão a mais que alunos alocados a um mau professor, o que representa 9,6 pontos da escala do Saeb, dado que foi comparado em relação à aceleração do aprendizado; isso quer dizer que os alunos alocados a um bom professor aprendem 68% mais que os alocados a um mau professor.

A preocupação de homogeneizar o resultado e traduzi-lo não só na escala do Saeb, mas como “aceleração do aprendizado” para o público em geral e para o gestor em particular, visa dar uma noção mais clara do impacto daquele conjunto de informações.

Em resumo:

<b>CAMINHOS PARA MELHORAR O APRENDIZADO</b>	
<b>Escopo do estudo</b>	
<ul style="list-style-type: none"> <li>• Identificar, isolar e quantificar os determinantes do aprendizado.</li> <li>• Dentre os 400 estudos nacionais e internacionais identificados com base em critérios de qualidade e relevância (qualidade da publicação e representatividade da amostra – amostra mínima de 1.500 alunos), 165 foram selecionados para análise.</li> <li>• Os resultados desses estudos foram (i) catalogados, (ii) organizados por temas, (iii) comparados em termos de magnitude e (iv) avaliados segundo:               <ul style="list-style-type: none"> <li>• validade interna da análise;</li> <li>• validade externa da análise; e</li> <li>• precisão e robustez das estimativas.</li> </ul> </li> </ul>	
<b>Relevância substantiva, comparabilidade e estimativas de impacto</b>	
<ul style="list-style-type: none"> <li>• Toda unidade de medida de impacto tem numerador e denominador:               <ul style="list-style-type: none"> <li>• aumento do aprendizado (numerador) por ano de experiência do professor (denominador);</li> <li>• aumento do aprendizado (numerador) por número de alunos a menos em cada turma (denominador).</li> </ul> </li> <li>• Impacto do quê?               <ul style="list-style-type: none"> <li>• Vários determinantes (características dos sistemas educacionais, escolas e professores).</li> </ul> </li> <li>• Sobre o quê?               <ul style="list-style-type: none"> <li>• Sobre o aprendizado.</li> </ul> </li> <li>• Sempre que possível, o impacto sobre o aprendizado foi padronizado para a escala do Saeb.</li> </ul>	

*Padronizou-se o impacto de uma intervenção por alguma medida, mas ainda assim restou uma dificuldade: não conseguimos padronizar as unidades dos fatores determinantes [...] Esse é o limite da meta-análise*

Obviamente, com isso tudo, padronizou-se o impacto de uma intervenção por alguma medida, mas ainda assim restou uma dificuldade: não conseguimos padronizar as unidades dos fatores determinantes. Por exemplo: o impacto de reduzir o tamanho da turma em 30% é duas vezes maior do que estudar com um professor com três anos de experiência em comparação com estudar com um professor sem experiência. De outro lado, o impacto de reduzir o tamanho da turma em 10% é menor do que o de estudar com um professor com cinco anos de experiência em vez de estudar com um professor sem experiência. O que tem mais impacto sobre o aprendizado: redução do tamanho da turma ou experiência do professor? Sempre acabamos desaguando nesse que é o dilema diário dos gestores. Esse é o limite da meta-análise. A seguir, darei exemplos de alguns resultados desse estudo.

Primeiro: sabemos que escola e professor são importantes para o grau de exposição do aluno ao aprendizado, porém, obviamente, essa exposição pode ocorrer de diversas maneiras: com redução do absenteísmo do professor ou do aluno, com aumento da jornada diária, com cumprimento efetivo do ano letivo, com redução do tamanho das turmas, entre outras. O aprendizado de um aluno alocado a um professor que não faltou durante o ano letivo é tipicamente 44% maior do que o que ele teria caso fosse alocado a um professor que faltou 10 dias. Vários fatores impactam o aprendizado, mas, se eu fosse selecionar um como prioridade para a gestão, seria este: redução do absenteísmo de professor e aluno. A literatura internacional mostra que os im-

pactos são muito robustos e em média levam a essa estimativa; isso aparece em nossa meta-análise também.

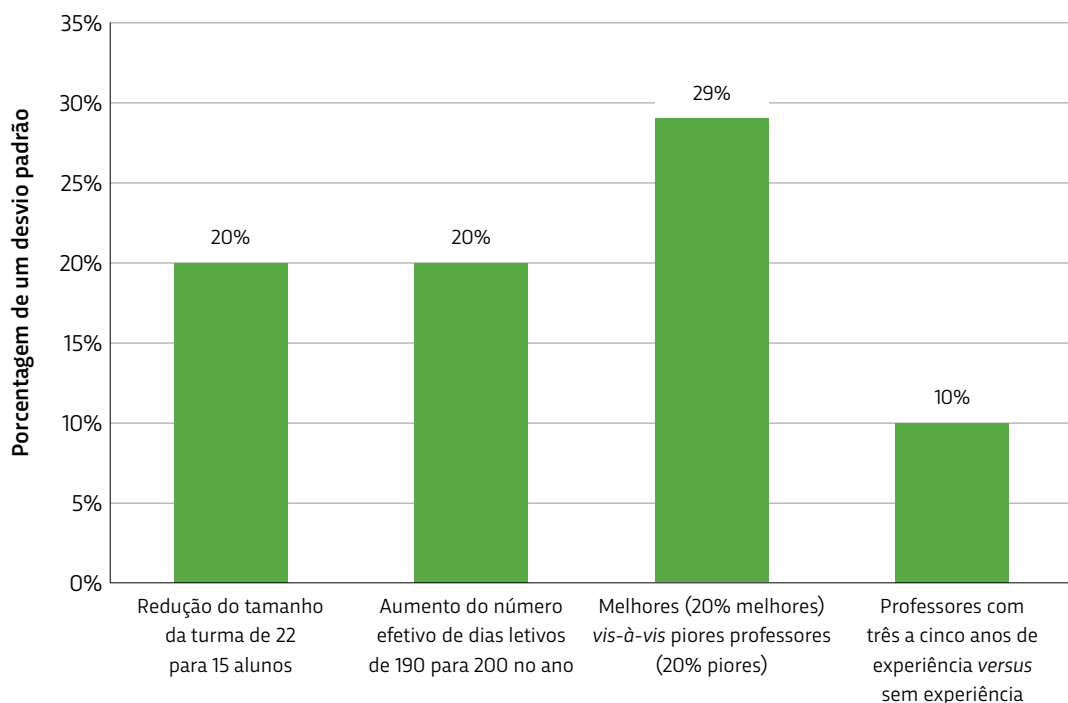
Segundo: o tamanho da turma também afeta o aprendizado – concluímos que o aprendizado de um aluno alocado a uma turma com 15 estudantes é em média 44% maior do que se ele fosse alocado a uma turma de 20 alunos.

Terceiro: aprender com um bom professor é melhor do que aprender com um professor não tão bom. Nesse caso, os alunos que estudam com os 20% melhores professores vão ter 29% de desvios padrão acima do que os que estudam com os 20% piores.

Quarto: o aumento do número de dias letivos de 190 para 200 no ano aumenta o aprendizado em 20%.

Quinto: outro aspecto bastante robusto na literatura é a experiência do professor, que de fato tem impacto significativo – em geral, com um professor com três a cinco anos de experiência *versus* um sem experiência, o aluno tem um ganho de 10% de desvios padrão de aprendizado.

### MAGNITUDE DO IMPACTO DE FATORES DETERMINANTES SELECIONADOS DO APRENDIZADO



Fonte: Ricardo Paes de Barros.

*Sabemos que vários fatores são alterados ao mesmo tempo, e a interação de tudo isso talvez não seja simplesmente a soma dos impactos que constatamos*

Esses dados sintetizam o esforço de dois anos, que envolveu cerca de 20 pessoas, organizadas em grupos e temas, para selecionar e homogeneizar os resultados, encontrar a métrica adequada para comunicar, entre outras tarefas. Dito isso, é claro que a metodologia tem limitações. Quais são elas?

## **LIMITAÇÕES**

Todo o esforço destinava-se a verificar o impacto de um fator sobre um dado de aprendizado – a variação da unidade desse fator sobre a escala do Saeb. Em todos os nossos estudos, alterava-se um fator e mantinham-se os demais constantes. Sabemos que na vida real vários fatores são alterados ao mesmo tempo, e a interação, a combinação concomitante de tudo isso talvez não seja simplesmente a soma dos impactos que constatamos aqui. Os resultados finais podem ser potencializados ou mitigados por essas interações.

Isso variou para cada tema e fator – os estudos são bastante categóricos em relação à experiência do professor e ao absentéismo, mas há outros temas sobre os quais encontramos poucos estudos. É preciso também levar em conta que todo estudo se refere ao impacto de uma mudança, de um fator, mantendo tudo o mais constante. Então, se um aluno estiver exposto a um professor com cinco anos a mais de experiência, será diferente se toda a rede tiver professores com cinco anos a mais de experiência – podem ocorrer efeitos gerais que anulam o efeito que imaginávamos para um indivíduo ou, ao contrário, magnificam o impacto observado.



*Pode-se chegar a conclusões que talvez não sejam desejadas nas políticas públicas. Por exemplo: concluímos que turmas homogêneas favorecem o aprendizado [...], mas a Educação pode ter outros objetivos, como tolerância diante da diversidade*

É claro que existem também impactos diretos e indiretos que devem ser levados em conta. Um exemplo clássico encontrado na literatura diz respeito à redução do tamanho das turmas. Alunos agrupados em turmas menores, mantendo as demais condições constantes, aprendem mais, em média, do que os que estudam em turmas maiores. Em termos de política pública, reduzir o tamanho das turmas implica, em muitos casos, aumentar o número delas, o que, por sua vez, significa contratar mais professores, e assim por diante. Porém, ao querer induzir uma política de redução do tamanho das turmas e aumento do número delas, pode ser necessário contratar professores menos experientes, de tal maneira que a política acaba tendo efeitos não almejados sobre o aprendizado.

O foco no aprendizado foi uma limitação que facilitou muito nosso trabalho de meta-análise. Pretendia-se saber o que impacta aprender matemática e língua portuguesa. No entanto, é claro que se pode chegar a conclusões que talvez não sejam desejadas nas políticas públicas. Por exemplo: concluímos que turmas homogêneas favorecem o aprendizado nessas disciplinas, mas, obviamente, a Educação pode ter vários outros objetivos para além dos conteúdos, como tolerância diante da diversidade, e, para isso, turmas heterogêneas talvez sejam melhores.

Também não há informações sobre os custos dessas intervenções ou das mudanças de fatores, portanto não fizemos análises de custo-benefício, inclusive para poder fazer avaliações de custo-efetividade entre essas políticas.

Todos os resultados estão disponíveis *on-line*. Esse é um campo a ser explorado e divulgado. Iniciativas como a Campbell Collaboration têm agora uma oportunidade muito grande no Brasil, porque há novos estudos sendo feitos no país.

## **CONSIDERAÇÕES FINAIS**

O FGV/EESP Clear (<<http://www.fgvclear.org>>) é um centro que faz parte de uma ação global de várias entidades e países que visa o aperfeiçoamento de políticas e programas por meio do fortalecimento de capacidades e sistemas de monitoramento e avaliação, uma plataforma que inclui estudos de meta-análise para políticas públicas em geral, não só educacionais. Há poucas plataformas como a Campbell Collaboration em português, e a ideia é poder transmitir esse conhecimento para todos os gestores das redes municipais e estaduais do Brasil e, se possível, da África lusófona.

# *POLÍTICA E PRÁTICAS BASEADAS EM EVIDÊNCIAS: DE ONDE VIERAM, PARA ONDE VÃO E POR QUE SÃO IMPORTANTES*

**HERB TURNER** é professor associado da Divisão de Métodos Quantitativos em Educação da Universidade da Pensilvânia. Leciona estatística e programação estatística nos cursos de graduação da mesma instituição e tem mais de 30 anos de experiência na condução de pesquisas empíricas em Educação e áreas relacionadas.

Nos Estados Unidos, a jornada da política baseada em evidência começou em 2002, com a International Cochrane Collaboration. Uma das perguntas que os educadores se faziam então era se esse movimento seria temporário – outra moda em Educação – ou duradouro e de longo prazo.

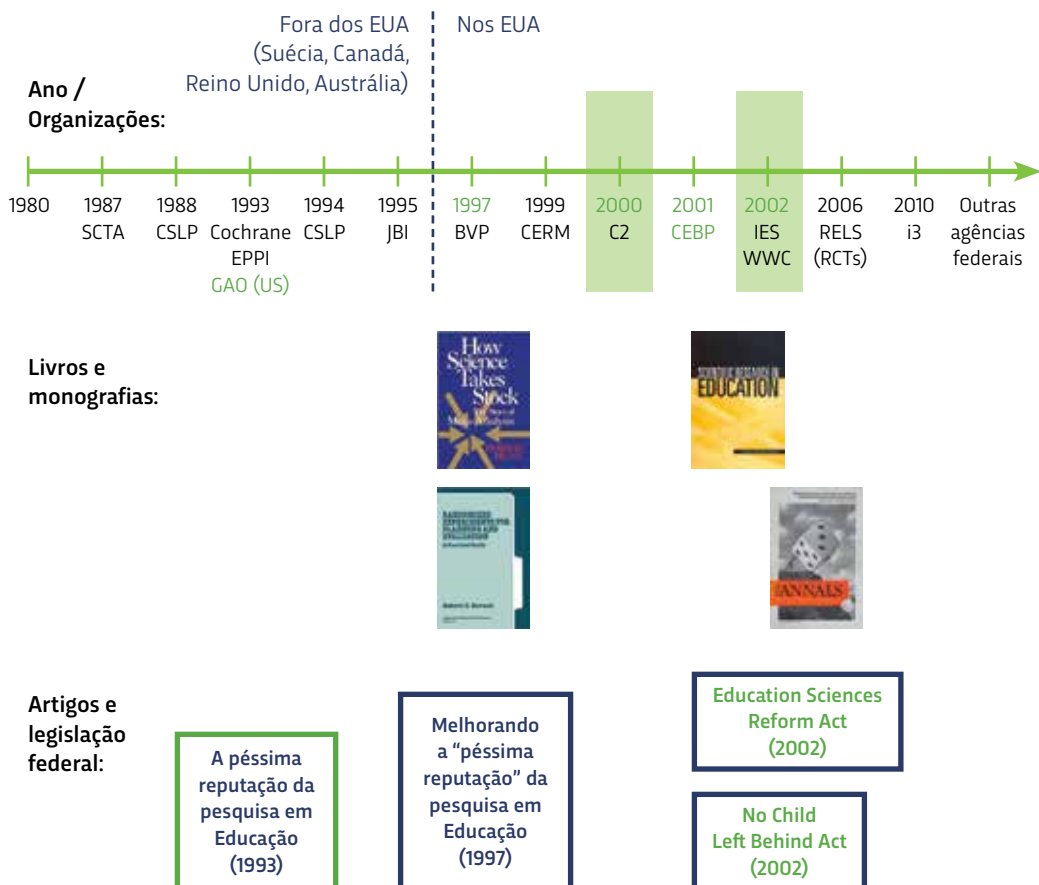
Neste artigo, abordarei, em primeiro lugar, as origens históricas das políticas e práticas baseadas em evidência (EBPP, sigla em inglês de *evidence-based policy and practices*). Em seguida, tratarei do impacto das EBPPs na pesquisa em Educação e das implicações para o Brasil e para o tipo de trabalho que educadores brasileiros realizam hoje. Finalmente, repassarei as lições aprendidas sobre as EBPP nos Estados Unidos e, mais uma vez, as implicações para o Brasil.

O movimento da política baseada em evidência cresceu nos Estados Unidos e o portfólio de evidências expandiu-se para levar em consideração a sabedoria resultante da prática.

## ORIGENS HISTÓRICAS

Começo pelas origens históricas das organizações que conduzem revisões sistemáticas. A figura abaixo apresenta uma visão longitudinal de políticas e práticas baseadas em evidência na pesquisa científica em Educação e algumas influências-chave. O movimento está em curso desde 1987. Foi então que uma organização realizou a primeira revisão sistemática, na Suécia. Assim, a política baseada em evidência, na medida em que se refere ao uso de revisões sistemáticas para sustentar decisões baseadas em evidência, tem origem europeia.

### PRINCIPAIS INFLUÊNCIAS NO CAMPO DAS ANÁLISES SISTEMÁTICAS E PESQUISA CIENTÍFICA EM EDUCAÇÃO



Fonte: elaboração do autor.

A linha do tempo também mostra que organizações muito influentes realizaram revisões sistemáticas. Em 1993, a Cochrane Collaboration fez revisões sistemáticas e meta-análises na área da medicina, e o EPPI-Centre, revisões sistemáticas de evidências no Reino Unido. Naquela época, a única organização que realizava revisões sistemáticas nos Estados Unidos era o General Accounting Office (GAO).

O terceiro aspecto da figura é que os Estados Unidos passaram década após década sem realizar trabalho algum sobre política baseada em evidência – a expressão nem sequer era empregada. Somente em 2000, quando a International Campbell Collaboration foi fundada, é que começaram a ser feitas revisões sistemáticas, e o termo “política baseada em evidência” tornou-se parte do vocabulário da Educação e do pesquisador.

Outro ponto a destacar é que uma série de influências exógenas resultou na migração da política baseada em evidência da Europa para os Estados Unidos. A partir da década de 1990, surgiram livros, monografias, artigos acadêmicos e leis para facilitar a política.

O primeiro artigo foi publicado por Carl Kaestle em 1993, com o título “The awful reputation of education research” (A péssima reputação da pesquisa em Educação, em tradução livre). Existe uma percepção de que os Estados Unidos sempre realizaram pesquisa de alta qualidade em Educação, mas o fato é que, em 1993, o que esse trabalho de fato dizia era que a área tinha péssima reputação. Em 1997, o artigo, “Improving the ‘awful reputation’ of education research” (Melhorando a “péssima reputação” da pesquisa em Educação, em tradução livre), de Gerald E. Sroufe, deu seguimento àquele. No mesmo ano, Morton Hunt escreveu *How science takes stock* (Como a ciência faz balanços, em tradução livre), primeiro livro na área das ciências sociais sobre estudos aleatorizados controlados.

Assim, livros, monografias, artigos acadêmicos e duas leis federais – o Education Sciences Reform Act e o No Child Left Behind Act, ambos de 2002 – foram as forças que lançaram, nos Estados Unidos, os fundamentos da política baseada em evidência utilizando revisões sistemáticas. O No Child Left Behind Act, que menciona a expressão “pesquisa científica” mais de

cem vezes, deu origem à divisão de pesquisa do Departamento de Educação do governo norte-americano, o Office of Education Research and Improvement, rebatizado como Institute of Education Sciences em 2002.

Por fim, a International Campbell Collaboration, fundada em 2000 como uma iniciativa internacional de cooperação de pesquisadores reunidos em torno de revisões sistemáticas, levou a revisão sistemática para o *mainstream*, com a fundação, em 2002, pelo Institute of Education Sciences, da What Works Clearinghouse (WWC). A WWC conduz revisões sistemáticas sobre vários tópicos de política educacional, em um esforço para entender o que funciona. Esse esforço em particular fomentou uma verdadeira explosão de agências federais que, na atualidade, financiam e apoiam revisões sistemáticas.

A implicação de tudo isso para o Brasil é reconhecer que o movimento da política baseada em evidência teve início na Europa, foi para a América do Norte e tem um padrão natural de migração que agora, por meio do Brasil, chegou à América do Sul. O trabalho que está sendo feito aqui – a pesquisa em política baseada em evidência e os estudos aleatorizados controlados – é parte de um esforço global para entender o que funciona em Educação para crianças e jovens em todo o mundo.

## **COMO A CIÊNCIA FAZ BALANÇOS**

As revisões sistemáticas são o modo como a ciência faz balanços. Foi esse o título do livro de Morton Hunt publicado em 1997, *How science takes stock*. É uma obra fascinante, porque foi escrita por um jornalista que queria entender o que acontece quando um campo do conhecimento gera tanta pesquisa que se torna muito difícil interpretar estudos individuais – alguns dizem que uma intervenção funciona, outros alegam que ela pode ser prejudicial, e há os que constatarem que a iniciativa não produz efeito algum.

Hunt debruça-se sobre a história da meta-análise. Gostaria que o leitor prestasse atenção a alguns tópicos explorados por ele, uma vez que podem ser relevantes para educadores e tomadores de decisão quando tentarem organizar o caos científico,

*As revisões sistemáticas são o método primário para organizar e analisar um corpo de evidências sobre determinado tópico – por exemplo, eficácia docente, prevenção de abandono escolar etc.*

que é o que começará a acontecer no Brasil, visto que o país cresce como espaço de evidência – haverá um grande número de estudos, e a questão que surge é como dar sentido a eles.

Alguns tópicos que Hunt analisa resolvem dúvidas sobre psicoterapia, esclarecem assuntos obscuros em Educação (capítulo 3 da obra) e tratam de quem deve tomar decisões quando médicos não chegam a um consenso (capítulo 4). Outros temas são dar estrutura às instáveis ciências sociais e, finalmente, “iluminar o caminho para formuladores de políticas sociais” (últimos dois capítulos). O autor também faz algumas previsões acerca do futuro da meta-análise.

Acredito que a conclusão de Hunt é que, por causa da era da informação e da proliferação da pesquisa, a meta-análise veio para ficar, porque é uma ferramenta básica para dar sentido às políticas que funcionam. Quero enfatizar, assim, que o seminário internacional “Caminhos para a qualidade da Educação pública: impactos e evidências”, realizado pelo Instituto Unibanco em setembro de 2016, confirma que a pesquisa baseada em evidência não é uma moda temporária em Educação no Brasil – é o início de uma jornada que incluirá, ao longo do tempo, a incorporação de evidência, à medida que o movimento se expande no país.

Sabemos, portanto, que a ciência realiza balanços por meio de revisões sistemáticas. Elas são o método primário para organizar e analisar um corpo de evidências sobre determinado tópico – por exemplo, eficácia docente, desenvolvimento de hábitos mentais matemáticos em crianças, prevenção de abandono escolar ou organização e governança da escola.



## ESTUDOS RIGOROSOS

Muitos artigos neste livro abordam o funcionamento dos estudos aleatorizados randomizados. No dia a dia, nas conversas entre educadores, aparece a expressão “estudo rigoroso” em frases como “É muito importante que façamos um estudo rigoroso”.

Por vários anos me perguntei qual o significado dela, porque as pessoas a utilizam com certo ar de superioridade: “Ah, é muito importante que façamos um estudo rigoroso!”. Penso nisso há bastante tempo, e o que vou apresentar são as lições que aprendi na WWC depois de revisar cerca de 12 mil estudos sobre diferentes assuntos: minha interpretação acerca de o que é um estudo rigoroso.

Em primeiro lugar, o leitor deve ter ouvido falar muito sobre grupo de controle, e, sem dúvida, um estudo rigoroso necessita de um. No entanto, refiro-me, aqui, a situações em que um pesquisador na área da Educação diz: “Quero fazer um estudo aleatorizado controlado”. Qual é a motivação subjacente e por que o grupo de controle é importante? Nos Estados Unidos, antes de 2002, o tipo predominante de design de pesquisa era o que está destacado no retângulo vermelho no quadro a seguir.

UM ESTUDO RIGOROSO TEM UM GRUPO DE CONTROLE					
Aproveitamento em matemática	Mediana do grupo de intervenção de matemática	Mediana do grupo de controle	Diferença estimada	Valor de $p$ -	Intervalo de confiança de 95%
Nota antes da intervenção	621,46	621,35	0,11	0,964	(-4,81, 5,03)
Nota depois da intervenção	647,41	646,60	0,81	0,734	(-3,82, 5,44)
<b>Ganho</b>	<b>25,95</b>	<b>25,25</b>	<b>0,70</b>	-	-

Falácia do design de prova prévia/posterior de um grupo

Fonte: elaboração do autor.

## *Analisar o grupo de controle ajuda a responder qual teria sido o resultado se o grupo de tratamento não tivesse recebido a intervenção e qual teria sido o resultado se a intervenção não ocorresse*

Esse é o resultado real de um estudo de intervenção computadorizada em matemática. Observe o que significa o grupo de intervenção de matemática no retângulo vermelho. No pré-teste, a nota da prova de matemática antes da intervenção foi 621,46, em uma escala de 200 a 800, ou seja, essa foi a nota média no pré-teste. No pós-teste, a nota média foi 647,41, um ganho de 25,95. Esse era o tipo predominante de pesquisa utilizado nos Estados Unidos – considerava-se apenas o ganho e constatava-se: “A intervenção funcionou”.

Agora, porém, analisar o grupo de controle ajuda a responder à seguinte pergunta: “Qual teria sido o resultado se o grupo de tratamento não tivesse recebido a intervenção?”. Perceba que a nota do grupo de controle também aumentou em um número de pontos semelhante, 25,25. Portanto, basicamente, se o pesquisador só tivesse usado o design pré-teste/pós-teste, teria concluído que a intervenção funcionou, ao passo que, ao utilizar o grupo de controle, a conclusão foi: “As crianças teriam obtido aproximadamente mais 25 pontos na nota mesmo que não tivessem recebido a intervenção, continuando com a aprendizagem de matemática como de costume”.

Esse é o motivo da ênfase que este artigo atribui aos estudos aleatorizados controlados. Também se constata a necessidade de um grupo de controle quando realmente se quer saber qual teria sido o resultado para as crianças se a intervenção não ocorresse. Uma vez que não se pode observar as duas situações a partir do momento em que o grupo de tratamento recebe a intervenção, usa-se um *proxy* – o *proxy* é o grupo de controle.

## *Em um estudo rigoroso o controle se dá por variáveis que poderiam explicar a diferença média entre os grupos de intervenção e de controle*

Em relação à evolução dos estudos aleatorizados controlados e à política baseada em evidência, se o Brasil seguir o padrão que se tem observado na Europa e ocorrer o que se verifica na WWC, creio que o país se encontra no estágio de prova de conceito (nos Estados Unidos, estávamos nesse ponto entre 2002 e 2006; queríamos demonstrar que é possível fazer estudos aleatorizados controlados). Então, o foco é obter o efeito no grupo de tratamento.

Porém, à medida que se realiza a prova de conceito, as pessoas passam a dizer: “Não quero saber apenas se a intervenção funciona; quero saber para quem e em quais circunstâncias”. É aí que se entra em uma área mediana em termos de fidelidade de implementação – o que acontece como parte do processo de ensino e aprendizagem no meio do estudo. Com o decorrer do tempo, surgirão novas perguntas, mas a estrutura permanecerá a mesma, e serão mencionadas expressões como “estudo aleatorizado controlado” e “estudo quase experimental”.

Outra lição extraída da WWC ao longo dos anos é que os dois tipos de estudo têm exatamente a mesma estrutura. A diferença entre eles encontra-se em como são constituídos: para um estudo randomizado controlado, o grupo de controle é formado por sorteio; para um estudo quase experimental, o que tipicamente se faz é começar com o grupo de tratamento e então encontrar um par para ele, de modo que as comparações são por aproximação. A estrutura desses dois tipos de design são essencialmente idênticas.

Então, um estudo rigoroso é aquele em que o controle se dá por variáveis que poderiam servir de fator explicativo da diferença média entre os grupos de intervenção e de controle no re-

sultado, e não pela intervenção em si. Ao receber os resultados, o pesquisador naturalmente perguntaria: “Como saber que foi o programa que causou o efeito positivo e não fatores alternativos, como a experiência dos professores? Poderia ter ocorrido de os gestores mais competentes terem sido alocados ao grupo de escolas de tratamento e isso ser a razão da diferença?”. A questão cuja resposta ele está tentando obter é: “Há fatores de confusão em como saber se o efeito ocorreu realmente no grupo de tratamento e não no outro?”.

A razão pela qual o estudo randomizado controlado e o estudo quase experimental são tão mencionados é que constituem os designs mais fortes para controlar as explicações alternativas ao grupo de controle. Alguns educadores desanimam porque têm a impressão de que só se fala desses dois tipos de pesquisa e não se reconhecem ou se marginalizam os outros. É importante ressaltar que esses dois tipos de design são levados em conta porque realmente permitem depurar a categoria de pesquisa até chegar à metodologia “que funciona”.

Assim sendo, estou tratando do que funciona. Outras questões – por exemplo, como as crianças desenvolvem hábitos de raciocínio matemático ou hábitos mentais de leitura – exigem estudo qualitativo, pesquisa no campo da neurociência, métodos e designs diferentes. No entanto, é universalmente aceito, nas áreas da medicina, da assistência médica e da Educação, que, no limite, quando se tenta responder à pergunta “Será que esta intervenção funciona?”, os dois designs preferenciais são, nessa ordem:

1. **Experimento (RCT, de *randomized control trial*):** aleatorização usada para formar grupos de intervenção e de controle.
2. **Quase experimento (QED, de *quasi experimental design*):** método não aleatório utilizado para formar grupos de intervenção e de controle.

Vale fazer uma observação adicional a respeito das abordagens experimental e quase experimental, para enfatizar o ponto de que esses dois designs são o resultado de anos de aplicação e observação sobre o que funciona.

Nos Estados Unidos, um surto de poliomielite teve efeito devastador na população até 1950. Jonas Salk tinha desenvolvido uma vacina, e o próximo passo, no início da década, era fazer o teste para poder responder à pergunta “A vacina funciona?”.

Havia, assim, uma questão específica relacionada à eficácia. Não se debateu usar outros designs além do experimental e do quase experimental, então ambos foram feitos. Obtidos os resultados, Salk e o diretor do projeto discutiram ferozmente sobre quais usar – os do experimento ou os do quase experimento. Como o experimento controlava características inobserváveis que podiam explicar por que a vacina funcionava, decidiram publicar os resultados dele. Esse fato deu origem a testes de estudos aleatorizados na medicina. Ocorre que, antes dos testes da vacina contra pólio de Salk, não havia, mesmo na medicina, uma preferência por estudos aleatorizados controlados ou quase experimentos – muitos outros designs eram empregados em pesquisas da área.

Há, portanto, uma longa história por trás da identificação e da escolha desses dois métodos para entender o que funciona. Não se trata de entrar aqui na história completa da política baseada em evidência nos Estados Unidos, que é muito interessante. Quero avançar para discorrer sobre a WWC.

## CLEARING HOUSES

Vale ressaltar que este tópico tem implicações para o Brasil. Não existiam *clearing houses* (tipo de organização que tem como focos a coleta, a organização e a divulgação de informações sobre assuntos específicos) para revisões sistemáticas acerca do que funciona em Educação até 2002, quando foi criada a WWC. Depois, surgiram *clearing houses* para outras áreas, como trabalho, criminalidade, juventude e atendimento domiciliar, além de uma base de dados sobre gravidez entre adolescentes.

No decorrer do tempo, o que aprendemos nos Estados Unidos foi que, quando se quer algo que faça sentido, que seja correto e apropriado, haverá muito debate acalorado a respeito, mas, no fim, prevalecerá uma ideia boa, consistente, com base cien-

## *É responsabilidade [do pesquisador] decidir se os resultados na WWC podem ser generalizados para a população com a qual se está trabalhando*

tífica. Foi o que ocorreu com os estudos aleatorizados controlados. Portanto, é provável que o leitor veja sentido nisso, porque o Brasil deve estar prestando atenção aos Estados Unidos, quando são somados os estudos conduzidos pelo Institute of Education e pela National Science Foundation. Havíamos produzido menos de 20 até 2002, e hoje se observa o tremendo aumento desse número. Pode-se traçar uma linha reta em termos de crescimento dos estudos aleatorizados controlados, para entender o que funciona, que é a pergunta específica ao longo dos anos.

Os tópicos destacados aqui mostram aonde ir quando se quer saber o que é importante nos Estados Unidos no nível federal. Trata-se de uma boa referência, pois eles são realmente transparentes sobre os métodos. É responsabilidade sua decidir se os resultados na WWC podem ser generalizados para a população com a qual você está trabalhando, mas a lista de tópicos é vasta: crianças e jovens com deficiência, carreiras universitárias, prevenção de evasão, organização e governança escolar.

Educadores podem ter a impressão de que há uma ênfase forte demais em estudos aleatorizados controlados e quase experimentos, a ponto de se poder imaginar os pesquisadores, na origem, sendo muito fervorosos acerca desses dois designs. O que o Brasil provavelmente terá de enfrentar à medida que se avança na direção de políticas baseadas em evidência é que não há muitas evidências em diversos assuntos, então a pergunta que se coloca é: como fazer uma revisão sistemática? Nos Estados Unidos, no nível federal, um estudo aleatorizado controlado custa a partir de US\$ 1 milhão e leva em média três anos para produzir resultados. Quem pode imaginar esperar três anos para tomar uma decisão em Educação?

O que aprendemos na WWC é que os educadores tomarão decisões com ou sem evidências. Sendo assim, o que precisa ser feito enquanto cresce a base de evidências é prover os educadores com informação, e, nesse sentido, os guias de práticas têm sido uma solução. O que o guia faz é começar com sabedoria prática, reunir um grupo de especialistas que provê recomendações possíveis de implementar e então pesquisar a literatura em busca de evidências, para depois atribuir nota às recomendações com base no nível da evidência.

Se as evidências são muitas, a confiança na recomendação é grande. Se há evidências mínimas para o cenário, não se pode ter tanta confiança, mas é possível avançar e fazer uma tentativa. Assim, o que se deve fazer é o seguinte: há especialistas que fazem as recomendações; você pesquisa na literatura e encontra tópicos lá – por exemplo: *Dropout Prevention* (Prevenção de abandono), *Reducing Behaviour Problems in the Elementary School Classroom* (Redução de problemas de comportamento nas classes de primário) e *Assisting Students Struggling with Reading* (Assistência para estudantes com problemas de leitura).

Outro aspecto que deve ser considerado no Brasil é o fato de que, não raro, mesmo nos Estados Unidos, são produzidas muitas pesquisas, mas poucas atendem aos padrões, de modo que ainda há um longo caminho a percorrer em termos do movimento de política baseada em evidência.

*O que aprendemos na WWC é que os educadores tomarão decisões com ou sem evidências. Sendo assim, o que precisa ser feito enquanto cresce a base de evidências é abastecê-los com informação, e, nesse sentido, os guias de práticas têm sido uma solução*

## COMENTÁRIOS FINAIS

A política baseada em evidência teve origem na Europa, na área da medicina, e migrou naturalmente para a América do Sul. A International Campbell Collaboration foi a força motriz dessa migração.

Não foram realizadas revisões sistemáticas por um longo período nos Estados Unidos e o fenômeno é relativamente recente lá. São de fato revisões sistemáticas em meta-análise, uma vez que se chega ao espaço de evidências; são o recurso inicial e básico para compor uma política baseada em evidência, porque esta se assentará em diversos estudos, e não em um só.

O ponto mais importante que os educadores devem perceber é que, nos Estados Unidos, a produção de revisões sistemáticas e a ênfase no estudo aleatorizado controlado e no design quase experimental se mantêm. A tendência de produção de meta-análises (revisões sistemáticas) com ênfase em avaliações de impacto experimentais e quase experimentais não é algo efêmero ou temporário; veio realmente para ficar e só faz crescer. Migrou para a América do Sul, e acredito que o mesmo acontecerá por aqui.

Outra abordagem pode ser utilizada enquanto se espera que a base de evidências cresça: o guia de práticas. Tanto o guia de práticas como a revisão sistemática se assentam na validade do estudo. Leva tempo gerar evidência e desenvolver uma massa crítica de estudos acerca de um tópico para subsidiar revisões sistemáticas. Demorou mais de uma década para que a WWC tivesse uma grande abrangência de tópicos. Desenvolver padrões de evidência para revisar o estudo também leva tempo.



# *LIÇÕES EXTRAÍDAS DE UMA REVISÃO SYSTEMÁTICA DO QUE FUNCIONA EM EDUCAÇÃO*

**EDOARDO MASSET**, economista, é diretor adjunto e chefe do escritório de Londres da 3ie, ONG internacional dedicada a promover avaliações de impacto sobre desenvolvimento mundial. Suas principais pesquisas incluem desenvolvimento rural, pobreza infantil, consumo de alimentos e avaliação do impacto de programas de desenvolvimento. O escritório de Londres é voltado para a produção e suporte de revisões sistemáticas de evidências de estudos de eficácia.

Este artigo trata de um estudo realizado pela organização não governamental (ONG) International Initiative for Impact Evaluation (3ie), uma revisão sistemática do que funciona em Educação em vários países, não em uma nação específica. Procurarei extrair algumas lições desse panorama.

A 3ie é uma ONG internacional que financia avaliações de impacto de iniciativas semelhantes ao projeto Jovem de Futuro em muitos países, bem como divulga resumos de evidências. Isso significa que realizamos e apoiamos equipes de pesquisa para sintetizarem evidências provenientes de diversas avaliações de impacto. Também fornecemos recomendações para subsidiar políticas públicas e tomadas de decisão. Algumas dessas evidências são o assunto deste artigo.

Concluimos recentemente um estudo que foi uma sistematização de iniciativas que trazem resultados em Educação em diversos países. Como ocorre em toda pesquisa rigorosa, o estudo continha perguntas, e a primeira delas era: “Quais são os efeitos de intervenções na área de Educação na aprendizagem e na frequência em escolas localizadas em países cuja população

tem renda baixa e média?”. A questão é muito ampla, e o relatório resultou em 900 páginas, o que fornece uma ideia do tempo que levamos para obter respostas a essa indagação. A pesquisa incluía questões secundárias, entre elas: “Quais são os fatores de sucesso?”. Como se sabe, em alguns casos, as intervenções são exitosas, em outros, não. Quais elementos ajudam um programa a ser bem-sucedido e quais levam ao fracasso?

Em termos de abordagem e metodologia, é isso que nos norteia na 3ie. Utilizamos uma metodologia desenvolvida originalmente no campo da medicina, por organizações como a Cochrane Collaboration e a Campbell Collaboration. Avaliamos todas as evidências disponíveis sobre determinado assunto, publicadas e não publicadas. Seleccionamos os estudos que cumprem certos requisitos para incluí-los em nosso levantamento – em geral, o fato de serem avaliações de impacto rigorosas. Analisamos, então, os resultados. Quando há diversos resultados em muitas avaliações, realizamos a chamada meta-análise, na qual resumimos os resultados em uma única medida do efeito. Quando, ao contrário, há poucos estudos, apenas fornecemos uma narrativa que contém um resumo das evidências (ver artigo de Herb Turner, p. 114).

Em nosso caso, encontramos muitos estudos, o que não surpreende, pois a pergunta era muito ampla. Começamos analisando 90 mil trabalhos que localizamos na literatura. Ao final, consideramos apenas 420 avaliações de impacto.

As avaliações precisavam ser rigorosas, assim como o projeto Jovem de Futuro. Os estudos vieram sobretudo de países do sul da Ásia, Paquistão e China; havia alguns da América Latina – Chile, Peru, Colômbia e México; e muito poucos eram da África.

*Utilizamos uma metodologia desenvolvida originalmente no campo da medicina. Avaliamos todas as evidências disponíveis e seleccionamos os estudos que cumprem certos requisitos*

Quando há diversos resultados em muitas avaliações, realizamos a chamada meta-análise, na qual resumimos os resultados em uma única medida do efeito

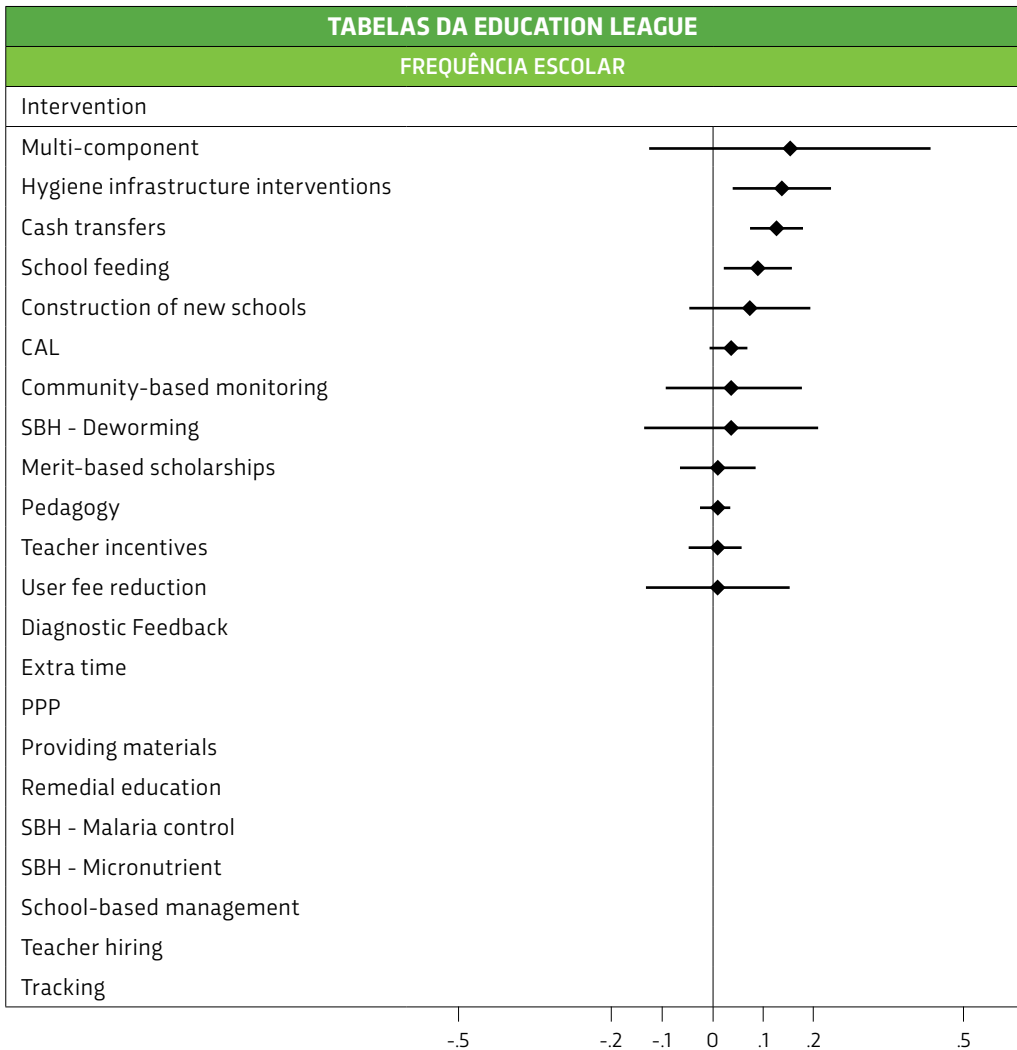
Em relação ao tipo dos estudos, 50% eram estudos aleatorizados controlados (*randomized control trials*, em inglês, ou RCTs), nos quais o local da intervenção é escolhido mediante sorteio, como ocorre no caso do Jovem de Futuro, e os demais, do tipo quasi-experimental, que não são estudos aleatorizados, mas procuram reproduzir o funcionamento de um estudo aleatorizado, isto é, contam com rigor similar.

Foi possível fazer meta-análise, pois encontramos um número grande de estudos. A meta-análise é, basicamente, uma média. Como essa média funciona?

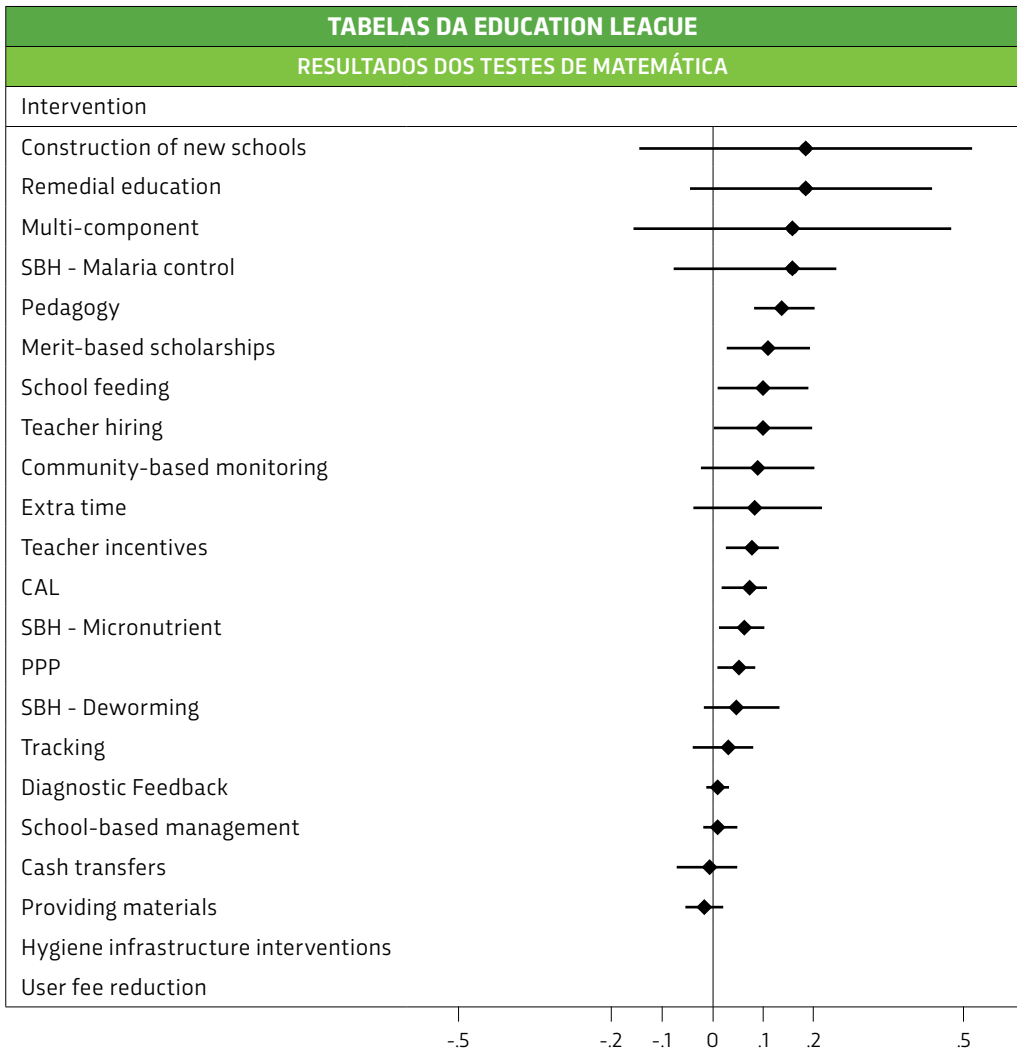
Classificamos as intervenções por grupos: intervenções em Educação, transferência condicionada de renda, alimentação na escola, aprendizagem assistida por computador, bolsas de estudo e assim por diante. Havia de dez a 30 estudos em países diferentes para cada uma dessas categorias.

A seguir, extraímos os efeitos de cada intervenção. Por exemplo, na Etiópia, a frequência escolar aumentou 5%; na China, 3%. Tomamos os resultados dos dois estudos e fizemos uma média – basicamente, uma meta-análise.

Os estatísticos gostam de mostrar os resultados no tipo de plotagem na figura a seguir, que se chama *forest plot*. A linha de base zero significa que não houve impacto. Se os resultados estão do lado direito da linha, houve impacto positivo; se estão do lado esquerdo, o impacto foi negativo. A linha pontilhada é o efeito, e a *band*, o intervalo de confiança que temos no resultado do ponto de vista estatístico.



*Os estatísticos gostam de mostrar os resultados em forest plot. A linha de base zero significa que não houve impacto. Se os resultados estão do lado direito, houve impacto positivo; se estão do lado esquerdo, o impacto foi negativo*



Ordenei as intervenções pelo tamanho do impacto: no topo, estão as de maior impacto, indicando os projetos exitosos; na base, os que falharam.

Os efeitos são expressos em termos de desvio padrão. Se tomarmos a média das médias, o impacto do conjunto das 420 intervenções que consideramos é próximo de 0,1 de desvio padrão. A pergunta que surge com frequência em apresentações como as do seminário internacional “Caminhos para a qualidade da Educação pública: impactos e evidências” é: o que isso significa?

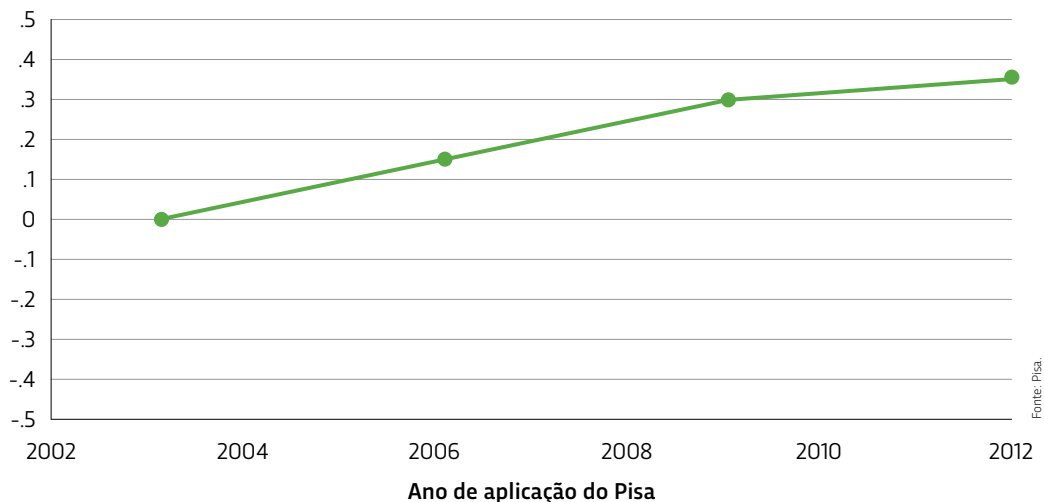
Minha visão é que não vale a pena responder a essa questão – trata-se de uma interpretação. No entanto, abordarei isso brevemente, porque o seminário trouxe vários exemplos de como se pode responder, mas o que se faz é selecionar uma mudança de referência ou um impacto de referência que se possa comparar com o impacto que se observa. Uma forma é, por exemplo, considerar os alvos. O que os formuladores de políticas ou os gerentes de projeto definem como alvo em uma intervenção? Isso seria o ideal, porém não é nada prático, porque, normalmente, formuladores de políticas não estabelecem alvos desse modo – com certeza não em Educação.

Outra maneira eficiente de interpretar é comparar aquele 0,1 com alguma referência. Um exemplo é uma tendência, como no Pisa (Programme for International Student Assessment ou, em português, Programa Internacional de Avaliação de Estudantes), algo mencionado muitas vezes nesta publicação. Selecionei os números apresentados na figura seguinte dos resultados de exames do Pisa de 2002 a 2012 em termos de desvio padrão. A mudança que verificamos no Brasil na nota desse exame no período é 0,3 de desvio padrão. A evolução do país é fruto de diversos programas implementados pelo governo e por ONGs e do avanço geral nas competências cognitivas e educativas, resultante da melhoria do padrão de vida de grandes segmentos da população. Suponhamos que uma intervenção está melhorando as notas das provas em 0,1 de desvio padrão em apenas um ano. Essa melhoria, com base na tendência de 0,3 no período de dez anos, levaria ao menos três anos para ser atingida.

*A pergunta que surge [em relação aos efeitos dos programas é]: o que isso significa? [...] O que se faz é selecionar uma mudança ou um impacto de referência que se possa comparar com o impacto que se observa*

### BENCHMARKING POR TENDÊNCIAS

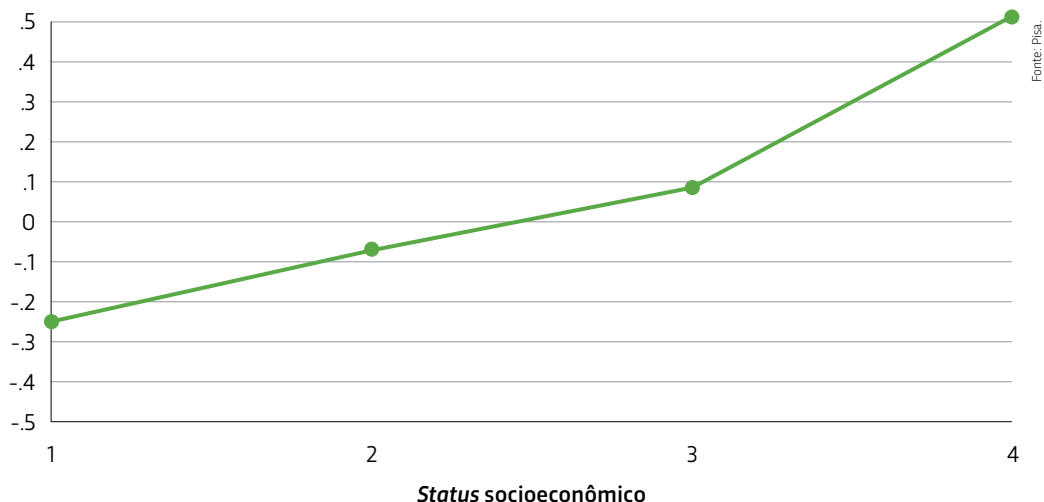
#### RESULTADOS DO BRASIL NO PISA DE MATEMÁTICA – 2002-2012



Um modo ainda melhor de fazer *benchmarking* é analisar em que medida uma intervenção está reduzindo a diferença entre segmentos da população. Abaixo, os resultados do Pisa foram desagregados por *status* socioeconômico: os da esquerda referem-se a estudantes pobres no quartil de renda inferior, e os da direita, a estudantes ricos no quartil de renda superior.

### BENCHMARKING POR HIATOS

#### RESULTADOS DO BRASIL NO PISA EM MATEMÁTICA – POR STATUS SOCIOECONÔMICO 2002-2012





Pode-se observar que há uma distância entre ricos e pobres de cerca de 0,8 de desvio padrão ( $0,5 - -0,3 = 0,8$ ). Se, por exemplo, a intervenção em estudantes em situação de desvantagem tem um impacto de 0,1, isso significa que se cobriu, talvez, cerca de 15% da distância. Tal informação dá uma ideia de quão longe se está do que seria o resultado ideal nessa intervenção.

Muitas pessoas têm dúvidas sobre o tamanho do efeito – se um desvio padrão de 0,1 é grande ou pequeno. Minha resposta é, em primeiro lugar, que é bom. Reclama-se que são gastos bilhões em Educação sem obter resultados, porém na verdade encontramos os efeitos das intervenções, o que é bastante positivo. Em relação ao tamanho do efeito, creio que essa avaliação depende dos objetivos, e isso é muito específico do contexto que se está analisando. Portanto, é preciso investir um bom tempo na interpretação dos resultados, mas não há efeito grande ou pequeno em termos absolutos – é questão de considerar o que está sendo comparado.

## **CASOS DE ÊXITO**

Como já mencionado, classificamos as intervenções em diversos grupos – transferência condicionada de renda, alimentação na escola, construção de escolas, contratação de professores, incentivo aos docentes etc. –, todas realizadas em Educação. Encontramos dois tipos de casos de êxito.

O primeiro foi a transferência condicionada de renda. Descobrimos que esse tipo de programa, que dá dinheiro à mãe para mandar a criança à escola (caso típico é o programa Oportunidades, no México), é o mais eficaz para aumentar a frequência. Não há dúvida a respeito: os efeitos são grandes e consistentes – muitos estudos encontram os mesmos efeitos.

Se observarmos o impacto na aprendizagem, normalmente medida por notas de provas nas disciplinas, a intervenção que tem o maior impacto é a mudança na metodologia em sala de aula, o que talvez não seja surpreendente. Métodos que procuram melhorar a maneira como as matérias são ensinadas em classe são, em geral, os que têm o maior impacto na aprendizagem.

Como o relatório é muito longo – 900 páginas, como citado antes –, é impossível mencionar o impacto de todas as intervenções, mas quero destacar os que encontramos no que chamamos de gestão escolar, porque esse é o assunto desta publicação.

Nessa categoria, incluímos estudos que eram um tanto diferentes: monitoramento na escola, no qual o desempenho do professor é monitorado no nível da comunidade; alguns similares à intervenção do Instituto Unibanco, que são os que realmente provocam mudanças no nível da escola; e parcerias público-privadas na gestão escolar.

Verificamos que o impacto desse tipo de intervenção é muito pequeno, quase zero. Pouquíssimas intervenções em gestão escolar conseguem produzir impacto, seja em aprendizagem, seja em frequência.

Na verdade, porém, há grande diversidade e variabilidade de resultados entre os estudos, e esse é um ponto bem importante. Encontramos algo como 12 estudos em gestão escolar, e o efeito médio de tais intervenções foi próximo de zero. Contudo, a dispersão dos resultados é imensa, portanto, alguns estudos têm um impacto muito grande no mesmo grupo.

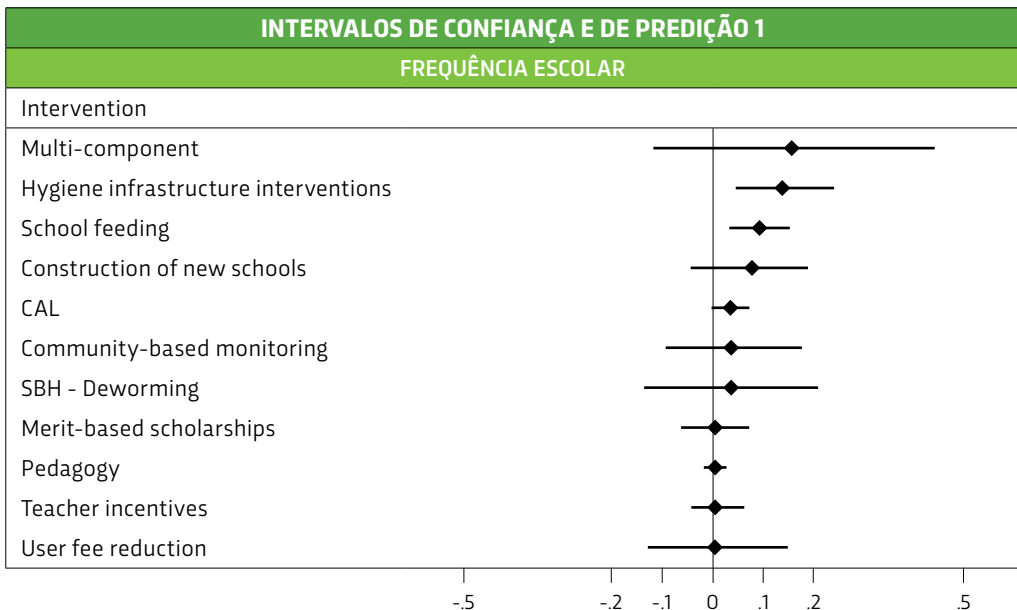
Também procuramos investigar a razão para o sucesso dos estudos. Concluímos que, normalmente, intervenções em gestão escolar adaptadas a contextos específicos e que tentam abordar diversas barreiras à Educação são as que apresentam mais êxito.

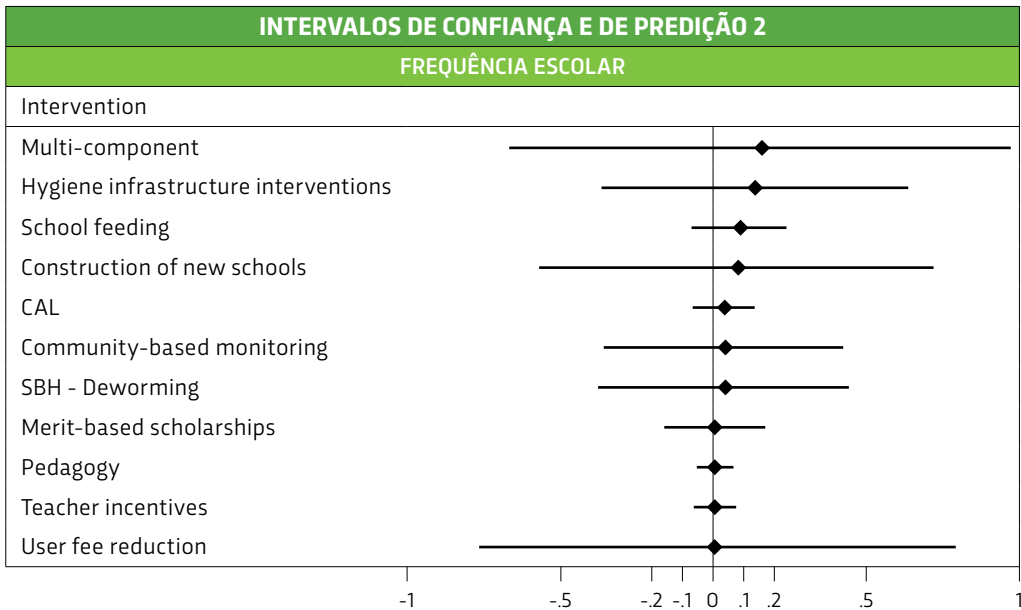
*Encontramos dois tipos de casos de êxito [...] A transferência condicionada de renda é o programa mais eficaz para aumentar a frequência, e métodos que procuram melhorar a maneira como as matérias são ensinadas são os que têm o maior impacto na aprendizagem*

## DIVERSIDADE

Há um comentário adicional em relação à diversidade e variabilidade de resultados. O que fiz foi mudar ligeiramente a meta-análise – em geral, meta-análises são realizadas da maneira que já mencionei neste artigo. Entretanto, esse tipo de meta-análise olha para a variabilidade de resultados com base em variações amostrais apenas. Quando lidamos com estudos conduzidos em diferentes países e contextos, cada um é um evento específico. Em outras palavras, há uma variabilidade contextual entre os estudos, portanto, a dispersão dos resultados é maior que a dispersão simples resultante da variabilidade de amostras.

Se consideramos a diferença contextual entre estudos ao lado da variabilidade de amostras, obtemos intervalos de confiança muito diferentes. Como se pode ver nas figuras seguir, as linhas do intervalo de confiança do quadro da próxima página são maiores que as do quadro abaixo. Quando observamos intervalos de confiança que levam em conta as diferenças de contexto da avaliação de impacto, deparamos com mais incerteza acerca de nossos resultados.





É preciso lembrar que, se uma linha de intervalo de confiança cruza a vertical zero, isso significa que, do ponto de vista estatístico, não há impacto. Nas intervenções do topo, há intervalos de confiança muito grandes (painel da direita de ambas as figuras). Se a mesma intervenção precisasse ser implementada em outro contexto, os resultados poderiam ser bem diferentes, obtendo-se muito mais incerteza acerca de seu impacto potencial. O que se conclui daí é que, em Educação, até onde se pode depreender com base na informação disponível, não há fórmula mágica. Nenhuma intervenção funciona em toda parte, todas as vezes. Não existem soluções prontas, nem sequer a transferência condicionada de renda. Uma intervenção pode ser bem-sucedida no México, na Colômbia e em São Paulo, mas a maioria dos programas educacionais não se comporta dessa maneira. Não há garantia de que o programa que funcionou no Quênia funcionará em outro país.

Parece óbvio, mas, em geral, não é assim que pensam as organizações internacionais; ao menos esse era o caso no passado. Por exemplo, a abordagem típica de organizações como o Banco Mundial ou o Fundo das Nações Unidas para a Infância (Unicef) na tomada de decisões seria a seguinte: “Este estudo funciona

bem em Bangladesh, vamos fazê-lo na Indonésia também”. Esse tipo de raciocínio ou de tomada de decisão pode ser adequado para algumas intervenções, mas não para as da área de Educação. A diferença de contexto é tamanha que não se pode replicar facilmente uma intervenção exitosa de um lugar para outro.

## **MULTICOMPONENTES**

Por fim, quero abordar intervenções que envolvem uma multiplicidade de componentes. Algumas intervenções, como a transferência condicionada de renda, tentam resolver problemas muito específicos, e o fazem empregando uma única ferramenta – nesse caso, dinheiro para as mães. Às vezes, utiliza-se um pacote de intervenções concomitantemente – por exemplo, treinamento de professores, construção de escolas, reforma de banheiros, fornecimento de livros – ou seja, muitos componentes ao mesmo tempo.

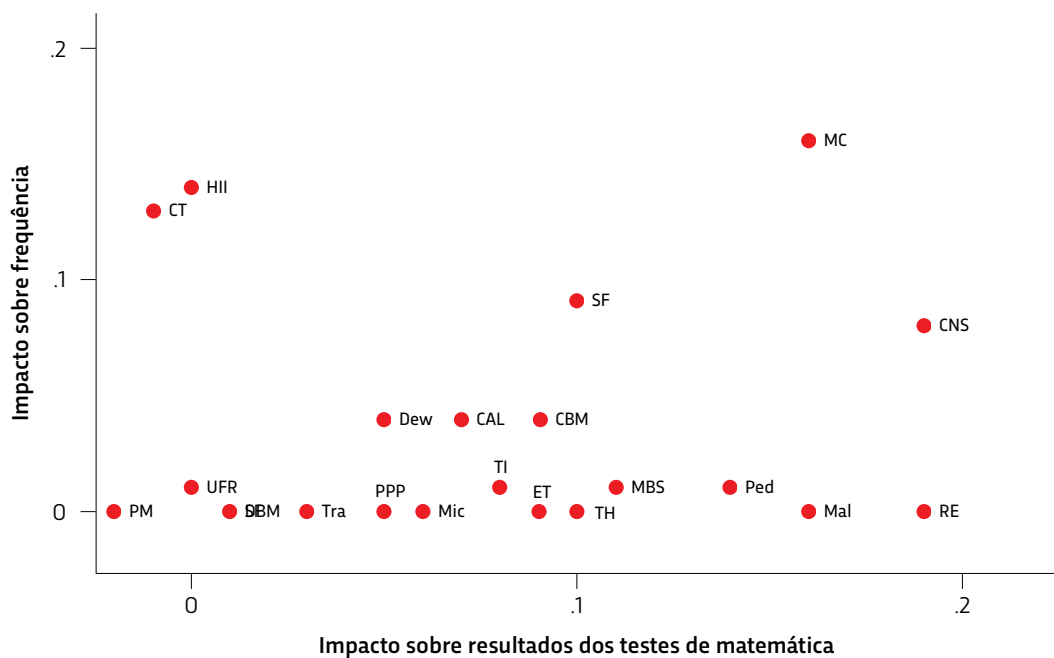
Tais intervenções foram muito bem-sucedidas. Não as mencionei antes porque os estudos são poucos, não atingindo, portanto, relevância estatística. Como a amostra é bastante pequena, seriam necessários mais estudos de intervenções compostas.

Intervenção múltipla é aquela que contém uma combinação de intervenções. Há mais problemas envolvidos na interpretação dos impactos desse tipo de intervenção. É quase impossível – isolar o impacto de cada componente.

Outras complicações vêm do fato de que tais intervenções apresentam efeitos múltiplos. Há uma camada inicial de complexidade, relacionada não apenas às intervenções múltiplas, mas também aos efeitos múltiplos. A intervenção pode afetar, por exemplo, a frequência e a aprendizagem ao mesmo tempo. Então, como levar isso em conta?

Para facilitar a compreensão, criei um quadro (ver p. 141), no qual coloquei no eixo vertical o impacto na frequência e, no eixo horizontal, o impacto na aprendizagem de matemática. Chamo isso de fronteira na política de Educação. Basicamente, as melhores intervenções serão as da parte superior do diagrama. Desse modo, do ponto de vista da formulação de políticas,

## IMPACTO SOBRE DOIS RESULTADOS A FRONTEIRA NA POLÍTICA DE EDUCAÇÃO



Fonte: elaboração do autor.

se se deseja aumentar a frequência, pode-se usar a transferência condicionada de renda; se o que se deseja é melhorar a aprendizagem, então deve-se investir em intervenções ao longo do eixo direito. Mas o que fazer se há interesse em produzir resultados em ambos os aspectos?

Pode ser que não seja esse o caso no Brasil, onde as taxas de matrícula são muito altas, mas, em muitos países da África, os governos não estão interessados apenas em subir as notas das provas, mas também em levar as crianças para a escola. Não surpreende, mais uma vez, que intervenções com componentes múltiplos – por exemplo, alimentação na escola – são as que obtêm resultados para os dois objetivos.

Como mencionado antes, é muito difícil avaliar essas intervenções. Em geral, isso é feito mediante a observação de multicomponentes como “pacotes” de intervenções. Existe alguma maneira de decompor o impacto de cada componente? Creio que em um estudo aleatorizado controlado a única maneira seria um

desenho fatorial para identificar uma variação diferente de cada intervenção, mas isso é quase impossível, porque o desenho de tal RCT seria muito complexo e exigiria testar diversas combinações dos componentes da intervenção – como quando se está tentando encontrar a receita certa. É muito difícil implantar um experimento desse modo.

Outra possibilidade é fazer isso por meio da análise de custo-eficácia, tomando resultados de estudos diferentes – de certa forma como na meta-análise –, e usando simulações. No entanto, nesse caso, o problema é que as simulações têm pouca credibilidade, porque não se baseiam em observação direta.

Ao pensar em termos do que funciona, penso se existiria algum outro modo de avaliar intervenções que procuram produzir muitos efeitos ao mesmo tempo. Recordo-me de quando estive envolvido em uma avaliação pela primeira vez, em 2004. Tratava-se de uma avaliação de impacto da reforma educacional em Gana. Era uma intervenção sistêmica, na qual quase tudo mudou – do currículo e da seleção e remuneração de professores ao idioma das aulas.

Isso seria considerado a arqueologia da avaliação de impacto. Utilizamos metodologias que consideravam, por exemplo, os orçamentos, os gastos governamentais, como o processo político ocorreu. Fizemos pesquisa qualitativa com diretores, professores e políticos. Também realizamos análise quantitativa. Chamamos o estudo de avaliação de impacto – embora não se possa utilizar a expressão pelos padrões atuais – e usamos análise regressiva para examinar o impacto – um recurso muito sofisticado –, mas sem experimento algum.

O que quero transmitir aqui é que a avaliação feita em Gana foi muito informativa, sem ser experimental. Não é o tipo de estudo que a *3ie* financeira hoje, nem que seria publicado ou citado, e talvez isso se justifique. Se eu tivesse de repeti-lo, provavelmente criaria uma avaliação quantitativa diferente – utilizaria um desenho experimental. No entanto, ao contrário de muitas avaliações realizadas na atualidade, o estudo em Gana captou uma visão mais ampla de políticas educacionais por meio de entrevistas com professores e formuladores de políticas e pela tentativa de entender a situação de modo mais amplo.

Isso é algo em que até a 3ie tem falhado na safra mais recente de avaliações de impacto. Muitas das avaliações que financiamos e apoiamos são microestudos muito rigorosos e bem desenhados, mas que não captam o cenário mais abrangente. Eu gostaria de ver no futuro mais estudos que utilizam metodologia rigorosa e conseguem isso.

## CONCLUSÕES

Em primeiro lugar, intervenções em Educação têm impacto. Esse é um ponto interessante, pois não é verdade que são gastos bilhões para nada. Em geral, intervenções em Educação em países de renda média e média-baixa têm impacto. O impacto é pequeno? Creio que não podemos julgar pelos números que encontramos em meta-análises. Trata-se de algo que precisa ser interpretado caso a caso, em relação à situação do país ou dos objetivos em uma nação em particular.

Também mencionei aqui que algumas intervenções são claramente exitosas, como a transferência condicionada de renda e as chamadas intervenções pedagógicas. Há numerosas evidências consistentes de que o primeiro tipo de projeto leva mais crianças à escola e que o segundo melhora a aprendizagem. No entanto, é preciso lembrar que se encontra grande variedade na dispersão dos resultados.

A última conclusão é que não há fórmula mágica. Não existem modelos prontos de intervenção que possam ser aplicados – do tipo “Use isto e seu sistema educacional mudará e será bem-sucedido”. Para que uma intervenção tenha êxito, deve ser ajustada ao contexto específico de cada país.

Finalmente, discorri sobre como avaliar intervenções em um nível mais sistêmico, intervenções múltiplas que buscam atingir diversos objetivos ao mesmo tempo. Essas são muito difíceis de avaliar, e creio que a maneira de fazer isso é combinar métodos rigorosos que usamos atualmente, como o estudo aleatorizado controlado, com outros tipos de análises que captam um panorama mais amplo da realidade no setor educacional.



**PRODUÇÃO E USO  
DE EVIDÊNCIAS  
EM EDUCAÇÃO  
NA VISÃO DOS  
PESQUISADORES  
BRASILEIROS**



# *PRODUÇÃO E USO DE EVIDÊNCIAS NA EDUCAÇÃO BÁSICA*

**JOSÉ FRANCISCO SOARES** é doutor em estatística pela Universidade de Wisconsin e pós-doutor em Educação pela Universidade de Michigan. É professor titular aposentado da Universidade Federal de Minas Gerais (UFMG), onde continua suas atividades de pesquisa no Grupo de Avaliação e Medidas Educacionais (Game). Na área de avaliação educacional, foi presidente do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). É membro do Conselho Nacional de Educação (CNE).

Tratar de produção e uso de evidências na Educação Básica requer, antes de mais nada, falar sobre a palavra “evidências”. Diz a lenda que Bertrand Russell<sup>1</sup> – conhecido como ateu ou agnóstico – morreu e encontrou-se com Deus. Russell explicou: “Desculpe, as evidências não eram suficientes”.

Na Educação também é preciso haver nitidez, pois é concretizada por meio de resultados tangíveis para o estudante: seu acesso e sua permanência em uma escola de Educação Básica e a aquisição dos aprendizados necessários para sua vida. A palavra “resultados” – que é muito pesada dentro da Educação – se concretiza com a palavra “aprendizados”, no plural, para deixar completamente claro que há, pelo menos, três níveis de aprendizado.

Em um primeiro nível, ficam as habilidades e os conhecimentos cognitivos, bem como a capacidade analítica. Em um se-

---

1. Bertrand Russell (1872-1970) foi um dos mais influentes filósofos britânicos do século 20. Foi escritor, ensaísta, crítico social e ativista político. Mais conhecido por seu trabalho de lógica matemática e filosofia analítica, ganhou o Prêmio Nobel de Literatura em 1950.

gundo nível, há a cultura e as atitudes – as características socioemocionais. Em um terceiro nível, a ideia de valores. O projeto educacional tem de contemplar essas dimensões, mesmo que as pedagogias sejam diferentes.

Com o objetivo de que a criança aprenda, gestores, pesquisadores e professores querem organizar, oferecer e pesquisar. Existem, porém, muitos aprendizados, e eles ocorrem em vários locais: na vida familiar, na convivência humana, no trabalho, nas instituições de pesquisa, nos movimentos sociais, nas organizações da sociedade civil, nas manifestações culturais, mas, principalmente, na escola.

O presente artigo fará esse corte – o da Lei de Diretrizes e Bases da Educação (LDB) – e se concentrará na “ESCOLA”, grafada assim mesmo, com letras maiúsculas, porque a escola tem uma responsabilidade específica e, hoje em dia, há um movimento para minimizar sua importância. Claro que é possível aprender em todos os lugares, mas, sem uma boa escola funcionando, haverá problemas com todos aqueles aprendizados.

Dessa forma, o debate educacional, formalizado ou não, é profundamente necessário. Afinal, os gestores tomarão decisões baseadas nas evidências que tiverem à disposição. Além dos gestores, os professores, como grupo, devem ser ouvidos e participar do debate educacional, porque eles serão afetados pelas escolhas.

No debate, surgirá uma primeira questão relativa – justamente, as evidências. Quais deverão ser consideradas? São basicamente três: as dos resultados, as das condições da escola, e, depois, as pedagógicas.

## **TRAJETÓRIAS E APRENDIZADOS**

Resultados, neste artigo, referem-se a “trajetórias e aprendizados”. Trajetória é saber em que ano escolar cada aluno está. Esse dado está disponível, atualmente, no Censo Escolar, que, desde 2007, é feito por aluno. Trata-se de um resultado bastante demográfico, por meio do qual é possível saber, por exemplo, em que ano escolar está hoje um estudante que cursava o primeiro ano em 2007 na cidade de Santana dos Montes (MG). Aprendi-

zados são medidos pela Prova Brasil<sup>2</sup> e pelo Exame Nacional do Ensino Médio (Enem)<sup>3</sup>, em que é possível fazer a avaliação de desempenho dos alunos. Além de usar evidências de resultados, é preciso utilizar as evidências da escola, que são as condições de funcionamento, as pessoas, o projeto pedagógico e os processos internos. Finalmente, é fundamental considerar as evidências do contexto, seja legal, social, gerencial, seja de infraestrutura, de pessoas, de opções pedagógicas.

Dentre todos, o que requer atenção especial é o contexto pedagógico.

Quando o gestor for sintetizar o que fazer, precisa decidir sobre quatro pontos: o que ensinar, como ensinar, como verificar se a criança ou o jovem aprendeu e como alinhar essas situações.

O primeiro ponto – o que ensinar – é muito importante. Para tal, é preciso que o gestor, o pesquisador, incorpore uma clareza pedagógica, porque temos, no Brasil, uma tradição pedagógica muito ambígua. No momento de definir um aprendizado cognitivo é preciso delimitar seu objetivo, requerendo uma ação passível de implementação, para prover um domínio daquele conteúdo ou conhecimento. O Brasil precisa de mais gente dedicada às opções pedagógicas, porque faltam especialistas, diferentemente do que ocorre em outros países, que publicam muitos artigos científicos sobre o tema.

Em Educação, fala-se muito que “o objetivo é compreender”, mas compreender não é um bom verbo, porque não se sabe o quanto alguém compreendeu. “Comparar” é um verbo melhor, porque dá para saber que um aluno foi capaz de comparar dois elementos ou informações explícitas sobre o mesmo assunto. Além de comparar, é importante que saiba reconhecer informa-

---

2. A Prova Brasil é uma avaliação em larga escala aplicada aos alunos de 5º e 9º anos do Ensino Fundamental, nas redes estaduais, municipais e federais, de área rural e urbana. Como resultado, a Prova Brasil fornece médias de desempenho com base na avaliação de conteúdos de língua portuguesa e matemática para cada uma das escolas participantes.

3. O Exame Nacional do Ensino Médio (Enem) foi criado em 1998 e é uma avaliação de desempenho dos estudantes de escolas públicas e particulares do Ensino Médio. Os resultados dos alunos geram dados que são analisados pelo governo, para que seja possível melhorar as políticas públicas de Educação, além de servir como um bom boletim de desempenho individual.

ções implícitas, distinguir *fato* de *opinião* e estabelecer relação entre *causa* e *consequência*.

Com o intuito de avaliar o desempenho de aprendizados cognitivos, evidências são coletadas. O que fazer com elas, porém? Antes de mais nada, é preciso traduzir essa evidência, ponderar o nível de aprendizado conquistado e, ainda, apurar desigualdades entre grupos sociais.

Na primeira coluna da tabela a seguir, há números, que são notas. Na segunda, há níveis, que oferecem um dado concreto para se trabalhar. Um nível insuficiente, por exemplo, diz muito mais do que apenas os números 152 ou 166. Mas falta ainda certa clareza, que vem descrita na coluna da direita. Ou seja, cada nível tem que estar suficientemente bem descrito para que, quando um aluno estiver no nível básico, seja possível dizer o que ele aprendeu e o que não aprendeu.

MAPA DE ITENS – MATEMÁTICA – 5º ANO		
Escala	Proficiência	Sentença descritora
152	INSUFICIENTE	Calcular, em situação, soma de números naturais de três dígitos.
166		Ordenar números naturais de oito dígitos dados em tabela de dupla entrada para identificar elementos iguais, maior ou menor.
175	BÁSICO	Estabelecer correspondência entre um sólido e sua planificação.
204		Inferir o padrão em reta numérica e identificar ponto correspondente a número ou vice-versa.
220		Resolver situação utilizando operações com números naturais.
234	ADEQUADO	Transformar medida de massa de uma para outra unidade padrão.
251		Decompor um número de quatro ou mais dígitos em ordens e expressar a decomposição em forma de soma de produtos.
277	AVANÇADO	Identificar lados perpendiculares e de mesma medida em polígono.
297		Comparar figuras em malha para determinar a relação entre seus perímetros e/ou áreas.
362		Interpretar percurso em malha para identificar ponto, dada sua posição relativa ao final do percurso.

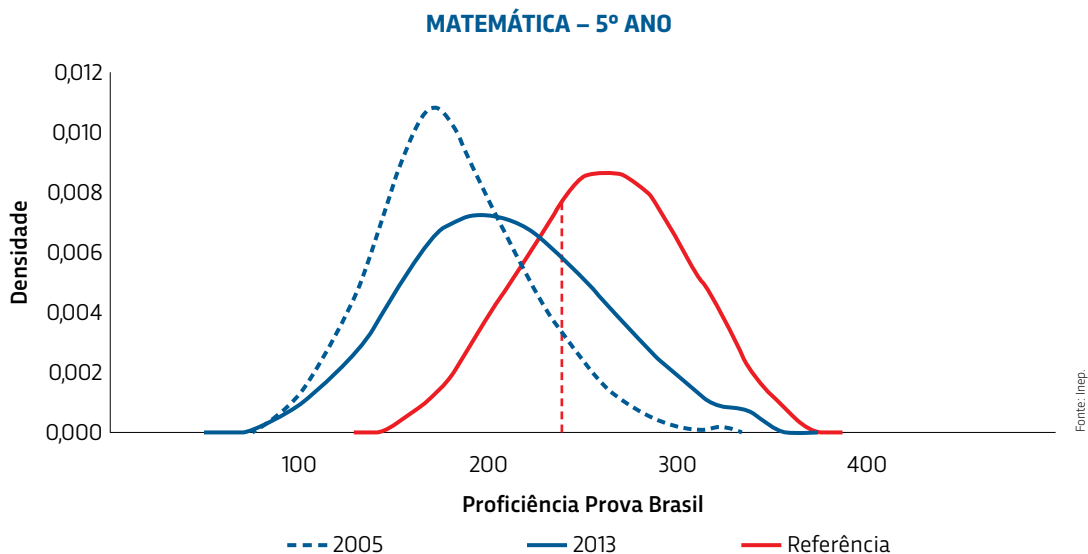
Fonte: Inep.

No Brasil, esse tipo de trabalho não está bem-feito. O país deveria ter dezenas, centenas de oficinas, em que se discutisse

como ensinar a distinção entre fato e opinião. Essa é a linguagem pedagógica que deveria ocorrer.

O segundo ponto que é preciso apresentar nas evidências é a ideia de onde estamos, bem como de onde deveríamos estar. Usemos como exemplo a leitura de alunos do 5º ano no gráfico a seguir. Observe que há três curvas: uma relativa a 2005 (pontilhada), uma relativa a 2013 e uma de referência (vermelha).

A curva mais à direita é a meta, à qual deveríamos chegar.



Por meio dessas curvas, é possível ver claramente que há problemas generalizados. Em 2013, todos os desempenhos ficaram abaixo da meta estabelecida. Quando a comparação é feita com outros países, uma coisa é ver se nosso melhor fica atrás do melhor deles; outra é se nosso pior fica atrás do pior deles. São desempenhos que exigem pedagogias diferentes.

<b>INTERVENÇÃO PEDAGÓGICA</b>	
Insuficiente	Recuperação
Básico	Expansão
Adequado	Aprofundamento
Avançado	Desafio



## LEITURA, MATEMÁTICA E DESIGUALDADE

Outro exemplo é a comparação de desempenho para leitura e matemática entre meninas autodeclaradas pretas, de nível socioeconômico baixo (grupo 1), e meninos autodeclarados brancos, de nível socioeconômico elevado (grupo 2), na próxima tabela. A primeira leitura que se extrai dela é que nem em 80 anos os grupos se aproximarão, porque se trata de 5º ano, a diferença é muito grande e, o principal problema, o ritmo de queda é muito desigual – cai muito mais lentamente para o grupo 1 do que para o grupo 2.

DESIGUALDADES – 5º ANO					
Grupo/Disciplina	2005	2007	2009	2011	2013
Grupo 1 – Leitura	1,31	1,31	1,17	0,97	0,93
Grupo 2 – Leitura	0,61	0,65	0,40	0,30	0,19
Grupo 1 – Matemática	2,13	1,80	1,54	1,48	1,36
Grupo 2 – Matemática	0,89	0,57	0,30	0,22	0,20
<b>Observação:</b> Grupo 1: meninas, pretas, NSE baixo. Grupo 2: meninos, brancos, NSE alto.					

Fonte: Inep.

Assim, para se falar em evidências, é preciso falar de evidências da situação da escola. Existe muita pesquisa, alguns dados estão disponíveis, muitos outros são necessários. O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) mediu alguns indicadores<sup>4</sup>, como a situação do docente, por exemplo.

A tabela a seguir trata de cinco indicadores: (i) nível socioeconômico, (ii) complexidade da gestão, (iii) adequação da formação, (iv) regularidade do vínculo e (v) Índice de Desenvolvimento da Educação Básica (Ideb)<sup>5</sup> médio das escolas.

4. É possível acessar esses vários dados das escolas no site do Ideb <<http://idebEscola.inep.gov.br/ideb/consulta-publica>> (acesso em novembro de 2016).

5. O Índice de Desenvolvimento da Educação Básica (Ideb) foi criado pelo Inep em 2007 e representa a iniciativa pioneira de reunir em um só indicador dois conceitos igualmente importantes para a qualidade da Educação: fluxo escolar e médias de desempenho nas avaliações. Ele agrega ao enfoque pedagógico dos resultados das avaliações em larga escala do Inep a possibilidade de resultados sintéticos, facilmente assimiláveis e que permitem traçar metas de qualidade educacional para os sistemas.

SÍNTESE DAS EVIDÊNCIAS EXISTENTES				
Nível socioeconômico	Complexidade da gestão	Adequação da formação	Regularidade do vínculo	Ideb médio das escolas
Baixo	Difícil	Baixa	Baixa	3,3
Baixo	Baixa	Alta	Alta	4,0
Médio	Difícil	Baixa	Baixa	3,9
Médio	Baixa	Alta	Alta	5,1
Alto	Difícil	Baixa	Baixa	5,3
Alto	Baixa	Alta	Alta	6,0

Fonte: Inep.

O nível socioeconômico é um limitador. Quanto maior a complexidade da gestão – escola grande, que funciona em vários turnos –, maior a dificuldade. Quanto mais adequada for a formação do docente – se possuir formação superior em licenciatura e na mesma disciplina que leciona –, melhor será o desempenho. Quanto maior a regularidade no vínculo com a escola – o professor que leciona hoje estava na mesma escola no ano anterior? –, melhor será o ensino.

Basta comparar escolas com alunos de baixo nível socioeconômico, alta complexidade de gestão, baixa adequação na formação dos docentes e baixa regularidade de vínculo dos professores com escolas em que há alto nível socioeconômico, baixa complexidade de gestão, alta adequação na formação e alta regularidade de vínculo. As primeiras obtiveram um Ideb médio de 3. As últimas, de 6. Um dado interessante é que, se os alunos excluídos de uma escola com nível socioeconômico alto migrarem para a melhor escola possível, o Ideb sobe para 4. Eles têm uma condição social ruim, mas são capazes de aprender.

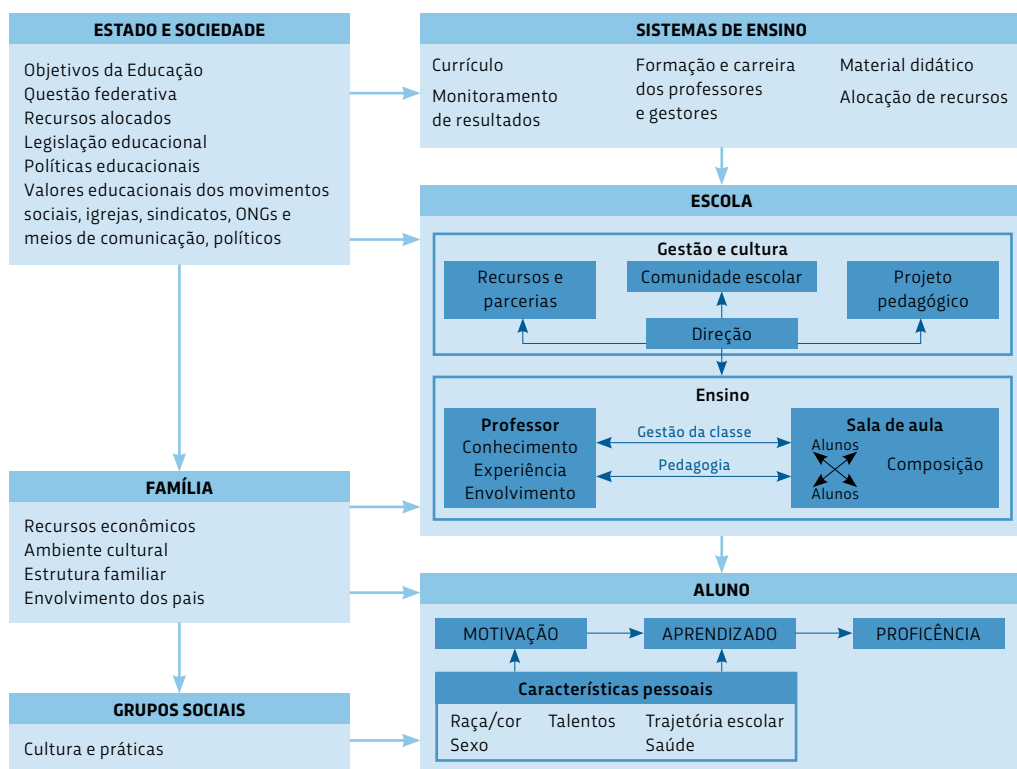
Tais dados permitiram uma síntese das evidências e um registro das associações entre elas, mas ainda falta uma explicação – falta definir o próximo passo. Seja qual for, a transformação só ocorrerá dentro da escola, dentro da sala de aula. O gestor não tem, necessariamente, que fazer essa transformação. Ele tem de cobrar. Demanda por competência gera competência.

Falta, na Educação, uma teoria pedagógica. Na economia, por exemplo, existe uma teoria econômica. O professor e pedagogo francês Clermont Gauthier<sup>6</sup> escreveu, inclusive, um livro sobre o assunto. Em português, chama-se *Por uma teoria da pedagogia* (editora Unijuí). Ele fala em “pedagogias”, porque acredita que não haja apenas uma.

## SISTEMAS DE ENSINO

O quadro a seguir aponta para a complexidade das relações entre Estado, sociedade e sistemas de ensino.

### RELAÇÕES ENTRE ESTADO, SOCIEDADE E SISTEMAS DE ENSINO



Fonte: elaboração do autor.

6. Clermont Gauthier é professor titular da Faculdade de Educação da Universidade de Laval, Quebec (Canadá). Entre suas linhas de pesquisa, estuda as mudanças nas práticas de professores em sistemas educacionais contemporâneos.

## John Hattie fez milhares de meta-análises e perguntou: “O que funciona?”. O que de fato funciona são as ações pedagógicas

Os sistemas de ensino compreendem currículo, formação e carreira de professores e gestores, material didático, alocação de recursos e monitoramento de resultados. São influenciados pelo Estado e pela sociedade, que definem os objetivos da Educação, as políticas educacionais, a legislação, e são envolvidos por movimentos sociais e religiosos, entre outros fatores. Por sua vez, os sistemas de ensino influenciam as escolas nos quesitos “gestão e cultura” e “ensino”, os quais, por sua vez, influenciam os alunos em relação a motivação, aprendizado e proficiência. A proficiência é a medida do aprendizado de um aluno, que também sofre influência da família e de suas características pessoais e sociais. O fato é que a Educação é complexa, e os grandes sucessos ocorrem a partir de intervenções no sistema, e não na escola.

Para conseguir realizar a intervenção adequada, é necessário ter um projeto pedagógico. E para um projeto pedagógico, é preciso dispor de evidências. O neozelandês John Hattie<sup>7</sup> fez milhares de meta-análises e perguntou: “O que funciona?”. O que de fato funciona são as ações pedagógicas (chamadas por ele de *visible learning*, “aprendizado visível”, em tradução livre), como se pode ver no diagrama a seguir.

O Brasil precisa evoluir. Precisamos melhorar a comunicação e não parar no resultado. Muitos dizem: “Basta o resultado”; outros argumentam: “Basta o processo”. Porém, há algo mais a considerar: “É preciso ciência na sala de aula”.

---

7. John Hattie é diretor do Instituto de Pesquisas Educacionais da Universidade de Melbourne (Austrália) desde 2011. Foi professor de Educação na Universidade de Auckland (Nova Zelândia). Realiza pesquisas sobre indicadores de desempenho educacionais, avaliação em Educação e modelos de ensino e aprendizagem.

# *O PAPEL DOS PESQUISADORES DA EDUCAÇÃO NA PRODUÇÃO E DIFUSÃO DE PESQUISA EDUCACIONAL*

**PAULA LOUZANO** é pedagoga formada pela Universidade de São Paulo (1994), com doutorado em política educacional pela Universidade de Harvard (2007) e mestrado em Educação internacional comparada pela Universidade de Stanford (1999). Foi professora da Faculdade de Educação (FE) da Universidade de São Paulo; atualmente, realiza pesquisa de pós-doutorado nessa instituição. É pesquisadora visitante na Faculdade de Educação da Universidade de Stanford e coordena o Programa de Especialização Docente (PED Brasil), baseado no modelo de formação da Universidade de Stanford (STEP).

Este artigo reflete minha participação no seminário “Caminhos para a qualidade da Educação pública: impactos e evidências”, organizado pelo Instituto Unibanco, e é fruto da reflexão sobre a seguinte pergunta norteadora: de que forma o campo da Educação/pedagogia pode contribuir para a produção e disseminação de boas pesquisas e, conseqüentemente, para o aprimoramento das políticas educacionais?

Este trabalho tem dois objetivos principais. O primeiro é discutir a importância que pesquisadores com formação em Educação têm para o desenho e a produção de pesquisas educacionais que deem suporte à tomada de decisões em políticas públicas. O segundo é propor uma discussão sobre a relevância na definição de que pesquisas fazer e sobre o rigor na escolha das perguntas que devem ser respondidas. O artigo usará como exemplo para ilustrar esses dois pontos o debate sobre elaboração de currículos na área de matemática e a sua relação com a pesquisa acadêmica.

Os pesquisadores da Educação – refiro-me àqueles oriundos da área educacional – têm uma grande responsabilidade. Maior,

inclusive, que a dos pesquisadores que vêm de outras áreas e que, eventualmente, pesquisam sobre o tema.

Os originários da área educacional deveriam ocupar lugar central no desenho, produção e disseminação de pesquisas em Educação porque eles têm o potencial de fazer perguntas relevantes e de realizar e disseminar pesquisas que estejam a serviço das demandas reais das escolas e redes de ensino, de modo a contribuir com o aprimoramento das práticas de sala de aula bem como das políticas educacionais. Isso porque há um diferencial relacionado a sua formação e ao lugar que ocupam. Primeiramente, esses pesquisadores dominam o conhecimento específico sobre o objeto de sua pesquisa – os fenômenos relacionados ao ensino e à aprendizagem. Em segundo lugar, detêm o conhecimento sobre o contexto educacional – onde estão e como funcionam nossos sistemas escolares e escolas. Seja por sua trajetória profissional, já que muitos foram professores ou profissionais da Educação Básica e, portanto, entraram em contato com a realidade da sala de aula. Seja porque têm a responsabilidade de, entre suas atribuições, formar professores e, nesse contexto, relacionam-se com o cotidiano das escolas e redes. Portanto, em tese, são esses profissionais que conhecem melhor do que ninguém quem são nossos alunos, nossos atuais e futuros professores e o contexto de nossas escolas e do nosso sistema educacional.

*É preciso fazer pesquisas com perguntas que, além de relevantes, sejam bem estruturadas. Ademais, os pesquisadores de Educação precisam se preocupar em utilizar métodos capazes de responder a essas perguntas e, independentemente da metodologia escolhida, é fundamental que sejam rigorosos*

Além disso, as faculdades de Educação têm o potencial de dialogar com as outras áreas de conhecimento dada sua natureza e constituição, já que, em seus quadros, congregam psicólogos, cientistas sociais, economistas e acadêmicos das distintas disciplinas relacionadas à Educação Básica. Faculdades de Educação também têm o potencial de ocupar lugar estratégico no contexto acadêmico ao interagir com outras faculdades por meio de suas licenciaturas.

Mas, para que esse potencial de pesquisa se transforme de fato em capacidade de influir no debate acadêmico e nos rumos da Educação, é necessário colocar em prática essa promessa e assumir essa responsabilidade que nos cabe. É preciso fazer pesquisas com perguntas que, além de relevantes, sejam bem estruturadas. Ademais, os pesquisadores de Educação precisam se preocupar em utilizar métodos capazes de responder a essas perguntas e, independentemente da metodologia escolhida, é fundamental que sejam rigorosos.

Em todo o mundo, e em nosso país também, a falta de rigor científico é apontada como uma das debilidades das pesquisas oriundas da área educacional. Além disso, a falta tanto de clareza sobre o objeto a ser pesquisado, como de um método rigoroso para responder à pergunta de pesquisa, faz com que pesquisadores de outras áreas ocupem esse espaço. Ainda que muitas pesquisas e achados careçam de relevância e de conhecimento profundo sobre o contexto do ensino, ao optar por enfatizar o rigor metodológico, estudos de outros campos ganham destaque no debate educacional no Brasil e no mundo. Mas muitas dessas pesquisas, por não contarem com a participação de pesquisadores oriundos da Educação, não conseguem atingir o núcleo pedagógico da mudança, que é a sala de aula e o ensino. Como cada vez mais o debate sobre como aprimorar a qualidade de um sistema educacional passa por esse núcleo pedagógico, na prática elas contribuem muito pouco para a busca de soluções.

Isso não quer dizer que pesquisas de outra natureza não sejam extremamente relevantes no debate da política educacional. A discussão sobre o currículo e a política curricular é um exemplo de como pesquisas de diferentes naturezas são fundamentais tanto para estabelecer onde está o problema, como para pensar alternativas para solucioná-lo.



## *Um estudo de impacto para medir o potencial de uma intervenção demanda um saber pedagógico, principalmente se utiliza alguma medida de aprendizagem*

Um estudo de impacto que vise medir o potencial de uma intervenção educacional, por mais que persiga e utilize os “golden standards”<sup>1</sup> da pesquisa (ver artigo de Herb Turner, p. 114), demanda necessariamente um saber pedagógico, principalmente se utiliza alguma medida de aprendizagem. Por exemplo, ao medir se o conhecimento matemático dos alunos aumentou após uma intervenção, os alunos são submetidos a questões matemáticas que, além de apontar a eficácia da intervenção, nos informam também sobre o desafio cognitivo e o tipo de matemática que se espera que os alunos aprendam em nossas escolas, os quais, supostamente, a intervenção está ajudando (ou não) a melhorar. São os educadores e pesquisadores oriundos das áreas de ensino, por meio de suas pesquisas, que deveriam informar sobre esta questão – assim como ajudar a pensar sobre o desenho da intervenção –, enquanto outros se preocupam em medir a eficácia. Todas essas evidências são extremamente relevantes para o debate das políticas públicas e devem andar de mãos dadas, sob a pena de terem o potencial de causar efeitos deletérios ao debate público e aos sistemas educacionais caso estejam isoladas.

Para exemplificar a importância do conhecimento estritamente pedagógico para a discussão das políticas educacionais, vamos tratar neste artigo especificamente do caso do debate da política curricular e do ensino de matemática, na perspectiva de meu trabalho internacional comparado com políticas curriculares e formação de professores.

---

1. Pesquisas que utilizam aleatorização para formar os grupos de intervenção e de controle.

## POLÍTICA CURRICULAR E ENSINO DE MATEMÁTICA

O estudo “The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003)”<sup>2</sup> de David Autor<sup>3</sup> e Brendam Price<sup>4</sup> (2013) atualizou a pesquisa anterior sobre as mudanças nos tipos de tarefas exigidas pelo mercado de trabalho americano. Essa metodologia é inovadora na medida em que, em vez de utilizar o indicador tradicional de “anos de escolaridade”, considerado uma *proxy*<sup>5</sup> para medir habilidades adquiridas na escola, utiliza a descrição das tarefas exigidas pelas ocupações ao longo do tempo. Isso permite entender não só a importância da escolarização no processo produtivo, mas também *qual* Educação é necessária para garantir o desenvolvimento econômico e social.

Ainda que o papel da Educação não seja somente o de preparar os indivíduos para o mundo do trabalho, mas também para a cidadania ativa e a participação crítica na sociedade, não podemos negligenciar essa dimensão do papel da escola. Pesquisas que a abordem devem ser analisadas e utilizadas pelos pesquisadores oriundos da Educação para produzir estudos que respondam a esses tópicos. Além disso, algumas das habilidades requeridas pelo mercado de trabalho na atualidade coincidem, em certa medida, com habilidades que apoiam a consecução do fim social da Educação; a saber, pensamento abstrato, raciocínio lógico, criatividade e trabalho coletivo. Embora não sejam suficientes, não há dúvida de que essas habilidades apoiam o desenvolvimento da pessoa e sua capacidade de pensar criticamente sobre o mundo.

---

2. “The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003)”. Disponível em: <<http://economics.mit.edu/files/9758>>. Último acesso em novembro de 2016.

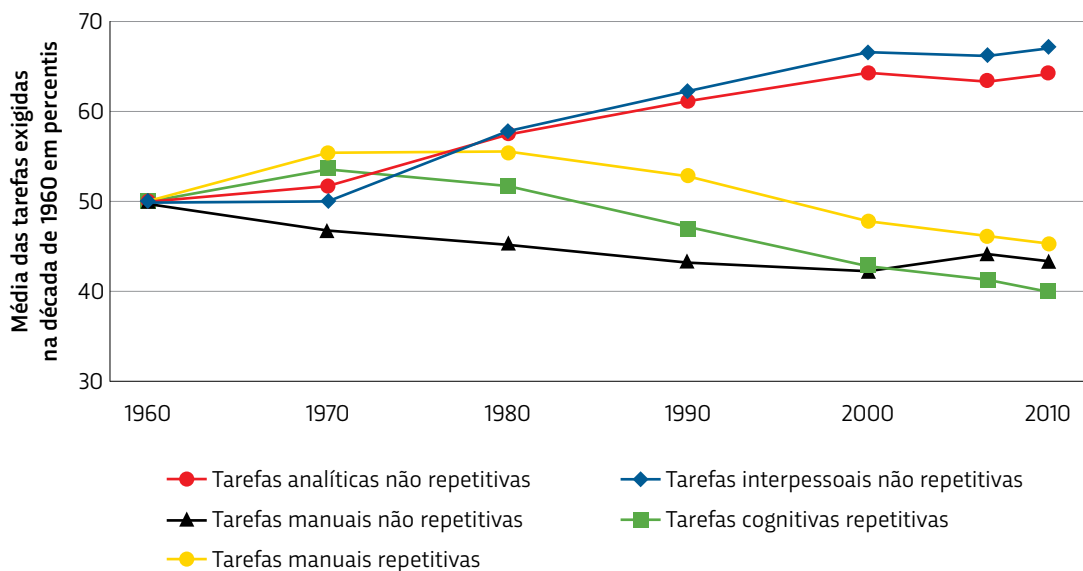
3. David H. é economista norte-americano e professor de Economia no Massachusetts Institute of Technology (MIT).

4. Brendam Price, bacharel em economia e ciência política pela Universidade de Columbia, era doutorando em economia pelo MIT em novembro de 2016, com Ph.D. previsto para junho de 2017.

5. Variável de medição indireta de objeto de difícil observação e aferição que o investigador pretende estudar.

## EVOLUÇÃO DOS TIPOS DE TAREFAS EXIGIDAS NO MERCADO DE TRABALHO

### TAREFAS DA FORÇA DE TRABALHO NA ECONOMIA DOS ESTADOS UNIDOS, 1960-2009



Fonte: elaboração da autora e Price, 2013.

O gráfico mostra como houve uma mudança significativa nos tipos de tarefa que são exigidas dos trabalhadores ao longo do tempo. Em primeiro lugar, observa-se um grande crescimento da demanda por tarefas interpessoais e analíticas que exigem raciocínio abstrato, resolução de problemas, criatividade e construção constante de conhecimento, além de trabalho em equipe. Chama atenção também o declínio acentuado de tarefas cognitivas repetitivas, cada vez menos demandadas pelo mercado de trabalho porque são passíveis de serem automatizadas e substituídas pelo computador – por exemplo, tarefas relacionadas à ocupação de bancário. Se na década de 1970 elas ocupavam a segunda posição nas tarefas mais demandadas pela sociedade, hoje estão em último lugar entre os cinco tipos de tarefas analisadas<sup>6</sup>.

6. As tarefas analisadas neste estudo foram: tarefas interpessoais não repetitivas, tarefas analíticas não repetitivas, tarefas manuais repetitivas, tarefas manuais não repetitivas e tarefas cognitivas repetitivas.

Esse estudo traz uma informação fundamental para quem está à frente dos sistemas educacionais, bem como para quem está pesquisando a Educação e pensando sobre currículos em nossas escolas. Ele suscita questões importantes sobre o tipo de currículo que as escolas devem oferecer com o objetivo de desenvolver habilidades interpessoais não repetitivas e habilidades analíticas.

No que diz respeito ao conhecimento matemático, índice (*Equations and Inequalities*, OCDE, 2016) de exposição de alunos a atividades escolares relacionadas aos principais conceitos de álgebra aferidos a partir dos dados coletados no Pisa (Programme for International Student Assessment ou, em português, Programa Internacional de Avaliação de Estudantes), mostra que o Brasil está em último lugar, atrás, inclusive, da Colômbia e da Tailândia. Coreia do Sul ocupa a primeira posição; Cingapura, a segunda; e os Estados Unidos, a décima-terceira. Mais do que a última colocação do Brasil, o que chama muito a atenção é que mesmo as nossas “melhores” escolas – muitas delas privadas – não alcançam a média da parte inferior da Austrália, penúltima colocada. O índice não trata de aprendizagem matemática, mas de acesso a um tipo de conhecimento por meio de um currículo desafiador.

Os dados apresentados acima poderiam resultar de pesquisas capitaneadas por profissionais de outras áreas que não a Educação. No entanto, há duas perguntas que emergem desses conjuntos de dados que são, sim, específicas de quem entende de ensino e aprendizagem. São elas: (i) qual é o currículo que predomina em nossas escolas para gerar o resultado do Pisa e quão distante o país está de ter outro perfil de aluno e (ii) que currículo seria capaz de ajudar os alunos a desenvolver habilidades analíticas e interpessoais não só requeridas pelo mercado de trabalho, mas também essenciais para o exercício da cidadania?

*Há perguntas que emergem desses conjuntos de dados que são específicas de quem entende de ensino e aprendizagem*

## FRAÇÕES: TAREFA DE BAIXA DEMANDA COGNITIVA E TAREFA ANALÍTICA

Com o objetivo de exemplificar o tipo de pesquisa que poderia ajudar a responder a essas questões, serão apresentados, a seguir, exemplos relacionados ao ensino de frações. Compararemos duas tarefas sobre o mesmo tópico matemático e analisaremos em que medida cada uma delas é capaz de desenvolver as habilidades analíticas referidas neste artigo.

O primeiro exemplo se refere a atividades tradicionalmente utilizadas em nossas escolas para o ensino de fração.

### TAREFAS DE BAIXA DEMANDA COGNITIVA

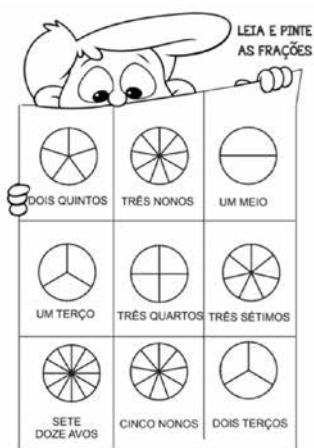


Figura 1

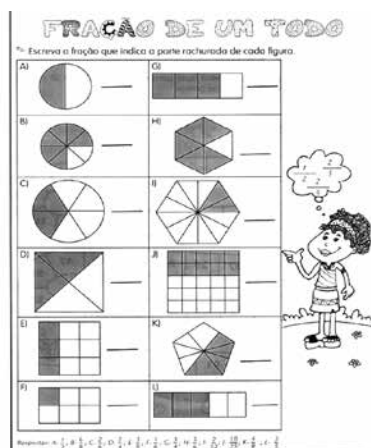


Figura 2

Ambas atividades se limitam a tarefas cognitivas repetitivas, como se pode ver nas figuras 1 e 2. O que se entende por matemática nesses dois modelos – que infelizmente predominam em nosso país – já não representa o tipo de matemática que deve ser ensinada. Aqui, o aluno não precisa compreender o conceito de fração, uma vez que ele pode perceber como preencher as lacunas da tarefa (“em cima vai o número de pedaços pintados e em baixo o número total de pedaços”) tanto na figura 1 como

na figura 2. As tarefas são procedimentais e repetitivas. O aluno não precisa pensar a relação entre os algorismos em uma fração, pois a tarefa não exige isso. O currículo não foi pensado para tal. Pede-se, exclusivamente, que ele relacione fração a “pedaços de uma pizza”. Não por acaso, se alguém pedir a alunos de 4º e 5º ano para mostrar o que é uma fração, muitos deles farão desenhos muito similares a esses dois exemplos (“uma pizza”) e darão explicações de como estabelecem o numerador e o denominador da fração a partir desse desenho.

### QUE TIPO DE CURRÍCULO PRECISAMOS TER EM NOSSAS ESCOLAS PARA DESENVOLVER HABILIDADES ANALÍTICAS NOS ALUNOS?

Se  =  $2/3$

Desenhe:

$1/3$



1

$4/3$

2



Por outro lado, a tarefa da figura acima é uma atividade que requer que a criança reflita sobre o conceito de fração, o que é fundamental para o avanço do aprendizado de conceitos como razão e proporção em matemática. Na pesquisa americana que serviu como base para informar o debate do *Common Core* (base nacional curricular americana), o conteúdo relacionado à fração foi apontado como um ponto fraco do ensino de matemática no país, responsável por dificuldades futuras dos alunos com a disciplina. A forma como esse tópico foi incluído nos novos padrões curriculares levou isso em consideração<sup>7</sup>.

Portanto, não se trata de incluir atividades como essa de maneira aleatória – como um desafio no final da prova para que os

7. Schmidt, William H.; Houang, Richard T. “Curricular Coherence and the Common Core State Standards for Mathematics”. *Educational Researcher* 41 (8), 2012, p. 294-308. Disponível em: <<https://www2.ed.gov/about/bdscomm/list/mathpanel/report/learning-processes.pdf>>. Último acesso em dezembro de 2016.

alunos ganhem pontos extras e ou em olimpíadas de matemática. Trata-se de um desenho curricular que objetiva desenvolver compreensões muito mais profundas sobre o tópico “frações”, e o material didático e as atividades derivadas desse currículo estão a serviço desse tipo de aprendizado.

O grau de demanda cognitiva da tarefa na figura anterior é claramente superior ao das tarefas anteriores, já que não há repetição no padrão de resolução de cada um dos exercícios e observa-se uma complexidade crescente. Esse aumento intencional de complexidade (“scaffolding”) apoia o aluno no caminho de uma abstração cada vez maior e faz parte do desenho do currículo.

Esse exemplo foi extraído do currículo de matemática para a Educação Básica desenvolvido pela Universidade de Chicago por meio de um projeto da faculdade de matemática dessa instituição (University of Chicago School Mathematics Project, em inglês)<sup>8</sup>. Esse projeto de pesquisa deu origem ao currículo e aos materiais didáticos *Everyday Mathematics*, voltados para alunos da pré-escola ao 6º ano. O que o torna um exemplo interessante é o fato de ele ser fruto de uma política pública com vistas a colocar a pesquisa no campo educacional a serviço de um ensino mais rigoroso da disciplina por meio do desenvolvimento de materiais didáticos para a sala de aula. Ou seja, foi um esforço deliberado do governo norte-americano por meio de editais de financiamento da National Science Foundation (NSF, Fundação Nacional de Ciência, em português) – comparável ao nosso Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – para melhorar o ensino de matemática nas escolas daquele país.

Desde 1998, o *Everyday Mathematics*<sup>9</sup> vem sendo aprimorado tanto por seu uso pelas escolas como pelas pesquisas realizadas sobre sua implementação e eficácia. Os recursos de sua venda são usados para financiar o trabalho de pesquisa, produção de conhecimento e formação de professores.

---

8. Disponível em: <<http://stemeducation.uchicago.edu/>>. Último acesso em dezembro de 2016.

9. É possível visitar o site da Universidade de Chicago específico desse projeto: <<http://everydaymath.uchicago.edu>>. Último acesso em novembro de 2016. Nele, há link para toda a pesquisa que embasou a elaboração do currículo: <<http://everydaymath.uchicago.edu/about/em-history/curriculum-research/>>. Último acesso em novembro de 2016.

## *Precisamos de pesquisas sobre como desenhar currículos, materiais didáticos e atividades que levem ao desenvolvimento de habilidades cognitivas analíticas*

O currículo foi desenvolvido seguindo uma metodologia de estudos comparados, a partir de livros internacionais traduzidos de países como Japão e Rússia. O material didático foi criado após pesquisas entre o conhecimento matemático e o contexto. Para isso, foi necessário envolver professores, educadores e pesquisadores da área de Educação.

Ainda que haja vasta pesquisa nacional sobre o ensino de matemática no país e que vários matemáticos e educadores matemáticos escrevam livros didáticos para editoras que servem ao Programa Nacional do Livro Didático (PNLD), é tênue essa sinergia entre concepção e visão de ensino, produção de material didático e de tarefas específicas, orientações curriculares nacionais e locais, além de pesquisa acadêmica – e, portanto, incapaz de gerar mudança na escala necessária. Além disso, precisamos de pesquisas sobre como desenhar currículos, materiais didáticos e atividades que levem ao desenvolvimento de habilidades cognitivas analíticas e de alta ordem baseadas nas mais recentes pesquisas no campo educacional, e esses materiais precisam chegar às escolas. Precisamos também de pesquisas sobre quais são as práticas pedagógicas e estratégias necessárias para ensinar esse currículo, e quais os processos de interação que devem acontecer na sala de aula e como colocá-los em prática. As pesquisas devem estudar as interações entre o aluno e o conhecimento e, portanto, focar na elaboração, seleção e implementação de atividades e tarefas escolares desafiadoras, e também a interação entre os alunos e desses com o professor, trazendo elementos concretos sobre como nossos professores podem estimular não só o pensamento analítico e crítico, mas também a convivência e o trabalho coletivo.



Referências sobre como materializar determinadas concepções e visões em sala de aula são raras em nosso país, e cabe aos pesquisadores da área de Educação encarar esse desafio tão necessário para a melhoria da nossa Educação. Estamos atrasados nessa tarefa.

Especificamente, o que diferencia o exemplo norte-americano apresentado é que ele busca materializar o diálogo com o campo da prática desenvolvendo instrumentos e ferramentas de trabalho para o professor da Educação Básica utilizar em seu dia a dia. Nesse modelo, a pesquisa acadêmica está claramente a serviço das necessidades de professores e alunos, fazendo com que estes tenham acesso muito rápido às últimas descobertas do mundo científico. Além disso, essa é a maneira como os pesquisadores se conectam com a realidade das escolas, ajudando no delineamento de boas perguntas de pesquisa. Outro exemplo a ser destacado é o projeto *Youcubed*, liderado pela educadora matemática Jo Boaler da Faculdade de Educação de Stanford. O objetivo do *Youcubed* é “inspirar, educar e empoderar professores de matemática transformando as mais recentes pesquisas sobre ensino de matemática em materiais práticos e acessíveis”<sup>10</sup>. O projeto oferece atividades de matemática referenciadas no *Common Core*, vídeos de aulas com boas práticas de ensino, cursos para alunos e professores de matemática e textos sobre pesquisas recentes relacionados a neurociência e ensino de matemática.

Um conceito importante da pesquisa de Boaler, materializado nos exemplos de currículo, atividades e cursos de formação que ela oferece para alunos de Educação Básica e professores, é o do “low floor, high ceiling” (“chão baixo, teto alto”, ou “entrada fácil, nível alto”)<sup>11</sup>, que busca ampliar o acesso a um currículo rigoroso de matemática visando a equidade e excelência. O “low floor, high ceiling” postula a importância de desenvolver atividades matemáticas que sejam fáceis de começar (entrada fácil), mas que tenham por objetivo chegar a um desafio maior (nível

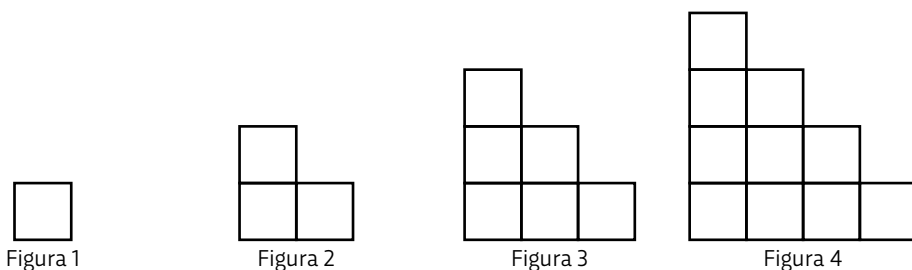
---

10. Disponível em: <<https://www.youcubed.org/ourmission/>>. Último acesso em setembro de 2016.

11. Disponível em: <<https://www.youcubed.org/grade/low-floor-high-ceiling/>>. Último acesso em setembro de 2016.

## “LOW FLOOR, HIGH CEILING”

### EXEMPLO DE ATIVIDADE COM ENTRADA FÁCIL E NÍVEL ALTO



Fonte: Youcubed/Stanford University School of Education.

1. Como você vê a sequência crescendo?
2. Quantos quadrados há na figura 10? Explique como você chegou a essa conclusão.
3. Quantos quadrados há na figura 55? Explique como você chegou a essa conclusão.
4. Você consegue usar 190 quadrados para fazer uma estrutura em formato de estrela? Justifique sua resposta matematicamente.
5. Como você consegue saber o número de quadrados de uma figura qualquer?

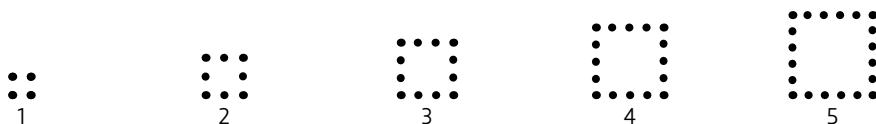
alto). Dessa forma, todos os alunos se sentem confortáveis para participar da tarefa e, portanto, para aprender.

A figura acima mostra um exemplo desse tipo de tarefa. Nela o aluno deve derivar a fórmula. No entanto, em vez de perguntar somente “Qual a fórmula para saber quantos quadrados haverá na figura 55”, a tarefa começa levando o aluno a descrever como ele enxerga o crescimento da sequência. A primeira pergunta da atividade – “Como você vê a sequência crescendo?” – pode ser respondida por qualquer aluno, independentemente do grau de conhecimento do conteúdo e do seu conforto com a matemática. Apesar da entrada fácil (“low floor”), ao final da tarefa o aluno haverá realizado uma atividade desafiadora.

Já na figura a seguir, a atividade pede que o aluno “determine 4 fórmulas diferentes (e equivalentes) para o total de bolinhas de uma figura genérica  $n$  dessa sequência”. A dificuldade está posta já no ponto de partida e poderá excluir, logo no início, um grande número de alunos que não são “bons de matemática” e não conhecem a fórmula – e que, muito provavelmente, podem nem tentar.

## “HIGH FLOOR”

### EXEMPLO DE ATIVIDADE QUE APRESENTA DIFICULDADE NO PONTO DE PARTIDA



“Cada figura da sequência está indicada por um número. Determine quatro fórmulas diferentes (e equivalentes) para o total de bolinhas de uma figura genérica  $n$  dessa sequência”.

Fonte: Material de Apoio ao Currículo.  
Caderno do Aluno, matemática, 3ºE-SP.

As tarefas das figuras anteriores requerem a mobilização do mesmo conteúdo matemático, mas a maneira como são propostas faz a diferença. Esse é um exemplo de saber sistematizado pelos pesquisadores da Educação (Boaler, 2002 e 1998)<sup>12</sup> e extremamente poderoso para pensar políticas educacionais que afetem o cerne da qualidade e da equidade educacional: a sala de aula.

Hoje há um grande debate sobre o tipo de conhecimento de que os gestores educacionais precisam para desenhar suas políticas e gerenciar seus sistemas. A avaliação de impacto, tema deste seminário, é necessária, mas não suficiente para essa tarefa.

Não podemos prescindir – e ousar dizer que precisamos mais do que nunca – do conhecimento específico sobre o objeto da pesquisa educacional: a pedagogia, o ensino e a aprendizagem. Essa demanda tem de ser preenchida pelos pesquisadores da Educação. As faculdades de Educação brasileiras precisam abrir linhas de pesquisa sobre desenho de currículos e material didático, e sobre como ensinar atividades como as apresentadas aqui para todos os alunos, mas especialmente para aqueles que necessitam mais da escola, os mais pobres e excluídos de nosso país. Além disso, elas precisam internacionalizar-se, referenciando-se no que há de mais avançado no mundo em termos de pesquisa educacional. Não podemos olhar somente para dentro,

12. Boaler, J. (2002). “Exploring the Nature of Mathematical Activity: Using theory, research and ‘working hypotheses’ to broaden conceptions of mathematics knowing. Invited Paper”. *Educational Studies in Mathematics*, 51(1-2), p. 3-21.

Boaler, J. (1998). “Open and Closed Mathematics: Student experiences and understandings”. *Journal for Research in Mathematics Education*, v. 29, n. 1, p. 41-62.

sob pena de sermos atropelados ou nos tornarmos irrelevantes. O desafio é imenso. Podemos aprender com os erros e acertos de outros países que já começaram a perseguir esse objetivo.

Minha contribuição tem sido buscar associar minha pesquisa na área de Educação internacional comparada a demandas práticas de nosso sistema educacional. Meu trabalho na Universidade de Stanford inclui o desenvolvimento de um curso de formação continuada para professores na área de ensino de matemática a ser implementado em sete instituições de Ensino Superior e duas secretarias de Educação estaduais no Brasil<sup>13</sup>. O Programa de Especialização Docente (PED Brasil) surgiu da pesquisa internacional que sistematiza a base de conhecimento necessária para a docência (Louzano e Moriconi, 2015)<sup>14</sup>, bem como do exemplo de sucesso do curso de formação de professores da Universidade de Stanford (STEP)<sup>15</sup> em criar um modelo institucional que, além de incorporar essa base de conhecimentos para o ensino, estabelece uma relação concreta e observável entre teoria e prática por meio do desenho de seu programa e de seu currículo. Além disso, conta com a pesquisa sobre ensino de matemática dessa instituição, já descrita neste artigo.

O resultado desse projeto de pesquisa de formação de professores de matemática foi materializado em módulos e atividades específicas para formação docente que perfazem 480 horas/aula (um “currículo”) e em cursos de formação de formadores das instituições envolvidas no projeto para que seus educadores possam implementar esse currículo. A pesquisa, princípios e premissas que embasaram o PED Brasil, bem como sua implementação, serão alvo tanto de novas pesquisas como de uma avaliação para que sigamos aprendendo com o conhecimento teórico e prático produzido.

---

13. Disponível em: <<https://lemanncenter.stanford.edu/events/programa-de-especializa%C3%A7%C3%A3o-docente-ped-brasil>>. Último acesso em setembro de 2016.

14. Louzano, Paula; Moriconi, Gabriela Miranda. “Visión de la Docencia y Características de los Sistemas de Formación Docente”. *Cadernos Cenpec* | Nova série, [S.l.], v. 4, n. 2, jun. 2015. ISSN 2237-9983. Disponível em: <<http://cadernos.cenpec.org.br/cadernos/index.php/cadernos/article/view/287/285>>. Acesso em 20 de março de 2017. doi:<http://dx.doi.org/10.18676/cadernoscenpec.v4i2.287>.

15. Disponível em: <<https://ed.stanford.edu/step>>. Último acesso em setembro de 2016.

# *EM BUSCA DA MELHORIA DA QUALIDADE DO CLIMA ESCOLAR*

**TELMA VINHA** é pedagoga, doutora em Educação na área de psicologia, desenvolvimento humano e Educação pela Faculdade de Educação da Universidade Estadual de Campinas (Unicamp) e professora do departamento de Psicologia Educacional da mesma instituição. Realiza pesquisas nas áreas de relações interpessoais e desenvolvimento moral, sendo membro do Laboratório de Psicologia Genética da Unicamp (LPG) e do Grupo de Estudos e Pesquisa em Educação Moral (Gepem) da Universidade Estadual de São Paulo (Unesp) e da Unicamp.

Vários olhares são necessários quando se fala de pesquisa em Educação, principalmente envolvendo avaliação escolar. É preciso ir além das pesquisas solitárias ou daquelas em que predomina uma única área; nestas, o diálogo é insuficiente e repetem-se os mesmos processos. Por isso, o Grupo de Estudos e Pesquisa em Educação Moral (Gepem), o qual coordeno atualmente, reuniu uma grande equipe de pesquisadores<sup>1</sup> de oito instituições

---

1. Pesquisadores: dra. Alessandra de Morais Shimizu (FCL/Unesp/Marília), vice-coordenadora; dra. Ana Maria Falcão de Aragão (FE/Unicamp/Campinas); dra. Adriana de Melo Ramos (PD/FE/Unicamp/Campinas); dra. Carmen Lúcia Dias (Unoeste/Presidente Prudente); dra. Juliana Ap. Matias Zechi (Dep. Psic./UFMS/Paranaíba); dra. Luciene Regina Paulino Tognetta (FCL/Unesp/Araraquara); dra. Maria Suzana de S. Menin (FCL/Unesp/Presidente Prudente); dra. Marialva Rossi Tavares (Fundação Carlos Chagas); dra. Patrícia Unger Raphael Bataglia (FCL/Unesp/Marília); dr. Raul Aragão Martins (Ibilce/Unesp/São José do Rio Preto); dra. Roberta Gurgel Azzi (FE/Unicamp/Campinas); dra. Carolina de A. Escher Marques (Gepem/Unesp/Unicamp); Lívia Maria Ferreira da Silva (Gepem – Unesp/Unicamp); Cesar Augusto Amaral Nunes (PD/FE/Unicamp/Campinas). Pós-graduandos: Adriano Moro (Fundação Carlos Chagas – doutorando – FE – Unicamp); Mariana Guimarães Wrege (doutoranda – FE – Unicamp); Terezinha Ferreira da Silva Colombo (doutoranda – Unesp/Marília); Fabrício Costa de Oliveira (doutorando – Unesp/Marília); Flávia Maria de Campos Vivaldi (doutoranda – FE – Unicamp); Mariana Tavares Almeida Oliveira (doutoranda – FE – Unicamp); Thais Leite Bozza (doutoranda – FE – Unicamp). Esse estudo teve o auxílio da Fapesp, Fundação Lemann e Itaú BBA.

distintas e vários pós-graduandos para realizar uma pesquisa sobre clima escolar, foco principal deste artigo.

O clima escolar é compreendido como o conjunto de percepções e expectativas compartilhadas pelos integrantes da comunidade escolar, decorrente das experiências vividas nesse contexto com relação aos seguintes fatores inter-relacionados: normas, objetivos, valores, relações humanas, organização e estruturas física, pedagógica e administrativa que estão presentes na instituição educativa. O clima corresponde às percepções dos docentes, discentes, equipe gestora, funcionários e famílias, a partir de um contexto real comum e, portanto, constitui-se de avaliações subjetivas. Refere-se à atmosfera psicossocial de uma escola, sendo que cada uma possui o seu clima próprio. Ele influencia na dinâmica escolar e, por sua vez, é influenciado por ela; desse modo, interfere na qualidade de vida e na qualidade do processo de ensino e de aprendizagem. A ideia de avaliação do clima escolar é mais ampla do que a de desempenho. O desempenho em língua portuguesa e matemática é importante, mas uma escola é muito mais do que isso; envolve toda uma complexidade. A análise do clima escolar permite, justamente, olhar para essa complexidade.

Nossos objetivos com a pesquisa realizada foram (i) construir, testar e validar instrumentos, adaptados à realidade das escolas brasileiras, para avaliar o clima escolar com alunos a partir do 7º ano, docentes e gestores; e (ii), a partir do diagnóstico do clima escolar, elaborar e desenvolver em escolas públicas um programa de formação de professores e transformações na escola com ações diferenciadas e complementares visando à melhoria da qualidade do clima escolar nas dimensões que envolvem as relações e os conflitos interpessoais, favorecendo a convivência ética. Neste artigo, será focado o primeiro objetivo.

Apesar de inúmeras pesquisas mostrarem que o clima escolar influencia as aprendizagens e o desempenho (rendimento escolar, motivação etc.), vários estudos também têm indicado que essa influência se estende para as relações sociais. Quando o clima escolar é negativo, pode representar um fator de risco da qualidade de vida na escola. A síntese das pesquisas mostra que o clima é determinante e está diretamente associado ao senti-

mento de autoconfiança e bem-estar geral para realizar o trabalho, à motivação, às aprendizagens e ao rendimento, à atitude em face da utilidade dos estudos, à identificação com a escola, ao desenvolvimento emocional e social dos alunos e professores, aos comportamentos, ao estresse, às interações com os pares e com os outros atores da instituição.

As pesquisas são importantes para explicar o “efeito da instituição”, pois muitas vezes, em um bairro com o nível socioeconômico muito baixo e com problemas de violência, se o clima escolar é positivo na visão dos alunos, encontra-se um melhor desempenho e maior qualidade nas relações.

Uma escola com um clima positivo apresenta bons relacionamentos interpessoais, qualidade no processo de aprendizagem, espaços de participação e decisão, proximidade com pais e comunidade, boa comunicação, senso de justiça (as regras são necessárias e obedecidas e as sanções são justas) e um ambiente de cuidado e confiança, estimulante e apoiador, centrado no aluno; nela os indivíduos sentem-se seguros, amparados, engajados, pertencentes à instituição e respeitosamente desafiados.

## **UM INSTRUMENTO BRASILEIRO**

Considerando que não há instrumentos validados e adaptados à realidade das escolas brasileiras que avaliem o clima, em nossa pesquisa elaboramos e testamos instrumentos para avaliar o clima escolar entre alunos, professores e gestores. Para isso, fizemos uma revisão da literatura na área e analisamos 62 pesquisas de diversos países. Depois dessa análise, percebemos que nenhuma delas dava conta da complexa realidade das escolas do Brasil. Então, foram construídos novos instrumentos capazes de captar a diversidade escolar brasileira, além das relações interpessoais (as pesquisas de clima escolar, no Brasil, geralmente focam relações interpessoais e conflitos).

Assim, elaboramos uma matriz formada por oito dimensões inter-relacionadas, consideradas constituintes do clima escolar:



1. As relações com o ensino e a aprendizagem.
2. As relações sociais e os conflitos na escola.
3. As regras, as sanções e a segurança na escola.
4. As situações de intimidação entre alunos.
5. A família, a escola e a comunidade.
6. A infraestrutura e a rede física da escola.
7. As relações com o trabalho.
8. A gestão e a participação.

A matriz foi utilizada como base para a construção de instrumentos para alunos a partir do 7º ano, docentes do Ensino Fundamental II e gestores da Educação Básica. São oito dimensões destinadas aos professores e gestores (dimensões 1 a 8) e seis relativas aos alunos (1 a 6). Os questionários foram compostos por seções contendo itens que medem e avaliam as percepções dos sujeitos a respeito dessas dimensões no formato da escala Likert com quatro pontos: não concordo, concordo pouco, concordo, concordo muito (quadro a seguir). Há itens exclusivos, ou seja, só para os alunos, só para os professores e só para os gestores, mas há também itens de relação, que são comuns aos três grupos. Os instrumentos foram submetidos à apreciação de especialistas na área que não conheciam os itens, os quais foram aperfeiçoados após tais análises. O tempo usado para responder foi em torno de 25 minutos.

*Analizamos 62 pesquisas de diversos países. Depois dessa análise, percebemos que nenhuma delas dava conta da complexa realidade das escolas do Brasil. Então, foram construídos novos instrumentos capazes de captar a diversidade escolar brasileira*

OS INSTRUMENTOS – EXEMPLO DE QUESTIONÁRIO				
MARQUE O QUANTO VOCÊ CONCORDA COM CADA UMA DAS AFIRMAÇÕES A SEGUIR	NÃO CONCORDO	CONCORDO POUCO	CONCORDO	CONCORDO MUITO
1. O que aprendo na escola é útil para a minha vida.				
2. Se eu pudesse, eu mudaria de escola.				
3. A quantidade de lição de casa é bem distribuída entre as matérias.				
4. Os professores propõem em sala de aula atividades em grupos que promovem a troca de ideias e a cooperação.				

Fonte: Ceppem/Unesp.

Esse é um exemplo de itens da dimensão “ensino e aprendizagem” (a questão 4 é válida para professor, gestor e aluno).

Inicialmente, aplicamos os questionários impressos e *on-line* a cerca de 1.500 respondentes e, a partir dos resultados, realizamos análises estatísticas e qualitativas, aperfeiçoando os instrumentos. Após esse aperfeiçoamento, com parcerias de diversas instituições, aplicamos a pesquisa novamente a 11.516 sujeitos de 62 escolas de diferentes estados brasileiros. Por volta de 9 mil eram alunos (do 7º ano do Ensino Fundamental ao 3º ano do Ensino Médio); e os demais, professores e gestores. Houve a participação de escolas privadas, mas em sua maioria eram escolas públicas.

Com os resultados dessa ampla coleta de dados, realizamos análises psicométricas em busca das evidências de validação para esses instrumentos. Também foram utilizadas outras análises qualitativas, como a comparação dos resultados encontrados com o que foi observado nas escolas pelos pesquisadores e com a percepção dos professores e gestores da instituição. É de fundamental importância refletir sobre o que fazer com os dados encontrados para que contribuam com a melhoria do clima nas unidades escolares.

*Uma boa avaliação voltada à melhoria das escolas tem de devolver-lhes os dados, contribuindo para que melhorem; ao mesmo tempo, é preciso fomentar o diálogo e o processo reflexivo*

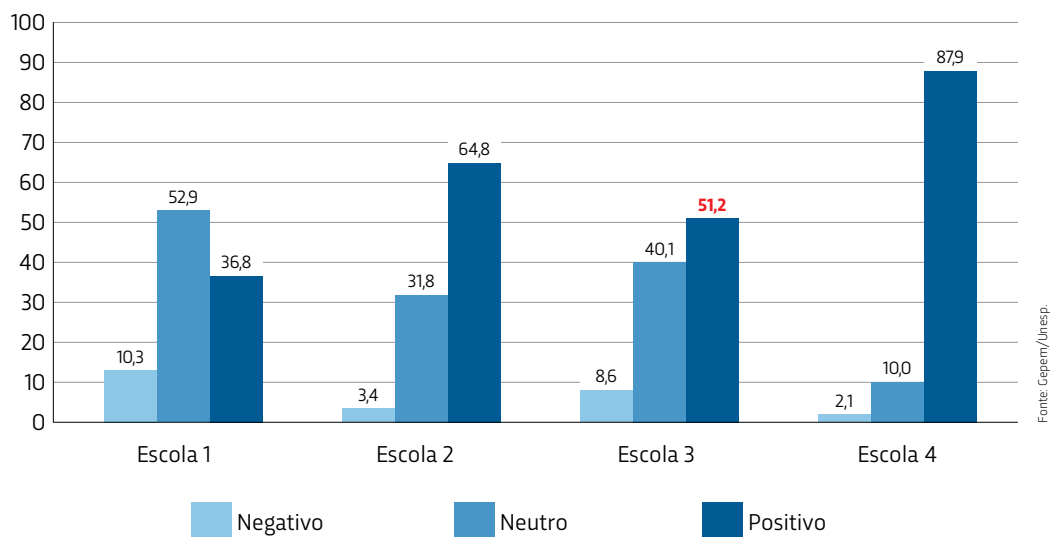
## **DEVOLVER PARA A ESCOLA**

Uma das pesquisas do Grupo de Estudos e Pesquisa em Educação Moral (Gepem) da Universidade Estadual de São Paulo (Unesp) envolve, atualmente, dez escolas públicas<sup>2</sup>. Mais importante do que simplesmente coletar os dados e avaliar é devolver esses resultados para as escolas, para que discutam e analisem os aspectos que estão bem e outros que pedem melhorias, na perspectiva dos vários grupos. Geralmente, uma escola toma decisões sobre como intervir e sobre o que precisa ser melhorado a partir da percepção dos gestores e professores, poucas vezes envolvendo diretamente os alunos (não apenas por representação). O Gepem defende que uma boa avaliação voltada à melhoria das escolas tem de devolver os dados para a instituição, contribuindo, assim, para que melhore; ao mesmo tempo, é preciso fomentar o diálogo e o processo reflexivo.

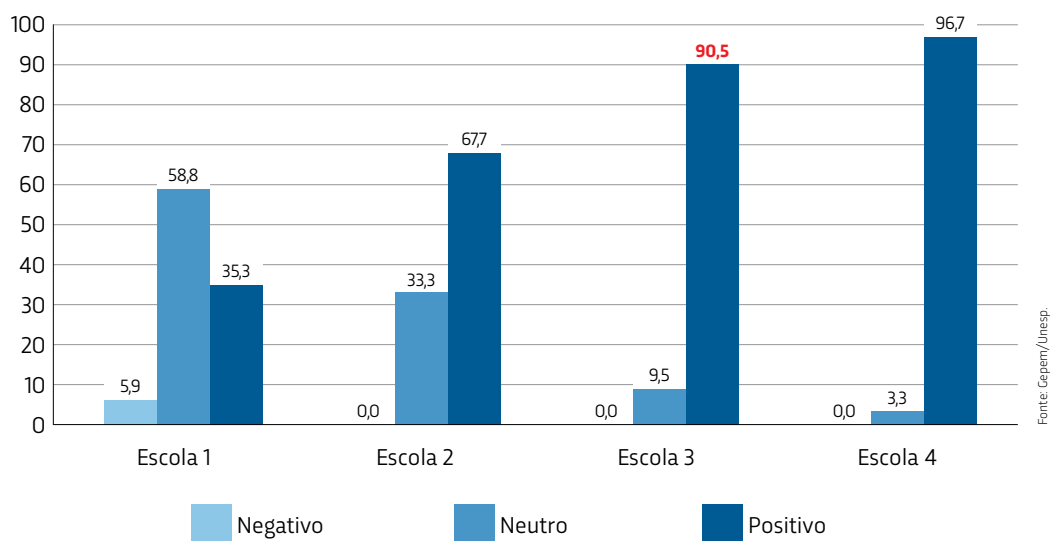
---

2. “A convivência ética na escola: promovendo a melhoria do clima escolar”. Esse projeto tem como objetivo descrever e avaliar um programa de intervenção que foi realizado em escolas públicas de Ensino Fundamental em duas cidades do estado de São Paulo. Tal programa visa favorecer o desenvolvimento da autonomia moral e intelectual, bem como a melhoria da qualidade do clima escolar nas dimensões que envolvem a construção do conhecimento e as relações e conflitos interpessoais, favorecendo a convivência respeitosa. O programa tem a participação de professores e gestores e busca atuar nas vias institucional, pessoal e curricular. Consiste principalmente na inserção, no currículo, de uma disciplina específica para a reflexão dos valores e da convivência; na criação de espaços para a resolução de conflitos; na implantação de sistema de apoio entre pares; no fomento ao protagonismo dos estudantes por meio do uso da tecnologia; na formação continuada para toda a equipe; na participação em um ambiente colaborativo *on-line*; e no emprego das avaliações responsiva e de clima escolar.

## DIMENSÃO 1 – AS RELAÇÕES COM O ENSINO E COM A APRENDIZAGEM – ALUNOS



## DIMENSÃO 1 – AS RELAÇÕES COM O ENSINO E COM A APRENDIZAGEM – PROFESSORES



Os quadros anteriores mostram dados de quatro escolas públicas que foram avaliadas somente na dimensão “As relações com o ensino e com a aprendizagem”, para alunos e professores. Interessante observar que, a escola 4 ficava em um bairro muito

violento, mas tinha um clima excelente em todos os itens que compõem essa dimensão. A escola 1 apresentava um clima ruim em todas as dimensões – não muito ruim, mas ruim –; e, nas outras, o nível era intermediário. Essas informações sobre a dimensão de forma geral não dizem quase nada à escola. O que ela faz com uma média de itens?

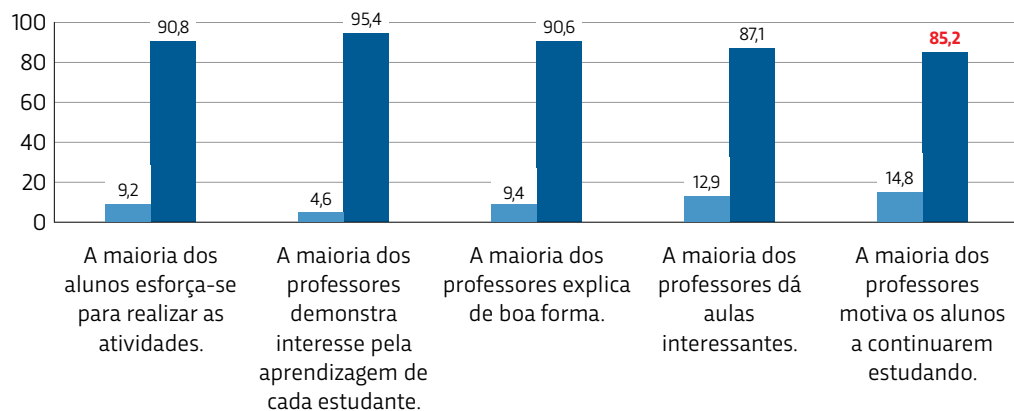
Para uma escola, a avaliação do clima só contribui se resultar em um processo de intervenção, em um projeto construído coletivamente, que pode ser com a ajuda da universidade ou com o compartilhamento de práticas e estudo. Nesse caso, a ideia de parceria está sempre presente. É importante que as intervenções atuem naquilo de que cada escola precisa; as necessidades são diferentes. Essas intervenções precisam ser sistêmicas, atuando como “remédio” e “vacina”.

Apenas avaliar o clima e dar uma devolutiva para a escola, mas não fazer nada com isso, ou seja, não planejar intervenções consistentes para os problemas identificados, não ajuda, só gera mais angústia. Todavia, é necessário haver informações consistentes para que a intervenção aconteça.

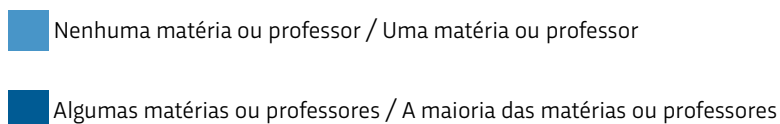
Nos quadros a seguir, é possível ver percepções positivas por parte dos alunos. Concordam que “a maioria dos professores motiva os alunos a continuarem estudando” 85% dos alunos. Nos quadros da página seguinte, para o mesmo item, 100% dos professores respondem que eles também acham que motivam os alunos a continuarem estudando.

*A avaliação do clima só contribui se resultar em um processo de intervenção, em um projeto construído coletivamente, que pode ser com a ajuda da universidade ou com o compartilhamento de práticas e estudo. Nesse caso, a ideia de parceria está sempre presente*

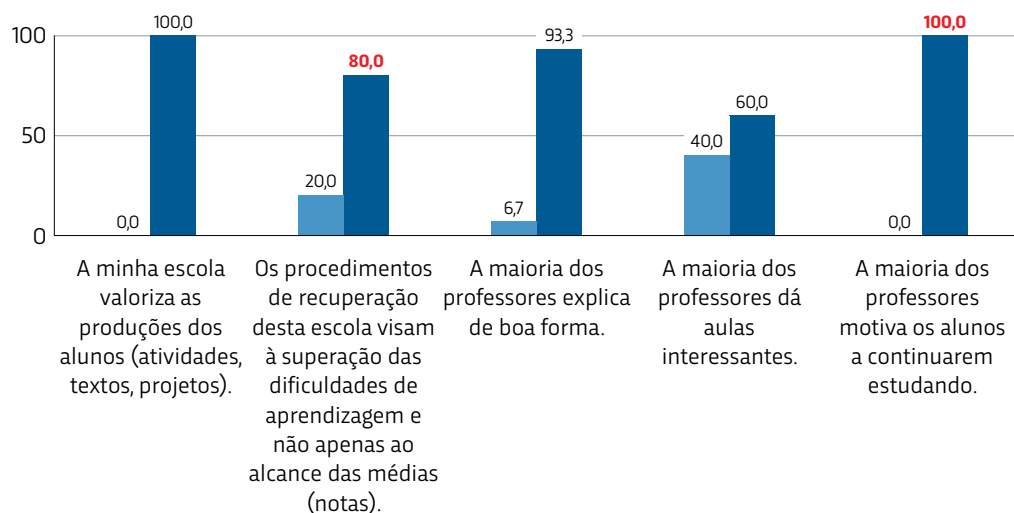
## PERCEÇÕES PREDOMINANTEMENTE POSITIVAS – ALUNOS



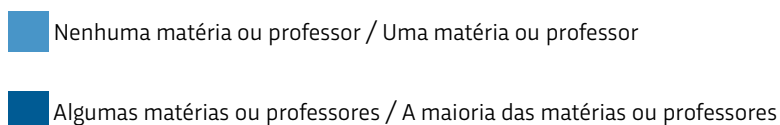
Fonte: Cepem/Unesp.



## PERCEÇÕES PREDOMINANTEMENTE POSITIVAS – PROFESSORES

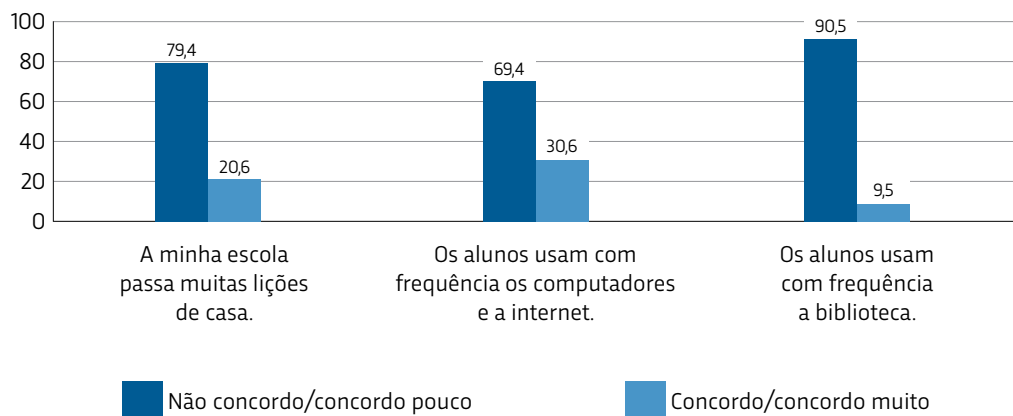


Fonte: Cepem/Unesp.

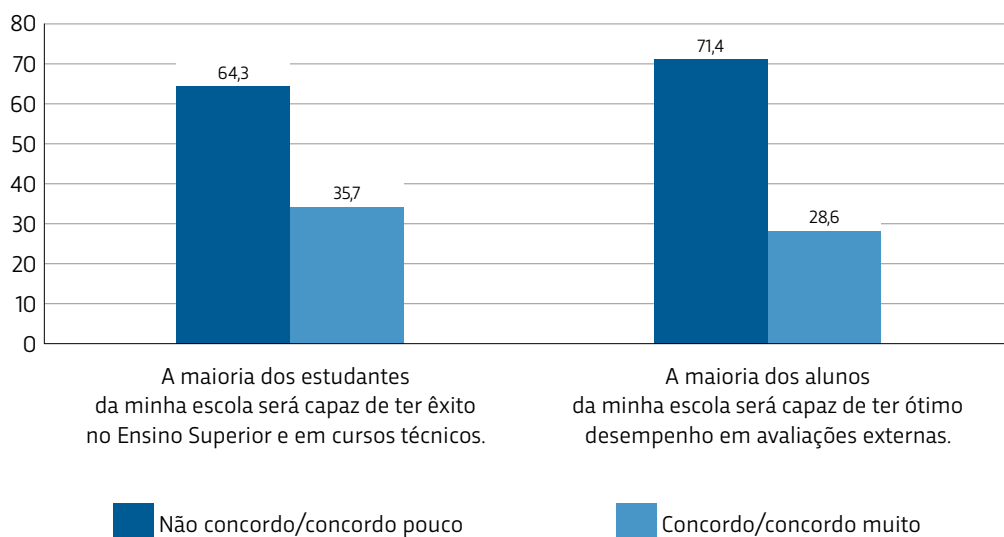


Há, também, perspectivas negativas, como se vê nos quadros abaixo. Quase 80% dos alunos disseram não concordar com a frase “A minha escola passa muitas lições de casa”. Para o item “Os alunos usam com frequência os computadores e a internet”, 69% disseram que não. E para “Os alunos usam com frequência a biblioteca”, 90% responderam que não.

### PERCEÇÕES PREDOMINANTEMENTE NEGATIVAS – ALUNOS



### PERCEÇÕES PREDOMINANTEMENTE NEGATIVAS – PROFESSORES



No quadro inferior (ver p. 182), é possível ver a percepção negativa dos professores. Para a afirmação “A maioria dos estudantes será capaz de ter êxito no Ensino Superior”, só 35% disseram concordar. Para a proposição “A maioria dos alunos será capaz de ter ótimo desempenho em avaliações externas”, 71% dos professores disseram não concordar.

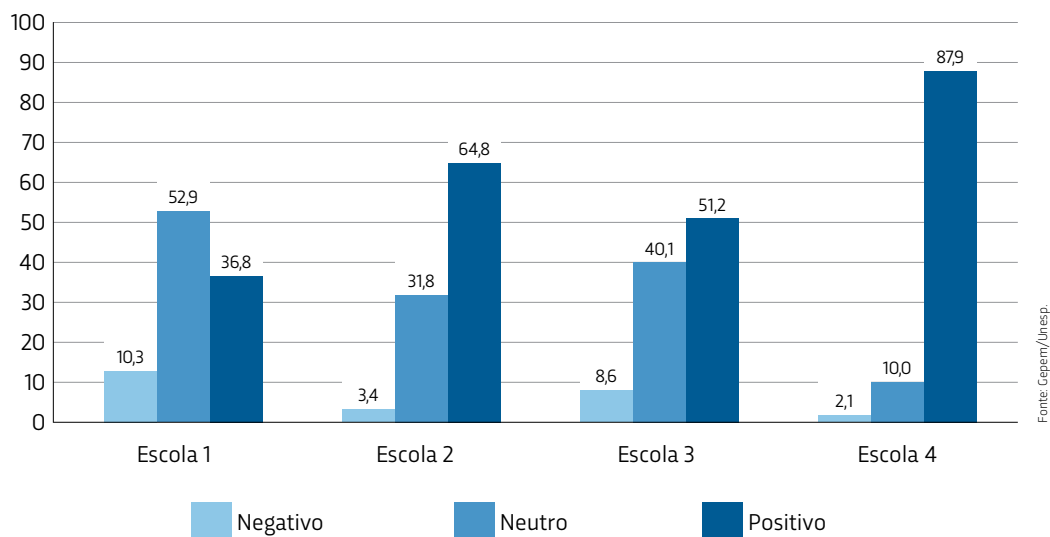
Muitas vezes, as descobertas mais interessantes da pesquisa de clima de uma escola são as discrepâncias surpreendentes entre as percepções dos vários grupos, as quais podem ser usadas como recurso importante de melhorias significativas, promovendo o diálogo para chegar a um acordo sobre missão e objetivos.

Nos quadros a seguir, por exemplo, é possível notar na escola 3 que metade dos alunos avaliou como positiva a dimensão de aprendizagem (quadro superior). Os professores fizeram uma avaliação muito mais positiva (quadro inferior).

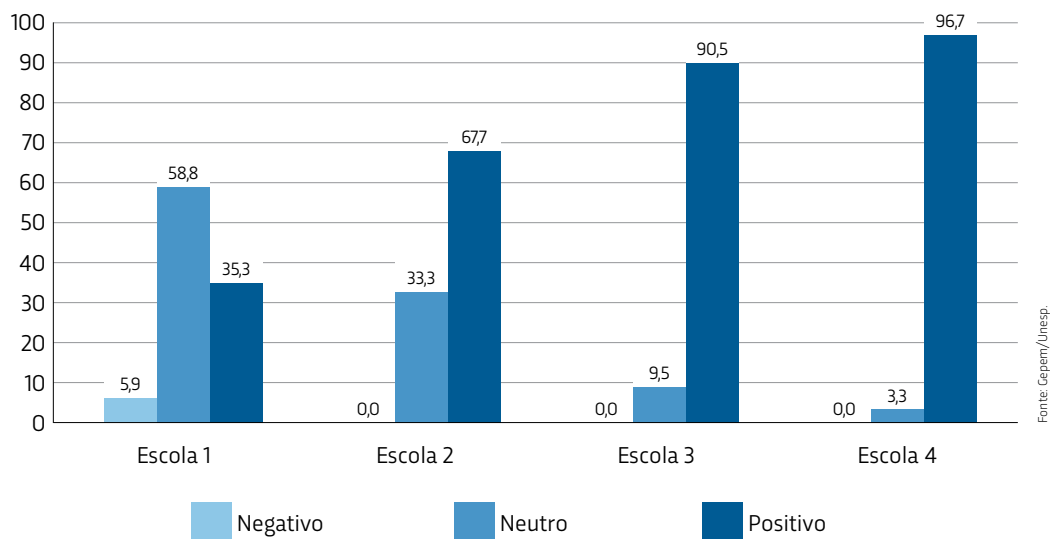
*Muitas vezes, as descobertas mais interessantes da pesquisa de clima de uma escola são as discrepâncias surpreendentes entre as percepções dos vários grupos, as quais podem ser usadas como recurso importante de melhorias significativas, promovendo o diálogo para chegar a um acordo sobre missão e objetivos*



## DIMENSÃO 1 – AS RELAÇÕES COM O ENSINO E COM A APRENDIZAGEM – ALUNOS

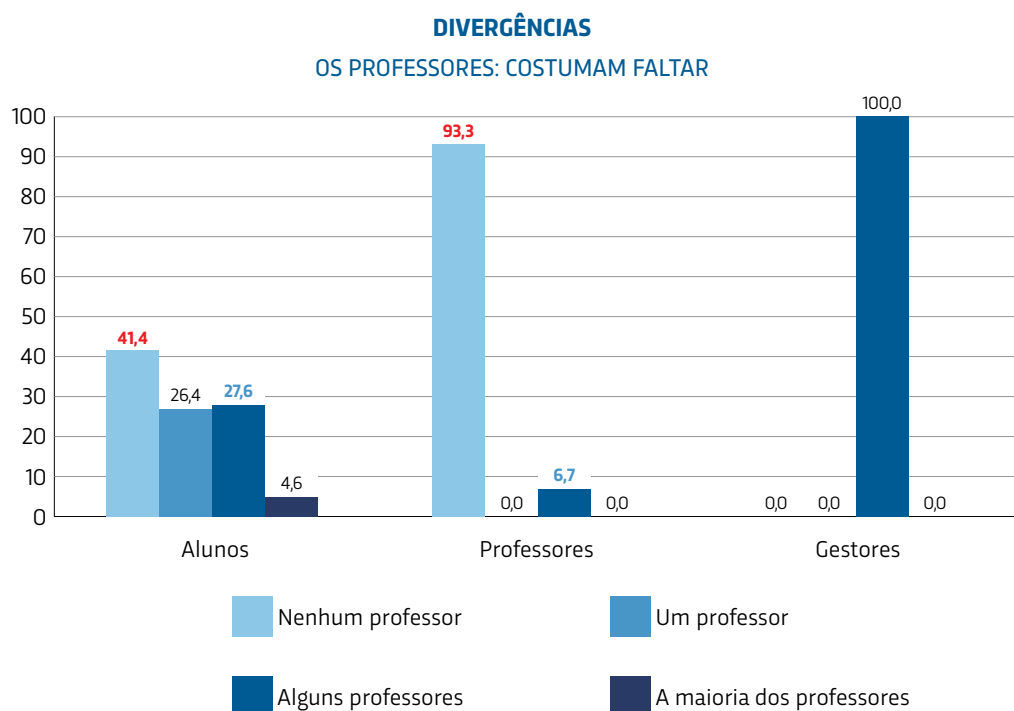


## DIMENSÃO 1 – AS RELAÇÕES COM O ENSINO E COM A APRENDIZAGEM – PROFESSORES



O próximo quadro trata da afirmação “Os professores costumam faltar”: 27% dos alunos disseram que alguns professores costumam faltar, e 26% que apenas um professor costuma faltar. Contudo, 93% dos professores disseram que nenhum professor da escola costuma faltar. Já para 100% dos gestores, alguns professores costumam faltar.

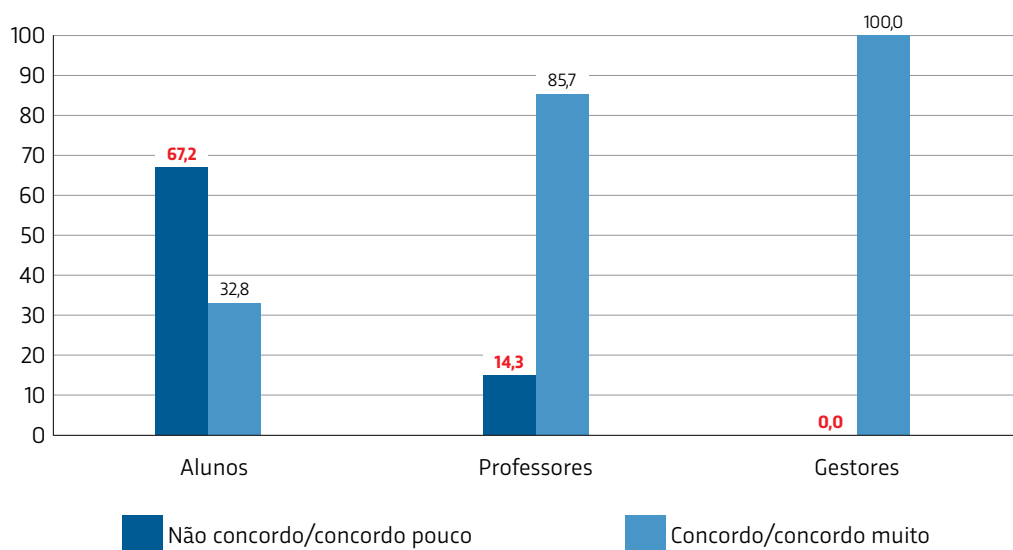
*Todas as escolas que participam de nossos trabalhos o fazem por adesão, não por imposição. Ainda assim, há resistências que fazem parte do processo e são esperadas em um ambiente onde predomina a diversidade*



Quando a afirmação é “Há poucos desentendimentos entre os alunos”, assunto do próximo quadro, cresce a discrepância: 67% alunos discorda – para eles há muitos desentendimentos. Professores e gestores veem de outra maneira: 85% dos professores avaliam haver poucos desentendimentos; 100% dos gestores dizem que há poucos conflitos.

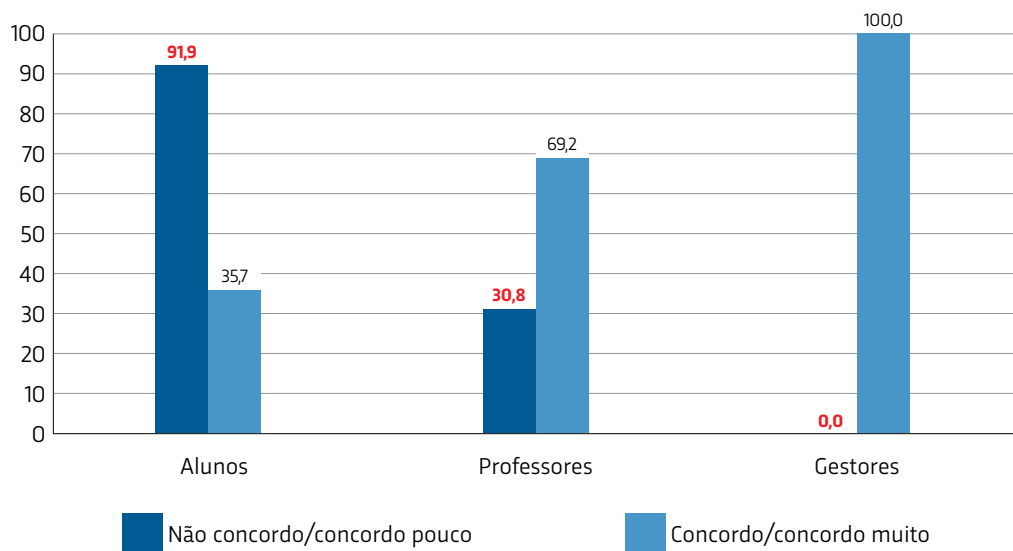
## DIVERGÊNCIAS

### HÁ POUCOS DESENTENDIMENTOS ENTRE OS ALUNOS



## PERCEPÇÕES DIVERGENTES – PROFESSORES, GESTORES E ALUNOS

### OS BANHEIROS ESTÃO EM BOM ESTADO E COM PAPEL HIGIÊNICO E SABONETE À DISPOSIÇÃO



*Esperamos que ocorra um diálogo que avalie o que a instituição já faz bem, quais são suas dificuldades e por que há discrepância de olhares. É preciso engajá-la em projetos de transformação a curto, médio e longo prazo*

Chama atenção o item “Os banheiros dos alunos estão em bom estado e com papel higiênico e sabonete à disposição”, no quadro anterior: 91% dos alunos dizem que não, enquanto 69% dos professores e 100% dos gestores avaliam como bom o estado dos banheiros. Ou seja, eles não usam o banheiro estudantil ou acreditam que apenas a limpeza é suficiente, ou não responderiam como responderam.

## **A AVALIAÇÃO É UM MEIO A SERVIÇO DO PROJETO**

Por que essas avaliações são importantes? Para que, ao serem devolvidas para a escola, haja uma discussão entre representantes de alunos, professores e gestores. Esperamos que ocorra um diálogo que avalie o que a instituição já faz bem, quais são suas dificuldades e por que há discrepância de olhares. É preciso engajar a escola em projetos de transformação a curto, médio e longo prazo. O diagnóstico contribui para identificar os problemas; o grupo os hierarquiza, estuda-os para encontrar formas diferentes e coordenadas de lidar com tais problemas e planeja as intervenções.

A avaliação do clima escolar não pode ser um fim em si mesma. Ela tem de ser o meio a serviço de um projeto. Em primeiro lugar, essa avaliação deve ser fomentada: a escola precisa querer olhar para si por meio das perspectivas diversas que a constituem e seus integrantes têm de ser sensibilizados a participar.

Todas as escolas que participam de nossos trabalhos o fazem por adesão, não por imposição. Ainda assim, há resistências que fazem parte do processo e são esperadas em um ambiente onde predomina a diversidade. Em segundo lugar, o processo tem de ser coletivo, não deve ser realizado apenas pelo gestor nem só pelo professor. Por fim, a aplicação e a devolução dos dados devem ocorrer em um processo de discussão coletiva, selecionando o que vai bem e o que pode ser melhorado, hierarquizando problemas e discutindo etapas para enfrentá-los.

## **FOTOGRAFIA DA ESCOLA**

Esses dados são previamente discutidos e analisados pelos pesquisadores, porém devolvemos os resultados para as escolas sem análises, de modo que os profissionais das instituições estudem os gráficos, discutam os dados e identifiquem o que julgarem ser mais relevante. Quando apresentam aos pares os resultados e itens que selecionaram de cada dimensão, contribuímos para o debate, colocando nosso olhar sobre os resultados.

O clima emerge de modo mais específico e pessoal e tem por finalidade a descrição de fenômenos em um determinado período. Diz respeito a situações subjetivas e mais imediatas, a partir das percepções dos atores da escola. A cultura, por sua vez, é mais profunda, resistente às mudanças, estando permeada pelos valores institucionais. Assim, a análise do clima tem limitações, pois fornece uma “fotografia” do ambiente socioeducativo em determinado período a partir do conjunto de percepções de todos, sem particularizar.

Contudo, em geral leva-se anos para transformar efetivamente o clima de maneira mais consistente, por isso é normal que se faça uma nova pesquisa após dois ou três anos, proporcionando um retrato tanto durante quanto após um processo de transformação na instituição.

Sabemos que realizar apenas a avaliação do clima é insuficiente. É preciso que se busque também outros dados, inclusive qualitativos, em avaliações externas e internas. Muitas vezes, a escola percebe como positivo o resultado de uma dimensão ava-

liada porque está presa em seu cotidiano e ainda não consegue perceber como problemática determinada questão. Um exemplo que ajuda a ilustrar essa ideia ocorreu na avaliação da dimensão “intimidações” em uma escola. Na primeira avaliação, alunos e professores diziam que não tinham problemas com essa questão, vendo como naturais agressões e provocações jocosas, entre outras. Após estudo e discussão sobre o tema, em uma nova avaliação tal dimensão foi vista pelo mesmo grupo como algo negativo e presente na escola, que precisava de intervenções para melhorar.

A partir da identificação de um problema a ser enfrentado, é preciso planejar ações coordenadas e sistêmicas. Vamos citar como exemplo uma escola que identificou enfrentar muitos conflitos. Além de formar os professores para que melhorem a qualidade das relações com os alunos, é preciso rever a qualidade e quantidade de regras e sanções. Também podem ser implantados momentos como assembleias ou rodas de diálogo, em que professores e alunos se reúnem periodicamente para conversar sobre dificuldades e êxitos, buscando entender uns aos outros e colocar suas perspectivas de modo a melhorar os trabalhos e a convivência (as questões abordadas devem ser coletivas, como apelidos pejorativos, escolha de times, ocorrência de *bullying*, entre outros). Também é preciso pensar em procedimentos “curativos”, de intervenção imediata, tais como os círculos restaurativos, em que as partes envolvidas em um conflito reúnem-

*Em geral leva-se anos para transformar efetivamente o clima de maneira mais consistente, por isso é normal que se faça uma nova pesquisa após dois ou três anos, proporcionando um retrato tanto durante quanto após um processo de transformação na instituição*

*Subjacente a um procedimento avaliativo há uma concepção de escola, de Educação. É preciso ser coerente com o tipo de indivíduo que se quer formar e com o que se espera dos professores*

-se na presença de um facilitador para falar e ouvir um ao outro, reconhecer escolhas e responsabilidades e chegar a um acordo (envolvem conflitos particulares, como brigas de namoro, maus-tratos entre pares, fofocas ou danos materiais). Há outros procedimentos que podem ser selecionados e implantados pela escola, tais como equipes de ajuda, comunicação não violenta, atividades que envolvam as habilidades sociais, entre outras.

Em resumo, a avaliação tem de contribuir para atender a escola no que ela precisa, mas, ao mesmo tempo, é necessário que se tenha uma formação de alto nível. Em nossa experiência, discutir os problemas identificados em uma avaliação de clima somente a partir do referencial da escola não promove avanços além dos que ela já conquista.

## **AVALIAÇÃO RESPONSIVA**

Em conjunto com a avaliação do clima, estamos realizando também a avaliação responsiva, que é uma forma de avaliação qualitativa do programa<sup>3</sup> em que os dados são devolvidos para as escolas durante o processo de implantação de nosso projeto. Está sendo desenvolvida pelo dr. Cesar Augusto Amaral Nunes, e é inspirada na que foi elaborada inicialmente por Robert Stake. É responsiva no sentido de envolver os participantes da escola no processo (responde aos anseios, conta como está

---

3. No caso, “A convivência ética na escola: promovendo a melhoria do clima escolar”.

sendo a transformação, acolhe sugestões etc.). Para tanto, utilizamos um ambiente colaborativo *on-line* para que relatos, opiniões, dúvidas, dificuldades e sugestões fiquem visíveis e sejam compartilhados, contribuindo para a transformação. Os relatos são formativos e também integram os dados da avaliação. Afinal, a busca de avanço (melhoria) no programa depende de avanço no nível da escola e de cada um.

Durante a implantação do projeto, faz-se a síntese e a devolutiva desses dados para a própria escola e incorporam-se sugestões para a melhoria do programa. Então, é possível, na medida em que as escolas façam relatos a partir dessa síntese, realizar os reajustes necessários durante o processo, e não somente depois, o que geralmente ocorre com as avaliações de programas.

Defendemos a necessidade de uma homologia dos processos. Subjacente a um procedimento avaliativo há uma concepção de escola, de Educação. É preciso ser coerente com o tipo de indivíduo que se quer formar e com o que se espera dos professores. Se queremos favorecer a autonomia, é preciso empregar procedimentos cooperativos e respeitosos ao entrar na escola e ao desenvolver qualquer ação. Por exemplo, se o desejo é que essas escolas implantem um ambiente de respeito mútuo, participação e colaboração, a chegada da equipe de pesquisadores também deve ser muito respeitosa. O processo de avaliação precisa ser coerente e mediado pelo debate. A avaliação do clima por diversos grupos e a devolutiva para a escola, assim como a avaliação responsiva por meio desse ambiente colaborativo, têm trazido maior coerência ao processo dentro das propostas de transformação da unidade.

Em síntese, é preciso ir além das avaliações quantitativas, inovando também nessa área. Refletir previamente sobre a concepção de escola e de Educação, assim como de coerência no processo e de respeito pela escola. Quanto maior a diversidade de olhares dos especialistas nesse trabalho, melhor. Acreditamos que uma avaliação deve contribuir para a transformação, para a melhoria da escola.



# ***PESQUISA SERVE PARA ALGUMA COISA?***

**CLÁUDIO DE MOURA CASTRO** é economista pela Universidade Federal de Minas Gerais (UFMG), tem mestrado pela Universidade de Yale e doutorado pela Universidade de Vanderbilt. É especializado em Educação. Foi professor da Pontifícia Universidade Católica do Rio de Janeiro, da Fundação Getúlio Vargas e das Universidades de Chicago, Brasília, Genebra e Borgonha. Trabalhou no Banco Mundial, como economista sênior de Recursos Humanos, e no Banco Interamericano de Desenvolvimento (BID), onde chefiou a Divisão de Programas Sociais.

Em geral, as pesquisas sobre Educação no Brasil tendem a ser frágeis. Entre os principais problemas, podemos citar a antipatia da maioria dos pesquisadores pelos números; as chamadas “pesquisas-sermão”, nas quais se prega como as coisas deveriam ser; as falhas sérias nos métodos utilizados; a redação obscura; e a escassa divulgação que é feita. A consequência é seu fraco impacto.

Aprofundemos cada um desses argumentos. Em seguida, examinemos o impacto da pesquisa de um autor.

## **EDUCAÇÃO BASEADA EM EVIDÊNCIA?**

Nos dias de hoje, fala-se muito em Educação baseada em evidência. Mas onde está a evidência para dar solidez a suas conclusões? De fato, em boa parte das pesquisas, há pouco material que possa, realmente, ser considerado factual e confiável.

*Se folhearmos periódicos de Educação, veremos que grande parte do que está lá só tem a teoria. O lado empírico não comparece. Ninguém se deu ao trabalho de averiguar se aquela teoria se confronta com a realidade*

Quando Francis Bacon<sup>1</sup> fundou a ciência moderna, ele disse, mais ou menos, o seguinte: invente a teoria que desejar, depois veja se o mundo real aceita ser explicado por ela. Essa é a regra básica da ciência moderna: primeiro teoria e depois empirismo, ou seja, a observação do mundo real.

Se folhearmos periódicos de Educação, veremos que grande parte do que está lá só tem a teoria. O lado empírico não comparece. Ninguém se deu ao trabalho de averiguar se aquela teoria se confronta com a realidade. Portanto, em Educação, estamos diante de muita pesquisa pré-baconiana.

## **OS DOZE PECADOS CAPITAIS DA PESQUISA EM EDUCAÇÃO**

Vários obstáculos limitam a divulgação e o impacto das pesquisas. Mas, obviamente, primeiro temos de analisar sua qualidade. Só então faz sentido perguntar que capacidade terão de influenciar a Educação. Examinemos as causas dessa relação tão complexa entre pesquisadores e pesquisa em Educação.

---

1. Francis Bacon (1561-1626), filósofo, ensaísta e político inglês, é considerado o fundador da ciência moderna. Bacon é chamado de pai do empirismo moderno por ter formulado os fundamentos dos métodos de análise e pesquisa da ciência moderna. Para ele, a verdadeira ciência é a ciência das causas, e seu método é conhecido como racionalismo experimental.

## ONDE ESTÃO OS NÚMEROS?

Como já sugerido, o primeiro problema dos pesquisadores da área da Educação é com os números: alguns não sabem lidar com eles. Outros deparam-se com o fato de que o número encontrado atrapalha de maneira espantosa a teoria que ele, pesquisador, vinha cultivando. Daí a tentação de abandonar o confronto com a realidade.

## FALHAS NA APLICAÇÃO DO MÉTODO CIENTÍFICO

Costuma-se encontrar sérios problemas com a aplicação do método científico. É possível citar como exemplo questões com o viés amostral, com os grupos de controle, com a maneira como foi formulada a análise de regressão.

Não se trata de um problema exclusivamente brasileiro. Nos países anglo-saxônicos, quando ocorreu o grande debate sobre método fônico *versus* método global na alfabetização<sup>2</sup>, foram identificadas cerca de 70 mil pesquisas sobre o tema. Após uma apuração rigorosa, restaram pouco mais de 80, porque as demais tinham falhas metodológicas. Aprofundando a análise, ficaram cerca de 15. Eram as que, metodologicamente, sobreviveram. Ou seja, não é só no Brasil que grande parte da pesquisa utiliza-se de métodos falhos.

## PROBLEMAS COM O GRUPO DE CONTROLE

Muitas vezes, um experimento tem grande sucesso. Por essa mesma razão, o grupo de controle o adota. Ótimo. Porém, do ponto de vista da avaliação em curso, não se sabe mais o que aconteceu, pois o controle e o experimental se aproximam. Essa

---

2. Trata-se da áspera controvérsia a respeito do melhor método para alfabetizar. O método fônico prescreve um aprendizado dos sons e letras correspondentes. No método global, mostra-se ao aluno a palavra inteira ou até uma frase completa, cabendo ao aprendiz decifrar o que vê diante de si.

adoção pode ser a marca do sucesso da nova ideia, mas a contaminação do grupo de controle gera uma dificuldade crescente para a validação dos resultados na avaliação.

## DEFEITOS NAS AMOSTRAS

É comum o uso de amostras insuficientes. Isso pode ocorrer quando o pesquisador tem um forte viés na análise de uma pesquisa. Pode também acontecer de a especificação ser inadequada ou insuficiente ou as medidas escolhidas serem muito frágeis. Isso ocorre mais com pesquisas qualitativas.

Jean Piaget formulou suas teorias a partir da observação de seus sobrinhos. Como fonte para sua inspiração criativa, nada contra. No método científico, não interessa de onde veio a ideia. O que ele fez foi levantar hipóteses, criar uma teoria nova. Completou apenas a primeira parte da díade de Bacon. Faltou testar a hipótese. Na verdade, isso foi feito por alguns de seus discípulos. Mas, por maior que seja a nossa admiração por Piaget, é preciso reconhecer que ele não completou o ciclo. Desde que alguém o faça, estamos bem. Mas o que não podemos fazer é tomar o trabalho de Piaget como mais do que a primeira metade do ciclo da ciência. O mesmo vale para qualquer outra pesquisa.

## A “PESQUISA-SERMÃO”

Muito do que se passa por pesquisa não é mais do que exortação: “Faça isso”; “Deve ser assim”; “O certo é como estou dizendo”. O que se esperaria que fosse pesquisa é uma tentativa de convencer os leitores que esta ou aquela explicação ou comportamento é o melhor. Os cursos de metodologia científica advertem contra a presença de julgamentos de valor onde se esperaria o peso da evidência empírica.

Um caso curioso aconteceu com minha aluna de mestrado. Queria demonstrar que os testes baseados nas teorias de Piaget seriam menos influenciados por fatores culturais do que os testes convencionais de inteligência. Foi construída uma amostra

*Uma falha comum entre pesquisas em Educação é sua redação. Costuma ser obscura, confusa ou técnica demais. A lição óbvia é que escrever corretamente é fundamental*

e montamos um teste empírico. Infelizmente, para suas crenças, as correlações mostraram o oposto. Os testes de Piaget eram ainda mais condicionados pelo *status* socioeconômico dos alunos. A moça teve uma reação emocional: não aceitava a evidência empírica. Chegamos a um impasse, e ela acabou perdendo o mestrado. A ironia do caso é que mostrei os resultados a Jacques Voneche, sucessor de Piaget na Universidade de Genebra. Para ele, os resultados eram interessantíssimos e seriam uma contribuição original aos debates. Em outras palavras, a aluna, por sua incapacidade emocional de traír o grande mestre, deixou de fazer uma tese que teria considerável impacto na literatura.

## **INDIGESTÃO SEMÂNTICA**

Muito do que passa por pesquisa é apenas uma polifonia de palavras, cujo significado não se conhece bem e que não tem a mais remota possibilidade de que seja traduzido em proposições rigorosas. Não há uma tradução sucinta e rigorosa do palavrório.

## **A PESQUISA ESTÁ ESCRITA EM JAVANÊS?**

Uma falha comum entre pesquisas em Educação é sua redação. Costuma ser obscura, confusa ou técnica demais. A lição óbvia é que escrever corretamente é fundamental. Ademais, é preciso levar em conta a audiência, para assegurar-se de que os resultados serão compreendidos.

Imaginemos o seguinte texto resumindo uma pesquisa: “Em uma amostra de alunos, escolhida por um processo aleatório simples, ao comparar o seu QI, medido pelo teste de Binet-Simon, com o seu *status* socioeconômico, avaliado em uma escala de prestígio ocupacional, encontramos um coeficiente de correlação positivo de 0,6, significativo a 1%”. Quantas pessoas no universo da Educação dominam as ferramentas requeridas para interpretar essa frase? No entanto, dependendo do leitor colimado, é possível optar por um texto mais claro e que atinja um público muito maior do que apenas o universo dos pesquisadores. O texto poderia ser “traduzido” assim: “Os filhos de pais mais ricos tendem a se sair melhor na escola”. Na “fábrica” da pesquisa, a primeira formulação é essencial. Mas, para divulgar os resultados para um público mais amplo, é preciso traduzi-los<sup>3</sup>.

Sejamos economistas, pedagogos ou engenheiros, o fato é que, enquanto nós, pesquisadores, não conseguirmos deixar de “falar javanês” e adotarmos uma escrita inteligível para quem toma decisões, nosso impacto será muito pequeno.

## **NÃO É BOA IDEIA ESTAR CERTO ANTES DA BOA HORA**

Não basta estar certo, é preciso estar certo na hora certa, no lugar certo, com a pesquisa certa, comunicada no momento certo. Sem todo esse sincronismo, não acontecerá nada. A pesquisa passa despercebida e cai no esquecimento, a despeito do valor ou originalidade que possa ter.

## **O AUTOR DA PESQUISA É QUEM VAI SER AFETADO POR ELA**

Estamos diante de um complicador quando o autor da pesquisa é dono do processo avaliado. Pode acontecer de o resultado não lhe convir, se a pesquisa diz que seu programa é ineficaz. Nesses casos, não é incomum que a pesquisa desapareça, como em um passe de mágica.

---

3. Esse comentário está explicitado por Jacques Barzun em seu livro *The House of Intellect* (New York: Harper Torchbooks, 1959), capítulo 1.

Se os resultados não lhe agradam, vão contra o que acredita, pode ser que engavete o estudo. Este é o pior cenário: uma pesquisa que simplesmente desaparece e sobre a qual nunca mais se ouve falar, se é que alguém ouviu falar alguma vez.

O que se observa na prática? As pessoas que estão em posição de decisão leem a pesquisa e decidem se vão ou não considerar seus resultados. São convenientes ou não? Ajudam sua posição ou não? São bons para sua carreira ou não? Há um detalhe adicional: raramente se utiliza a pesquisa como um todo. As partes boas para os tomadores de decisão são destacadas, enquanto as prejudiciais são descartadas.

No entanto, também há vantagens nessa proximidade – quando o dono da pesquisa é também o dono do processo. Quando quem faz ou quem encomenda a pesquisa está próximo ao que está sendo avaliado, há condições para que tal estudo seja realmente levado a sério, e que mudanças resultem daí.

Para entender o que pode ocorrer com a pesquisa, devemos indagar qual é o custo (político) de aceitar, e qual o de ignorar? Um político que defendeu ardorosamente um programa não vai gostar de ver uma pesquisa que o revela irrelevante. Provavelmente, fará tudo para que o estudo seja ignorado.

Há uma outra consideração curiosa. Na pira das vaidades humanas, reivindicar a autoria de uma ideia ou de uma pesquisa pode ser fatal para sua aceitação, pois se o responsável pela implementação das ideias geradas na pesquisa for discretamente elevado à posição de autor, as chances de que venham a ser implementadas aumentam. Ou seja, se o verdadeiro autor insistir em ser reconhecido, isso pode militar contra a aceitação da pesquisa.

## O MARKETING DA PESQUISA

Um forte obstáculo ao impacto de uma pesquisa podem ser as falhas em sua divulgação. Em boa medida, a disseminação depende da capacidade de o pesquisador vender bem ou mal seu trabalho. Tal diferença de “marketing” pode influenciar o efeito de uma pesquisa ou de uma teoria.



*Em certos momentos, os resultados começam a convergir, ou seja, todos vão descobrindo a mesma coisa. Ainda assim, a comunidade acadêmica ignora o achado, pois está acomodada à teoria anterior. Chega um momento, porém, em que não dá mais para ignorar a importância do achado*

Vale a pena lembrar um exemplo no campo médico. Quando, do alto de seus dois prêmios Nobel, Linus Pauling disse que vitamina C curava gripe, todos acreditaram, instantaneamente. Só que pesquisas subsequentes mostraram que não cura.

Quando um médico obscuro da Austrália disse que úlcera não passava de uma infecção curável com antibióticos, ninguém acreditou. Mas era verdade.

Nesses casos, a reputação de um vendeu instantaneamente a mensagem. A falta de reputação do outro trouxe um grande ceticismo inicial. A virtude do método científico é poder corrigir esse tipo de engano. Mas nem sempre.

Mario Vargas Llosa cita o caso de Raymond Aron e Jean-Paul Sartre <sup>4</sup>, que nasceram no mesmo ano. Comemorando o centenário desses dois autores, a Biblioteca Nacional Francesa realizou uma exposição sobre cada um.

Notou o autor peruano que Sartre errou em tudo o que podia sobre o comunismo, pois, inexplicavelmente, continuou

---

4. Raymond Aron (1905-1983) foi um filósofo, sociólogo e comentarista francês, crítico do marxismo, para ele o ópio dos intelectuais. Ficou conhecido por seu livro de 1955, *O Ópio dos Intelectuais*. Jean-Paul Sartre (1905-1980) foi um filósofo francês marxista, escritor, ensaísta, ativista político, crítico literário francês. Foi uma das figuras centrais do existencialismo.

apoiando o Partido Comunista russo quando já havia ampla evidência de seus desmandos. Em contraste, Raymond Aron, no livro *O ópio dos intelectuais*, acertou em tudo em sua crítica ao Marxismo. No entanto, na exposição, a sala para homenagear Sartre era muito maior do que a dedicada a Raymond Aron.

Por quê? Parece que a resposta é a diferença no “marketing”. Sartre se vendia muito melhor do que Aron. Daí se depreende que é fundamental saber vender a pesquisa.

## O IMPACTO DA PESQUISA É CUMULATIVO

Sabemos que o impacto da pesquisa é um processo evolutivo; as pesquisas vão migrando para a cabeça das pessoas, mas de maneira lenta.

A Lei de Mateus (da vantagem acumulada), proposta pelo sociólogo Robert Merton, diz que “a quem tem, mais lhe será dado”<sup>5</sup>. Quando uma descoberta é feita simultaneamente por dois pesquisadores e um é mais conhecido do que o outro, a opinião pública atribuirá a autoria da pesquisa ao mais famoso. Pode ser extremamente injusto, mas o mundo é assim.

Na maioria dos casos, o impacto da pesquisa raramente resulta de uma contribuição individual. Não é como atribuir a Pedro Álvares Cabral o descobrimento do Brasil. Um tema importante atrai pesquisadores durante períodos de tempo extensos. Em certos momentos, os resultados começam a convergir, ou seja, todos vão descobrindo a mesma coisa. Ainda assim, a comunidade acadêmica ignora o achado, pois está acomodada à teoria anterior. Chega um momento, porém, em que não dá mais para ignorar a importância do achado. A essa altura, talvez já tenha saído de cena o autor da primeira pesquisa, tecnicamente a pioneira. A fama vai para o estudo que coincidiu com a mudança de percepção dos outros pesquisadores. Mais uma injustiça sem remédio.

---

5. Robert Merton, “The Matthew Effect in Science”, *Science* 159 (1968), p. 56-63.

## ONDE ENCONTRAR AS PESQUISAS?

Onde se encontram as pesquisas que desejaríamos que tivessem impacto? As revistas de Educação, em geral, têm tiragens de dois mil a três mil exemplares, o que é quase nada. Só nos edifícios do Ministério da Educação (MEC) há cinco vezes mais funcionários, potenciais consumidores de pesquisas.

Pode-se argumentar que o problema poderia ser resolvido com o advento da internet. Mas não é o caso. Antes, encontrávamos pesquisas de menos; agora, há demais. O problema passa a ser discernir o que é relevante. Já mencionamos como é difícil encontrar pesquisas confiáveis em Educação. Voltando a um exemplo anterior, quem se dá ao trabalho de avaliar milhares de pesquisas comparando métodos de alfabetização?

## UM ESTUDO DE CASO DE UM ÚNICO AUTOR

Para exemplificar o caráter aleatório do impacto e da atribuição de autoria, usarei meu próprio trabalho como pesquisador. Como já venho fazendo pesquisa sobre Educação há muito tempo, posso perguntar o que sei sobre o impacto das pesquisas que fiz ao longo de minha vida.

Não foram poucas cujo impacto foi nulo, ou quase. Vejamos um primeiro exemplo. Desde 1990 escrevo sobre Ensino Médio, sempre um trabalho novo, com um título diferente, mas com

*Antes, encontrávamos pesquisas de menos; agora, há demais. O problema passa a ser discernir o que é relevante. Já mencionamos como é difícil encontrar pesquisas confiáveis em Educação*

*Tive o azar de meter-me em temas ainda fora do radar da sociedade. Escrevi livros e artigos, gastei tempo e desenvolvi metodologias para calcular o custo da Educação por aluno pelo menos 20 anos antes da hora*

uma mensagem muito parecida. A recente mudança da lei do Ensino Médio reflete minhas pregações desde então. Mas ninguém se lembra de que foram meus os ensaios pioneiros sobre o tema.

Desde 1970, escrevo sobre Ensino Superior. Concretamente, o único impacto que registro foi o de convencer a Secretaria de Educação Superior (SESu) a não cobrar para imprimir uma página do relatório de faculdades em busca de uma autorização. A taxa era de 30 centavos, mas obrigava os representantes das faculdades a gastar um dia adicional em Brasília apenas para ir ao Banco do Brasil, do outro lado da Esplanada, pagar a guia e voltar ao MEC. Consegui essa vitória!

Muitas vezes, é imperdoável estar certo antes da hora. Em meu caso, tive o azar de meter-me em temas ainda fora do radar da sociedade. Escrevi livros e artigos, gastei tempo, desenvolvi metodologias maravilhosas para calcular o custo da Educação por aluno. Porém, fiz isso tudo, pelo menos, 20 anos antes da hora. Hoje, quando finalmente o tema volta ao primeiro plano, ninguém ouviu falar do que havia escrito a respeito de custos. Ou seja, perdi meu tempo. Pior, perdi a paciência para voltar ao tema.

Em 1975, comecei uma avaliação do rendimento escolar em cinco países da América Latina. Na verdade, era um estudo muito parecido com o Programa Internacional de Avaliação de Estudantes (Pisa). Porém, naquela época, nem mesmo o Banco Mundial acreditava em tais estudos. Nos dias de hoje, apesar de precursora, quem já ouviu falar da minha pesquisa?

*Quando cheguei à Capes e comecei a introduzir mudanças na avaliação, vali-me de tudo o que já estava sendo feito anteriormente e consegui, de maneira razoável, estruturar a avaliação da pós-graduação. Deu certo*

Gostaria de relembrar outra história pessoal, só que no sentido inverso. Quando cheguei à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e comecei a introduzir mudanças na avaliação, vali-me de tudo o que já estava sendo feito anteriormente e consegui, de maneira razoável, estruturar a avaliação da pós-graduação. Deu certo. Hoje é um sistema robusto e consolidado. Mas voltemos a R. Merton. Muitos acham que fui eu quem criou a avaliação da pós-graduação da Capes. Não foi assim. A iniciativa inicial foi de Darcy Closs, meu antecessor, presidente entre 1974 e 1979. Ou seja, a opinião pública atribui a mim a avaliação da pós-graduação porque sou mais conhecido do que Darcy. Funcionou a Lei de Mateus, a meu favor.

Tive alguns casos de sucesso. O Telecurso 2000 foi resultado de uma pesquisa feita por mim e por João Batista Araujo e Oliveira<sup>6</sup>. Sucesso quase imediato, pois a Fundação Roberto Marinho a perfilhou, obtendo um generoso financiamento da Federação das Indústrias do Estado de São Paulo (Fiesp).

---

6. João Batista Araujo e Oliveira é referência nacional em Educação. Graduado em psicologia pela Universidade Católica de Minas Gerais, com mestrado pela Universidade Tulane e doutorado em pesquisa educacional pela Universidade Estadual da Flórida, é o fundador e presidente do Instituto Alfa e Beto, ONG promotora de políticas práticas de Educação que priorizam a alfabetização. Foi secretário-executivo do MEC (1995) e idealizou o programa Acelera Brasil, que visa acelerar e corrigir o fluxo escolar com o apoio do Instituto Ayrton Senna. Já publicou dezenas de livros, entre eles: *A pedagogia do sucesso*, *A escola vista por dentro*, *ABC do alfabetizador*.

Em 1985, escrevi uma nota sobre o Ensino Técnico e a enviei ao então ministro da Educação. Nela, propunha separar a parte técnica da parte acadêmica, porque ocorria um fenômeno que até hoje se observa nas escolas técnicas federais: elas se tornam escolas de elite, com vagas muito disputadas. Quem realmente gostaria de ser técnico não encontra vagas. Propus a separação do currículo profissional do acadêmico. Com isso, os alunos buscando um vestibular competitivo não precisariam fazer as disciplinas técnicas. Dessa forma, sobriariam vagas para aqueles desejando se profissionalizar. Minha sugestão foi ignorada. Dez anos depois, quando o então ministro da Educação, Paulo Renato Souza (1945-2011<sup>7</sup>), negociava um empréstimo junto ao Banco Interamericano de Desenvolvimento (BID) para o Ensino Técnico, houve um impasse, e a ideia de separação foi adotada, pode-se dizer, do dia para a noite.

Em suma, em pesquisa é “uma no cravo, outra na ferradura”. Tempo e energia são perdidos em pesquisas ignoradas, talvez, por aparecerem antes da hora. Ou suas implicações são bloqueadas pelas casamatas políticas desta ou daquela burocracia do MEC. Em contraposição, tive pesquisas que influenciaram de modo claro as políticas educativas. Jamais teria sido capaz de antecipar o impacto delas. Sempre uma surpresa.

---

7. Paulo Renato Souza, economista formado pela Universidade Federal do Rio Grande do Sul, com mestrado na Universidade do Chile e doutorado pela Universidade Estadual de Campinas (Unicamp). Um dos fundadores do PSDB, foi ministro da Educação (1995-2003) do governo Fernando Henrique Cardoso, secretário de Educação dos governos José Serra (2009-2010) e Franco Montoro (1984-1986) e reitor da Universidade Estadual de Campinas – Unicamp (1987-1991).

***GESTORES  
EDUCACIONAIS  
E O USO DE  
EVIDÊNCIAS:  
EXPERIÊNCIAS  
E REFLEXÕES***





# ***FINLÂNDIA: EVIDÊNCIAS PARA CONSTRUIR UM DOS MELHORES SISTEMAS EDUCACIONAIS DO MUNDO***

**ILKKA TURUNEN** é conselheiro especial do Ministério da Educação e Cultura da Finlândia. Especialista em políticas baseadas em evidência, já atuou como secretário-geral de Pesquisa e Inovação e foi conselheiro da divisão de Educação Superior e Ciência do órgão.

Iniciarei este artigo tecendo alguns comentários sobre a situação atual da Finlândia. Em muitos aspectos, a evolução dos sistemas de Educação, pesquisa e inovação da Finlândia no pós-guerra é uma história de sucesso espetacular – devido, em grande parte, às instituições e às políticas. No entanto, a crise global atingiu gravemente o país em 2009. O Produto Interno Bruto (PIB) caiu 8,3%, e a recuperação tem se mostrado difícil desde então. A demanda fraca no mercado externo nos últimos sete anos levou a economia à recessão. A produtividade média do trabalhador caiu de maneira acentuada. Em meio à crise, a confiança que a nação depositava na importância da pesquisa e da inovação e no paradigma de Educação – reconhecido e difundido internacionalmente – para o desenvolvimento e o crescimento do país foi abalada. As dificuldades econômicas que vêm persistindo há algum tempo de certo modo afetaram a capacidade de o país lidar com a concorrência internacional cada vez mais acirrada e de manter o Estado de bem-estar social no nível a que nos acostumamos. A Finlândia buscou crescimento econômico e bem-estar social com base em conhecimento e exper-

tise. Não há nada no horizonte que torne necessário abrir mão dessa estratégia – ao contrário. No contexto global, a produção relacionada a cadeias de valor vem se tornando cada vez mais dependente de conhecimento e investimento. Na atualidade, a maior parte do crescimento da produtividade se baseia em conhecimento, pesquisa e inovação.

O passado da Finlândia tem implicações para o futuro e a governança do sistema. Educação, pesquisa e inovação contam com apoio político e popular sem equivalente no mundo. Os bons resultados em comparações internacionais não teriam sido possíveis sem uma base de valores sólida que favoreceu a Educação. A importância da Educação para o país também ficou evidente em um estudo realizado em 2009 sobre a opinião dos finlandeses acerca do passado. Para nós, o elemento mais importante de nossa história é ter tornado a Educação Básica acessível a todos. Mais de 75% elegeram a Educação como uma das cinco mais importantes conquistas históricas; o percentual de respostas para a Guerra de Inverno, o voto universal, o Estado de bem-estar social, o ingresso na União Europeia ou a guerra civil foi muito menor.

A valorização da Educação conta com uma base sustentável. Em geral, aparece em primeiro lugar em rankings sobre o valor que os finlandeses atribuem às instituições públicas, como as das áreas de segurança, polícia, defesa e sistema judicial. Em 2014, o sistema de Educação foi apontado como a instituição mais confiável, pouco acima da polícia.

*Educação, pesquisa e inovação  
contam com apoio político e popular  
sem equivalente no mundo. Os  
bons resultados em comparações  
internacionais não teriam sido possíveis  
sem uma base de valores sólida*

## NOSSA MAIOR FORÇA: OS RECURSOS HUMANOS

WEF Índice de Capital Humano 2016	WEF Índice de Competitividade Global 2015-2016	WEF e Insead Índice de Prontidão para Networking 2016	Insead, WIPO Índice Global de Inovação 2015 Efetividade do governo	OCDE Qualidade da produção científica	Pisa OCDE Médias em leitura, matemática e ciência	Richard Florida, Charlotta Mellanger, Karen King Índice Global de Criatividade 2015
1. Finlândia	1. Suíça	1. Cingapura	1. Finlândia	1. Suíça	1. China	1. Austrália
2. Noruega	2. Cingapura	2. Finlândia	2. Cingapura	2. Holanda	2. Cingapura	2. Estados Unidos
3. Suíça	3. Estados Unidos	3. Suécia	3. Dinamarca	3. Dinamarca	3. Hong Kong	3. Nova Zelândia
4. Japão	4. Alemanha	4. Holanda	4. Suécia	4. Islândia	4. Coreia do Sul	4. Canadá
5. Suécia	5. Holanda	5. Noruega	5. Noruega	5. Suécia	5. Japão	5. Dinamarca
6. Nova Zelândia	6. Japão	6. Suíça	6. Suíça	6. Bélgica	6. Finlândia	6. Finlândia
7. Dinamarca	7. Hong Kong	7. Estados Unidos	7. Canadá	7. Estados Unidos	7. Estônia	7. Suécia
8. Holanda	8. Finlândia	8. Reino Unido	8. Holanda	8. Reino Unido	8. Canadá	8. Islândia
9. Canadá	9. Suécia	9. Luxemburgo	9. Nova Zelândia	14. Finlândia	9. Polônia	9. Cingapura
10. Bélgica	10. Reino Unido	10. Japão	17. Hong Kong	15. Israel	10. Holanda	10. Holanda

Fonte: elaboração do autor.

Como se pode ver no quadro acima, a Finlândia encontra-se muito bem posicionada na área de políticas sociais em termos internacionais. É um dos países que se destacam, mas isso não é algo a ser propagandeado. Sabemos que esse tipo de quadro não é apenas ciência ou apenas entretenimento, mas tende a impactar as políticas educacionais.

Os dados mostram que a Finlândia tem bom desempenho em comparações de indicadores sobre potencial de crescimento e competitividade estrutural no médio e longo prazos. O país desfrutou de forte progresso econômico nas últimas décadas, o que é evidenciado pelo alto padrão de vida, destacando-se pela alta percepção subjetiva acerca de bem-estar, Educação e capacitação, qualidade do meio ambiente e segurança pessoal. A desigualdade (medida pelo coeficiente de Gini de desigualdade

social) é uma das mais baixas entre os países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) e tem permanecido constante desde a virada do milênio. A pobreza absoluta (medida por privação material e de moradia) está entre as mais baixas da União Europeia.

A Educação encontra-se certamente entre nossos pontos fortes, e o que quero discutir aqui é se existe relação entre isso e as políticas baseadas em evidências. De início, discorrerei sobre a abordagem finlandesa da Educação e dos recursos humanos.

## **PROFESSORES**

Um dos conceitos centrais é a confiança nos professores. Os professores são os especialistas, os profissionais que utilizam e desenvolvem as evidências disponíveis. A formação de professores é realizada por oito universidades, que são os laboratórios onde ocorrem as invenções e os experimentos. Os estudantes desses cursos se envolvem com o dia a dia do trabalho na escola.

O conceito de recursos humanos também é de particular importância. Acreditamos que atingiremos crescimento econômico e bem-estar social sustentáveis se utilizarmos as reservas de talento contidas na nação como um todo.

Cultivamos a ideia de que a criatividade é diversa e pertence a todos e que é uma parte constituinte do dia a dia, da atividade social, da ciência, da arte, da Educação e da vida profissional. A criatividade não é reservada apenas aos melhores e mais inteligentes – é a substância de cada ser humano. Ao estimular a criatividade nos primeiros anos de vida e na Educação Básica, estabelecemos as bases do sucesso futuro na vida das pessoas.

O princípio central da política educacional finlandesa tem sido o cuidado em relação ao capital humano, o que também criou um bom cenário para uma política de inovação ampla. Investimentos em Educação e pesquisa promoveram avanços tecnológicos que foram colocados em prática de maneira eficiente. Em nosso ponto de vista, a melhor maneira de conquistar crescimento econômico, bem-estar e renovação social sustentáveis é mobilizar as competências e os talentos dos cidadãos. De acordo

*O aspecto mais importante é que a Educação no país baseia-se em uma cultura de confiança em relação aos professores e no alto nível da formação desses profissionais. Docentes e diretores têm grande autonomia local para criar*

com essa abordagem, o bem-estar, a equidade, a criatividade e a competitividade estão interligados. Bem planejado, o sistema de seguridade social permitiu ao país assumir riscos calculados e renovar a economia. Além de assegurar o funcionamento apropriado da sociedade, as instituições públicas e privadas do país fizeram uma contribuição significativa ao capital social, o que se revela no amplo respeito da sociedade à aplicação da lei e à baixa corrupção. À medida que a economia se abria cada vez mais, as redes de segurança social e as instituições fortes e confiáveis reduziram o receio dos cidadãos.

O aspecto mais importante é que a Educação no país baseia-se em uma cultura de confiança em relação aos professores e no alto nível da formação desses profissionais. Docentes e diretores têm grande autonomia local para criar e desenvolver os melhores métodos para ajudar os estudantes a aprender. Professores qualificados são constantemente atualizados para melhorar suas competências e contam com o apoio de outros profissionais bem formados, o que lhes permite trabalhar com uma boa dose de independência e atender às necessidades individuais dos alunos. Os professores também contam com um grau considerável de liberdade para escolher os métodos de ensino. Como os objetivos a serem atingidos não estão especificados no currículo, os procedimentos didáticos e pedagógicos são definidos por eles de acordo com o que consideram melhor para cada grupo. Avaliações qualitativas buscam dar suporte ao desenvol-

vimento e à aprendizagem e não têm caráter de controle ou sanção. Os professores também dão suporte àqueles alunos que têm dificuldades de aprendizagem, e há muita expertise e recursos de apoio para atender àqueles que apresentam algum problema no processo de aprendizagem. Focamos soluções flexíveis e sob medida.

A profissão de professor é muito valorizada e respeitada na Finlândia – de acordo com alguns estudos, é uma ocupação que as pessoas buscam em um companheiro. Sendo assim, é possível selecionar os melhores estudantes para seguir a profissão. Apenas 10% dos candidatos são aceitos nas universidades que formam docentes ou nos programas educacionais, o que explica por que essa é uma profissão tão atrativa. Os docentes tendem a gostar do trabalho e se veem como desenvolvedores que exploram informações e conhecimentos muito diversos.

## **EVIDÊNCIAS**

Otto von Bismarck, chanceler alemão de 1871 a 1890, dizia: “Deem-me boas estatísticas e criarei boas políticas”. Em meu trabalho para o Ministério da Educação e Cultura, tendo a concordar com ele: as evidências são quase tudo. No entanto, quero destacar aqui que necessitamos de diferentes tipos de evidência.

É muito importante contar com uma base de dados sólida para realizar diferentes tipos de monitoramento de políticas e acompanhamento de práticas. Isso pode parecer um pouco antiquado, mas estatísticas são muito importantes. Existem diversos tipos de avaliação: a autoavaliação, a avaliação pelos pares etc. No campo da Educação, temos uma longa tradição de cooperação entre o ministério e a comunidade acadêmica. Sendo assim, não é incomum que o ministério convoque uma mesa-redonda para obter panoramas e análises quando alguma questão emerge.

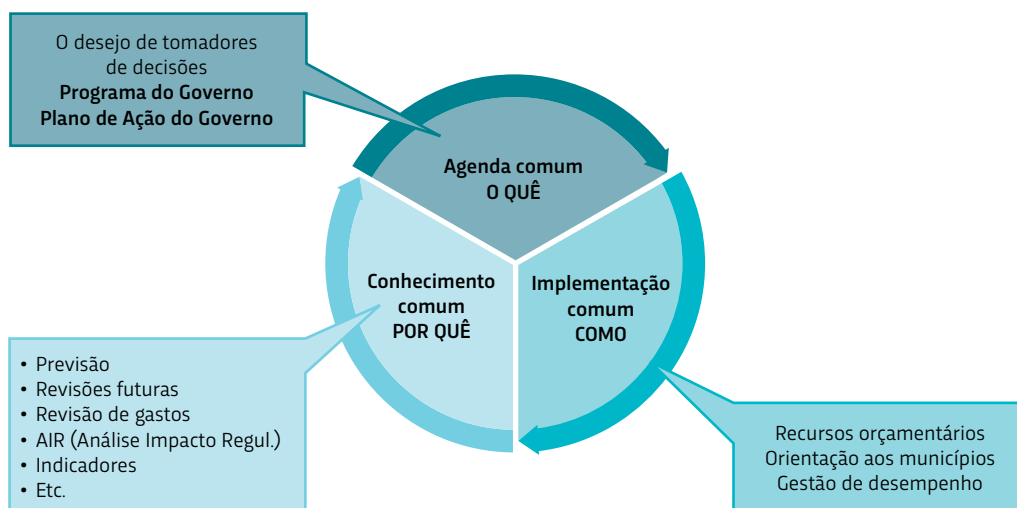
Ainda em relação a evidências, existem também projetos de desenvolvimento e programas de boas práticas nacionais e internacionais como o Pisa (Programme for International Student Assessment ou, em português, Programa Internacional de Ava-

liação de Estudantes) e o PIACC (Programme for the International Assessment of Adult Competencies ou Programa para a Avaliação Internacional das Competências dos Adultos).

A política relacionada à tomada de decisão envolve necessariamente uma combinação de ciências, valores, preferências e avaliações práticas acerca de viabilidade e legitimidade. Isso precisa ser levado em consideração.

As duas próximas figuras descrevem o processo estratégico do governo da Finlândia.

### EVIDÊNCIAS E PENSAMENTO ESTRATÉGICO

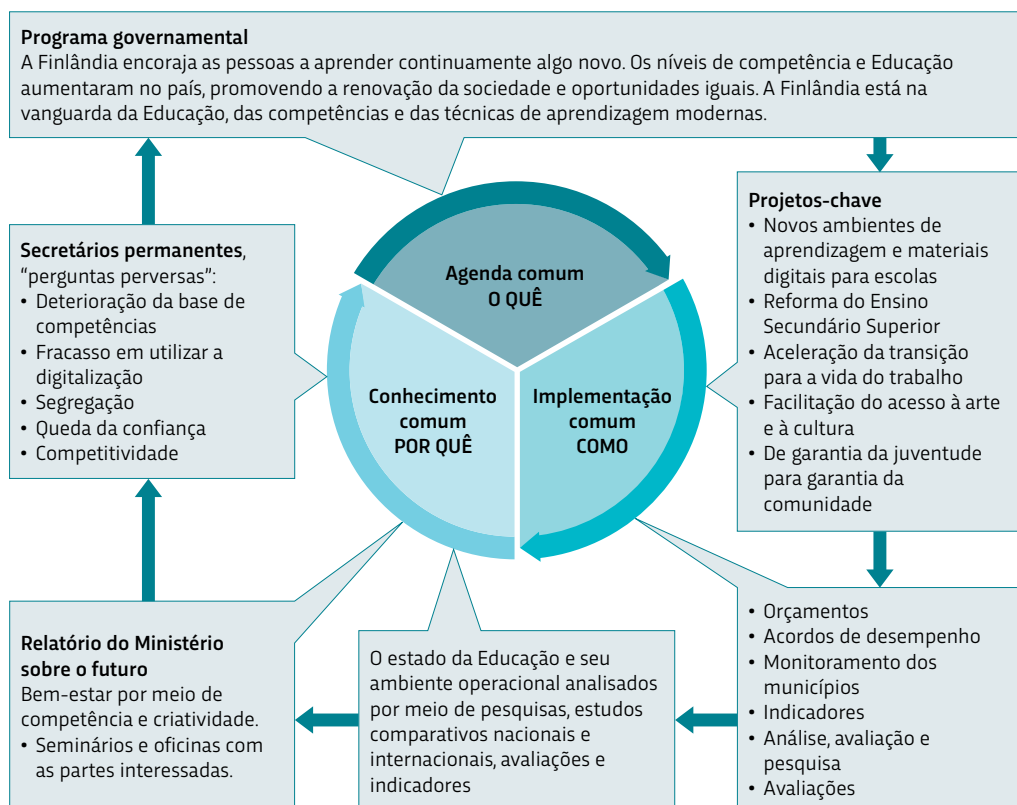


A agenda governamental é constituída por dois passos: 1) a compilação de um programa governamental estratégico e 2) uma fase em que um plano de ação governamental mais específico é criado, incluindo um plano fiscal geral governamental. O conteúdo e os prazos do plano de ação governamental e do plano fiscal geral governamental são integrados. O processo estratégico sustenta-se de maneira consistente em informação e em evidências. Quando os objetivos-chave do governo são claros, e os processos são direcionados a apoiar a conquista desses objetivos, obtêm-se os recursos e os poderes para conduzir as mudanças necessárias.



Em nossa visão, o conhecimento é um recurso estratégico para o desenvolvimento da sociedade e a tomada de decisão. Do ponto de vista do direcionamento, um objetivo essencial deveria ser o de garantir que os tomadores de decisão tenham à disposição informação confiável da melhor qualidade. A informação a que me refiro deveria cobrir ao menos os seguintes aspectos: objetivos (programa governamental, estratégias variadas, objetivos de desempenho etc.), ações (projetos, grupos de trabalho, processos etc.), recursos (organizações, recursos humanos e financeiros etc.) e capacidades (know-how etc.). Sem estruturas de conhecimentos padronizadas, é impossível ou complicado compor uma visão geral sobre o cerne do próprio governo.

## EVIDÊNCIAS E PENSAMENTO ESTRATÉGICO (CONTINUAÇÃO)



Fonte: elaboração do autor.

*A função do relatório era fornecer conteúdo com base em evidências para subsidiar discussões políticas que antecedem a eleição ao Parlamento e também como contribuição estratégica para o programa do novo governo*

A figura anterior descreve o processo estratégico que antecedeu as últimas eleições para o Parlamento, realizadas em 2015. Durante a preparação do Programa de Governo, os políticos necessitavam de informação acerca do desenvolvimento futuro da sociedade e da situação presente. O Ministério das Finanças produz regularmente informação relacionada à situação econômica. Outros instrumentos para o trabalho do governo no futuro incluem o Relatório sobre o Futuro para o Parlamento, a descrição conjunta do ambiente operacional preparado pelos ministérios, os relatórios específicos de cada ministério e o trabalho nacional de prognóstico.

Iniciamos o trabalho no Ministério da Educação e Cultura analisando o *status* do sistema finlandês de Educação na base existente de pesquisas. A seguir, organizamos seminários com pesquisadores e representantes da sociedade com o objetivo de capturar as forças motrizes e as tendências para o futuro. O resultado desse processo foi o relatório do ministério sobre o futuro intitulado “Bem-estar por meio da competência e da criatividade”. Então, os secretários permanentes de cada ministério criaram um relatório conciso abrangendo todos os campos das políticas, com foco em “questões problemáticas” – aspectos muito complexos. A função do relatório era fornecer conteúdo com base em evidências para subsidiar as discussões políticas que antecedem a eleição ao Parlamento e também como uma contribuição estratégica importante para o programa do novo governo.

## PROGNÓSTICOS

Como mencionei antes, temos tentado combinar pesquisa nos campos dos prognósticos e da avaliação.

A Finlândia e as organizações finlandesas encontram-se na liderança mundial na produção de prognósticos. O trabalho de prognóstico fundamenta a preparação para o futuro; não se trata de estratégias de previsão ou de construção de cenários. O objetivo é construir um futuro para todos com base em informações e perspectivas bem fundamentadas. Na Finlândia, muitos atores dos campos privado e público se envolvem nas atividades de prognóstico. No setor público, temos o Relatório sobre o Futuro e as previsões preparadas pelos ministérios; as atividades de prognóstico também são conduzidas em muitos ministérios e setores do governo – por exemplo, para atender a necessidades do mercado de trabalho e de treinamento. No setor privado, são envolvidos pesquisadores, empresas e organizações não governamentais, que também produzem previsões.

Para fins de tomada de decisão pelo governo, é crítico que os prognósticos sejam fundamentados em uma ampla base de fontes da mais alta qualidade, captadas na sociedade como um todo. Informação e interação organizadas promovem a cooperação entre os setores público e privado.

*O trabalho de prognóstico fundamenta a preparação para o futuro [...] O objetivo é construir um futuro para todos com base em informações e perspectivas bem fundamentadas. Na Finlândia, muitos atores dos campos privado e público se envolvem nas atividades de prognóstico*

## *Os temas do prognóstico do governo envolvem a reconfiguração e o futuro da força de trabalho finlandesa*

O gabinete do primeiro-ministro e o Sitra, o Fundo Finlandês de Inovação, coordenam e apoiam as atividades de prognóstico e de trabalho em rede no país. A rede congrega as entidades que produzem dados e constitui um fórum de discussão e coordenação dos atores desse campo. O objetivo é promover o uso da informação relacionada às perspectivas de futuro na tomada de decisão e fomentar a discussão pública, a pesquisa e as decisões sobre desafios e oportunidades para a sociedade finlandesa.

O Grupo Governamental de Prognóstico indicado pelo gabinete do primeiro-ministro em 21 de janeiro de 2015 é responsável por liderar e coordenar os esforços nacionais de prognóstico, além de dar visibilidade à operação. Esse organismo apoia o trabalho da rede nacional de prognóstico, facilitando o uso das visões que ela produz na tomada de decisão e em outros processos-chave.

O governo submete ao Parlamento um Relatório sobre o Futuro com foco em perspectivas de longo prazo uma vez a cada período eleitoral. O documento restringe-se a temas estratégicos relacionados às decisões sobre políticas a serem adotadas em um período de 10-20 anos. A construção do relatório envolve o governo e o Parlamento. O objetivo é estimular um amplo debate na sociedade.

Os temas do prognóstico do governo do primeiro-ministro Juha Sipilä envolvem a reconfiguração e o futuro da força de trabalho finlandesa. A preparação do relatório de prognóstico mais recente teve início em março de 2016, e o trabalho será concluído em duas partes, uma em 2017 e outra em 2018. O objetivo do documento é responder a perguntas amplas acerca de como trabalharemos no futuro, gerar informação sobre o significado da reconfiguração do trabalho e orientar sobre como a Finlândia pode se adaptar às mudanças com êxito.

A primeira parte, a ser entregue em 2017, examinará o futuro do trabalho, os fatores que o impactam e o que essas mudanças representarão para o país. Abordará como o mercado se transformará; quais serão as características da força de trabalho e das estruturas dos ambientes de trabalho; e as maneiras como se espera que o trabalho mudará. Os resultados devem gerar uma atualização detalhada acerca do futuro do trabalho ao longo de cerca de duas décadas. A segunda parte, que será um Relatório Governamental para o Parlamento em 2018, tem como objetivo formular uma visão que contemple as soluções necessárias para que a Finlândia e os trabalhadores finlandeses tenham sucesso no futuro.

## **PESQUISA**

Pesquisa, prognóstico e atividades relacionadas à avaliação compõem a base para a tomada de decisão proativa e pertinente. Ao mesmo tempo, promovem a qualidade, a eficácia e a produtividade da tomada de decisão. Em relação à pesquisa, introduzimos novos instrumentos para financiar estudos que tenham relevância social.

Um dos novos instrumentos disponíveis são as investigações financiadas pelo Conselho Nacional de Pesquisa Estratégica, que trabalha com um horizonte de tempo mais longo que um mandato do Parlamento. O conselho financia pesquisa científica de alta qualidade que tenha impacto social considerável. A cada ano, a instituição submete uma proposta sobre temas estratégicos prioritários para a pesquisa que deve ser aprovada pelo governo finlandês. O governo determina as necessidades e os temas a serem pesquisados, o que o conselho transforma em programas de pesquisa e concursos. O processo de preparação dos temas envolve uma pesquisa *on-line*, *workshops* temáticos, *workshops* de prognóstico e uma consulta aberta. Nesta, pesquisadores e usuários finais dos dados têm a oportunidade de comentar e discutir as propostas de temas. Os programas do conselho duram de três a seis anos, e o orçamento anual é de cerca de 55 milhões de euros.

*O governo adota anualmente um plano [...] O objetivo é criar uma base para o uso sistemático e amplo de informação gerada por pesquisa na tomada de decisão e na condução e operacionalização de práticas*

Além do financiamento proveniente do Conselho Nacional de Pesquisa Estratégica, o governo adota anualmente um plano para análise, avaliação e pesquisa relacionadas às decisões tomadas, para realizar estudos em áreas prioritárias específicas eleitas pelo próprio governo. O objetivo do plano é criar uma base para o uso sistemático e amplo de informação gerada por pesquisa na tomada de decisão e na condução e operacionalização de práticas. O plano para 2016 tinha oito elementos principais, sendo dois muito relevantes aqui: conhecimento e Educação, e digitalização, experimentação e desregulação.

O gabinete do primeiro-ministro e os ministérios designados aos diferentes temas são responsáveis por implementar o plano em colaboração com o grupo de trabalho governamental para a coordenação das atividades de pesquisa, prognóstico e avaliação. Os critérios de seleção para os estudos encomendados incluem relevância, efetividade, usabilidade, qualidade e viabilidade.

## **DESAFIOS DA EDUCAÇÃO NA FINLÂNDIA**

Foi registrado um declínio nos resultados de aprendizagem dos jovens. Houve uma queda acentuada no nível de competências básicas. A parcela de jovens com nível insuficiente de competências básicas em diferentes coortes cresceu. Embora essa deterioração seja mais evidente em matemática, existe uma ten-

dência de queda também em leitura. A maneira de viver e pensar dos jovens mudou. Comportamentos que dificultam a aprendizagem se fortaleceram, e a diferença entre meninos e meninas é maior. Na Finlândia, a correlação entre fatores como autoconceito, autoeficiência e motivação do estudante e seus resultados de aprendizagem em matemática é maior que a média da OCDE. A equidade, que tem sido um ponto forte do país, parece ter sido corroída, e o histórico socioeconômico é um fator cada vez mais relevante para a aprendizagem. As competências dos imigrantes são uma preocupação.

Também notamos que, por volta do início da década de 2020, a tendência de crescimento dos indicadores relacionados às conquistas educacionais se interromperá, o que representa uma ameaça real se quisermos criar empregos de alto valor agregado, pois o país poderá se ver sem talentos, capacitação e habilidades para propiciar esse tipo de ocupação. Em 1991, a Finlândia e a Coreia eram as nações que tinham a nova geração mais bem preparada em termos de Educação de nível superior. Agora, porém, a Finlândia encontra-se nas posições 19 ou 20 nesse tipo de comparação. Algo ocorreu nesse campo. Outro tema é a subexploração de nosso ambiente altamente tecnológico. Muitos estudos indicam que a Finlândia encontra-se entre os melhores países do mundo em termos de benefícios decorrentes da expansão e do aprofundamento da digitalização. No entanto, o país não aproveitou as oportunidades que a digitalização oferece, o que também tem reflexos na Educação. Algumas escolas abraçaram a tecnologia digital de diversas maneiras para dar suporte ao ensi-

*A tendência de crescimento dos indicadores relacionados às conquistas educacionais se interromperá, o que representa uma ameaça real se quisermos criar empregos de alto valor agregado*

## *Nem todos os professores contam com as competências necessárias para usar a tecnologia em sala de aula*

no e à aprendizagem, mas, na média, os estudantes finlandeses usam as tecnologias de informação e comunicação (TICs) menos que outros colegas europeus. Os jovens não transportam para a escola as competências que adquirem no tempo livre. Nem todos os professores contam com as competências necessárias para usar a tecnologia em sala de aula, e a pedagogia não evoluiu na mesma velocidade que os equipamentos.

### **AVALIAÇÃO**

É preciso registrar que uma política de Educação sábia não pode se pautar apenas por tendências internacionais prevalentes. Às vezes, é preciso contrariar a tendência. O papel e o *status* da avaliação de estudantes e da Educação na Finlândia diferem radicalmente do que se faz na maioria dos países. Uma das características interessantes de nosso modelo é o fato de que conduzimos estudos baseados em amostras, em vez de provas nacionais. O principal objetivo de uma avaliação estudantil é promover a aprendizagem e estimular o estudante. Os estudantes não devem ser comparados entre si. Avaliações qualitativas destacam a importância da evolução em detrimento do monitoramento. Lutamos contra a tentação de produzir listas de melhores escolas. No lugar de inspeções escolares, escolas e professores fazem autoavaliação. Apoiamo-nos em confiança e cooperação, não em controle e punição. Há confiança na capacidade de professores muito bem formados de atingir objetivos de aprendizagem em cada situação e atender às necessidades dos estudantes empregando os métodos e procedimentos que considerarem mais adequados. Os critérios do currículo proporcionam um enquadramento sólido, porém flexível.



*É preciso registrar que uma política de Educação sábia não pode se pautar apenas por tendências internacionais prevalentes. Às vezes, é preciso contrariar a tendência. O papel e o status da avaliação de estudantes e da Educação na Finlândia diferem do que se faz nas maioria dos países*

Nossa ideologia é a da condução por meio da informação, do apoio e do financiamento, e os objetivos são estabelecidos pela legislação e por um currículo central nacional.

O currículo central nacional é o documento-chave de nosso sistema educacional. Recentemente redigimos um novo currículo central nacional para a Educação Infantil, a Educação Básica e a Educação Secundária.

O plano de avaliação é o documento-chave para a entidade responsável pelo planejamento da avaliação, o Centro Finlandês para a Avaliação da Educação (KARVI, na sigla em finlandês, ou FINEEC, na sigla em inglês)<sup>1</sup>. No passado, havia no país duas instituições, uma para a Educação Básica e outra para a Educação Superior, mas foram fundidas. O novo organismo é mais forte que os anteriores.

O Centro Finlandês para a Avaliação da Educação é a entidade que determina os objetivos das avaliações, uma vez que cria a proposta para a avaliação nacional para o Ministério da Educação. O plano de avaliação abrange a totalidade do sistema educacional, da Educação Infantil à Educação Superior, e também as várias áreas da escola, como avaliação sobre a funcionalidade do

---

1. Disponível em: <<https://karvi.fi/en>>. Último acesso em setembro de 2017.

sistema educacional, avaliações temáticas e sistemáticas, avaliação de resultados de aprendizagem e auditorias de sistemas da qualidade.

A figura a seguir mostra o último plano de avaliação, último tópico de minha apresentação. Ele abrange áreas como desenvolvimento de avaliações de aprendizagem e competências e funcionalidade e desenvolvimento do sistema educacional, incluindo tópicos de gestão escolar como liderança, entre outros estreitamente relacionados ao seminário “Caminhos para a qualidade da Educação pública: impactos e evidências”.

### ÁREAS FOCAIS RELATIVAS À AVALIAÇÃO DA EDUCAÇÃO, 2016-2019



Fonte: elaboração do autor.

# *A COMPLEXIDADE DA AVALIAÇÃO EM EDUCAÇÃO: A EXPERIÊNCIA MEXICANA*

**MIGUEL SZÉKELY** é diretor do Centro de Educação e Estudos Sociais do México e consultor de organizações internacionais como Banco Mundial, Banco Interamericano de Desenvolvimento (BID), Programa das Nações Unidas para o Desenvolvimento (PNUD) e Fundo das Nações Unidas para a Infância (Unicef). Atuou recentemente como diretor do Instituto para Inovação em Educação do Instituto Tecnológico e de Estudos Superiores de Monterrey, México. Além disso, foi subsecretário de planejamento e avaliação do Ministério do Desenvolvimento Social e subsecretário de Ensino Médio no Ministério da Educação. Lecionou na Universidade de Oxford, no Centro de Estudios Economicos – El Colegio de México e no Instituto Tecnológico Autónomo de México.

É possível existir um sistema educacional excelente. Abordarei neste artigo um quadro mais problemático, mas é importante saber que há experiências que estabelecem uma boa referência para quem se interessa por políticas educacionais. É nessa direção que devemos mirar.

Apesar dos idiomas diferentes, México e Brasil têm muitas similaridades. O que apresentarei aqui é uma visão da América Latina, pois compartilhamos problemas similares em termos de Educação.

## **O SETOR COM MAIOR PODER DE TRANSFORMAÇÃO**

A Educação é o setor que tem o maior poder de transformação em uma sociedade. Infelizmente, também é o mais difícil de modificar, porque empreender qualquer mudança é um processo muito complexo. Olhar para um país como a Finlândia, que é modelo para muitos outros, é encorajador (ver artigo de Ilkka Turunen, p. 208). Obviamente existem peculiaridades, mas, ape-

sar dos problemas, o país mostra que é possível chegar a um estágio mais desenvolvido do sistema educativo.

Este artigo abordará dois temas. Primeiro, as complexidades da avaliação em Educação. Segundo, dois exemplos do México e suas falhas, em parte porque há muito mais problemas que conquistas no país, mas também porque é ilustrativo analisar os desafios na implementação de uma avaliação – eles fazem parte da vida real. Conhecer os problemas é muito informativo, pois permite evitar erros que surgem no caminho.

## **AS PEÇAS DO QUEBRA-CABEÇA**

A Educação como prestação de serviço é muito complexa. Em primeiro lugar, é necessário saber qual é o objetivo de um sistema de Educação. Em alguns casos, os sistemas os explicitam. Por exemplo, na Finlândia o objetivo é o desenvolvimento de algumas competências ou o atingimento de padrões específicos. Em modelos educativos do século passado o objetivo era acumular informação e conhecimento.

O objetivo do sistema de Educação precisa ser muito claro. Como mostra a figura a seguir, há pelo menos seis elementos importantes em torno desse elemento central, como se fossem peças de um quebra-cabeça:

1. O conteúdo educativo ou currículo: o que e como ensinar.
2. As modalidades de entrega: Educação a distância, presencial etc.
3. Os recursos humanos: professores, diretores, assistentes etc.
4. Os materiais: infraestrutura necessária para a entrega do serviço.
5. A gestão escolar: elementos facilitadores e deflagradores.
6. Os mecanismos de suporte: elementos de que o sistema precisa dispor para facilitar a entrada, como bolsas de estudo, alimentação, transporte, segurança etc.

## APRENDER PARA QUÊ?



Fonte: elaboração do autor.

## A EVOLUÇÃO PARA UM SISTEMA EDUCACIONAL MODERNO

Países como a Finlândia se encontram em um estágio no qual os seis elementos estão implementados, mas os sistemas em países da América Latina, Brasil inclusive, ainda se encontram no século passado: a escola foca o currículo. Porém, como mostra a próxima figura, é necessário evoluir para um sistema que tenha como principal objetivo desenvolver competências, o que é algo muito diferente.

COMPARAÇÃO ENTRE MODELOS EDUCACIONAIS DO PASSADO E DO PRESENTE		
	PASSADO	PRESENTE
Modelo pedagógico	<ul style="list-style-type: none"> <li>• Memorização, repetição, acumulação de fatos, dados, informações</li> </ul>	<ul style="list-style-type: none"> <li>• Compreensão e utilização de informação para objetivos concretos, sendo criativo, analítico, proativo</li> </ul>
Modelo de suprimento	<ul style="list-style-type: none"> <li>• Privilegia modelo central único</li> <li>• Modelo de expansão para capitalizar economias de escala e atingir mais estudantes rapidamente</li> </ul>	<ul style="list-style-type: none"> <li>• Oferta adaptada à cultura, ao contexto socioeconômico e ao meio ambiente para promover aprendizagem</li> </ul>
Modelo de recursos humanos	<ul style="list-style-type: none"> <li>• Professor ensina, estudantes aprendem</li> <li>• Instrutor dá ordens e mostra o caminho a ser seguido</li> </ul>	<ul style="list-style-type: none"> <li>• Professor facilita, indica, guia</li> <li>• Parceiro na aprendizagem, promove desenvolvimento de competências e cria estratégias para garantir o progresso</li> </ul>
Modelo de sala de aula	<ul style="list-style-type: none"> <li>• Sala de aula fixa</li> <li>• Todos devem prestar atenção a instruções</li> </ul>	<ul style="list-style-type: none"> <li>• Flexível em espaço e tempo</li> <li>• Ambiente colaborativo e interativo</li> <li>• Permite ritmos diferentes</li> <li>• Interage com o ambiente, promove colaboração e interação</li> </ul>
Modelo de gestão escolar	<ul style="list-style-type: none"> <li>• Diretor como administrador</li> <li>• Administração de recursos para que o ensino aconteça</li> </ul>	<ul style="list-style-type: none"> <li>• Diretor envolvido com o desenvolvimento do professor, a adaptação do espaço da escola, a incorporação de tecnologia e de inovação</li> <li>• Gestão para que a aprendizagem aconteça</li> </ul>
Modelo de acesso	<ul style="list-style-type: none"> <li>• Foco no aumento da oferta (mais estudantes na escola)</li> </ul>	<ul style="list-style-type: none"> <li>• Foco em ambientes de aprendizagem (adaptação, flexibilidade, acessibilidade etc.)</li> </ul>

Pode-se decorar o currículo, mas, ao medir isso, obtém-se uma pontuação muito baixa nos testes atuais que estão estruturados para aferir competências, não porque os alunos não sabem os fatos do currículo, mas porque o sistema ainda não evoluiu para transformá-lo em questões práticas, como visa o modelo.

O ensinar também está mudando. Os professores eram basicamente instrutores – entravam na sala e diziam aos alunos o que fazer. Porém, os conteúdos e o currículo em si estão evoluindo para se transformarem em competências.

Mais do que tudo, talvez, o papel dos professores está mudando dramaticamente. Um professor que no passado foi treinado para entregar o currículo hoje tem que fazer algo totalmente diferente. Um professor do século 21 precisa ser um apoio para o aluno. Ele terá que ser um parceiro do processo educativo e terá que tratar cada estudante de acordo com suas necessidades. Até a organização da sala de aula e os materiais utilizados estão mudando no novo sistema de competências.

Para desenvolver as competências, o velho modelo do professor na frente da lousa tornou-se inadequado. O uso de tecnologia é muito mais frequente, o que significa que os alunos nem sequer precisam estar em uma sala de aula.

Naturalmente, a gestão escolar, tema central do seminário “Caminhos para a qualidade da Educação pública: impactos e evidências”, está mudando. No modelo antigo, os diretores eram administradores – papel que lembra o do contador, informando ao sistema quantos recursos foram recebidos e como foram utilizados. O novo papel do diretor é diferente.

*Um professor do século 21 precisa ser um apoio para o aluno. Ele terá que ser um parceiro do processo educativo e terá que tratar cada estudante de acordo com suas necessidades*



O gestor escolar precisa assegurar que os professores estejam preparados para ajudar os alunos. Ele tem de organizar a escola de maneira a promover a aprendizagem. Trata-se de um papel totalmente diferente. O projeto Jovem de Futuro do Instituto Unibanco atinge o cerne da transformação do papel do gestor escolar, que precisa ser mais um líder que um administrador.

O apoio aos estudantes também é muito importante no âmbito dos seis elementos aos quais me referi antes, porque desenvolver competências e assegurar que todos atinjam o mesmo nível de tais competências requer um serviço escolar que ofereça suporte diferenciado para estudantes com necessidades variadas.

Em resumo, há uma série de elementos que precisam estar em ação para apoiar os alunos. Países como a Finlândia talvez já estejam operando de acordo com tal modelo. México e muitos outros países da América Latina, incluindo o Brasil, estão no modelo antigo, o que torna muito complexo avaliar qualquer intervenção no sistema escolar.

Em minha experiência, a gestão escolar é o mais importante dos seis elementos, porque se a gestão falha, em geral a escola também falha.

## **UM QUEBRA-CABEÇA COMPLEXO**

Um aspecto importante a destacar são as interações complexas entre os seis elementos em torno do objetivo do sistema de Educação.

Pode haver um ótimo diretor e uma ótima gestão, mas os resultados serão muito limitados sem bons professores, um bom currículo e um bom objetivo do sistema, em comparação com uma situação em que todos os elementos atuam em prol do atingimento do objetivo. Do mesmo modo, os resultados de professores excelentes sob um péssimo gestor serão muito diferentes de um caso em que os dois elementos estão sintonizados.

Os principais elementos da complexidade são apresentados a seguir:

## TAMANHO DO SETOR

- A Educação é tipicamente um setor muito grande em termos econômicos e demográficos.
- Há grandes interesses envolvidos, com grande número de pessoas empregadas, o que implica em sindicatos e organizações poderosas (a complexidade da aglomeração).

## MEDIÇÃO DE SUCESSO E/OU FRACASSO

- É difícil medir no curto prazo.
- A família tem dificuldade de optar por outra escola, pois, normalmente, usuários do serviços de Educação pública não escolhem a unidade que frequentam: são clientes “cativos”, obrigados a se matricular.
- O sistema de responsabilização é difuso (com sindicatos poderosos envolvidos).
- Há dificuldade de responsabilizar e cobrar: usuários do sistema não podem reivindicar seus direitos retrospectivamente.

## DURAÇÃO LONGA DO PROCESSO

- São 12 anos de Educação Básica, com função de produção complexa.
- Requer permanente esforço estruturado que vai além dos ciclos políticos.
- Há dificuldade para medir o efeito dos componentes (professores, tecnologia, materiais): muitos atores envolvidos e muitas partes interessadas ao longo do processo.

## MULTIPLICIDADE DE OBJETIVOS

- Alcançar todo mundo x qualidade da Educação.
- Educação Básica x Educação Superior.
- Prover boas condições para quem trabalha no sistema.

- Lidar com questões que são relevantes para grupos diferentes.
- Em geral, confusão entre privilegiar o processo e privilegiar o objetivo final.

## RESISTÊNCIA NORMAL A MUDANÇAS

- Requer mudar comportamentos.
- É mais difícil quando os elementos do sistema não estão alinhados.
- Exige ajustes contínuos e detalhados para a implementação.
- Existe o problema dos incentivos desalinhados (por exemplo, direitos dos trabalhadores x qualidade).

Um dos primeiros elementos da complexidade é que os componentes interagem entre si de tal modo que se atuamos em um deles – por exemplo, a gestão escolar –, o efeito dependerá do que é feito nos outros cinco.

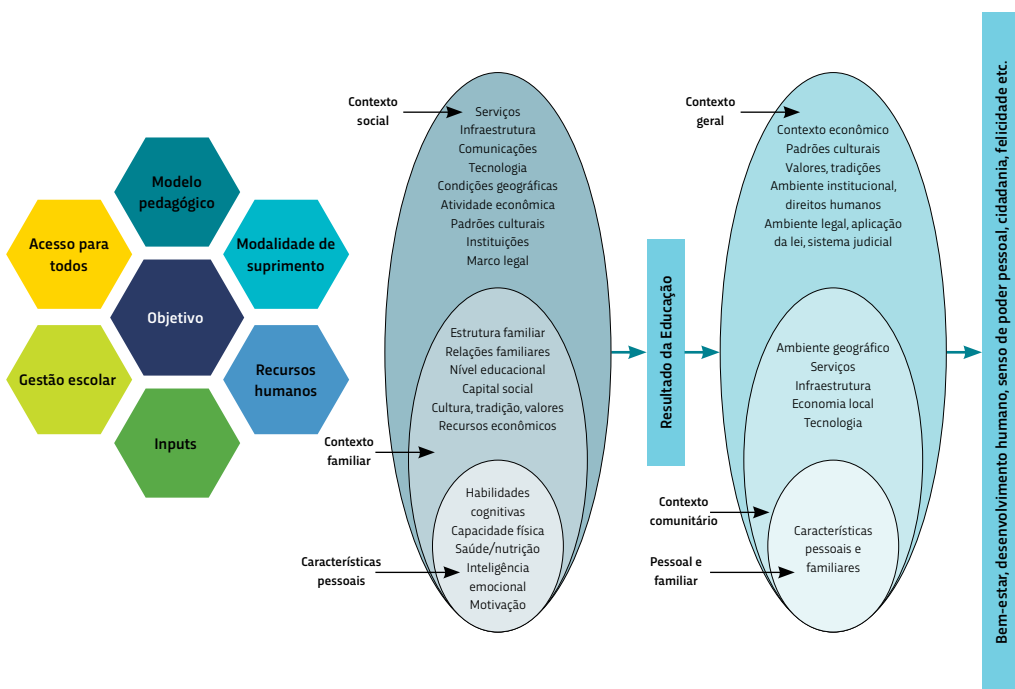
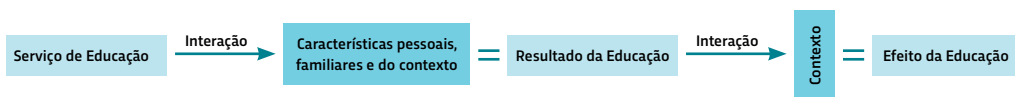
Outra característica é que, como vivemos em um mundo em transformação, os elementos – e até mesmo os objetivos do sistema escolar – mudam constantemente. Mesmo um ótimo projeto como o que o Instituto Unibanco promove, se ele atuar apenas sobre a gestão escolar e os outros elementos se mantiverem iguais, deve-se esperar um impacto limitado.

É claro que as intervenções podem ser concebidas para impactar os seis elementos do sistema ao mesmo tempo. Isso seria ótimo, mas às vezes é preciso selecionar onde atuar primeiro. O desafio então é escolher onde fazê-lo, uma vez que se concentrar em um elemento não gera um grande impacto se os demais não mudarem.

O terceiro elemento da complexidade é que a avaliação dos sistemas escolares também é algo complexo. Pode ser confuso estabelecer o que significa avaliação em Educação e o que ela avalia. Digamos que temos os seis elementos do sistema escolar e temos as escolas. Poderíamos avaliar, ao menos em teoria, se os seis elementos estão alinhados com os objetivos, mas isso seria muito limitado.

É preciso entender que a Educação interage com os estudantes, as características pessoais deles e a situação de suas famílias, como demonstra a próxima figura.

## INTERAÇÕES EM EDUCAÇÃO



**1**  
**Qualidade do serviço de Educação**  
 Medida em que a oferta é exibida e adaptada.

**2**  
**Qualidade dos resultados da Educação**  
 Medida em que a oferta a que os alunos são expostos gera os objetivos educacionais com equidade.

**3**  
**Qualidade dos efeitos da Educação**  
 Medida em que o resultado da Educação gera bem-estar, desenvolvimento humano, senso de poder pessoal, felicidade.

Fonte: elaboração do autor.

Qualidade na Educação

*A comunidade também importa –  
tradições, valores, cultura, infraestrutura.  
Há um conjunto de elementos  
que intervêm na interação entre o  
suprimento de Educação e o resultado*

Se uma criança vai para a escola mal nutrida, sua capacidade de se beneficiar de um ótimo modelo de Educação será limitada. Se uma criança mexicana subnutrida for exposta ao sistema de Educação finlandês, é quase certo que seu desempenho será pior que o de um estudante do país nórdico, porque ambos vêm de contextos diferentes em termos socioeconômicos – sendo o caso mexicano desfavorável.

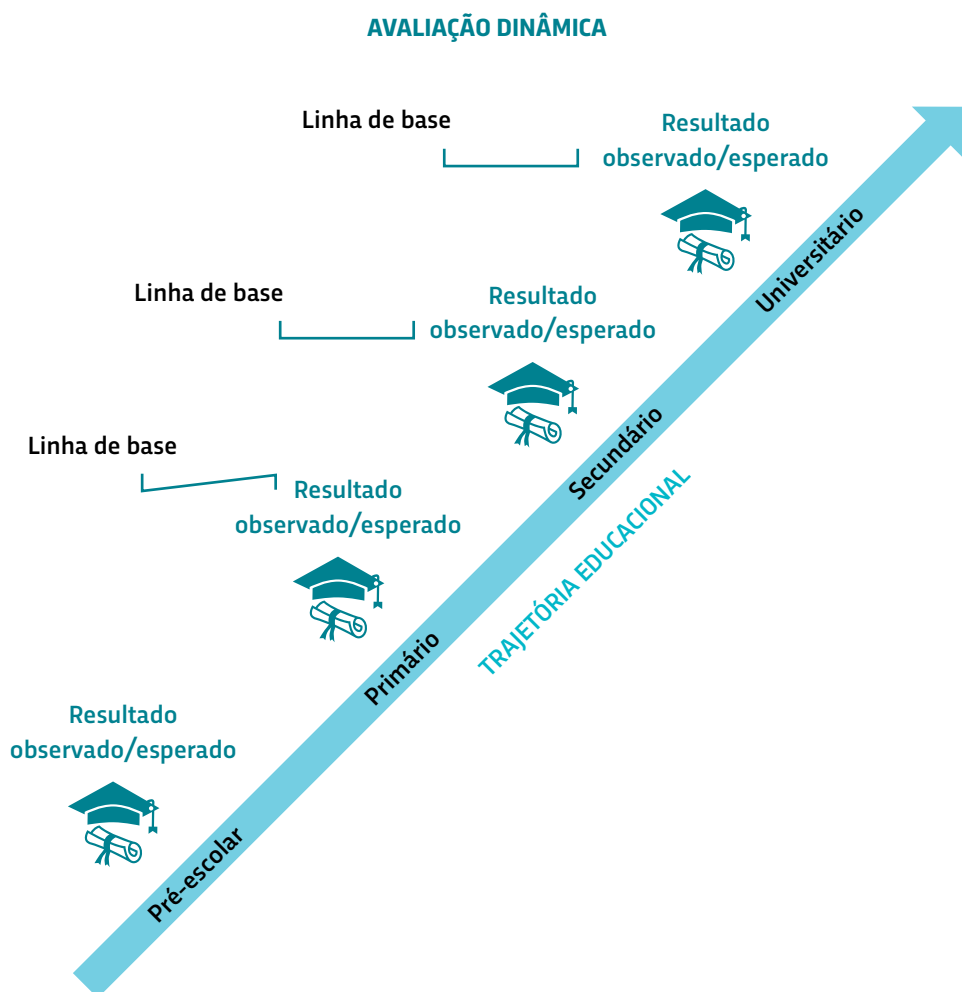
A comunidade também importa – tradições, valores, cultura, infraestrutura. Há um conjunto de elementos que intervêm na interação entre o suprimento de Educação e o resultado. Por exemplo, compare-se um estudante que precisa caminhar duas horas para chegar à escola no México com outro que anda apenas um quarteirão. Obviamente a capacidade do primeiro será limitada pelo cansaço, de tal modo que se os demais itens fossem idênticos, espera-se que estudantes inseridos nesse tipo de ambiente tenham um desempenho pior, mesmo se expostos a um sistema escolar de qualidade. Portanto, um ótimo modelo de Educação terá efeitos diferentes dependendo das características pessoais e familiares dos alunos.

Existe também um contexto social que tem a ver com comunicação, tecnologia, oportunidades de emprego, nível de desenvolvimento da sociedade e do país. Sendo assim, poderíamos avaliar o serviço da Educação em si – por exemplo, se o conteúdo e o currículo são adequados, se a gestão escolar está alinhada com os objetivos.

Porém, poderíamos avançar e medir não apenas o serviço da Educação, mas o efeito de sua interação com os estudantes, para gerar o que chamamos de um resultado na Educação.

Digamos que uma escola seja muito ruim, mas os estudantes são excelentes em termos de qualificações para absorver a Educação. Talvez o resultado não seja tão ruim, porque os alunos têm características positivas. Ou, ao contrário, que a escola seja muito boa, mas os alunos enfrentam desafios imensos – os resultados da Educação talvez não sejam tão bons, não porque o serviço prestado não é bom, mas por causa dos outros elementos.

Portanto, em avaliação, é possível tentar medir o efeito de uma mudança no serviço em relação a um resultado específico. Também existe uma ampla discussão acerca do que significa exatamente um resultado em Educação.



Uma das complexidades é definir resultado educacional. Poderia-se dizer que o objetivo da Educação é tornar as pessoas felizes e contribuir para o bem-estar e para a cidadania. Grandes objetivos como esses são importantes, mas, nesse caso, deveríamos medir o efeito da Educação na felicidade, no bem-estar etc.

Como isso é muito complexo, é normal ater-se a uma forma mais reduzida. Por exemplo, medir o efeito em matemática, letramento, ciência. Assim, quando entramos em detalhes específicos dos resultados, acabamos sendo muito menos ambiciosos que o propósito ou objetivo geral do sistema educacional.

Na prática, o que acontece é que, quando uma pessoa termina sua formação – com 18 anos, se completou a Educação Básica, ou com cerca de 24 anos, se seguiu para a Educação Superior –, a Educação que adquiriu continuará interagindo com características pessoais e familiares que serão diferentes do primeiro conjunto de características.

No primeiro conjunto, um elemento importante seria a relação entre a Educação da criança e a da mãe, mas, no segundo momento, quando o resultado da Educação é exposto ao contexto pós-Educação, há elementos mais importantes, como empreendedorismo, liderança e outros que determinam como cada pessoa utiliza a Educação que recebeu para gerar outro resultado.

*Quando uma pessoa termina sua formação – com 18 anos, se completou a Educação Básica, ou com cerca de 24 anos, se seguiu para a Educação Superior –, a Educação que adquiriu continuará interagindo com características pessoais e familiares que serão diferentes do primeiro conjunto de características*

Também há que se considerar as características da comunidade e do país. Digamos que exista um sistema excelente e que as pessoas sejam muito capazes de absorver essa Educação excelente, mas o país está mal em termos econômicos. É possível que o resultado final não seja tão bom, não por causa da Educação, mas devido à interação da Educação com o ambiente.

No limite, se desejamos realmente aferir felicidade, bem-estar, cidadania, é preciso esperar até o fim do ciclo de vida de uma pessoa para medir o quanto desses aspectos se deve somente aos primeiros elementos da Educação, o que, obviamente, é muito complexo.

## SÍNTESE

Então, em Educação, dada a complexidade, tendemos a simplificar, medindo alguns resultados e, em circunstâncias muito boas, tentando acompanhar uma mudança no serviço que afeta aquele resultado em particular. É um desafio imenso.

Nesse contexto, podemos pensar em ao menos três tipos de avaliação: (1) avaliação do serviço em si, independentemente do que ele gera adiante; (2) avaliação dos resultados da Educação; e (3), se formos muito ambiciosos, a qualidade do efeito final da Educação, mas, nesse caso, é preciso esperar muito tempo.

Assim, mesmo exames como o Pisa (Programme for International Student Assessment ou, em português, Programa Internacional de Avaliação de Estudantes), que são amplamente utilizados, focam um conjunto pequeno de indicadores, que são os resultados educacionais aos 15 anos de idade. Uma das desvantagens de usar apenas os números do Pisa é que, em geral, não se qualifica o efeito das características. O Pisa é interpretado linearmente como resultado do primeiro dos três elementos, o serviço da Educação, mas na realidade não é apenas isso.

Um elemento adicional de complexidade na avaliação em Educação é que se trata de um setor enorme, com uma vasta diversidade em países como os nossos. Em países como a Finlândia e em outras economias desenvolvidas onde a desigualdade é baixa, a diversidade não é muito grande, mas ainda existe e envolve muitos interesses.



## *Como mensurar o efeito de cada professor no resultado obtido por cada estudante? Ainda mais difícil é medir o efeito no resultado final da Educação*

Também é difícil medir sucesso em um sistema educativo. Como mensurar o efeito de cada professor no resultado obtido por cada estudante? Ainda mais difícil é medir o efeito no resultado final da Educação.

Um aspecto importante da Educação pública é que as pessoas não podem simplesmente buscar outra escola se o atendimento que recebem é ruim, o que seria um bom indicador sobre o atingimento ou não dos objetivos educacionais. Esse é outro elemento da complexidade: muitos alunos estão em uma escola porque não têm outra alternativa, não importando se a Educação é boa ou ruim.

Também ocorre que a responsabilização nos sistemas educativos é muito baixa por causa dos diferentes *inputs* envolvidos no processo de produção. Mais uma vez, surge a questão do professor – como medir o efeito de um professor em particular no resultado final, dado que ele interage com muitos outros elementos? Algumas avaliações de impacto permitem chegar perto da resposta, mas ainda assim trata-se de uma situação muito complexa.

Outro problema é que são necessários 12 anos para concluir um ciclo educativo. Dispor de medidas ao longo do caminho pode proporcionar informações interessantes, mas não são o resultado final.

Há a questão dos objetivos múltiplos. Muitos pensam que a Educação tem objetivos totalmente diferentes. Às vezes existe um debate intenso sobre isso.

O elemento final é que há uma enorme resistência à mudança em Educação. Ainda que se tenha um desenho de avaliação excelente, nem sempre é possível aplicá-lo.

## ALGUNS CASOS ILUSTRATIVOS

Os exemplos que abordarei não são casos de sucesso, mas o oposto. Este artigo tratará de duas avaliações e quais foram as falhas e os problemas em cada uma.

### AVALIAÇÃO DE DIRETORES DE ESCOLAS

Tivemos uma iniciativa no Ensino Fundamental no México que foi muito similar ao desenho do projeto Jovem de Futuro, do Instituto Unibanco – abordar diretamente os gestores escolares e tentar transmitir-lhes habilidades que lhes permitissem gerir melhor as escolas.

Tínhamos para isso uma plataforma eletrônica na internet e pedimos a cada diretor para se inscrever e fornecer 50 indicadores diferentes, da quantidade e sexo dos alunos em cada série a questões mais complexas, como a formação ou a ausência de formação dos professores. A simples inscrição foi um desafio imenso para eles, porque nunca haviam sido solicitados a fornecer informações desse tipo.

Em um segundo momento, os diretores foram chamados a priorizar: precisavam indicar um objetivo para uma mudança que desejassem implementar no ano acadêmico. Foram informados de que, ao final do ciclo, os resultados seriam mensurados e a performance da escola seria comparada com os objetivos informados. Desse modo, poderíamos verificar quais diretores haviam cumprido as metas e avaliá-los em relação a elas.

*Tivemos uma iniciativa no México que foi muito similar ao Jovem de Futuro – abordar os gestores escolares e tentar transmitir-lhes habilidades que lhes permitissem gerir melhor as escolas*

Obtivemos quatro resultados muito interessantes. Novamente, lembro que o caso é mais uma medida de problemas do que de êxito.

Cerca de 20% dos diretores não preencheram de início a lista de indicadores, então achamos que havia algo errado com a tecnologia da plataforma, com os equipamentos, com a conexão etc. Insistimos que completassem a tarefa: “Por favor, vocês são os únicos que faltam. Por gentileza, façam o exercício”. A resposta foi simplesmente: “Não quero. O que vocês vão fazer a respeito?”.

Estávamos mudando muitos diretores na época, em um processo delicado, com muitas implicações. Esses diretores acharam que simplesmente pediríamos de novo: “Por favor, cumpram a tarefa”, e que, se não o fizessem, nada aconteceria. Foram informados então de que seriam demitidos, mas, ainda assim, não entregaram os indicadores. Nossa resposta foi: “Não podemos demiti-los agora, porque precisamos esperar. Pedimos que reconsiderem. Não os demitiremos agora, vamos dar-lhes um pouco mais de tempo”. Não podíamos conceder muito mais porque os demais haviam cumprido o prazo e não podíamos dispensar a esse grupo um tratamento muito diferente. Porém, esse contingente declarou desde o início: “Não importa o que façam, não vamos fornecer a informação”.

*A lição dessa experiência é que mesmo um processo simples como esse que relatei leva muito tempo para ser internalizado pelo sistema de Educação – imagine-se o que aconteceria com uma avaliação de impacto muito complexa. As pessoas levam tempo para se adaptar a uma nova cultura. É inevitável falhar no início*

Cerca de 40% dos diretores fizeram o que foi pedido e forneceram os dados solicitados, mas percebemos que os objetivos eram muito baixos. Como nunca lhes pediram que a melhora atingisse certo nível, não pudemos fazer nada a respeito e tivemos que aceitar as metas que propuseram.

Houve outro grupo, de cerca de 30% dos participantes, que estabeleceram metas imensas. Desejavam parecer muito competentes e declararam: “Quero melhorar 100%”.

O último grupo, uns 10% dos gestores, fizeram um exercício que se aproximava do que esperávamos.

No final do ano, tivemos que demitir os 20% que não participaram. Também analisamos o perfil daqueles que haviam estabelecido objetivos muito baixos e muito altos e as razões relacionadas.

Notamos que os diretores mais velhos, com cerca de 55 anos de idade, se propuseram objetivos muito baixos, mas sabiam exatamente o que estavam fazendo: “Vocês me disseram para estabelecer um objetivo. Coloquei a meta de um milímetro e cumpri-a em 100%, portanto o objetivo foi atingido”.

O grupo que estabeleceu objetivos muito ambiciosos era formado por diretores muito jovens que queriam ser bem avaliados. A estes respondemos: “Sensacional, vão em frente”.

Os 10% restantes fizeram o que esperávamos.

A lição dessa experiência é que mesmo um processo simples como esse que relatei leva muito tempo para ser internalizado pelo sistema de Educação – imagine-se o que aconteceria com uma avaliação de impacto muito complexa. As pessoas levam tempo para se adaptar a uma nova cultura. É inevitável falhar no início.

Na rodada seguinte, tivemos que informar aos 40% que eles haviam estabelecido objetivos muito baixos: “Não calibramos a proposta muito bem da primeira vez, então agora estabelecemos um avanço mínimo”.

Os diretores que queriam ser bem-vistos e definiram objetivos muito ambiciosos perceberam nessa segunda rodada que não haviam conseguido cumprí-los. Foi um processo de aprendizagem para eles. Cinco anos depois, chegamos a um estágio em que o sistema funcionava como esperávamos.

## RESULTADOS DA AVALIAÇÃO

Este caso é uma avaliação de impacto que fizemos em uma intervenção com gestores escolares muito similar ao projeto do Instituto Unibanco. A estratégia foi totalmente diferente, e os resultados foram muito interessantes.

A ideia era fazer algo que gerasse uma mudança mensurável. Não podíamos medir felicidade, cidadania – aspectos inalcançáveis –, mas podíamos medir algo mais objetivo – competência matemática nos três anos do Ensino Médio.

Tínhamos uma linha de base formada por escolas selecionadas, três tipos de tratamento e um grupo de controle. Aplicamos uma prova de matemática em todos os alunos participantes de primeiro, segundo e terceiro anos do Ensino Médio. A diferença em relação à intervenção do Instituto Unibanco é que neste caso buscávamos o resultado final.

Havia uma linha de base e, se a pontuação dos estudantes melhorasse um certo percentual na segunda rodada de coleta de dados, pagaríamos aos diretores uma quantia em dinheiro diretamente na conta bancária.

A segunda intervenção concedia exatamente a mesma recompensa, mas somente aos professores e se suas classes atingissem um certo percentual de melhora.

O terceiro modelo incluía toda a escola: diretores, professores e estudantes receberiam um pagamento se houvesse melhora. O valor era quase o mesmo, mas a filosofia era diferente. Dissemos: “Não nos importamos como a melhoria será atingida; se conseguirem, obterão o prêmio”.

Convidamos três especialistas de renome internacional. Eles criaram o desenho da avaliação, seguida de sua implementação e interpretação, e forneceram os indicadores. Eles já haviam publicado um artigo no *Journal of Political Economy*, então o projeto cumpriu seu propósito acadêmico e também o propósito que desejávamos no sistema de Educação.

Os três modelos apresentaram efeitos, porém diferentes.

O maior e mais impressionante foi o caso em que toda a escola era premiada. O segundo foi o caso dos diretores, também com um efeito muito grande. O terceiro foi o dos professores,

*Geramos uma expectativa de algo que era inviável, porque, em primeiro lugar, não deveríamos pagar as pessoas – inclusive os alunos – para fazer o que deveriam fazer*

com um efeito pequeno. Esse experimento gerou dois conjuntos de informações muito importantes.

O primeiro foi relacionado ao dinheiro, que foi pago no final. Não pagamos nada durante um ano acadêmico inteiro. Dissemos aos participantes no início que seriam avaliados no final e o experimento durou um ano e meio.

Isso nos diz que, mesmo que não pagássemos nada, a simples promessa mobilizava a escola para obter um resultado que esperávamos que acontecesse de qualquer modo, com os recursos disponíveis. Era questão de alinhar motivação e incentivo para obter o resultado. Seria interessante comparar isso com o experimento do Instituto Unibanco.

A segunda informação que obtivemos é que também existem possíveis incentivos perversos ao fazer isso. Antes de mais nada, é preciso dizer que as pessoas no grupo de tratamento não estavam reclamando, elas queriam estar no experimento. Quem reclamou foram os indivíduos do grupo de controle, que desejavam estar no experimento para também obter o incentivo. Houve, portanto, muita pressão relacionada ao dinheiro, mas no final todos diziam: “Fico feliz que este programa esteja em curso. Ele acontecerá para sempre, certo?”. Geramos uma expectativa de algo que era inviável, porque, em primeiro lugar, não deveríamos pagar as pessoas – inclusive os alunos – para fazer o que deveriam fazer.

Foi um ótimo experimento de interesse acadêmico, mas, no final, não foi útil para a tomada de decisões. Gerou conhecimento – pode-se provocar um impacto a partir de motivação –, mas, ao considerar o tamanho do esforço, não gerou nada que modificasse o sistema escolar.

# *SENSO COMUM VERSUS EVIDÊNCIAS: RISCOS DA CULTURA DO “ACHISMO”*

**EDUARDO DESCHAMPS** é secretário de Estado da Educação de Santa Catarina. Possui graduação, mestrado e doutorado em engenharia elétrica pela Universidade Federal de Santa Catarina (UFSC). Desde 1990 é professor efetivo do departamento de engenharia elétrica e de telecomunicações da Universidade Regional de Blumenau (Furb), da qual já foi reitor. Em 2011, foi nomeado secretário-adjunto de Educação de seu estado, assumindo a pasta no ano seguinte. Foi presidente do Conselho Nacional de Secretários de Educação (Consed) até novembro de 2016 e atualmente preside o Conselho Nacional de Educação (CNE).

Embora seja engenheiro, atuo na Educação, principalmente no nível superior, desde 1990. Era pesquisador de laboratório, e naquele ambiente as experiências que realizava tinham resultados claros, inquestionáveis. Por isso, quando cheguei à Secretaria de Estado da Educação de Santa Catarina, em 2011, o que mais me impactou foi a dificuldade – não apenas por parte dos gestores, mas de todos que atuam na área – de trabalhar com dados e utilizá-los a favor da construção de ações. Observei isto tanto no âmbito das redes quanto, mais tarde, à frente do Conselho Nacional de Secretários de Educação (Consed): as equipes enfrentam muitos obstáculos para, a partir dos dados e informações que têm, traçar as políticas adequadas. É comum utilizarem apenas os dados que confirmam suas teorias (ver também artigo de Cláudio de Moura Castro, p. 192), e não aqueles que poderiam, eventualmente, interferir no cenário e produzir ações diferenciadas do ponto de vista das políticas educacionais.

Verdade é que utilizar bem os dados não constitui tarefa simples. É um processo que exige dos gestores do âmbito da Educação não apenas capacidade gerencial – ou seja, aplicar so-



luções técnicas que já existem –, mas também o que chamamos de liderança adaptativa: se, por um lado, pode ser válido utilizar experiências testadas anteriormente, por outro talvez não haja fórmulas prontas. Nesse cenário, o uso de evidências pode ser decisivo para o sucesso de uma rede (ver também artigo de Ilkka Turunen, p. 208).

O principal embate para um gestor público no campo da Educação, e talvez no das políticas públicas do Brasil, como um todo, se dá entre o senso comum e as evidências. É muito frequente, ao tentar trabalhar com evidências, que o gestor se depare com o senso comum envolvendo os tópicos em discussão. Por exemplo: qual é o número ideal de alunos em sala de aula? Qual evidência aponta aquele número como ideal? Seriam 40, 30, 20, 15, cinco alunos? Isso faz toda a diferença na hora de moldar e gerir um sistema ou uma rede, pois, dependendo do número de alunos definido como o que dará o melhor resultado educacional em uma determinada escola, a rede terá um tamanho ou outro, com mais ou menos professores. Ocorre que o recurso público para organizar esse sistema, proveniente dos impostos pagos pela sociedade todos os dias, é finito. Sem a utilização de evidências explícitas e detalhadas sobre o melhor número de alunos em sala, corre-se o risco de dimensionar a rede de um tamanho irreal apenas por ouvir dizer que o ideal é 20 alunos. A consequência é o desperdício de recursos.

Outro questionamento diz respeito ao uso de novas tecnologias nas escolas. Afinal, elas funcionam ou não? É necessário

*Sem a utilização de evidências explícitas e detalhadas sobre o melhor número de alunos em sala, corre-se o risco de dimensionar a rede de um tamanho irreal apenas por ouvir dizer que o ideal é 20 alunos*

haver um computador por aluno? Um laboratório de informática é indispensável? A lousa digital faz diferença na aprendizagem? Uma pesquisa publicada recentemente aponta que a simples disponibilidade de tecnologia nas escolas não traz resultados; na melhor das hipóteses, produz impacto muito baixo. Se o professor não estiver preparado para utilizar tecnologia, e se a tecnologia empregada não for adequada para o aluno nem para o tipo de ação pedagógica pretendida, muito dinheiro será investido e o impacto será zero. É por isso que evidências são tão importantes.

## **PONTOS SENSÍVEIS: SALÁRIO DO PROFESSOR E TEMPO INTEGRAL**

As evidências são igualmente necessárias em questões profundamente sensíveis, como o salário dos professores. Esse salário é adequado? Uma pesquisa, ainda inédita, do Ministério da Educação (MEC) a partir de informações da ferramenta Relação Anual de Informações Sociais (Rais), implementada pelo Ministério do Trabalho, desmente dados que circulam pela sociedade acerca dos salários dos professores. Os gestores estão devidamente preparados para avaliar os impactos em relação à questão salarial dos professores? O que está em jogo não é se o professor deve ganhar mais e melhor: indaga-se, isso sim, a que valor é preciso chegar, o que está ao alcance fazer e quais evidências há a respeito do impacto da melhoria de salário sobre os resultados educacionais. Várias pesquisas apontam que não há correlação direta. Obviamente, o salário de professor é uma questão básica, mas não é apenas por meio de ações nesse âmbito que o resultado educacional melhorará.

Em 2011, a Secretaria de Educação de Santa Catarina teve acesso a um estudo realizado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) em Santa Catarina<sup>1</sup>. O estudo analisava as políticas educacionais do estado e

---

1. OCDE: *Avaliação de Políticas Nacionais de Educação – Estado de Santa Catarina, Brasil, 2010.*

concluía haver excessiva preocupação com insumos, e não com resultados.

O tempo integral é mais um componente importante dessa equação. Vários dados, inclusive resultados recentes do Índice de Desenvolvimento da Educação Básica (Ideb) em Pernambuco, apontam que, se as crianças passarem mais tempo na escola, cresce a probabilidade de melhorar o resultado. Esse tempo ampliado, porém, não pode se traduzir simplesmente em empilhamento de disciplinas. É preciso integrar formas de atuação com as disciplinas, os conteúdos e o desenvolvimento de competências.

Finalmente, o que dará resultado efetivo? Trabalhar conteúdo ou desenvolver conhecimento com o estudante? Nosso atual modelo de escola é prussiano e data do século 19, articulando um sistema de tempos e métodos: é preciso oferecer determinados conteúdos em certo intervalo de tempo. Quem aprendeu, aprendeu; quem não aprendeu fica para trás. Sabe-se, igualmente, que esse modelo não funciona mais, e que é preciso modificá-lo. Em vez de privilegiar a quantidade imensa de conteúdo, melhor seria perguntar qual é o conhecimento que será trabalhado com o aluno. Recentemente, o Consed e a União Nacional dos Dirigentes Municipais de Educação (Undime) promoveram vários seminários estaduais para debater a Base Nacional Comum Curricular. Deles resultou um relatório entregue ao ministro da Educação. Um dos pontos principais foi a recomendação de trabalhar com menos amplitude e mais aprofundamento. É isso que tem feito a diferença no mundo inteiro; no entanto, o Brasil está na contramão, investindo demais em conteúdo.

## **IMPORTÂNCIA DE AVALIAÇÕES PRÓPRIAS**

Felizmente, já dispomos de algumas sinalizações importantes para organizar nossas escolas, redes e sistemas. O Ideb é um deles. No entanto, para os estados que não dispõem de sistemas próprios de avaliação do Ensino Médio, ele tem pouca serventia, porque o Sistema de Avaliação da Educação Básica (Saeb) é realizado por amostragem. Este ano, a amostra do Ideb do Ensino

*Se não houver gestores realizando esse balanço em cada escola, enquanto a secretaria se concentra naquelas que efetivamente têm problemas maiores, o sistema como um todo não avançará*

Médio de Santa Catarina praticamente não teve alunos de tempo integral, que correspondem a 10% da rede. Apesar de essa ser uma das principais políticas para avançar, não foi possível transformar tal informação ou dado em uma ação efetiva. Por isso é tão importante que cada estado também tenha um sistema próprio de avaliação.

Tempos atrás, estivemos em uma escola no interior de Santa Catarina e encontramos a equipe reunida em uma sala, avaliando todos os resultados do Ensino Fundamental, cruzando dados do Saeb, Ideb e do Censo Escolar, bem como números de reprovação e abandono. A partir de todas aquelas informações, desenhava-se um plano de ação para aquela escola. Quantas outras efetivamente estão preparadas para fazer isso, ou já fazem? Não basta dizer que isso é trabalho da Secretaria de Estado da Educação. Nossa rede conta com cerca de mil escolas; se não houver gestores realizando esse balanço em cada uma delas, enquanto a secretaria se concentra naquelas que efetivamente têm problemas ou dificuldades maiores para se organizar, o sistema como um todo não avançará.

Ao analisar a importância de basear-se em evidências para melhorar resultados de aprendizagem, não há como deixar de lado a reformulação do Ensino Médio. Esse segmento é a maior prova de que, no Brasil, não se trabalha com evidências. Sabe-se que 80% dos estudantes brasileiros não vão para a universidade; no entanto, o Ensino Médio é todo focado no preparo do aluno para ingressar no Ensino Superior. O Brasil é o único país no mundo que tem esse modelo. É preciso utilizar as várias informações disponíveis para realizar uma mudança.

*Falta nas redes pessoal qualificado no uso de evidências [...] Quando o dado chega à rede, os gestores, professores e demais atores não sabem utilizá-lo*

## **O PERIGO DO “ACHISMO”**

Nesse percurso, o maior obstáculo é a cultura do “achismo”. Todo gestor da Educação já viveu muitas situações em que, mesmo com um dado à mão para justificar certas proposições, ouvia de seus interlocutores: “Acho que isso não vai dar certo”. Se perguntava por que, escutava: “Porque sempre foi assim; já tentamos algo semelhante e não funcionou”. No entanto, a busca constante deve ser por dados e informações. Um exemplo dessa cultura envolveu um debate recente em Santa Catarina sobre manter ou não os exames finais; esses exames não contam como dias letivos por determinação da lei. Para mantê-los, é preciso comprimir o calendário escolar. Em defesa dele, os gestores argumentavam que haveria reclamações de alunos reprovados por margens mínimas. “Sem o exame final, teremos aumento na taxa de reprovação”, afirmaram. Ora, quantos casos são esses? Quantas vezes isso ocorreu? Onde? São perguntas simples, para as quais não havia respostas. A rede de Santa Catarina tem 500 mil alunos. Se houver mil casos assim, não existe justificativa para manter um sistema que não traz nenhum benefício – apenas comprime o calendário escolar e impede o professor de oferecer a recuperação paralela ao aluno, porque permanece à espera do exame final.

É preciso que os gestores comecem a pensar em informações dessa natureza. Mas a verdade é que falta nas redes pessoal qualificado no uso de evidências. Há pesquisadores, mas, quando o dado chega à rede, os gestores, professores e demais atores não sabem como utilizá-lo (ver também artigos de Paula Louzано, p. 156, e de Telma Vinha, p. 172).

Outro obstáculo é a existência de barreiras políticas. É comum que os resultados não sejam imediatos, mas venham em médio e longo prazo. Separar um grupo de controle, aplicar uma pesquisa, trabalhar na elaboração das evidências e avaliar resultados são condutas que exigem tempo de maturação (ver também artigo de Miguel Székely, p. 226). Começando pelo município de Sobral, o Ceará, que hoje se beneficia de excelentes resultados educacionais, conseguiu estabelecer uma continuidade de políticas que já dura mais de 15 anos. É difícil conquistar tal estabilidade.

Existe ainda grande dificuldade na articulação entre sistemas. Há a escola, a rede de escolas (que pode ser estadual, municipal ou privada) e o sistema, que é o conjunto dessas redes; nem sempre a rede municipal dialoga com a estadual, e nenhuma delas dialoga com a rede privada. É preciso pensar que se trata de um sistema de ensino, e não é questão de ser a favor do público ou do privado: todos congregam estudantes brasileiros. Todos são professores brasileiros. E, por incrível que pareça, muitos dos professores que atuam nas escolas privadas trabalham também nas públicas, de maneira que não é possível fazer tal distinção: é tempo de pensar em uma Educação brasileira em sua totalidade.

*Outro obstáculo é a existência de barreiras políticas. É comum que os resultados não sejam imediatos, mas venham em médio e longo prazo. Separar um grupo de controle, aplicar uma pesquisa, trabalhar na elaboração das evidências e avaliar resultados são condutas que exigem tempo de maturação*

## HUMILDADE PARA AVANÇAR

As universidades têm grande contribuição a dar em relação à coleta e ao uso de evidências. Porém, muitas vezes revelam-se profundamente ideologizadas. Em Santa Catarina, a Secretaria de Educação cultiva uma boa relação com elas, mas outros estados nem tanto. Certa vez, perguntei a uma professora sobre o andamento de sua formação para o Pacto Nacional pela Alfabetização na Idade Certa (Pnaic). Ela relatou dificuldades iniciais. “O pessoal da universidade achava que sabia tudo, mas não tinha a mínima ideia de como se organizava uma sala de aula para trabalhar a alfabetização”, disse essa professora. “Felizmente eles tiveram a humildade de construir conosco a melhor metodologia.” Ou seja: se não houver humildade por parte de quem está na universidade, o sistema não mudará e não será possível aplicar resultados de pesquisa; aliás, nem sequer se conseguirá realizar pesquisas nas escolas.

É necessário lembrar ainda que, no território das avaliações, principalmente nacionais, há imensa dificuldade para conseguir microdados que poderiam fazer a diferença. Sem eles, torna-se mais difícil para o gestor promover avanços em sua rede.

Estes são alguns dos maiores desafios que se colocam aos gestores da Educação:

*É necessário lembrar ainda que, no território das avaliações, principalmente nacionais, há imensa dificuldade para conseguir microdados que poderiam fazer a diferença. Sem eles, torna-se mais difícil para o gestor promover avanços em sua rede*

- Necessidade de adaptar pesquisas, resultados ou dados ao contexto de determinada região, pois o Brasil é muito desigual.
- Importância de priorizar problemas ou soluções: é preciso desenvolver um plano de ação muito bem estruturado.
- Tempo de produção do conhecimento e sua adaptação para aplicação. O Brasil tem indicadores muito frágeis na Educação, e precisamos de resultados consistentes, e rapidamente. Enquanto o país não conseguir desenvolver todas as ações necessárias do ponto de vista da pesquisa, uma saída possível talvez seja aplicar algumas ações básicas que deram certo em outras nações (ver também artigo de André Portela, p. 100).
- Redesenho para resultados não esperados. O que fazer se o resultado de determinada ação não foi bom? Esconde-se e encerra-se o processo? Em nossa opinião, não; é preciso, isso sim, redesenhá-lo e seguir reavaliando.

Finalmente, escola que funciona no mundo inteiro tem currículo, professor bom e boa gestão de unidade, rede e sistema. Porém, para que isso traga bons resultados, é indispensável que tenhamos um pacto político envolvendo toda a sociedade: é tempo de parar de reinventar a roda no Brasil e utilizar os dados à mão para fazer a diferença na Educação.



# *PERNAMBUCO: GESTÃO E EVIDÊNCIAS*

**FREDERICO AMANCIO** é secretário de Estado da Educação de Pernambuco. É auditor fiscal de carreira pela Secretaria da Fazenda de Pernambuco. Em 2008, integrou a equipe responsável pela aplicação de um novo modelo de gestão na saúde pública do estado, enquanto estava à frente da pasta. Também dirigiu as secretarias estaduais de Desenvolvimento Econômico e Planejamento e Gestão.

Quando foram divulgados pelo Ministério da Educação (MEC) os resultados do Índice de Desenvolvimento da Educação Básica (Ideb) relativos a 2015, o estado de Pernambuco despontou como o que possui o melhor Ensino Médio do Brasil, com nota 3,9, a mesma de São Paulo. A rede festejou o resultado, porém, logo depois, já se perguntava como atuar para seguir avançando. Por “avançar”, entendemos construir caminhos, dar passos e galgar degraus, com a consciência de que a trajetória é longa e não há acomodação possível, pois existe muito a fazer, ainda, pela Educação no Brasil.

É necessário revisitar a história recente do estado de Pernambuco para compreender esse resultado, considerando, inclusive, o encontro entre a gestão e a evidência e o impacto desse encontro sobre o desempenho da rede.

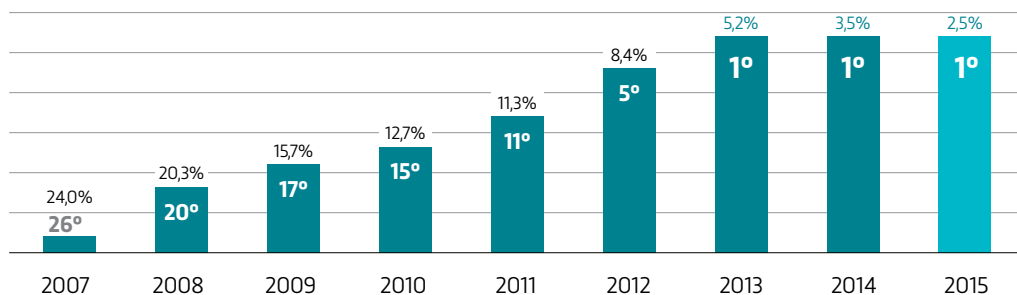
A história desse sucesso no Ideb que o estado comemorou em 2016 havia começado em 2007. Naquele ano, Pernambuco tinha alguns dos piores indicadores educacionais, se comparados aos dos demais estados do Brasil. Tornava-se inaceitável que não reagisse a isso, iniciando uma mudança em seu processo de

construção de políticas. Em um primeiro momento, a opção foi transformar a gestão da rede escolar. Com ela, veio a elaboração de um plano estratégico a partir de um diagnóstico que utilizou a produção científica existente no país e no estado.

Pernambuco escolheu enfrentar o ponto mais objetivo: a relação entre o baixo desempenho no Ideb e as altíssimas taxas de abandono escolar. O estado situava-se em 26º lugar em abandono escolar no Ensino Médio e em 21º lugar no Ideb. Do ponto de vista da análise subjetiva, havia a questão da ausência de protocolos para qualificar a atividade em sala de aula, que envolve o próprio trabalho pedagógico no dia a dia, e também o fato de que uma parte dos estudantes do Ensino Médio acha a escola “chata”. Muitos a abandonam porque gestores e professores não conseguiram construir, ao longo do tempo, caminhos para tornar o ensino mais atrativo para o estudante. Em 2007, essa preocupação tornou-se central, e o estado avançou também nesse ponto. Naquela época, de cada cem estudantes que iniciavam o Ensino Médio no estado, a rede perdia 24, a maior parte no primeiro ano. No entanto, desde 2013, Pernambuco registra a menor taxa de abandono escolar do Ensino Médio no Brasil. Trabalhar abandono escolar em um estado que ainda tem alguns dos piores indicadores sociais do Brasil é tarefa de extrema complexidade. Todos sabem que os aspectos sociais têm impacto direto no dia a dia da escola, o que dá a dimensão do desafio que enfrentamos, com sucesso.

### TAXA DE ABANDONO 2015 – ENSINO MÉDIO

PERNAMBUCO TEM A MENOR TAXA DE ABANDONO DO ENSINO MÉDIO – REDE ESTADUAL



Fonte: Inep/MEC

Outra face do resultado ruim que existia na época dizia respeito à eficácia da aprendizagem. Para equacionar esse ponto, o estado construiu alguns eixos estratégicos que deveriam ser implantados em toda a rede, começando pela melhoria da infraestrutura escolar, de maneira a tornar a escola mais capacitada para trabalhar com estudantes secundaristas. Naquele momento, já havia uma experiência que trouxera bons resultados: a adoção da Educação em tempo integral em algumas escolas desse segmento.

Em 2004, o estado de Pernambuco iniciou uma experiência pioneira no Brasil de Educação em tempo integral envolvendo uma escola de Ensino Médio, o Ginásio Pernambucano. Os primeiros números dessa experiência foram muito positivos; a título de comparação, os gestores compilaram experiências em outros países, todas relatando os bons resultados da exposição dos estudantes a um período pedagógico mais amplo.

Não bastava, porém, ampliar o tempo na escola oferecendo mais do mesmo. Ao longo das discussões sobre o melhor modelo para um Ensino Médio em tempo integral, chegou-se à conclusão de que não era suficiente olhar apenas para componentes curriculares; seria importante incorporar ao currículo alguns aspectos da formação socioemocional, tão discutidos e demandados hoje pelos estudantes.

Foi decidido, então, que o estado ampliaria a experiência do Ensino Médio em tempo integral, trabalhando a melhoria dos insumos pedagógicos como forma de tornar a escola mais atrativa e de aprimorar o processo de aprendizagem. Seria necessário, ainda, que houvesse uma política voltada ao acompanhamento permanente de toda a rede.

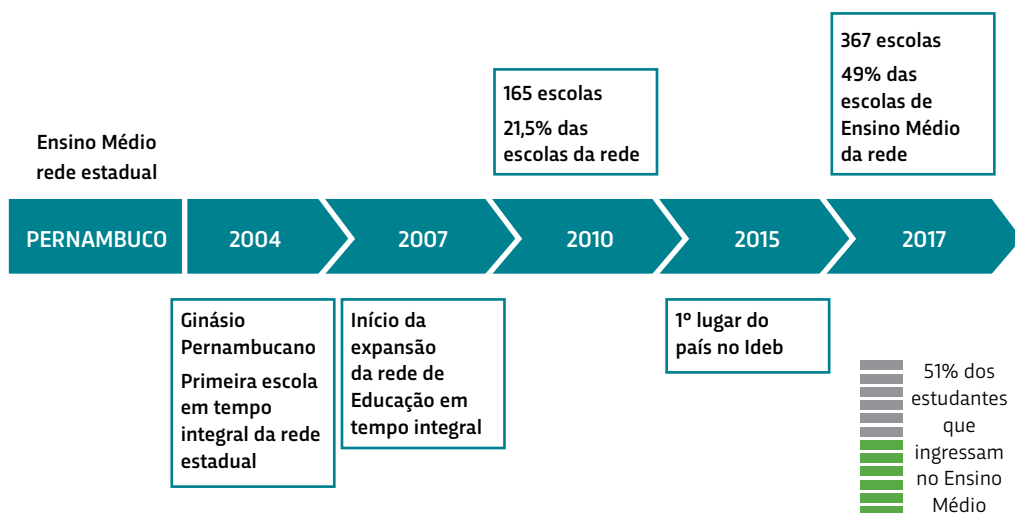
## **EDUCAÇÃO EM TEMPO INTEGRAL E PROFISSIONAL**

No primeiro governo de Eduardo Campos (2007-2010), 165 escolas, correspondendo a 21,5% da rede do Ensino Médio, passaram a oferecer ensino em tempo integral. Disso já resultou um salto expressivo no desempenho do estado no Ideb. O passo seguinte foi implantar o modelo em 300 escolas desse segmento, visando

## Pernambuco uniu duas estratégias [...] com resultados muito positivos: Educação Profissional e tempo integral

ter pelo menos uma unidade desse ciclo em tempo integral em cada município do estado. Pernambuco tem 184 municípios, e essa meta foi cumprida em 2016, inclusive no distrito estadual de Fernando de Noronha. Na realidade, esse número é maior: o estado dispõe atualmente de 367 escolas operando nessa modalidade, sendo 35 Escolas Técnicas Estaduais (ETEs). Com o objetivo de criar novas perspectivas para os estudantes, a Educação Profissional e tempo integral é intensamente debatida no âmbito do Ensino Médio. Em 2007, o estado tinha apenas seis escolas técnicas estaduais. Hoje já são 35 e há mais dez em construção neste momento. Pernambuco uniu duas estratégias que tendem a ser vitoriosas, com resultados muito positivos: Educação Profissional e tempo integral. Todas as escolas técnicas do estado funcionam em tempo integral, portanto, na modalidade técnico integrado.

### EDUCAÇÃO INTEGRAL – HISTÓRICO DA IMPLANTAÇÃO



Fonte: elaboração do autor.

Outro aspecto importante da história de sucesso do Ensino Médio de Pernambuco está relacionado à adoção de uma política que recebeu o nome de “Pacto pela Educação”. Trata-se de uma iniciativa que adapta ferramentas de gestão por resultados para o contexto da rede estadual. Adota conceitos e práticas gerenciais consolidados, especialmente melhoria contínua, monitoramento intenso e foco na resolução de problemas. Pacto é uma palavra muito importante para a rede nesse contexto, porque há um entendimento de que uma política na Educação – e talvez em qualquer área – não alcança resultados se não houver envolvimento de todos. Resultados não acontecem por decreto do governador nem por decisão do secretário: é preciso, sim, que haja a decisão política do governador e o comprometimento do secretário, mas, acima de tudo, é imprescindível envolver toda a rede, de diretores de escolas a professores e estudantes. É assim que o estado trabalha sua política de gestão por resultados; não basta realizar apenas um acompanhamento de números e indicadores: é vital envolver toda a rede no processo. Daqui a alguns anos, quem visitar uma das escolas de Pernambuco perceberá como essa cultura foi incorporada. O Jovem de Futuro não está implantado ainda em Pernambuco, mas o estado tem uma estratégia de atuação muito próxima do trabalho do projeto. Acompanha-se um conjunto de indicadores que analisa não só os resultados gerais do estado, mas também a correlação entre eles, bem como seu impacto sobre o aprendizado em cada escola.

Ao longo dos anos, Pernambuco conseguiu construir e implantar um sistema de avaliação próprio, que segue a metodologia do sistema nacional. Graças a essa conquista, o estado dispõe de informações de todas as escolas desde 2008, o que permite elaborar uma base de dados para transformar impressões em evidências a partir da própria base primária da rede. A Secretaria de Educação sabe o que se passa em cada estabelecimento e utiliza os dados recolhidos para definir estratégias e políticas. Acima de tudo, porém, esses dados são trabalhados internamente nas escolas. Cada gestor conhece os resultados, bem como os indicadores, e dialoga com professores e estudantes para que possam avançar.

## AVANÇO NAS ESCOLAS REGULARES

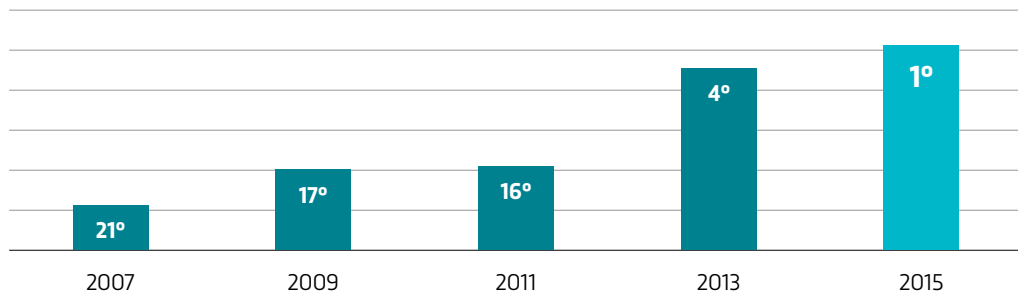
Se por um lado é verdade que houve avanços muito expressivos nas escolas em tempo integral em Pernambuco, caracterizando essa política como extremamente poderosa para o alcance de bons resultados, por outro é indiscutível que houve avanços também notáveis nas escolas regulares. Isso ocorreu porque, da mesma forma, apropriando-se de seus indicadores, a escola passa a conhecer a própria dinâmica de funcionamento. Cada uma das 761 escolas de Ensino Médio tem em seus computadores acesso ao sistema que monitora os indicadores do Pacto pela Educação. Essas informações permitem o acompanhamento e correção das ações planejadas pelas diferentes instâncias, desde a escola até a Secretaria de Educação.

Recentemente, esse trabalho foi estendido às 473 escolas que ofertam o Ensino Fundamental na rede estadual de Pernambuco, de maneira a configurar uma estratégia maior do estado, e também para criar um plano de ação em cada escola. O objetivo é permitir que cada uma, individualmente, possa enxergar qual caminho deve trilhar para melhorar seus resultados. O planejamento estratégico foi possível com a união da gestão da rede e da gestão escolar, ambas empenhadas no alcance de resultados bastante positivos ao longo dos anos. Houve crescimento consistente do desempenho de Pernambuco, que, em 2007, estava no 21º lugar no Ideb do Ensino Médio e, segundo o Ideb de 2015, chegou ao primeiro lugar da rede pública, ao lado do estado de São Paulo. Pernambuco é um dos dois estados que apresentaram melhora no Ideb todos os anos, ao longo dos últimos dez anos.

*Pacto é uma palavra muito importante para a rede nesse contexto, porque há um entendimento de que uma política na Educação [...] não alcança resultados se não houver envolvimento de todos*

## ÍNDICE DE DESENVOLVIMENTO DA EDUCAÇÃO BÁSICA PERNAMBUCO CHEGA AO PRIMEIRO LUGAR NO IDEB

Ensino Médio – Rede estadual



Fonte: elaboração do autor, com dados do Inep/MEC.

Tão importante quanto o número final é a percepção de que a rede como um todo está avançando. Isso é importante não apenas para o secretário ou para o governador; é vital para a própria rede, porque atua como estímulo para seguir nessa trajetória ascendente.

### EM BUSCA DA EQUIDADE

Dois outros dados são bastante relevantes e devem ser mencionados com o objetivo de mostrar que o trabalho realizado na rede pública também tem impacto significativo em comparação com a rede privada. Em 2007, a diferença de resultados entre o Ideb da rede pública e o da rede particular do estado era de 2,6. Hoje estamos apenas a um ponto de diferença, de maneira que Pernambuco é o estado do Brasil com a menor diferença entre os resultados de rede pública e privada no Ensino Médio. O estado ainda tem um longo caminho a percorrer, porém não se pode perder de vista a meta de buscar qualidade com equidade. Qualidade de ensino não significa apenas melhoria de resultados de aprendizagem, mas também garantia de oportunidades iguais para os estudantes de rede pública em comparação com os estudantes da rede privada.



## DIFERENÇA ENTRE REDE PRIVADA E REDE ESTADUAL

IDEB ENSINO MÉDIO



Fonte: elaboração do autor, com dados do Inep/INEC.

A rede estadual de Pernambuco tem muito a comemorar, e certamente os avanços foram festejados, porém o estado precisa continuar nessa trajetória – afinal, em Pernambuco e no Brasil, estamos ainda muito distantes da tão sonhada Educação de qualidade. Agora, o desafio se torna mais difícil: embora o estado tenha melhorado sob o ponto de vista de gestão, a próxima etapa exigirá que galguemos alguns degraus, e o esforço será ainda maior.

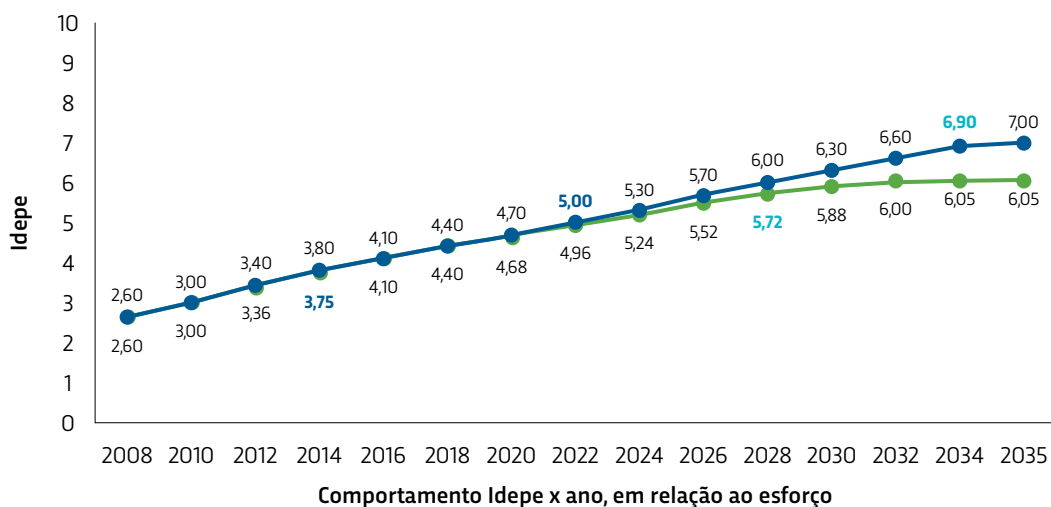
É comum pensar na Educação de qualidade como uma montanha muito alta cujo topo o país jamais conseguiria alcançar, tamanhos os desafios até lá. Como gestores, é preciso conclamar as equipes a juntas construir uma escada para chegar ao cume, subir um degrau de cada vez e comemorar cada etapa vencida.

Para avançar, além de seguir com o trabalho de gestão que o estado vem realizando, será fundamental utilizar as evidências para construir políticas consistentes. É preciso haver o encontro da gestão com as evidências. Em 2015, à frente da secretaria, começamos um trabalho intenso nesse sentido. À época, estavam disponíveis alguns estudos, baseados em dados do estado e em conexão com evidências, sobre qual seria o ponto de saturação no que diz respeito a resultados do Ensino Médio; em outras palavras, a partir de que nota teríamos dificuldade para obter progressos. Esses estudos eram preocupantes. Em primeiro lugar, porque indicavam que, no contexto do estado, seria penoso

*Como gestores, é preciso conchamar as equipes a juntas construir uma escada para chegar ao cume, subir um degrau de cada vez e comemorar cada etapa vencida*

avançar a partir da nota 6,90 no Ideb. Ocorre que tais estudos não haviam considerado a base de dados construída ao longo dos últimos anos em Pernambuco; quando se olhava para ela a realidade era ainda mais complexa e indicava como ponto de atenção a nota 5, um patamar muito baixo. Nossa equipe detectou que o ponto de saturação do estado se daria a partir da nota 5,72. Mas não poderia ser esse o teto: era preciso almejar mais.

### COMPREENSÃO DOS DESAFIOS PARA MELHORIA CONTÍNUA



● Projeção com base nos indicativos atuais de saturação  
Ponto 1: 2014 (Ideb 3,75)

**Crescimento torna-se saturado a partir de 5,72**

● Crescimento desejado: PE2035 com Ideb/Idepe 7,0

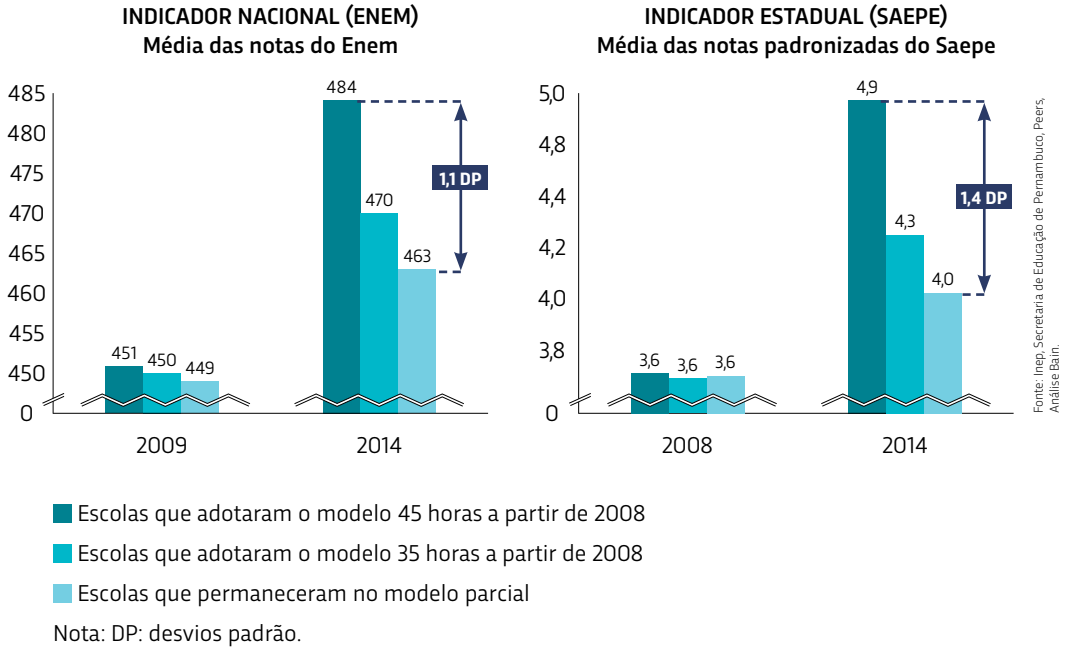
Fonte: Sistema de Avaliação da Educação Básica de Pernambuco (Saape).

Havia enorme preocupação com o baixo desempenho, e essa preocupação começava no Ensino Fundamental. No estado de Pernambuco, a maior parte das escolas desse ciclo é municipal. Nos anos iniciais, só há escolas municipais; nos anos finais, 60% das matrículas são municipais, todas com desempenho muito baixo. Portanto, era sabido que os estudantes chegavam ao Ensino Médio com muitas lacunas no processo de aprendizagem. Para suprir essas carências, uma das estratégias mais importantes adotadas no Ensino Médio foi o reforço escolar. Embora necessário, isso é muito preocupante. Nosso diagnóstico concluiu que, para avançar, era vital também transformar nosso Ensino Fundamental para garantir avanços cada vez mais efetivos e consistentes. Isso envolveria construir estratégias para apoiar os municípios.

Hoje, o estado vem trabalhando fortemente nesse aspecto e conta com as evidências. Pernambuco atua em duas frentes: primeiro, analisa os resultados colhidos na própria rede ao longo dos anos. Recentemente, o estado desenvolveu um trabalho em parceria com o Instituto Natura e o Instituto Sonho Grande, com participação de duas empresas de consultoria, a Peers e a Bain; considerando o ponto de vista do resultado, além dos custos e da eficiência do Ensino Médio. Esse estudo pode ser de grande valia para outros estados, ajudando-os a conhecer um pouco mais sobre a Educação em tempo integral, pois Pernambuco dispõe de uma base de dados importante, não só pela série histórica, mas porque já não se trata mais de um pequeno experimento. Temos 49% das escolas do Ensino Médio em tempo integral. Deixou de ser uma experiência isolada para tornar-se algo presente em toda a rede. As evidências comprovam que essa estratégia é poderosa.

*Nosso diagnóstico concluiu que,  
para avançar, era vital também  
transformar nosso Ensino Fundamental  
[...] Isso envolveria construir estratégias  
para apoiar os municípios*

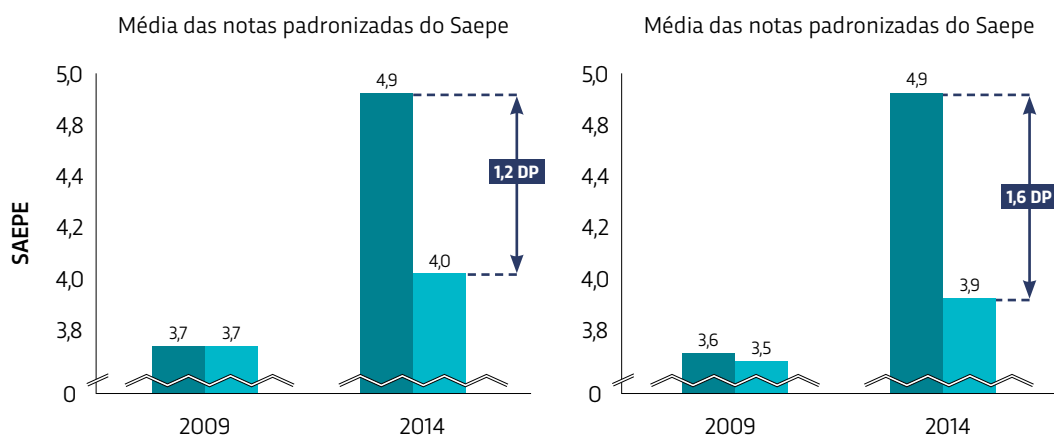
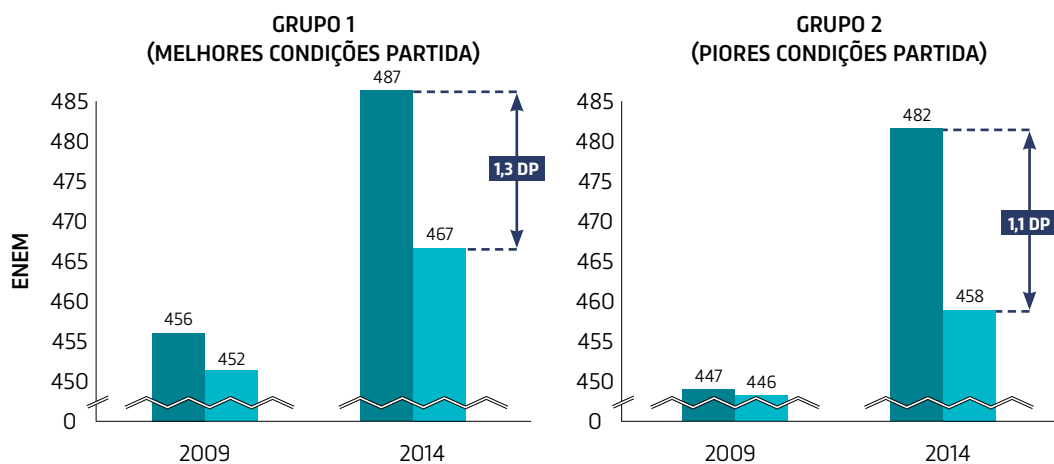
## IMPACTOS DA EDUCAÇÃO INTEGRAL



Sabemos que qualquer estudo de impacto que demonstre um desvio padrão de 0,1, 0,2 ou 0,3 já implica um resultado bastante interessante. A implantação de escolas em tempo integral em Pernambuco observou impactos da ordem de 1,1 até 1,4 na diferença perante as escolas regulares, o que é extremamente positivo, seja em relação aos resultados do Exame Nacional do Ensino Médio (Enem), seja em relação aos resultados no nosso Sistema de Avaliação da Educação Básica de Pernambuco (Saepe), e, inclusive, avaliando escola a escola.

Também vale mencionar que a rede estadual realizou um estudo observando grupos submetidos a piores condições sociais e de formação dos pais, comparando-os a outros grupos com melhores condições sociais e financeiras, além da formação dos pais. Constatamos que as políticas que vimos adotando também induzem à equidade. O resultado final é muito próximo. As evidências mostram que a escola em tempo integral é indutora de equidade.

## IMPACTOS DA EDUCAÇÃO INTEGRAL POR GRUPO, SEGUNDO CONDIÇÕES DE PARTIDA



■ Escolas que adotaram o modelo 45 horas a partir de 2008

■ Escolas que permaneceram no modelo parcial

Nota: DP: desvios padrão.

Fonte: Inep, Secretaria de Educação de Pernambuco, Peers, Análise Bain.

*A implantação de escolas em tempo integral em Pernambuco observou impactos de 1,1 até 1,4 na diferença perante as escolas regulares, o que é extremamente positivo*

*Muitos foram os avanços no Ensino Médio em Pernambuco, mas precisamos seguir em frente. O estado finalmente compreendeu que avanços maiores no Ensino Médio virão quando houver melhora substancial no Ensino Fundamental, que hoje, em sua maior porção, é responsabilidade dos municípios*

Os resultados trazem evidências de que é importante dar continuidade ao processo em curso no estado de fortalecer a Educação em tempo integral. A rede também vem trabalhando com dados de custos que permitem começar a desenhar uma ampliação com vistas à universalização do acesso às escolas em tempo integral.

## **FORTALECER O ENSINO FUNDAMENTAL**

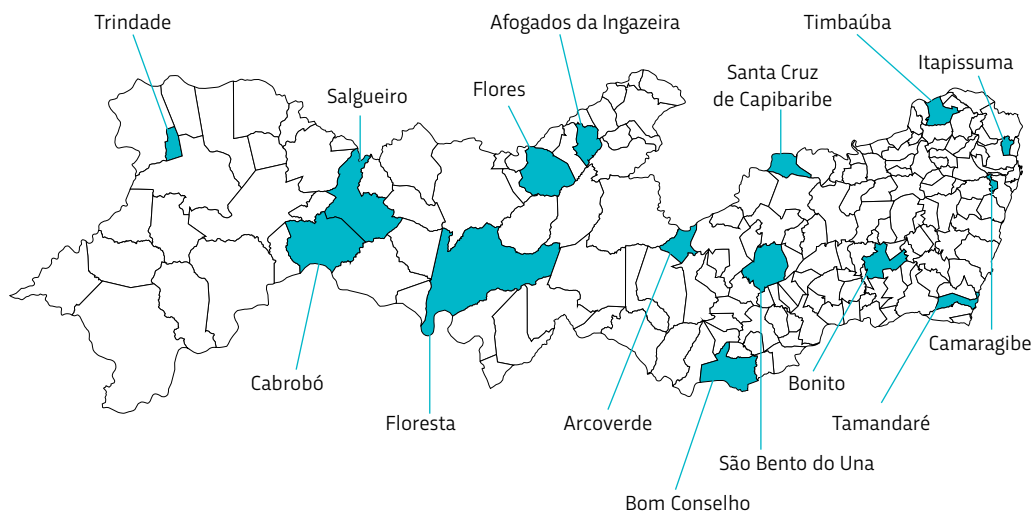
Muitos foram os avanços no Ensino Médio em Pernambuco, mas precisamos seguir em frente. O estado finalmente compreendeu que avanços maiores no Ensino Médio virão quando houver melhora substancial no Ensino Fundamental, que hoje, em sua maior porção, é responsabilidade dos municípios. Portanto, fica claro que é imprescindível construir uma política para apoiar o Ensino Fundamental. Isso é muito mais do que entregar a eles um documento. Em muitos deles, não existe sequer uma equipe estruturada para realizar esse trabalho; é comum haver somente um secretário municipal e alguns professores, sem qualquer estrutura de apoio. Após essa constatação, a Secretaria Estadual está implementando um acordo entre o governo do Estado, municípios e um conjunto de parceiros com o objetivo comum de

melhorar esse segmento. Isso se dará a partir da melhor qualificação da Educação Infantil, de um Programa de Alfabetização na Idade Certa inspirado no modelo do Ceará, da Educação em tempo integral e de maior suporte oferecido à gestão municipal, chamado de Educação Integrada.

Essa é uma política que se debruça sobre evidências e prevê um grupo de tratamento e um grupo de controle. O desenho inicial contou com o apoio de uma equipe de professores, estudantes (mestres e doutorandos) do Lemann Center, da Universidade de Stanford. Há várias etapas a cumprir envolvendo diversos eixos que estabeleceram, a partir de uma combinação de indicadores educacionais e socioeconômicos, quais municípios formariam um grupo de tratamento e, conseqüentemente, qual seria o grupo de controle dessa política. O objetivo era assegurar a participação de localidades de portes e contextos econômicos e sociais diferentes, de todas as regiões do estado de Pernambuco no processo.

### FORTALECIMENTO DO ENSINO FUNDAMENTAL

ATUAÇÃO EM PELO MENOS UM MUNICÍPIO DE CADA REGIÃO DE DESENVOLVIMENTO



Fonte: Secretaria de Educação de Pernambuco.

A estratégia do governo de Pernambuco é aproximar cada vez mais evidência e gestão, com a consciência de que toda e qualquer pesquisa que pretenda chegar ao chão da escola e impactar as redes de ensino deve considerar alguns aspectos muito importantes: o pedagógico, claro, porque o objetivo final é a aprendizagem; a gestão; os aspectos legais; e os custos, pois estes influenciam qualquer política em escala. Lembrando, acima de tudo, que no centro de toda política educacional deve estar o estudante.



# *A IMPORTÂNCIA DAS RELAÇÕES COM A ACADEMIA PARA A CONSTRUÇÃO DE EVIDÊNCIAS NA REDE DO ESPÍRITO SANTO*

**HAROLDO CORRÊA ROCHA** é secretário de Estado da Educação do Espírito Santo, cargo que já havia ocupado entre 2007 e 2010. Formou-se em ciências econômicas pela Universidade Federal do Espírito Santo (UFES), onde atuou como professor, técnico do Instituto Jones dos Santos Neves (IJSN) e pró-reitor de administração. Ocupou a Secretaria de Planejamento da Prefeitura de Vitória e foi diretor-presidente do Banco de Desenvolvimento do Espírito Santo (Bandes).

Os gestores públicos brasileiros da Educação guiam os sistemas às cegas. Sabem aonde querem chegar: à tão sonhada Educação de qualidade para todos. Sabem, ainda, do valor de iniciativas, como o projeto Jovem de Futuro, para alcançar esse objetivo. No entanto, o caminho a percorrer é longo e nem sempre se desenha com nitidez. Três questões mais gerais estão na gênese desse cenário e, antes de tratar da relação da rede do estado do Espírito Santo com a coleta e o uso de evidências, é importante discorrer sobre esses três pontos.

Em primeiro lugar, nós, brasileiros, precisamos de um sistema político melhor. Como sociedade, deve ser de todos a preocupação crescente em constituir bons governos. Um governo ruim faz estragos em todas as políticas públicas. Na Educação, porém, esse prejuízo é muito maior, pela importância e pelo espaço que o tema ocupa em qualquer governo.

Na década de 1990, o estado do Espírito Santo passou por algo semelhante ao que o Brasil tem vivido ultimamente, com corrupção e populismo. Havia, então, o que chamo de governo de qualidade ruim. Uma das consequências de um governo de

má qualidade é que geralmente se chega a uma crise fiscal, a qual, por sua vez, desorganiza o estado. Na Educação, que é um sistema grande e envolve muitas pessoas, o descalabro fiscal quase sempre leva à paralisação de atividades. Esta compromete o trabalho educativo, afetando o resultado da aprendizagem.

Não se trata, aqui, de responsabilizar sindicato e categoria profissional: trata-se de incorporar ao cotidiano de governo a busca pela qualidade da gestão, inclusive do aspecto fiscal. Da mesma forma, o gestor educacional deve assumir o compromisso inarredável de garantir o funcionamento da escola, solucionando questões inerentes a ela (ver artigo de Miguel Székely, p. 226). Embora seja condição fundamental, nem sempre se vê esse compromisso na gestão pública, com gestores assumindo uma postura clara de confronto. “Parou? Vamos ver como fica” – é o que muitas vezes se ouve. Não deveria ser assim: todo gestor deveria honrar seu compromisso.

No entanto, só pode cobrar o comprometimento do outro quem faz seu dever de casa. Obviamente, governo que não paga remuneração a quem trabalha não tem moral para exigir que as pessoas o façam. Governo que não lida com as categorias profissionais de forma profissional não pode exigir profissionalismo.

Houve um período em que o governo do Espírito Santo não pagava salários, e as categorias faziam paralisações frequentes. Não é incomum que a Educação, pela quantidade de servidores que emprega, seja usada para costurar pactos políticos, na tentativa de acomodar várias tendências e demandas. Porém, isso é um desastre. O estado chegou a ter 37 mil pessoas trabalhando, número que, em três anos, reduziu-se a 26 mil, sendo a diferença constituída por pessoas sem função efetiva. Esse trabalho de ajuste fiscal deu-se já nos anos 2000, quando finalmente foi possível normalizar o funcionamento das escolas do estado.

*Na Educação, o descalabro fiscal quase sempre leva à paralisação de atividades. Esta compromete o trabalho educativo, afetando o resultado da aprendizagem*

## A ACADEMIA E O INÍCIO DAS AVALIAÇÕES

O segundo ponto refere-se à relação com a academia, que, no Espírito Santo, resume-se basicamente à universidade federal.

Enquanto o estado se debatia em meio a graves problemas fiscais, obviamente não havia avaliação. A primeira foi realizada em 2000. Já existiam as avaliações nacionais, mas o Espírito Santo ainda não tinha a sua. O que se via até então era uma escola que vivia realizando paralisações porque os servidores não recebiam o que lhes era devido, que não avaliava seus resultados e que não demonstrava um comprometimento efetivo com a aprendizagem.

Aprendi muito nos quatro primeiros anos como secretário, entre 2007 e 2010, e, afinando nosso trabalho ao contexto do Brasil, começamos, no Espírito Santo, a fazer avaliação. Para torná-la mais efetiva, busquei estabelecer uma relação com a academia. Na prática, essa relação tinha duas frentes: uma no campo da produção de pesquisa, avaliação e evidências; a outra em um setor absolutamente decisivo, que é o de formação de professores. Em ambas, deparamos com dificuldades (ver artigo de Paula Louzano, p. 156).

Naquele momento, a secretaria criou um programa para melhorar a alfabetização. Não existia ainda o Pacto Nacional pela Alfabetização na Idade Certa (Pnaic), que sairia do papel somente em 2012, porém já eram conhecidas as experiências do Ceará e de São Paulo. Para atingir seus objetivos, o programa capixaba previa formação de professores, que ficaria a cargo do Centro de Alfabetização, Leitura e Escrita (Ceale), vinculado à Universidade Federal de Minas Gerais (UFMG), a mesma que fazia avaliação para o estado. Nosso plano era que a formação envolvesse a Universidade Federal do Espírito Santo (Ufes) e o Ceale. No entanto, a parceria não se concretizou, e a Ufes coordenou sozinha o processo de formação.

Isso ocorreu há quase uma década. Ao longo dos anos, houve certa mudança no entendimento do papel da avaliação externa e da formação, mas persiste uma dose de arrogância e intransigência, além de certa recusa em acompanhar o que acontece no mundo.

Outra tentativa de aproximação com a academia envolveu a Fundação Carlos Chagas. A ideia, à época, era pesquisar os currículos de formação de professores nas áreas de pedagogia, letras, matemática e ciências biológicas, indicando as lacunas. O objetivo era consolidar uma relação produtiva com a academia. A pesquisa, realizada pela Fundação, deu origem a extensas discussões que evoluíram para a melhoria dos currículos das faculdades. Foi interessante, mas muito difícil.

## **AS PERGUNTAS CORRETAS**

No que diz respeito à pesquisa, a universidade, como instituição, está pouco programada para realizar estudos aplicados na área da Educação e formar professores melhores. No entanto, esse é um ponto relevante, pois muitas universidades federais formam os professores que vêm para nossa rede.

De outra forma, os gestores precisam aprender a fazer as perguntas corretas (ver artigo de Cláudio de Moura Castro, p. 192). É necessário indicar exatamente qual é sua necessidade. Foi por isso que, no passado, a secretaria contratou a Fundação Carlos Chagas para estudar os currículos das faculdades; só com essa informação em mãos seria viável nortear a mudança desejada. Esse é um problema a ser enfrentado por todos os envolvidos na Educação, porque não há possibilidade de termos sucesso sem evidência, sem muita pesquisa (ver artigo de José Francisco Soares, p. 146). Isso depende das universidades, dos institutos de pesquisa e, de maneira indissociável, da formação de professores.

Ter professores bem formados é condição absolutamente decisiva. Certa vez, em visita a uma escola de primeiro ciclo do Ensino Fundamental, conversei com duas jovens, futuras professoras de pedagogia, inscritas no Programa Institucional de Bolsa de Iniciação à Docência (Pibid). Esse é um programa extremamente importante; temos no Espírito Santo uma versão nossa, estadual e complementar. Perguntei a elas como avaliavam o programa, e a resposta não poderia ter sido mais entusiasmada: “Está sendo sensacional, porque na universidade não temos a menor ideia do que seja uma escola. E mais: estamos levando

para a universidade o que aprendemos aqui e quase exigindo dos professores que trabalhem conosco nesse campo”.

EVOLUÇÃO DAS RELAÇÕES COM A ACADEMIA		
2007-2010	2011-2014	2015-2018
<ul style="list-style-type: none"> <li>• Abertura do banco de dados para pesquisa (sem consequências)</li> <li>• Programa Ler, Escrever e Contar – Formação de professores alfabetizadores – Ufes<sup>1</sup>/Ceale<sup>2</sup></li> <li>• Pesquisa sobre formação inicial de professores de língua portuguesa, matemática, ciências biológicas e pedagogia no Espírito Santo com a Fundação Carlos Chagas: “Formação de professores para o Ensino Fundamental: instituições formadoras e seus currículos no Espírito Santo”</li> </ul>	<ul style="list-style-type: none"> <li>• Secretaria de Educação e Ufes se alinham totalmente aos programas/projetos do Ministério da Educação (MEC)</li> <li>• Programas do MEC para o Ensino Médio ProEMI<sup>1</sup> e PNEM<sup>2</sup></li> <li>• Prioridade é a expansão dos direitos e a inclusão de diversos segmentos sociais</li> <li>• Melhoria da aprendizagem e qualificação da gestão ficam em segundo plano</li> <li>• Resultado: queda na proficiência no Saeb<sup>3</sup> e Paebes<sup>4</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Retomada da busca pela melhoria da qualidade da aprendizagem</li> <li>• Retomada espetacular do crescimento do Ideb<sup>1</sup> do Ensino Médio: 11º para o 4º lugar</li> <li>1º no Saeb Matemática</li> <li>3º no Saeb Língua Portuguesa</li> <li>• Melhoria da proficiência também é observada no Paebes<sup>2</sup></li> </ul>
<ol style="list-style-type: none"> <li>1. Universidade Federal do Espírito Santo.</li> <li>2. Centro de Alfabetização, Leitura e Escrita.</li> </ol>	<ol style="list-style-type: none"> <li>1. Programa Ensino Médio Inovador.</li> <li>2. Pacto Nacional pelo Fortalecimento do Ensino Médio.</li> <li>3. Sistema de Avaliação da Educação Básica.</li> <li>4. Programa de Avaliação da Educação Básica do Espírito Santo.</li> </ol>	<ol style="list-style-type: none"> <li>1. Índice de Desenvolvimento da Educação Básica.</li> <li>2. Programa de Avaliação da Educação Básica do Espírito Santo.</li> </ol>

Fonte: elaboração do autor.

Essas duas jovens serão professoras alfabetizadoras, e se queixam, por exemplo, de não aprender matemática na faculdade que cursam. O ensino de matemática é um grande problema. A Secretaria de Educação do Espírito Santo tem dedicado especial atenção a isso em nosso estado, e comemora uma melhoria significativa no Ensino Médio, passando do décimo primeiro lugar no ranking nacional para o quarto. Houve uma colaboração expressiva do Instituto Unibanco nessa caminhada.

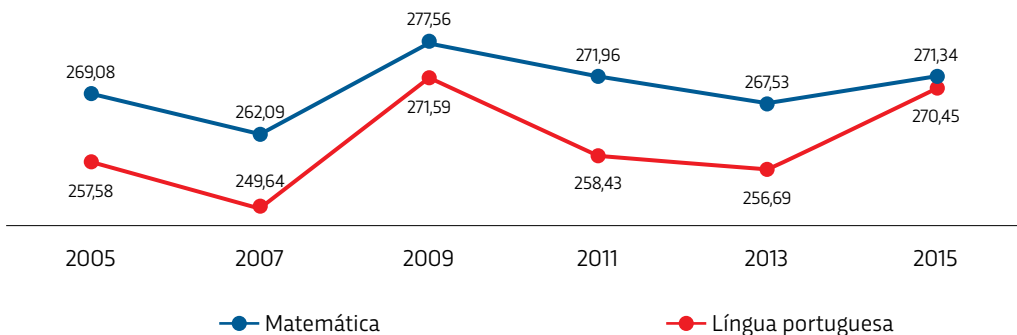
## RELAÇÕES SISTÊMICAS

Por fim, é do maior interesse retratar as relações sistêmicas do Brasil no terreno da Educação, sobretudo no que tange ao Ministério da Educação (MEC), mas considerando, igualmente, o que se faz, como sistema, nas escolas. É preciso ter maior clareza sobre os conceitos de unidade e sistema, e sobre como estabelecer relações para caminhar em uma direção que promova efetivamente a melhoria da qualidade da Educação.

Até há pouco tempo, o MEC adotava uma visão, a nosso ver equivocada, de que bastava enviar fundos para a escola para melhorar a Educação. Esse era o eixo central de programas como o Mais Educação, o Programa Ensino Médio Inovador (ProEMI) e o Plano Nacional de Educação (PNE). Não é o que ocorre: quando o MEC atua diretamente na escola, ele desorganiza o sistema.

Em termos de relações institucionais, a secretaria vem trabalhando tais aspectos com o ministério, de modo que os programas ganhem outro tipo de desenho. A escola não pode ter duas gestões: uma para o sistema e outra para o ministério, o que atrapalha seu funcionamento.

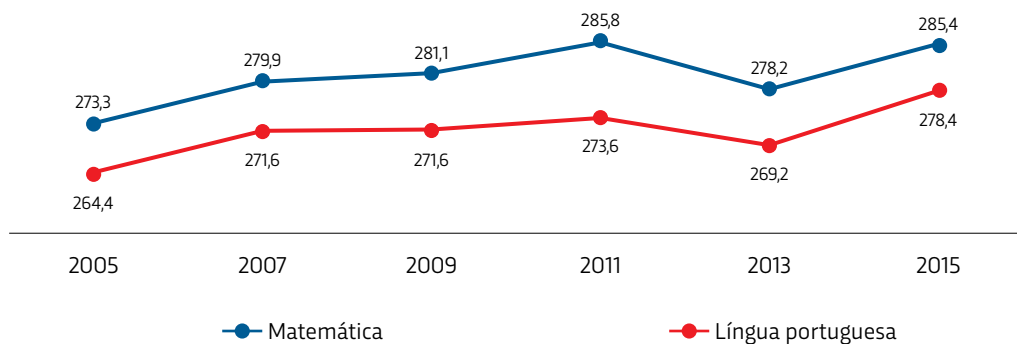
**EVOLUÇÃO DA PROFICIÊNCIA NO ENSINO MÉDIO**  
PROFICIÊNCIA MÉDIA NO SAEB



Fonte: Secretaria de Estado da Educação do Espírito Santo.

## EVOLUÇÃO DA PROFICIÊNCIA NO ENSINO MÉDIO

### PROFICIÊNCIA MÉDIA NO PAEBES



Fonte: Secretaria de Estado da Educação do Espírito Santo.

## EVIDÊNCIAS CIENTÍFICAS RECENTES

Como consequência da aproximação entre a secretaria e a academia, algumas pesquisas muito interessantes foram realizadas no Espírito Santo. Uma delas estudou os determinantes para ingresso de alunos das escolas públicas estaduais da região da Grande Vitória em instituições de Ensino Superior e mostrou muito sobre as relações com o programa Bolsa Família, a autoestima de alunos e os motivos da evasão.

Dessa pesquisa, é importante destacar o quadro sobre os motivos da evasão, que, de maneira geral, se repetem em todo o Brasil: 31% dos alunos deixam a escola para trabalhar.

Outra dissertação traz dados interessantes sobre a relação entre custos diretos e características das escolas, envolvendo o desempenho cognitivo no Ensino Médio.

Vale destacar, ainda, outro trabalho desenvolvido na Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças (Fucape) para avaliar o impacto das redes mistas.

É importante reiterar a importância de as universidades federais, incentivadas pelo MEC, assumirem o propósito claro de pesquisar a Educação Básica para gerar evidências. As secretarias estaduais contam com elas, também, para contribuir na elaboração de currículos atualizados à luz dos desafios que a escola pública vem enfrentando.



# *EXPERIÊNCIA DA REDE DE ENSINO DE GOIÁS COM O USO DE EVIDÊNCIAS: IMPORTÂNCIA DA GESTÃO ESTRATÉGICA*

**RAQUEL TEIXEIRA** é secretária de Estado de Educação de Goiás. Doutora em linguística pela Universidade da Califórnia (EUA), possui pós-doutorado em língua e cultura pela Escola de Altos Estudos de Paris. Foi deputada federal por dois mandatos (2003-2006 e 2007-2010).

A Educação é a ferramenta para a construção de uma sociedade na qual os jovens possam se preparar para uma inserção plena na vida adulta. É de todos – do poder público, da sociedade e da mídia – o compromisso de fazer chegar a cada estudante brasileiro a oportunidade de aprender tudo a que ele tem direito. O objetivo deste artigo é compartilhar, do ponto de vista do poder público, um pouco da experiência da rede de Goiás com o uso de evidências em busca desse objetivo.

Os gestores públicos da área da Educação têm uma angústia em comum: a de constatar o mau desempenho das escolas. Convivemos com estatísticas terríveis, como a de que somente 10% dos alunos saem do Ensino Médio dominando o que deveriam saber em matemática. Há muita pressão por resultados e pela melhoria da rede. Essa pressão é tão intensa que, muitas vezes, o gestor se esquece do longo caminho percorrido até chegar ao momento atual. O fato é que, em apenas 20 anos, houve progressos notáveis, e temos que manter isso como pano de fundo, inclusive como estímulo para continuar avançando.

Em 1999, fui secretária de Educação de Goiás pela primeira

## *Foi construído em Goiás um sistema equilibrado que tem feito avanços de maneira sistemática*

vez. Naquela época, tudo era muito mais precário. Um exemplo é o processo de escolha do livro didático. O diretor da escola escolhia o livro, que era enviado pelo Fundo Nacional de Desenvolvimento da Educação (FNDE). Um professor, eventualmente recém-saído da universidade, recebia aquele livro e tomava-o como currículo para sua disciplina. Não havia acompanhamento. Ao final de cada ano letivo, não se sabia se o conteúdo havia sido aplicado nem em que medida.

Foi justamente no final dos anos 1990 que se começou a definir um norte para a Educação. Houve dois grandes avanços naquele período: o estabelecimento dos parâmetros curriculares nacionais, em 1997, e a criação do sistema de avaliação, a partir de 1998. A lembrança que se tem dessa época são os boicotes ao provão, mas, se não fosse por ele, bem como por todo o sistema concebido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), hoje seria impossível fazer qualquer debate sobre a qualidade da Educação. Atualmente o Brasil tem um dos sistemas mais sofisticados de avaliação. Isso é muito importante, porque permite ao país avançar e recuperar um pouco o atraso e a dívida históricos para com a população brasileira.

A rede de Goiás integra-se a esse movimento de resgate histórico da dívida do poder público para com a Educação pública de qualidade. Possui 1.152 escolas (57% de Ensino Fundamental), totalizando 551.310 alunos, distribuídos em 40 regionais de 246 municípios.

Trata-se de uma rede complexa. Goiás é um estado prioritária e economicamente rural. No entanto, possui uma das maiores redes de escolas urbanas do país e uma das menores no que diz respeito a número de escolas rurais. Uma observação mais atenta, porém, revelará que isso não corresponde à verdade. O fato é que várias escolas urbanas contam com extensões distan-

tes na zona rural, criando enormes desafios de logística e transporte e caracterizando um modelo inadequado. Lamentavelmente, temos ainda 32,7% dos nossos alunos de Ensino Médio matriculados no curso noturno, sendo que, das 625 escolas de Ensino Médio, 84% oferecem esse ciclo à noite. Ainda não conseguimos mudar esse indicador.

Apesar da complexidade do sistema, observando os dados do Índice de Desenvolvimento da Educação Básica (Ideb) de 2007 até agora, talvez o fato mais constante seja o cumprimento das metas em todas as etapas e em todos os anos, exceto por um único ano no Ensino Médio. Para os anos iniciais do Ensino Fundamental, de 2007 a 2015, Goiás não só atingiu todas as metas propostas como ainda as superou; o mesmo ocorreu nos anos finais, de 2007 até agora; e em relação ao Ensino Médio, também atingimos a meta proposta.

Esses resultados posicionaram a rede do estado em uma classificação sustentável, até mesmo confortável: ocupa o segundo lugar no Ensino Médio, o primeiro nos anos finais do Ensino Fundamental e o quarto nos anos iniciais. Ou seja: foi construído um sistema equilibrado que tem feito avanços de maneira sistemática.

## **O PAPEL DO JOVEM DE FUTURO**

O projeto Jovem de Futuro teve, e segue tendo, papel significativo nas conquistas da rede de Goiás. Em 2011, data do início da implementação do projeto, o estado estava abaixo da linha média do Brasil. Desde então, o desempenho superou a média nacional – o que não é pouco em se tratando de uma rede que ainda tem escolas com quatro turnos, sendo o quarto aquele que chamamos “da morte”, impossível de ser atendido com qualidade. Lembro que o entorno de Brasília, um cinturão de pobreza, é um desafio imenso para a rede estadual. Nem mesmo a inauguração recente de 29 novas escolas na região permitiu acabar com o quarto turno.

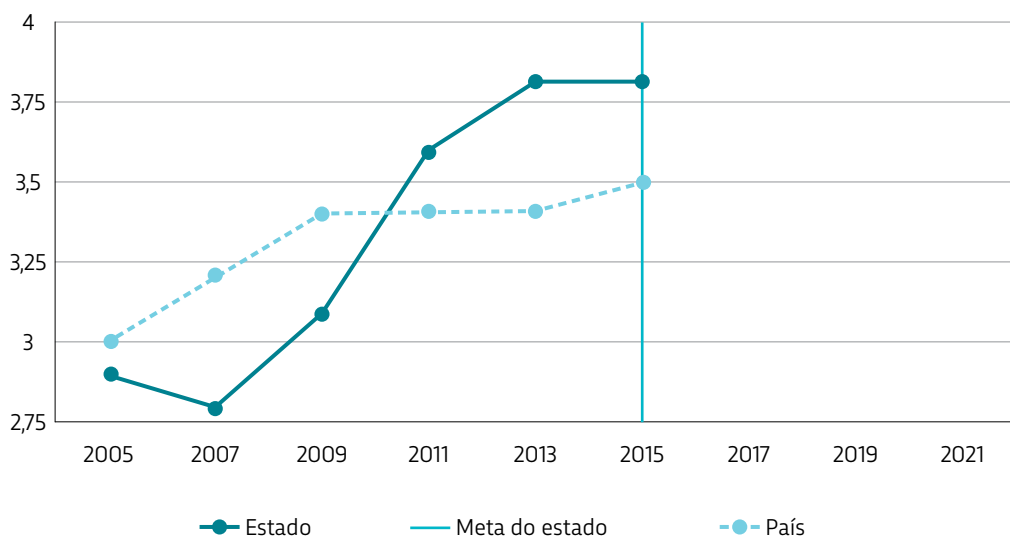
Apesar de tantas condições adversas – escolas com quatro turnos, noturnas, com salas de extensão longe da sede –, há avanços. Essa constância de resultados pode ser atribuída, em

parte, à cultura de avaliação que foi se instalando na rede, aos poucos, desde o início dos anos 2000. Na ocasião, foi criado um sistema de avaliação próprio, que funcionava com apoio do Inep, mas com quadros próprios da secretaria. Dois anos depois, por dificuldades técnicas e financeiras, decidiu-se desativá-lo e utilizar apenas o Inep. Em 2011, o Sistema de Avaliação Educacional do Estado de Goiás (Saego) foi recriado, com a sustentação técnica do Centro de Apoio à Educação a Distância (Caed), o que prevalece até hoje.

O currículo de referência é uma proposta de bimestralização de conteúdos, cujo objetivo é contribuir com as unidades educacionais para melhor compreensão dos componentes curriculares e sua utilização em sala de aula. O Currículo Referência da Rede Estadual de Ensino de Goiás é um instrumento pedagógico, resultante de ampla discussão, em diferentes momentos, entre centenas de servidores de toda a rede, para orientar os professores, de maneira clara e objetiva, quanto a competências e habilidades que devem ser adquiridas pelos estudantes em cada componente, ano de escolaridade e bimestre.

No entanto, mais do que isso, o Jovem de Futuro contribuiu mesmo quando houve queda, como em 2007.

### EVOLUÇÃO DO IDEB DE ENSINO MÉDIO NA REDE ESTADUAL DE GOIÁS



Fonte: Secretaria de Estado de Educação de Goiás.

Entre as escolas de controle e as escolas de tratamento, as de tratamento avançavam mais; mesmo quando caíam, essa queda era menor. Entre 2011 e 2015, o Ideb da rede estadual de Goiás ultrapassa a média brasileira.

ESTIMATIVAS PRODUZIDAS COM BASE NO SISTEMA DE AVALIAÇÃO EDUCACIONAL DO ESTADO DE GOIÁS (SAEGO) DE 2011 A 2013								
Escolas	Língua Portuguesa				Matemática			
	2011	2012	2013	Var (2013-2011)	2011	2012	2013	Var (2013-2011)
Goiás	260,2	262,3	268,6	8,4	268,5	265,4	262,3	-6,2
Escolas – Controle	259,2	257,9	264	4,8	263,1	259,3	253,7	-9,4
Escolas – Tratamento	256,3	256,3	266,8	10,5	261,7	258	256,4	-5,3

Fonte: Secretaria de Estado de Educação de Goiás.

No ano passado, ao assumir novamente a rede do estado de Goiás, o Ensino Médio já estava sob a segunda geração do Jovem de Futuro. A primeira havia deixado um legado importantíssimo de planejamento e execução, mas ainda há avanços necessários para a terceira geração, que é a atual. Esses avanços dizem respeito a não esperar até o final do ano para realizar a avaliação das diversas etapas intermediárias.

Outro diagnóstico importante para a rede diz respeito ao Ensino Médio. Ela era beneficiada por outros projetos além do Jovem de Futuro, como a Tutoria Pedagógica e o Programa Ensino Médio Inovador (ProEMI); tínhamos pelo menos três pessoas

*Entre as escolas de controle e as escolas de tratamento, as de tratamento avançavam mais; mesmo quando caíam, essa queda era menor*

*O Circuito de Gestão mudou profundamente a rotina e a cultura da escola, estimulando-a a participar [...] O trabalho é dividido com ela*

físicas diferentes – o tutor pedagógico, o tutor de área e o supervisor de ProEMI – fazendo com que as escolas acumulassem tarefas burocráticas em programas que, às vezes, conflitavam.

A secretaria, então, tomou a iniciativa de unificar tudo. Hoje, em paralelo à terceira geração do projeto Jovem de Futuro, criamos a Tutoria Educacional.

A Tutoria Pedagógica, assim intitulada de 2011 a 2014, desenvolvia-se por meio de tutores pedagógicos, sendo estes o tutor pedagógico das Subsecretarias Regionais de Educação, Cultura e Esporte, o qual atuava junto aos coordenadores pedagógicos das unidades escolares; o tutor pedagógico do Programa de Acompanhamento e Suporte Pedagógico (Pasp); e ainda os supervisores do ProEMI.

A partir de 2015, com o avanço dos resultados nessas ações formativas, bem como de monitoramento, concentraram-se em um único interlocutor as atribuições referentes à formação em serviço, tanto a do coordenador pedagógico quanto a do gestor. Assim, o tutor pedagógico passa a ser chamado de tutor educacional. Nessa proposta, a tutoria avança para além da formação, promovendo o monitoramento e a avaliação das ações realizadas nas unidades escolares, contribuindo, de maneira mais efetiva, para a aprendizagem dos estudantes.

O Circuito de Gestão do projeto Jovem de Futuro introduziu planejamento estratégico, protocolos e processos, o que mudou profundamente a rotina e a cultura da escola, estimulando-a a participar. Uma grande vantagem é que não se chega impondo nada: o trabalho é dividido com as escolas, e essa corresponsabilização que a atual gestão estratégica trouxe criou uma cultura de avanço que, em Goiás, está bem consolidada. Trabalha-se em

diversos níveis: na escola, na regional, na secretaria e com os estudantes, estimulando o protagonismo juvenil – há hoje oito mil agentes jovens em nossos programas de Ensino Médio.

Outro ponto importante no debate sobre os resultados é a existência de escolas em tempo integral, pois isso, claramente, faz diferença nas oportunidades educacionais oferecidas ao aluno. Em Goiás, apenas 3% da rede de Ensino Médio conta com ensino em tempo integral; mesmo assim, os resultados de aprendizagem têm sido muito interessantes.

## **INTRODUZINDO AS AVALIAÇÕES AMOSTRAIS**

Dentro das iniciativas de mobilização e engajamento criadas pela secretaria, cabe destacar a Avaliação Diagnóstica Amostral (ADA). Até então, a experiência, mesmo com o Jovem de Futuro, era que todo o processo de avaliação de resultados levava muito tempo. A equipe da secretaria planejava e executava, porém precisava esperar a avaliação final para consolidar mudanças. A ADA permitiu atuar com mais agilidade e eficácia. Trata-se de um mecanismo que, a cada dois meses, realiza uma avaliação amostral com 48 mil alunos nas áreas de língua portuguesa, matemática e ciências da natureza. Essas avaliações mostram lacunas e carências da rede naquele momento, e a equipe pedagógica tem sete dias para dar a devolutiva. Ou seja, sete dias depois de avaliadas, todas as escolas recebem o resultado da avaliação diagnóstica e sugestões de reforços e exercícios, além de orientações para o professor e para o aluno.

Isso exigiu a criação do Grupo de Inteligência Pedagógica (GIP), muito estruturado dentro da secretaria, que proporciona formação presencial com orientação quinzenal. Se antes havia quatro tutores educacionais na escola, cada um deles atento aos resultados dos programas sob sua coordenação, hoje existe um só, que tem uma visão global das questões envolvendo não apenas os diferentes programas, mas tudo o que está ocorrendo na escola. Isso tem feito uma diferença muito grande. A avaliação amostral tem contribuído para a superação das dificuldades que se originam da complexidade da própria rede.



Outro mecanismo importante é a plataforma Goiás 360. Trata-se de um projeto tão rico que a Microsoft pediu autorização à secretaria para apresentar o caso a outros países da América Latina. É essa plataforma que permite a devolutiva em uma semana. Com um clique, o diretor da escola e o professor acessam todos os dados, de todas as avaliações, na hora.

## **A CONSTRUÇÃO DO FUTURO**

Ao longo deste texto, foram enumeradas as várias iniciativas da Secretaria de Educação de Goiás para evitar que lacunas de conhecimento se acumulem e garantir que a aprendizagem avance. Entre elas, a unificação dos programas em um tutor com visão global, que recebe formação semanal, a ADA e a plataforma Goiás 360. No entanto, quero registrar neste texto meu depoimento como secretária de um estado beneficiário do Jovem de Futuro. Mais do que um projeto, o Jovem de Futuro é um modelo de gestão que, uma vez assimilado pelas escolas, pelas secretarias regionais e pelas redes, torna-se um poderoso instrumento de avanço.

Haverá quem argumente que não é a gestão que faz o aluno aprender. O que leva à aprendizagem, dirão esses, é o olho no olho na sala de aula, o entusiasmo do professor diante do aluno cujos olhos brilham ao fazer uma descoberta. Há certa verdade nisso: de fato, a gestão não gera alegria nem brilho no olho. No

*A sensação que se tem em Goiás é que, ao construir um caminho para os jovens aprenderem o que têm condição de aprender, cada um dos envolvidos também se modifica, cresce, inova e aprende, em rede*

entanto, sem ela, dificilmente a aprendizagem, a alegria e o brilho no olho aconteceriam. O que move a rede de Goiás, hoje, é o entusiasmo para avançar, a ponto de a própria rede propor caminhos: tanto a ADA quanto a plataforma Goiás 360 foram sugeridas por ela. O desejo de avançar e criar mecanismos resulta de um modelo que mudou a rotina e a cultura das escolas da rede. A gestão voltada para os resultados de aprendizagem do aluno, sem sombra de dúvida, acelerará a busca pela qualidade.

A sensação que se tem em Goiás é que, ao construir um caminho para os jovens aprenderem o que têm condição de aprender, cada um dos envolvidos também se modifica, cresce, inova e aprende, em rede. Vale mais do que nunca esta frase de Antoine de Saint-Exupéry: “O futuro não é um lugar aonde estamos indo, mas um lugar que estamos criando. O caminho para ele não é encontrado, mas construído, e o ato de fazê-lo muda tanto o realizador quando o destino”.

# ***O PROJETO JOVEM DE FUTURO***

## **INTRODUÇÃO**

O projeto Jovem de Futuro apoia secretarias estaduais de Educação na transformação do modelo de gestão educacional vigente nas escolas públicas de Ensino Médio, tornando-o mais orientado para três resultados: aumento da aprendizagem, expansão do número de concluintes e redução das desigualdades educacionais. Isso significa colocar o estudante no centro da gestão escolar e, a partir dessa centralidade, organizar os processos da escola, das regionais e da secretaria.

Para tanto, é fundamental reconhecer que o dia a dia das escolas, que envolve tarefas variadas como coordenar pessoas, cuidar da infraestrutura e relacionar-se com a comunidade escolar, deve estar a serviço daquela que é a rotina mais importante de todas: o processo pedagógico. Será que todos os problemas enfrentados pela escola afetam igualmente a permanência dos estudantes ou seu aprendizado? É preciso criar um senso de priorização nos gestores para que cuidem primeiro daquilo que

## *O projeto Jovem de Futuro apoia secretarias estaduais de Educação na transformação do modelo de gestão educacional vigente nas escolas públicas de Ensino Médio*

realmente pode fazer a diferença, considerando o objetivo de manter o aluno como foco da gestão.

O ponto de partida do Jovem de Futuro é a definição de metas para o estado, desdobradas para regionais e escolas, que enfocam a permanência dos estudantes no Ensino Médio e a aprendizagem. Metas servem para sintetizar desafios comuns, direcionar ações e integrar o trabalho de todos os envolvidos. São como um farol que guia os passos dos atores e geram, portanto, mobilização e coesão. A construção da meta deve basear-se em uma visão realista, mas ao mesmo tempo ousada, entendendo ser possível transformar a realidade. O horizonte de tempo do projeto é de quatro anos ao todo, sendo que a cada ano é preciso subir um degrau.

Para organizar o trabalho de todos os gestores educacionais, adota-se um método denominado “Circuito de Gestão”, que envolve técnicas para planejar, executar, avaliar, aprender com a prática e corrigir rotas. Nesse fluxo de trabalho, estão previstas interações estruturadas das escolas com as regionais e das regionais com o órgão central da secretaria, bem como trocas de experiência entre escolas e entre supervisores.

As regionais e o órgão central da secretaria apoiam as escolas na operação do Circuito de Gestão. Também são corresponsáveis por investir na solução de problemas que estão fora da alça-

---

1. O Circuito de Gestão adaptou o método PDCA, que tem como origem a produção de Edwards Deming (1950), para ser aplicado na gestão escolar. Esse método estabelece o contínuo controle e melhoria de processos em quatro etapas sucessivas: planejar (plan), executar (do), monitorar (check) e agir (act). Pelo PDCA, o planejamento deve ser constante e sistemático. Para outras aplicações do método PDCA no campo educacional, ver: Murici, Izabela e Chaves, Neuza. *Gestão para resultados na Educação*. Editora Falconi, 2013.

da das escolas; por exemplo, a falta de pessoal ou a realização de reformas nos prédios.

Preocupado com a elevada desigualdade educacional brasileira, o Jovem de Futuro também identifica escolas com resultados muito baixos e dá a elas metas bastante desafiadoras. Essas são escolas denominadas “prioritárias”, que precisam de suporte adicional para a implementação do Circuito de Gestão e uma responsabilidade maior da secretaria em termos de investimentos.

A equipe responsável por implementar o Circuito de Gestão na escola é a dupla formada pelo gestor escolar e pelo coordenador pedagógico. Na regional, estão envolvidos os dirigentes e mais um grupo de profissionais de apoio ao Circuito de Gestão<sup>2</sup>. Estes últimos têm dedicação exclusiva ao Jovem de Futuro e visitam de cinco a sete escolas, ao menos quinzenalmente, cumprindo pautas pré-especificadas. As visitas auxiliam a dupla gestora da escola a instaurar o Circuito de Gestão. No órgão central da Secretaria de Educação, é constituída uma equipe de técnicos liderada por um coordenador do projeto.

Todos os atores envolvidos nas três instâncias (escolas, regionais e órgão central da secretaria) são formados pelo Instituto Unibanco e recebem assessoria técnica durante as etapas do Circuito de Gestão.

## **CIRCUITO DE GESTÃO**

O Circuito de Gestão é um método bastante intuitivo. Sua lógica se assemelha ao funcionamento do processo de ensino-aprendizagem, em que o professor planeja, dá aulas, avalia os estudantes e identifica lacunas de aprendizagem. Na sequência, replaneja suas estratégias de ensino a fim de garantir que todos efetivamen-

---

2. Em diversas redes de ensino existe o cargo de supervisor escolar, com atribuições variadas. São profissionais geralmente lotados nas regionais de ensino, que visitam as escolas com fins de apoio e acompanhamento. São elos que conectam melhor as escolas ao órgão central. O projeto Jovem de Futuro prevê que as escolas sejam visitadas quinzenalmente por profissionais de apoio ao Circuito de Gestão (ACGs) para que elas incorporem essa metodologia em sua rotina. Em geral, o papel de ACGs é desempenhado por supervisores.

te aprendam. No campo da gestão, os procedimentos são análogos e referem-se a uma alternativa para organizar e articular o plano tático das escolas, regionais e secretaria em um ciclo anual orientado para a aprendizagem dos estudantes. As metas definidas para cada instância são codependentes e, quando alcançadas, indicam melhoria de desempenho do sistema de ensino.

Os gestores são convidados a refletir sobre os resultados recentes de aprendizagem, identificando causas explicativas para uma performance insatisfatória. Nesse processo de planejamento, o trabalho de priorização é fundamental, uma vez que seguramente há problemas que atrapalham mais o aprendizado dos estudantes do que outros. Com foco nos problemas prioritários, é possível pensar em ações que vão dando corpo ao esforço necessário para a superação do desafio proposto pela meta.

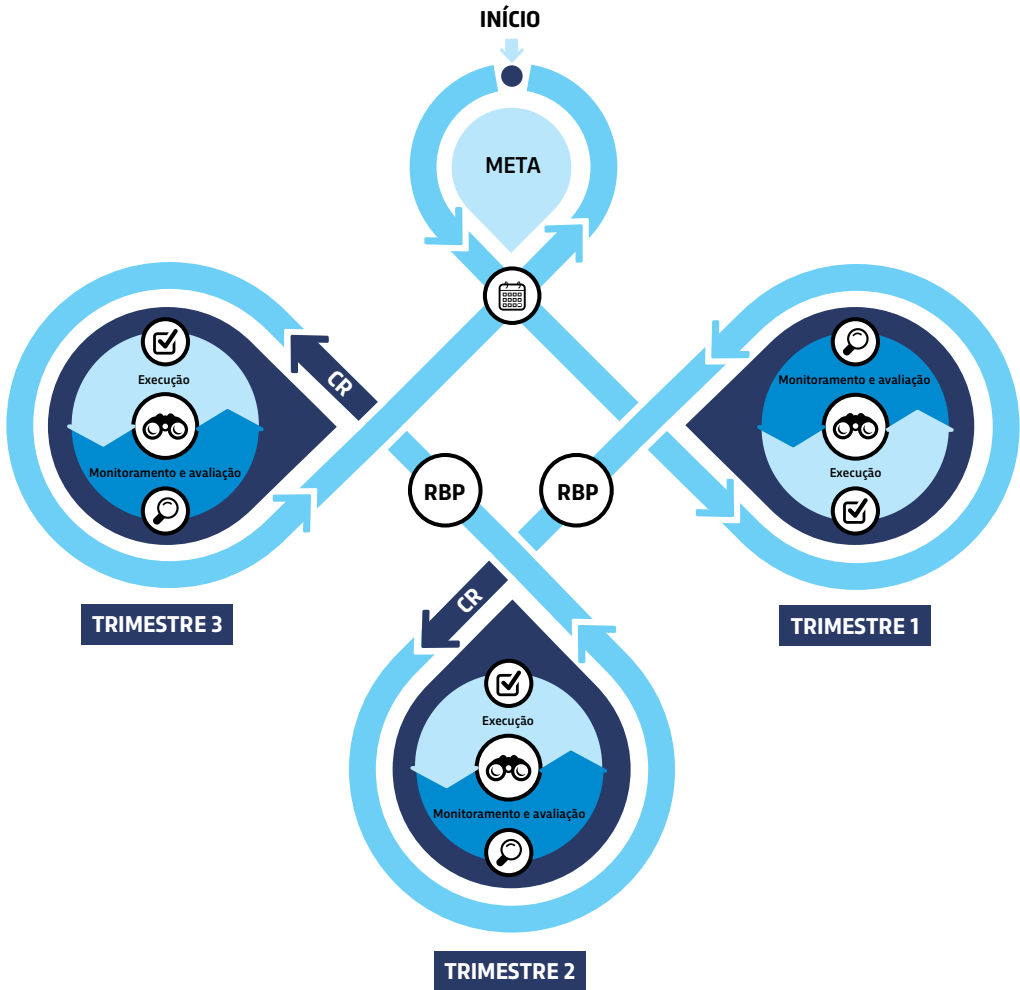
Com um plano de ação bem elaborado, a escola o executa, administra as contingências que surgem e acompanha se a melhora deverá ou não acontecer. Se estiver convencida de que está tomando o rumo errado e não conseguirá melhorar a aprendizagem dos estudantes com o plano construído, a escola precisa corrigi-lo. O mesmo raciocínio vale para as regionais e o órgão central da secretaria: quais são os problemas enfrentados pelas escolas que cabem a essas instâncias resolver e o que mais pode ser ofertado para impactar positivamente a aprendizagem dos estudantes?

As metas são anuais e, para atingi-las, é muito importante que se avalie no decorrer do ano, mais de uma vez, a adequação dos planos de ação. O Circuito de Gestão propõe três rodadas de avaliação no ano, incentivando ajustes frequentes nos planos, o que aumenta as chances de sucesso. A figura abaixo ilustra o ciclo anual descrito pelo Circuito.

*Os procedimentos organizam e articulam o plano tático das escolas, regionais e secretaria em um ciclo anual orientado para a aprendizagem*

## CIRCUITO DE GESTÃO

CICLO ANUAL



### PLANEJAMENTO

À luz de seu diagnóstico, a escola elabora seu Plano de Ação.



### EXECUÇÃO

Elaborado o Plano de Ação, é hora de colocar em prática as ações planejadas, não perdendo de vista o monitoramento de rotina, as adequações e o registro.



### MONITORAMENTO E AVALIAÇÃO DE RESULTADOS

Hora de checagem! Nesta etapa ocorrerá a Sistemática de Monitoramento e Avaliação de Resultados (SMAR) – reuniões trimestrais para verificar se a aprendizagem dos alunos está melhorando.

**CR**

### CORREÇÃO DE ROTAS

Momento de avaliar a Execução e verificar os resultados que estão sendo alcançados, com a intenção de aprimorar o Planejamento.



Logo após os períodos de avaliação<sup>3</sup>, estão previstos momentos de troca de experiências entre escolas e também entre ACGs. São as Reuniões de Boas Práticas. Dessa forma, a cada rodada, todos têm a chance de aprender com o que foi praticado na rodada anterior e, com isso, aumentar as chances de sucesso.

O Circuito de Gestão é um poderoso método para organizar o planejamento tático educacional, integrando a atuação da escola com a da regional e a do órgão central da secretaria. Mas a inteligência do processo está nos atores. Cabe a eles identificar prioridades e criar soluções. Ao final, o Circuito fortalece a autonomia e a corresponsabilização dos gestores da Educação.

## **META: MOBILIZAÇÃO QUE GERA SINERGIA**

Em 2007, o Brasil divulgou metas de aprendizagem bianuais para cada etapa de ensino. Essas metas foram formuladas com base no indicador de qualidade educacional denominado Índice de Desenvolvimento da Educação Básica (Ideb), que combina informações de proficiência (medida por avaliações padronizadas externas) e taxa de aprovação das escolas (fluxo escolar). O Ideb pode variar de 0 a 10; quanto mais alto, melhor.

Para o Ensino Médio da rede estadual, convencionou-se que o país deveria atingir em 2028 um Ideb de 6,0, que, segundo projeções, equivaleria à média de desenvolvimento educacional que os países integrantes da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) apresentavam em 2003.

O Ideb do Ensino Médio utiliza informações sobre proficiência oriundas do Sistema de Avaliação da Educação Básica (Saeb), que apura a cada dois anos a aprendizagem dos estudantes da última série do Ensino Médio. As disciplinas avaliadas são língua portuguesa e matemática. Todos os estados da federação participam. A prova do Saeb é aplicada por amostra e, portanto, não é possível desdobrar as metas estaduais de Ideb em metas por escola.

---

3. Os períodos de avaliação são estruturados a partir de uma Sistemática de Monitoramento e Avaliação de Resultados (SMAR).

## *Precisamos contar com metas anuais por escola. O Instituto Unibanco trabalha com o Ideb que usa a mesma fórmula de cálculo do índice original*

No entanto, para implantar o Circuito de Gestão, precisamos contar com metas anuais por escola. Por essa razão, o Instituto Unibanco trabalha com um indicador similar ao Ideb, denominado Ideb\* (lê-se Ideb estrela). Ele usa a mesma fórmula de cálculo do índice original, substituindo as informações sobre proficiência advindas do Saeb por informações das avaliações externas estaduais, realizadas anualmente e de forma censitária. Os dois indicadores estão correlacionados.

Cada estado parceiro realiza junto com o Instituto Unibanco um exercício de projeção do Ideb de Ensino Médio estadual para quatro anos à frente. A tarefa se encerra com a proposição de uma meta que seja desafiadora e factível. Uma vez que o cenário brasileiro do Ensino Médio mostra estagnação de resultados, em geral as metas bianuais de Ideb propostas pelo Ministério da Educação (MEC) já são extremamente desafiadoras. Por essa razão, não é raro que o estado parceiro do Jovem de Futuro assuma como meta estratégica a própria projeção do MEC e esta seja desdobrada em metas anuais para o estado, regionais e escolas, com o uso do Ideb\*.

Assim, o passo a passo para a definição de metas envolve: (1) definir qual nível de Ideb a rede de ensino quer atingir nos próximos quatro anos; (2) desdobrar essa meta final de Ideb em metas anuais de Ideb\* da rede, obedecendo a uma lógica escalonada; (3) desdobrar as metas anuais de Ideb\* da rede em metas para as escolas e regionais.

As metas não são recalculadas de um ano para outro após a verificação da performance do ano anterior. Eventuais desvios negativos são acumulados e devem ser neutralizados até o quarto e último ano de implementação do programa.

A compreensão da meta e a adesão a ela deixam os gestores e a comunidade escolar prontos para desenvolver a primeira etapa do Circuito de Gestão: o Planejamento.

## *PLANEJAMENTO: DIAGNÓSTICO E PLANO DE AÇÃO*

O planejamento tem início com a conscientização dos gestores de todas as instâncias da magnitude do desafio trazido pelas metas. Para isso, confrontam a situação atual do Ideb\* com o que é proposto alcançar ao término de quatro anos e também a cada ano.

O passo seguinte é buscar as causas que impedem as escolas de alavancar seus resultados, olhando para cada um dos cinco componentes do Ideb\* (proficiência em língua portuguesa e matemática, assim como as taxas de aprovação em cada uma das três séries do Ensino Médio) e refletindo sobre o que pode ser feito para modificar a situação atual.

As escolas identificam as causas sobre as quais podem atuar e também outras que demandam a corresponsabilização de outras instâncias. Todas as informações sobre causas identificadas pelas escolas são registradas no Sistema de Gerenciamento de Projetos (SGP)<sup>4</sup> e utilizadas como insumos para o planejamento da regional e da secretaria.

Cada gestor, seja da escola, da regional ou da secretaria, juntamente com suas equipes de trabalho e membros da comunidade escolar, elegem as causas que todos acreditam ser as mais impactantes no resultado. A partir daí, constroem seus planos de ação<sup>5</sup>, nos quais são elencados prazos e responsáveis para cada uma das atividades, bem como possíveis custos.

Quanto mais participativo e engajador for esse processo, maior será a chance de sucesso do plano. Afinal, transformar a escola não é tarefa de um, mas de todos.

---

4. O Sistema de Gerenciamento de Projetos (SGP) é uma ferramenta oferecida pelo Jovem de Futuro.

5. Plano de Ação é um documento estratégico da escola em que constam as tarefas e atividades definidas para enfrentar os problemas detectados e alcançar a meta anual de aprendizagem proposta. Nele estão reunidos mapas com suas respectivas ações, produtos e resultados esperados.

## EXECUÇÃO DOS PLANOS

A etapa de Execução é o momento em que as ações planejadas são colocadas em prática, e normalmente são exigidos ajustes para lidar com imprevistos do Planejamento. É necessário que todas as pessoas envolvidas no plano tenham clareza sobre seus papéis e tarefas. E quanto mais participativa e engajadora tiver sido a forma como o planejamento foi feito, maiores serão a integração da equipe e as chances de êxito das ações.

O acompanhamento da Execução deve ser sistemático, de forma a possibilitar ao gestor intervir sempre que identificar entraves; por exemplo, incluindo tarefas não previstas inicialmente e até atuando junto aos responsáveis que não estão conseguindo cumprir com suas ações.

Para tanto, é crucial estabelecer momentos de checagem do *status* das ações previstas no plano de ação e propor ajustes tão logo sejam identificados os problemas.

A Execução é a fase mais exigente em termos de gestão de pessoas. Os gestores precisam atuar como fortes lideranças, discernindo os pontos que merecem intervenção e garantindo sempre que a equipe esteja engajada no processo que levará ao atingimento da meta anual.

## SISTEMÁTICA DE MONITORAMENTO E AVALIAÇÃO DE RESULTADOS (SMAR)

A Sistemática de Monitoramento e Avaliação de Resultados (SMAR) é o momento de balanço e ocorre três vezes ao ano<sup>6</sup>. A pauta é avaliar quanto das ações planejadas foi executado até o momento e também se estão sendo alcançados bons resultados intermediários.

As reuniões da SMAR acontecem em três instâncias: nas escolas, nas regionais e na secretaria. Tanto as regionais quanto a

---

6. Aproximadamente nos meses de maio, agosto e novembro.

*O momento de balanço ocorre três vezes ao ano. A pauta é avaliar quanto das ações planejadas foi executado até o momento e também se estão sendo alcançados bons resultados intermediários*

secretaria analisam o conjunto de escolas e também seu próprio plano de ação e seus resultados intermediários específicos.

O método do Circuito de Gestão gera uma categorização das escolas após cada SMAR em função de seus níveis de execução e dos resultados intermediários alcançados. Aquelas com melhores resultados são identificadas como “em destaque”. Aquelas com boa execução, mas resultados intermediários fracos, são chamadas de “em alinhamento”. As demais são as “escolas em atenção”. Tal categorização é orientadora para a etapa seguinte do Circuito de Gestão: as Reuniões de Boas Práticas (RBPs).

Até meados do ano corrente é divulgado o Ideb\* do ano anterior, que permite a análise do atingimento da meta anual. Esse processo evidentemente tem início a partir do segundo ano de implementação do Jovem de Futuro no estado. É somado ao conteúdo da primeira SMAR<sup>7</sup> uma reflexão adicional que trata do atingimento da meta proposta para o ano anterior. Escolas, regionais e secretaria discutem se devem despender esforço adicional para o atingimento da meta no ano corrente, dependendo da necessidade de compensação de um resultado desfavorável.

7. A divulgação do Ideb\* exige que os resultados de proficiência medida pelas avaliações externas estaduais tenham sido disponibilizados, o que em geral ocorre no primeiro trimestre do ano, e também que as taxas de aprovação por escola tenham sido divulgadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), o que acontece em meados do ano. Sendo assim, normalmente a segunda SMAR é aquela em que se avalia o atingimento das metas propostas para o ano anterior.

## REUNIÃO DE BOAS PRÁTICAS (RBP): COMPARTILHANDO EXPERIÊNCIAS

Após a etapa SMAR, antes de efetivamente se fazer a Correção de Rotas nos planos de ação, completa-se um ciclo de reflexão e aprendizado prático.

A SMAR é um marco para a reflexão, e os debates ensejados por ela potencializam o aprendizado prático dos gestores sobre o que fazer e o que não fazer. Afinal, o plano de ação que levará ao sucesso não será perfeitamente entregue logo na primeira tentativa. As ações mais potentes aparecerão como soluções após os gestores se aproximarem sucessivamente de seus reais problemas. E cada SMAR aproxima os gestores mais e mais da compreensão do que podem fazer para transformar sua realidade.

O processo de aprendizado sobre a prática culmina com as Reuniões de Boas Práticas (RBPs). Trata-se de encontros com pauta estruturada, envolvendo gestores escolares e ACGs de uma mesma regional, organizados em grupos de até 30 escolas. Algumas escolas “em destaque” da regional são chamadas a apresentar suas melhores experiências no que concerne tanto à viabilização do Circuito de Gestão<sup>8</sup> quanto a algumas ações incluídas nos planos e avaliadas como capazes de impactar positivamente o atingimento da meta. Há discussões em grupo, nas quais todos são convocados a falar sobre suas experiências. O objetivo desses encontros é a ampliação de repertório e a inspiração.

## CORREÇÃO DE ROTAS

É neste momento que toda a rede de ensino retoma o Planejamento anterior, ainda dentro do ano letivo corrente, e o modi-

---

8. Viabilizar o Circuito de Gestão na escola demanda muitos esforços, uma vez que a equipe escolar precisa estar mobilizada e ter se apropriado do processo, e as informações sobre a execução dos planos precisam ser registradas periodicamente no sistema, assim como os dados sobre os indicadores estruturantes. É importante que as escolas troquem suas experiências para cumprirem melhor essa ritualística.

fica, a fim de aumentar sua capacidade de atingimento da meta anual proposta. As ações do plano podem ser reformuladas, canceladas ou substituídas à luz do que se vivenciou. Novas ações podem ser incluídas.

E por que é preciso aguardar essa etapa para alterar o escopo do plano de ação? Por que não se pode ajustá-lo antes da SMAR ou durante as RBPs? Porque é muito importante persistir por um tempo na estratégia definida. Afinal, as mudanças não acontecem da noite para o dia. As próprias ações precisam, muitas vezes, ser praticadas até que a execução fique bem-feita.

Além disso, é preciso confiar na capacidade da escola, da regional e da secretaria de identificar seus próprios problemas e pensar em soluções. O aprendizado é um acúmulo. Alterações erráticas no plano de ação só atrapalham. Portanto, trocar o escopo do plano antes dos três meses que antecedem a etapa Correção de Rotas não permite aprender sobre o que funciona e o que não funciona, o que falta ao plano e o que já está bom.

A última SMAR acontece sempre no final do ano, momento em que já não há mais tempo hábil para ajustar o plano corrente em prol de melhores resultados. Contudo, recomenda-se que o ritual seja seguido e que as lições aprendidas sejam registradas, completando o ciclo de aprendizado prático daquele ano. No ano seguinte, o planejamento inicial partirá seguramente de um nível mais elevado de apropriação dos gestores sobre suas reais lacunas e desafios.

*Antes de se fazer a Correção de Rotas nos planos de ação, completa-se um ciclo de reflexão e aprendizado prático [...] É um marco para a reflexão, e os debates ensejados por ela potencializam o aprendizado prático dos gestores sobre o que fazer e o que não fazer*

## **ESCOLAS PRIORITÁRIAS**

Tão importante quanto o estado atingir sua meta é reduzir a desigualdade entre as escolas. Portanto, as escolas com menor desempenho recebem metas mais desafiadoras e, ao mesmo tempo, maior apoio de sua regional e da equipe central da secretaria. Tal medida de corresponsabilização mais intensa pretende gerar as condições necessárias para que as escolas mais vulneráveis alcancem suas metas.

Uma preocupação pura com a eficiência resultaria em desafios maiores para escolas que já possuem bons resultados e em metas pouco ambiciosas para aquelas com baixos resultados, posto que estas últimas têm geralmente dificuldades grandes com o contexto familiar dos estudantes, um entorno violento ou ainda uma alta complexidade de gestão (escolas grandes, com várias modalidades etc.). Seguindo essa mentalidade, pediríamos a cada escola “um passo do tamanho de suas pernas”.

Para o Instituto Unibanco, qual o problema com esse raciocínio? É que, se permitirmos que metas pouco ousadas sejam estabelecidas para as escolas com resultados menores de Ideb\*, elas ficarão definitivamente para trás.

A escola precisa ser boa para todos os alunos e precisa fazê-los aprender sempre. Para ser justo, o sistema precisa ser igualitário. Não basta ser eficiente.

Cientes das dificuldades maiores enfrentadas por algumas escolas, não adianta crer que a simples designação de metas ousadas para elas fará todo o serviço. A solução está em tratá-las como escolas prioritárias.

Nas escolas prioritárias, a implantação do Circuito de Gestão é assistida com mais proximidade. As visitas técnicas realizadas pelos ACGs para apoiar a instauração do Circuito de Gestão são mais frequentes. Além disso, o planejamento das regionais e da secretaria deve levar em conta o tratamento customizado dessas escolas, atacando o conjunto de problemas enfrentados por elas que precisam da intervenção dessas instâncias para serem solucionados.

Além disso, essas escolas devem receber prioridade nos investimentos gerais feitos pela rede. É claro que a política educa-



cional não poderá restringir-se a esse grupo, mas ele certamente é merecedor de prioridade. Sem o tratamento diferenciado de escolas prioritárias, não será possível levá-las a níveis mais elevados de desempenho nem reduzir a enorme desigualdade educacional que assola as redes de ensino.

## **APOIO À IMPLEMENTAÇÃO DO CIRCUITO DE GESTÃO – JOVEM DE FUTURO**

Para instaurar o Circuito de Gestão em toda a rede de ensino, o Instituto Unibanco oferece três tipos de ações para diferentes públicos: formação, assessoria técnica e avaliação de impacto.

### **FORMAÇÃO**

É oferecido um programa de formação com conteúdo e formato diferenciado para quatro atores distintos: gestores escolares, profissionais de apoio ao Circuito de Gestão (ACGs), dirigentes regionais e profissionais da secretaria. Todas as formações estão organizadas em torno das etapas do Circuito de Gestão: Planejamento, Execução, Monitoramento e Avaliação de Resultados, Reunião de Boas Práticas e Correção de Rotas.

O curso para gestores escolares chama-se Gestão Escolar Orientada para Resultados de Aprendizagem (GEpR), com carga horária de 48 horas presenciais e 160 horas de educação a distância (EAD) em plataforma própria do Instituto Unibanco<sup>9</sup>. O conteúdo é mobilizador e convida o público a refletir sobre como a escola está organizada para melhorar os resultados de aprendizagem dos alunos e também como vem sendo conduzido o planejamento estratégico. O Circuito de Gestão é apresentado e discutido como solução possível para melhorar a gestão escolar. O curso aborda aspectos conceituais e práticos.

---

9. O Instituto Unibanco está aberto para viabilizar, junto com as secretarias, estratégias de certificação a partir da formação oferecida pelo Jovem de Futuro.

*É oferecido um programa de formação com conteúdo e formato diferenciado para quatro atores distintos: gestores escolares, profissionais de apoio ao Circuito de Gestão (ACGs), dirigentes regionais e profissionais da secretaria*

Os profissionais de apoio ao Circuito de Gestão (ACGs) participam de oficinas formativas intituladas Protocolos para Visitas Técnicas, que totalizam 48 horas presenciais. Nesses encontros são trabalhados as diretrizes e os materiais completos orientadores das visitas técnicas às escolas, recomendadas pelo projeto com o objetivo de instaurar o Circuito de Gestão. São oficinas de trabalho com apresentação de conteúdos e espaço reservado para problematizações e esclarecimentos. Há discussões em grupo e simulação de situações típicas a serem enfrentadas durante as visitas às escolas.

Ao serem visitados por esses profissionais, os gestores escolares colocam em prática os conteúdos discutidos no curso. As atividades desenvolvidas pelos ACGs seguem os protocolos propostos pelo Instituto Unibanco e, com isso, instrumentalizam a nova prática de gestão nas escolas.

Para os dirigentes regionais e gestores da Secretaria de Educação, também são oferecidas oficinas de trabalho intituladas Protocolos para Regionais e Secretaria de Educação, com carga horária de 48 horas. Há apresentação de protocolos específicos para essas instâncias, instrumentalizadores da rotina de trabalho em torno da instauração do Circuito de Gestão na rede. Estão incluídos tanto os processos de acompanhamento do trabalho dos ACGs e regionais como o próprio planejamento das regionais e da secretaria, que se corresponsabilizam, junto com as escolas, pelo atingimento de metas. Trata-se da verticalização do Circuito de Gestão.

Ao longo dos quatro anos de implantação, o Jovem de Futuro organiza com a secretaria seminários sobre temas de interesse para os gestores e também promove eventos formativos e mobilizadores com os jovens estudantes. O objetivo com esse último grupo é trabalhar o protagonismo juvenil e habilidades como ferramenta para resolução de problemas e trabalho colaborativo.

Por fim, Metodologias Pedagógicas e de Mobilização também são oferecidas na plataforma de ensino a distância para gestores escolares e professores<sup>10</sup>. Trata-se de opções de ações a serem incluídas nos planos anuais pelas escolas.

## ASSESSORIA TÉCNICA

O Jovem de Futuro conta com uma equipe técnica de gestores do Instituto Unibanco alocada nas secretarias de Educação parceiras. Também disponibiliza produtos que ajudam a instituir o Circuito de Gestão em todas as três instâncias (escolas, regionais e órgão central da secretaria), como sistemas de informação para gerenciamento das ações e análise de dados, além de práticas estruturadas de governança.

A equipe alocada na secretaria é composta de quatro gestores, que permanecem no local durante os quatro anos de projeto. Sua função é identificar dificuldades da secretaria para instaurar o Circuito de Gestão e ajudar a contorná-las com agilidade, esclarecendo temas e sistematizando demandas do órgão estadual para a equipe central do Instituto Unibanco. Portanto, fazem a conexão entre os parceiros e garantem uma comunicação ágil e eficiente, focada na resolução de problemas.

---

10. Destacam-se a metodologia pedagógica Entre Jovens e a de mobilização Agente Jovem. A primeira tem como objetivo contribuir para que os estudantes da 1ª e da 3ª séries do Ensino Médio possam rever conteúdos fundamentais em língua portuguesa e matemática e assim possam melhorar seu desempenho nessas disciplinas. É composta de oficinas de aprendizagem e de rodas de conversa, desenvolvidas como ações complementares no contraturno escolar ou como atividades eletivas. A segunda é uma estratégia de mobilização estudantil para diminuir a evasão escolar e estimular o protagonismo juvenil. O objetivo é estimular os jovens a propor ações para a melhoria do ambiente escolar, agindo como corresponsáveis nesse processo.

Essa equipe segue um intenso cronograma de trabalho e apoia tecnicamente os departamentos envolvidos para garantir as condições necessárias à instauração do Circuito de Gestão. Além disso, organiza eventos formativos temáticos e reuniões de trabalho customizadas de acordo com as demandas de ACGs, dirigentes regionais e gestores da secretaria.

Para dar suporte às práticas de análise de estatísticas educacionais, registro de informações do Circuito de Gestão e produção de informações sistematizadas para tomada de decisão, são disponibilizados dois sistemas:

- **Sistema Instituto Unibanco de Análise de Indicadores das Redes Estaduais do Ensino Médio (SITUA):** sistema via web baseado em protocolo de análise de indicadores, que conduzem os gestores na tarefa de realizar um diagnóstico mais preciso de sua realidade. Orienta a reflexão por meio de perguntas e dicas sobre como buscar as respostas nos dados disponíveis. É direcionado para profissionais do órgão central da secretaria e faz uso de indicadores de aprendizagem<sup>11</sup> e do contexto escolar<sup>12</sup>.
- **Sistema de Gerenciamento de Projetos (SGP):** sistema desenhado como suporte para os protocolos de implementação do Circuito de Gestão, no qual gestores das escolas, regionais e equipe central do Ensino Médio registram suas reuniões, seus planos de ação e a execução destes. Além disso, o SGP traz uma primeira linha de inteligência analítica, disponibilizando tabelas e gráficos com indicadores que ajudam na tomada de decisão orientada por evidências.

Estabelecer mecanismos de governança em processos de mudança é condição *sine qua non* para ampliar a assertividade

---

11. Os indicadores de aprendizagem considerados são: resultados do Ideb e seus componentes, Saeb, taxas de rendimento e avaliação estadual.

12. Em relação ao contexto escolar, são disponibilizadas informações sobre: localização da escola, tamanho, serviços públicos, estrutura física, equipamentos, perfil do aluno e qualificações quanto ao atendimento (modalidade de ensino, etapa, turno, jornada etc.).

*Esse conjunto de produtos e procedimentos tem como princípio a ideia de que informação é essencial para o sucesso da gestão. Um bom plano exige evidências que tornem os diagnósticos mais precisos*

das ações em direção aos resultados almejados. Para tanto, é realizado mensalmente o Comitê Operacional, momento em que os gestores locais do Instituto Unibanco e da secretaria se reúnem para resolver entraves, equacionar questões do contexto regional e também analisar as informações sobre a qualidade da implementação do Circuito de Gestão na rede.

No decorrer de cada ano são realizados ainda quatro Comitês Estaduais de Governança, com objetivo de capturar a análise estratégica dos dirigentes da Secretaria de Educação e do Instituto Unibanco. Nesses eventos são apresentadas informações sobre a implementação do Circuito de Gestão e os resultados alcançados, trazendo luz para o que precisa ser fortalecido na parceria. Discute-se de que forma os apoios do Jovem de Futuro podem potencializar mais a melhoria na gestão e no aprendizado dos estudantes e quais as medidas complementares que precisam ser tomadas pela secretaria.

Esse conjunto de produtos e procedimentos tem como princípio a ideia de que informação é essencial para o sucesso da gestão. Um bom plano exige evidências que tornem os diagnósticos mais precisos. Avaliar a qualidade da execução depende de registros atualizados no sistema de informações e sistematização inteligente dos dados, compondo um bom monitoramento. Os resultados conquistados também precisam ser acessados com agilidade.

Dessa forma, gera-se sinergia entre os atores, garantindo que todos conheçam diretrizes, orientações e resultados. Mais do que isso, trabalhamos para que todos conheçam e reconheçam a produção e o valor agregado por cada um.

## AVALIAÇÃO DE IMPACTO

A fim de se conhecer a verdadeira contribuição do projeto Jovem de Futuro para a melhoria na aprendizagem nas escolas, o projeto é submetido a uma avaliação rigorosa com desenho experimental, que é um “padrão ouro” das avaliações.

De acordo com esse desenho de avaliação, um conjunto relevante de escolas (escolas de tratamento) ingressa no projeto no início da parceria entre a Secretaria de Educação e o Instituto Unibanco, e outro conjunto (escolas de controle) ingressa no projeto somente a partir do terceiro e do quarto ano de implementação. Essas escolas de controle ajudam a saber o que aconteceria com as escolas que receberam o Jovem de Futuro caso o projeto não existisse. Afinal, muitos são os fatores que podem estar por trás da melhoria de aprendizagem nas escolas. Como saber o quanto veio dos efeitos de uma melhor gestão potencializada pelo projeto Jovem de Futuro?

Entendemos que o impacto do projeto é quanto a mais as escolas que receberam o projeto melhoraram a aprendizagem ao final da 3ª série do Ensino Médio comparadas com as escolas de controle. Se todas melhoraram igualmente, não há evidências de que o projeto teve impacto. As escolas que receberam o projeto precisam ter melhorado mais para que afirmemos com segurança que vale a pena fazer o Jovem de Futuro.

*O impacto do projeto é quanto a mais as escolas que o receberam melhoraram a aprendizagem ao final da 3ª série do Ensino Médio comparadas com o grupo de controle [...] Elas precisam ter melhorado mais para afirmar com segurança que vale a pena fazer o Jovem de Futuro*

Portanto, antes de universalizar o projeto para toda a rede, é prudente ter certeza sobre sua capacidade real de transformar a realidade. Afinal, o Jovem de Futuro exige a dedicação e o esforço de muitos atores. Produzir conhecimento sobre o impacto real do projeto é ação gestora de caráter ético e responsável. Nosso desafio é mobilizar também os gestores para uma gestão ética e responsável, que valoriza o conhecimento.

Além da avaliação de impacto, o Instituto Unibanco também realiza avaliações de processo e análise dos componentes, que buscam explicar os avanços ou recuos dos resultados alcançados. Essa agenda de produção de conhecimento está articulada ao modelo lógico do programa, que explicita as hipóteses da cadeia de impacto e permite o aperfeiçoamento contínuo ajustado às especificidades da realidade educacional de cada estado, possibilitando a identificação e a superação de desafios, a melhoria das conquistas parciais e o fortalecimento dos acertos.

## **PERSPECTIVA DE SUSTENTABILIDADE**

Todo o processo de implementação foi formulado para gerar raízes nas práticas dos profissionais e nos mecanismos institucionais. Traz consigo o “vírus” da sustentabilidade. A aposta é que, quanto mais efetivas forem as conexões na implementação, mais viral será a sustentabilidade do Circuito de Gestão.

É por isso que o Jovem de Futuro investe em produção e disseminação de conhecimento sobre gestão, pois é dessa forma que serão desenvolvidas competências individuais e coletivas. Institui um processo de gestão focado em resultados de aprendizagem e direcionado por protocolos de ação, pois só assim podemos renovar ou reinventar um “saber fazer”. Oferece um sistema via web para registro e tabulação de informações que permanece na secretaria. Promove ações de comunicação focadas na mobilização e no engajamento dos diversos atores em todas as instâncias.

Contudo, existem tarefas técnicas que viabilizam o Circuito de Gestão e que exigem conhecimento específico, como cál-

## *A partir do terceiro ano de Jovem de Futuro, o Instituto Unibanco passa a monitorar, com a rede, um indicador de sustentabilidade do Circuito de Gestão*

culos estatísticos, definição de critérios de seleção de escolas, definição de competências e atribuições dos atores, assim como normativas e legislação. Durante os primeiros dois anos de implementação, o projeto Jovem de Futuro se ocupa dessas tarefas.

Tendo em vista ser necessário transferir o conhecimento para que a rede seja completamente autônoma na operação do Circuito de Gestão, são implementadas ações específicas para a promoção da sustentabilidade. São ações formadoras de multiplicadores, desenvolvedoras de quadros técnicos e também de automatização de procedimentos, principalmente cálculos estatísticos.

Além de contribuir para a transferência de conhecimento, essas ações e estratégias buscam enfrentar práticas de descontinuidade das políticas públicas, questão estrutural que dificulta o alcance dos resultados esperados. Reconhecemos que uma parceria só é bem-sucedida quando, de modo efetivo, o corpo dirigente e técnico da Secretaria de Educação e das escolas se apropriam de forma estruturada, crítica e consistente dos conteúdos e dos modos de fazer do programa.

Com tudo isso, esperamos que sobrevivam os esforços e a dedicação do grupo de gestores que começou a implantar o Circuito de Gestão na rede com o apoio do Instituto Unibanco. A fim de acompanhar os acontecimentos a mais longo prazo, a partir do terceiro ano de Jovem de Futuro, o Instituto Unibanco passa a monitorar, em conjunto com a rede, um indicador de sustentabilidade do Circuito de Gestão. Trata-se de uma medida que sintetiza o quanto os processos mais cruciais estão sendo executados e com qual qualidade. Tal indicador continua sendo acompanhado pelo Instituto Unibanco mesmo após o fim do Jovem de Futuro no estado.



# *O IMPACTO DO PROJETO JOVEM DE FUTURO SOBRE A APRENDIZAGEM EM ESCOLAS PÚBLICAS*

**RICARDO PAES DE BARROS**

Economista-chefe do Instituto Ayrton Senna e professor no Insper.

**ROSANE MENDONÇA**

Universidade Federal Fluminense (UFF).

**SAMUEL FRANCO**

OPE Sociais e Instituto de Estudos do Trabalho e Sociedade (IETS).

**GABRIELA GALL ROSA**

Insper.

## **INTRODUÇÃO**

O Jovem de Futuro é um projeto do Instituto Unibanco que tem como objetivo melhorar a aprendizagem dos estudantes em escolas públicas estaduais de Ensino Médio por meio de uma gestão voltada para resultados. Iniciado em 2008, já foi implantado em dez estados brasileiros<sup>1</sup> e até 2015 tinha beneficiado quase 2,5 mil escolas estaduais e 1,6 milhão de estudantes de Ensino Médio.

Neste estudo, apresentamos a avaliação de impacto do Jovem de Futuro, que possui desenho experimental e tem sido conduzida em todos os estados que o adotaram. O projeto transforma práticas de gestão das escolas, e, portanto, o impacto investigado é uma medida agregada da aprendizagem dos alunos que frequentavam as escolas durante o período em que a avaliação foi realizada. O texto não se limita a descrever os resulta-

---

1. Os estados que têm ou já tiveram o projeto são: Ceará, Espírito Santo, Goiás, Mato Grosso do Sul, Minas Gerais, Pará, Piauí, Rio de Janeiro, Rio Grande do Sul e São Paulo.

dos encontrados e inclui uma apresentação do método utilizado para isolar o impacto e também propõe uma reflexão que nos permite julgar o quão elevado é o efeito apresentado.

## **METODOLOGIA**

### *IMPACTO SOBRE O QUÊ?*

Uma vez que a unidade de análise da avaliação de impacto do projeto Jovem de Futuro é a escola, a pergunta a ser respondida é o quanto “ter participado do projeto” garantiu às escolas um melhor desempenho acadêmico.

O desempenho acadêmico de uma escola, por sua vez, pode ser traduzido por diversos tipos de resultado, como proficiências em disciplinas específicas, taxas de aprovação, reprovação ou abandono etc. Selecionado o resultado de interesse, é necessário ainda especificar que vamos expressá-lo como uma *característica* específica da distribuição dos alunos que frequentavam uma determinada *série* dessa escola *ao fim da janela de tempo da avaliação*.

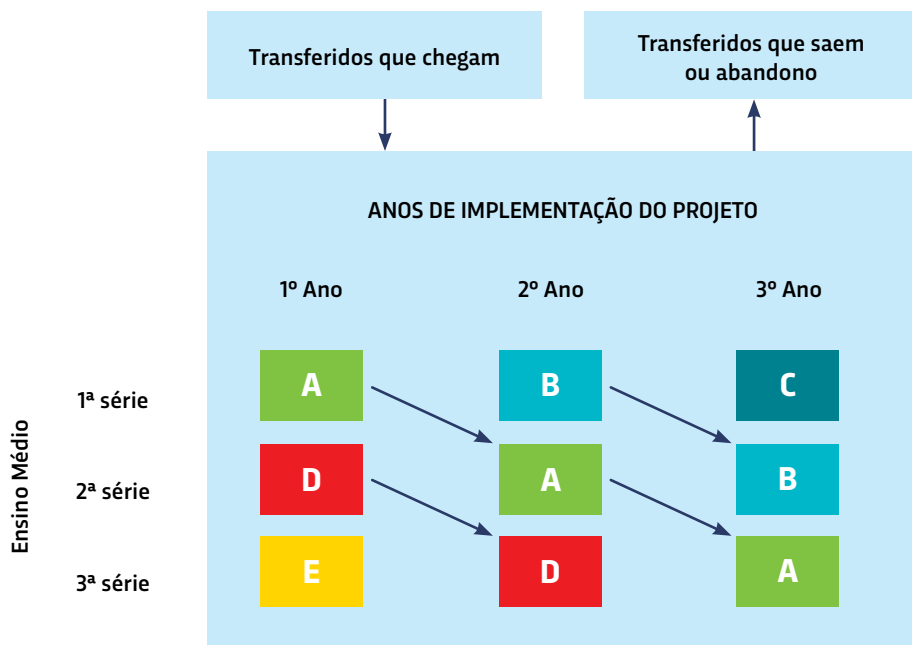
Nossa opção foi avaliar o impacto do Jovem de Futuro sobre dois resultados tradutores do desempenho acadêmico: as proficiências em língua portuguesa e matemática. Assim sendo, a formulação completa do que está sendo estudado é: o impacto sobre as *médias* (características da distribuição) das distribuições dos alunos da escola segundo as *proficiências em língua portuguesa e matemática* (resultados tradutores do desempenho acadêmico dos alunos) *ao final da 3ª série* (série selecionada), *três anos após o início da implantação do projeto* (janela de tempo da avaliação).

O fato de a avaliação ser realizada com os alunos da 3ª série do Ensino Médio após três anos da data de início do projeto é de particular importância. Afinal, esse é o primeiro momento em que se pode capturar o efeito de “cursar todas as séries dessa etapa de ensino” em uma escola com as práticas de gestão instituídas pelo Jovem de Futuro. Esse seria o impacto pleno. Avaliações em qualquer outra série ou em uma janela de tempo de avaliação inferior a três anos captariam o impacto de uma exposição parcial dos estudantes ao projeto, pois não consideraríamos aque-

les estudantes que completaram as três séries do Ensino Médio na escola modificada pelo projeto.

O diagrama abaixo nos ajuda a ilustrar essas ideias. Note que, após três anos de implantação do projeto, estarão cursando a 3ª série três tipos de estudantes. Primeiro, aqueles que no início do projeto estavam na 1ª série e foram sempre progredindo (caixinha A – três anos em uma escola com gestão modificada, sendo que em cada ano de implantação cursaram uma série diferente). Quando esse grupo chega ao final da 3ª série, estamos medindo o impacto pleno. Segundo, os estudantes que no início do projeto estavam na 2ª série e que alternaram reprovação e aprovação (caixinha B – três anos em uma escola com gestão modificada, porém cursaram apenas duas séries distintas no período). Terceiro, os que no início do projeto estavam na 3ª série e a repetiram duas vezes consecutivas (caixinha C – três anos em uma escola com gestão modificada, porém cursaram a mesma série em todo o período). Os estudantes representados pelas caixinhas D e E sequer chegaram à 3ª série ao final dos três anos de intervenção.

### FLUXO NA ESCOLA DURANTE OS 3 ANOS DE IMPLEMENTAÇÃO DO PROJETO

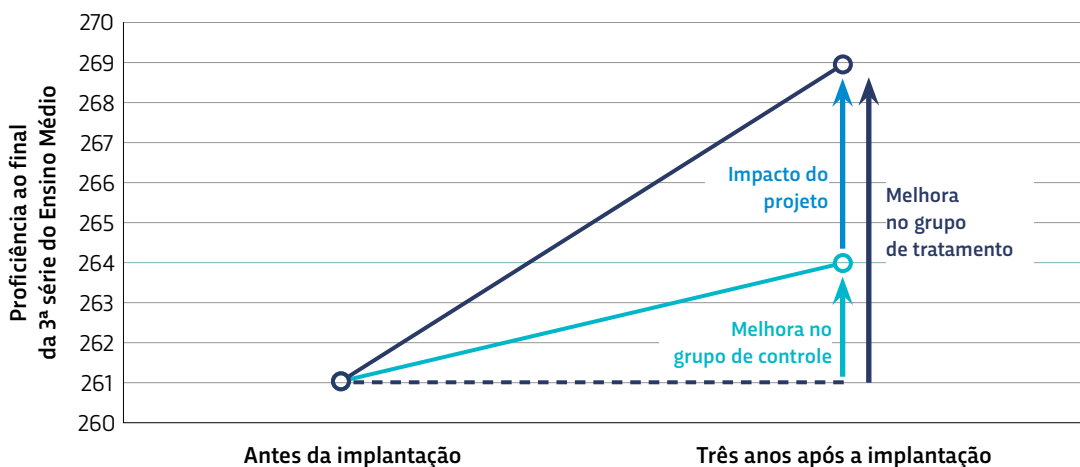


## CONSTRUINDO O GRUPO DE CONTROLE

Toda avaliação de impacto com desenho experimental baseia-se no contraste entre a melhoria no desempenho de um grupo de escolas beneficiadas pelo projeto (grupo de tratamento), antes e depois da intervenção (janela de tempo da avaliação), e a correspondente melhoria em um grupo de escolas não beneficiadas (grupo de controle) – ver gráfico abaixo. Assim, é fundamental em qualquer avaliação de impacto a especificação de como os grupos de tratamento e controle foram obtidos. Esse é o objetivo desta seção.

Em termos conceituais, o impacto de um projeto sobre um conjunto de escolas é, por definição, o contraste entre o que seria o desempenho dessas escolas na presença do projeto e o que teria acontecido caso não tivessem sido beneficiadas. Como, ao final, cada escola vivenciará apenas uma das duas situações, a outra será sempre uma conjectura, um contrafactual, impossível de ser observado.

### REPRESENTAÇÃO GRÁFICA DO CONCEITO DE IMPACTO COM DESENHO EXPERIMENTAL



Fonte: elaboração dos autores.

Tipicamente, o que se deseja estimar é o impacto de uma intervenção sobre um grupo de escolas beneficiadas, e, portanto, o que *não* se observa é o que teria acontecido com elas caso não fossem beneficiadas. Exatamente com o objetivo de preencher essa lacuna informacional é que necessitamos de um grupo de controle, o qual evidentemente deve ser formado por escolas não beneficiadas. Mas essa condição de não estar entre as beneficiadas não é suficiente. Esse grupo precisa também ser selecionado de forma muito especial, uma vez que a evolução do desempenho de um verdadeiro grupo de controle precisa ser idêntica à evolução que o grupo de escolas beneficiadas teria caso não tivesse sido atendido.

Embora ainda exista debate sobre como grupos de controle podem ser obtidos, uma questão já se encontra bem difundida há pelo menos um século: quando a seleção dos beneficiários é feita com base em um sorteio, os não beneficiados formam um grupo de controle ideal.

No caso da avaliação de impacto do projeto Jovem de Futuro, os grupos de tratamento e controle foram obtidos a partir de um processo que pode ser descrito em três etapas:

**1ª etapa:** em cada um dos estados participantes do projeto, um grupo ou a totalidade das escolas foi convidada a participar.

**2ª etapa:** as escolas que declararam interesse em participar foram agrupadas por semelhança. Esses grupos deveriam ser homogêneos com relação à evolução do desempenho que as escolas do grupo teriam na ausência do projeto. O ideal seria que os grupos fossem formados apenas por duas escolas (pares de escola, portanto). A maior parte dos grupos acabou atendendo a esse critério, embora em alguns casos o número de escolas do grupo chegue a 12 (ver tabela p. 318).

**3ª etapa:** organizou-se um sorteio público que, para cada grupo, selecionou aleatoriamente quais escolas seriam atendidas de imediato e quais teriam atendimento postergado. Em cada grupo, pelo menos uma escola teve o atendimento postergado para o fim da janela de tempo da avaliação.

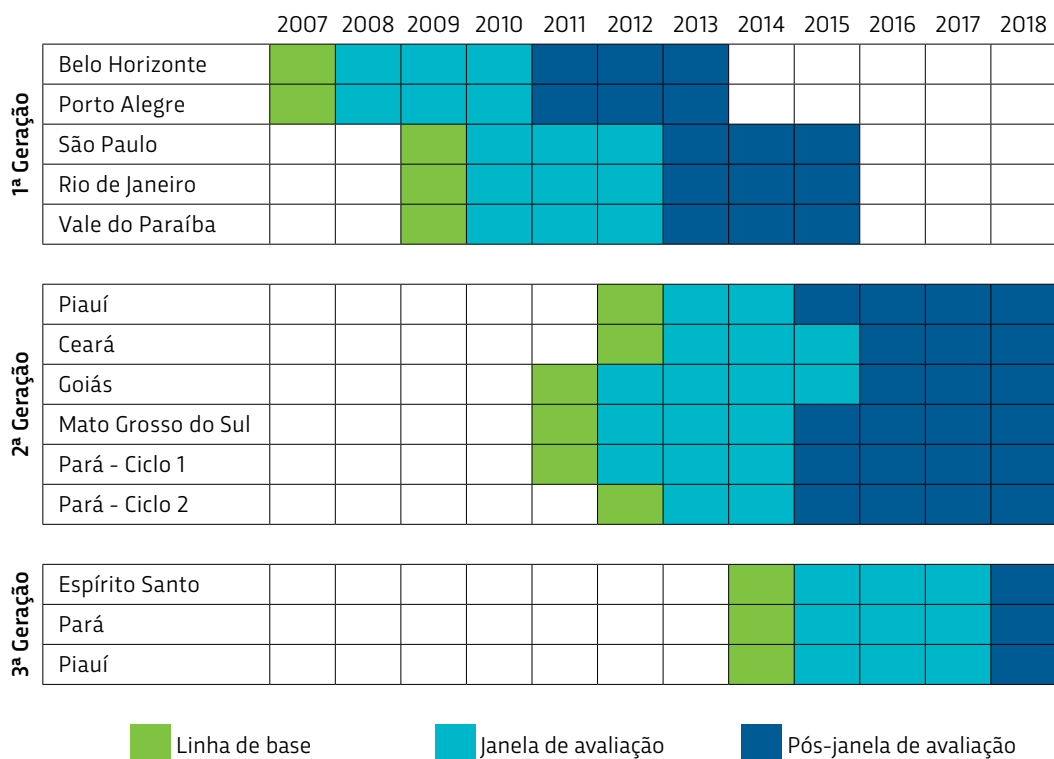
Esse processo vem sendo realizado desde o início do projeto em 2008. O diagrama (ver p. 321) apresenta o momento em que as escolas envolvidas na avaliação foram inicialmente sorteadas em cada um dos estados que são ou já foram parceiros do Instituto Unibanco. Ao longo desses sete anos de avaliação, muitas escolas foram sorteadas a partir dos grupos formados, e, com isso, foram criadas oportunidades para a avaliação de impacto do projeto. Temos hoje 409 grupos de escolas e, destes, 141 foram utilizados na avaliação de impacto do projeto (ver tabela abaixo).

<b>NÚMERO E TAMANHO DOS GRUPOS DE ESCOLAS UTILIZADOS NOS SORTEIOS PARA SELEÇÃO DE TRATAMENTO E CONTROLE</b>					
<b>GERAÇÃO</b>	<b>ÁREA</b>	<b>NÚMERO DE GRUPOS</b>		<b>TAMANHO DOS GRUPOS (NÚMERO DE ESCOLAS)</b>	
		<b>TOTAL</b>	<b>UTILIZADOS NA AVALIAÇÃO</b>	<b>GRUPO DE TRATAMENTO</b>	<b>GRUPO DE CONTROLE</b>
<b>1<sup>a</sup></b>	Minas Gerais (RM* de Belo Horizonte)	4	4	5	7
	Rio Grande do Sul (RM* de Porto Alegre)	25	0	1	1
	São Paulo (RM* de São Paulo)	20	17	1	1
	São Paulo (Vale do Paraíba)	20	20	1	1
	Rio de Janeiro (RM* do Rio de Janeiro)	15	12	1	1
<b>2<sup>a</sup></b>	Pará	44	15	1	1
	Piauí	25	0	7	1
	Ceará	26	25	5	1
	Goiás	25	25	5	1
	Mato Grosso do Sul	25	23	4	1
<b>3<sup>a</sup></b>	Pará	42	0	1	1
	Piauí	134	0	1	1
	Espírito Santo	70	0	2	1
<b>TOTAL</b>		<b>475</b>	<b>141</b>	<b>–</b>	<b>–</b>

\*Região metropolitana.

Fonte: elaboração dos autores a partir de definições com o Instituto Unibanco.

## JANELA DE TEMPO DE AVALIAÇÃO



Fonte: elaboração dos autores a partir de definições com o Instituto Uimbanco.

## IMPACTO MEDIDO DE QUE FORMA?

Para medir o desempenho acadêmico das escolas, entendido nessa avaliação de impacto como proficiências em língua portuguesa e matemática, optamos por utilizar somente medidas derivadas de fontes secundárias, independentes do projeto Jovem de Futuro. Em quase todos os estados onde o projeto foi implantado, estavam em funcionamento “sistemas locais de avaliação da aprendizagem dos alunos” que abrangem *todas* as escolas da rede<sup>2</sup>. O fato de a proficiência ser aferida para todas as escolas e não apenas para aquelas envolvidas na avaliação de impacto do

2. O único estado que não contava com avaliação externa para a 3ª série do Ensino Médio durante o período de avaliação de impacto do Jovem de Futuro era o Rio Grande do Sul e, por essa razão, não foi possível incluir suas escolas no presente estudo.



## *O fato de a proficiência ser aferida para todas as escolas [...] é muito importante para garantir a validade interna da avaliação*

projeto Jovem de Futuro é muito importante para garantir a validade interna da avaliação, isto é, para assegurar que o contraste entre as escolas de tratamento e controle, em termos de proficiência, seja um estimador não enviesado do impacto do projeto.

Quando são utilizadas medidas de proficiência especialmente desenvolvidas para a avaliação de impacto do projeto, mesmo sem a intenção, estaremos introduzindo algum tipo de viés na estimativa. Seja porque as escolas de controle e tratamento podem ter incentivos diferenciados para ir bem ou mal nos testes, seja porque os dois grupos podem ter os mesmos incentivos para se comportar de maneira atípica ao saberem do experimento, por exemplo, colocando excesso de empenho ou excesso de desmotivação na participação nos testes. Se os vieses fossem diferentes para tratamento e controle, estaríamos anulando a vantagem do sorteio, pois o que se observaria no grupo de controle não corresponderia ao que ocorreria com as escolas beneficiadas na ausência do projeto. Consequentemente, o impacto estimado captaria tanto o efetivo benefício do projeto como também diferenças, devido à imperfeita comparabilidade das medidas de proficiência entre tratamento e controle.

Não queremos com isso dizer que os “sistemas locais de avaliação da aprendizagem dos alunos” medem de forma não enviesada a real proficiência dos estudantes. Existem diversas razões não relacionadas com o projeto Jovem de Futuro para que essas medidas sobre-estimem (em particular, quando os alunos são treinados para fazer a prova) e para que subestimem (quando existe desinteresse dos alunos pelo resultado de um exame que não fará parte de seu histórico escolar) a real proficiência dos estudantes. No entanto, como essa mensuração é desvinculada do projeto, não existe razão para que esses vieses sejam diferentes nas escolas de tratamento e controle.

## **ESTIMANDO O IMPACTO**

### *O PROJETO JOVEM DE FUTURO TEM IMPACTO?*

Estatisticamente falando, nosso problema consiste em verificar se existe suficiente evidência para rejeitar a hipótese de que o projeto não tem qualquer impacto sobre a proficiência dos alunos em escolas avaliadas.

Caso o projeto não gere impacto, então a melhoria no desempenho das escolas beneficiadas, considerando o período antes e depois da intervenção, não deve ser estatisticamente diferente da correspondente melhoria observada nas escolas de controle. Ou seja, se o projeto não tem impacto, a chance de as escolas beneficiadas melhorarem mais do que as escolas de controle deve ser 50%. Isso pode ser entendido da seguinte forma: se “melhorar mais a proficiência” é algo independente do Jovem de Futuro, então é natural que, de forma aleatória, algumas escolas beneficiadas melhorem mais que as de controle e outras de controle melhorem mais do que as beneficiadas. Se esse é um efeito aleatório, uma escola beneficiada pelo Jovem de Futuro tem 50% de chance de ir melhor do que uma escola de controle. Por isso, independentemente do número de escolas em cada grupo e da divisão interna entre beneficiadas e controle, caso o projeto não tenha impacto, os dois grupos têm as mesmas chances de melhoria.

Dos 475 grupos de escolas formados nos estados e participantes do sorteio, foram considerados para a avaliação de impacto apenas 141. Isso porque foi necessário excluir grupos em quatro estados, seja porque não dispunham de um sistema externo de avaliação da aprendizagem dos alunos na 3ª série do En-

*Estatisticamente falando, o problema é verificar se existe suficiente evidência para rejeitar a hipótese de que o projeto não tem qualquer impacto*

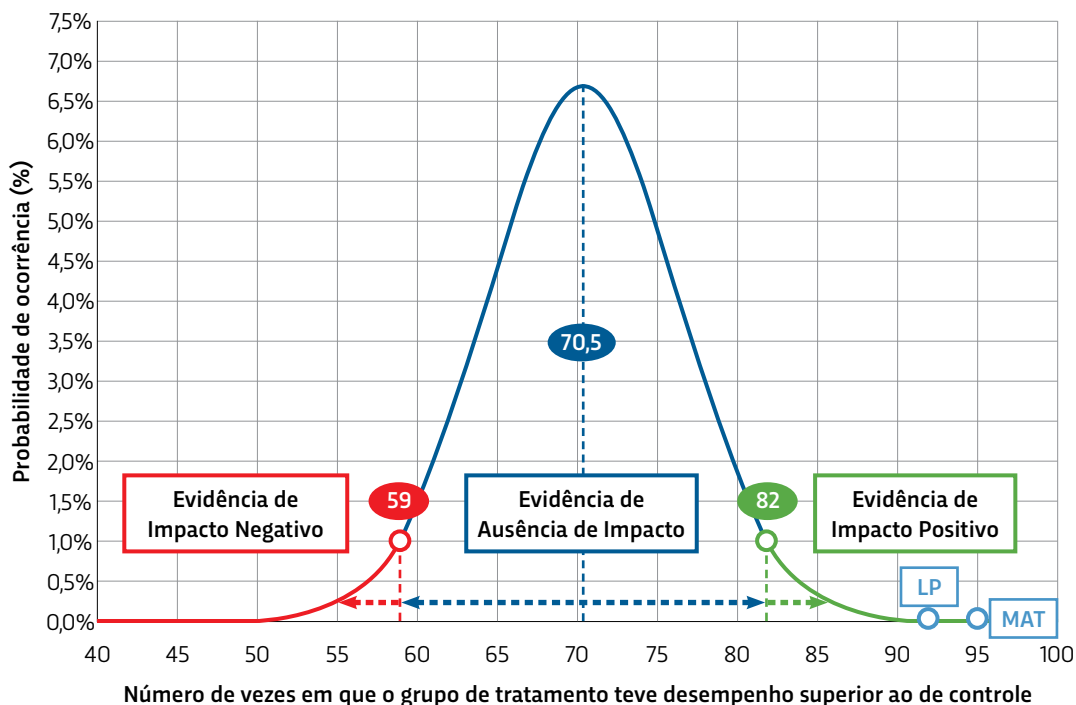
*Independentemente do número de escolas em cada grupo e da divisão interna entre beneficiadas e controle, caso o projeto não tenha impacto, os dois grupos têm as mesmas chances de melhoria*

sino Médio, seja porque o Jovem de Futuro ainda não completou os três anos de implementação.

Consideremos, por exemplo, a proficiência média em língua portuguesa ao final da 3ª série. Em cada um dos 141 grupos de escolas considerados na avaliação de impacto, medimos se a melhoria nas beneficiadas (antes e depois do projeto) foi maior do que nas de controle. Se o projeto não tem qualquer impacto, é esperado que, em metade dos grupos, escolas beneficiadas tenham se saído melhor do que as de controle. Portanto, o esperado é encontrar 70 grupos em que há impacto (beneficiadas se saindo melhor).

Mas é evidente que, da mesma forma que alguém que lança 141 moedas não deveria esperar obter exatamente 70 caras ou coroas, não deveríamos esperar que, na ausência de impacto do projeto, obtivéssemos exatamente 70 casos em que as escolas beneficiadas melhoraram mais do que as de controle. Na verdade, a probabilidade de se obter exatamente 70 caras e 71 coroas em 141 lançamentos é de apenas 6,7%. Assim sendo, utilizando um intervalo de confiança para tornar o argumento mais preciso, se o projeto não tem impacto, existiria 95% de chance de se observar entre 59 e 82 grupos em que as escolas beneficiadas se saem melhor do que as de controle. Para rejeitar a hipótese nula (hipótese de que o projeto não tem impacto), o número de grupos em que se observa que beneficiadas se saem melhor do que controles precisa ser superior a 82 (ver gráfico à direita).

**DISTRIBUIÇÃO DAS VEZES, EM 141 EXPERIMENTOS, EM QUE O GRUPO DE TRATAMENTO TEM DESEMPENHO SUPERIOR AO DO GRUPO DE CONTROLE, CASO O PROGRAMA NÃO TENHA IMPACTO**



Observamos que, três anos depois do início do Jovem de Futuro, em 92 dos 141 grupos de escolas, a melhoria na proficiência em língua portuguesa dos alunos na 3ª série das escolas beneficiadas foi superior à das escolas de controle. Existe, portanto, sólida evidência para rejeitarmos a hipótese de que o projeto não tem qualquer impacto sobre o desempenho acadêmico das escolas.

É mínima a chance de olharmos para esses 92 grupos em que as beneficiadas se saíram melhor e concluirmos equivocadamente que o projeto tem impacto. A chance desse erro ocorrer seria de uma em 10.000, portanto, extremamente improvável.

No caso da matemática, vimos um resultado ainda melhor. Em 95 dos 141 grupos, as escolas beneficiadas se saíram melhor do que as de controle. A chance de estarmos cometendo um equívoco ao concluirmos que o Jovem de Futuro tem impacto na proficiência de matemática é de uma em 100.000.

*Três anos depois do início do Jovem de Futuro, houve melhoria na proficiência em língua portuguesa em 92 dos 141 grupos. Em matemática, vimos um resultado ainda melhor: 95 dos 141 grupos*

Para completar a consistência dessa argumentação, vamos olhar para a situação das escolas antes da implantação do projeto. Conhecemos a proficiência de todas elas no ano imediatamente anterior ao da data de início do Jovem de Futuro. Como o projeto não pode ter impacto sobre o passado e na medida em que o sorteio foi bem conduzido, devemos esperar que em 70 grupos as escolas beneficiadas tenham um ponto de partida melhor do que as de controle. Como visto anteriormente, estatisticamente falando, deveríamos considerar um intervalo de confiança em torno dos 70 casos, que vai de 59 a 82. Os resultados encontrados confirmam integralmente a expectativa de que não há diferenças iniciais relevantes entre tratamento e controle. Com relação à proficiência média em língua portuguesa, há 65 dos 141 grupos em que as beneficiadas tinham situação de partida melhor. No caso da proficiência média em matemática, encontramos que em 68 dos 141 grupos o ponto de partida das escolas beneficiadas era melhor.

Podemos concluir com elevadíssimo grau de certeza que o projeto Jovem de Futuro possui impacto na proficiência média de língua portuguesa e matemática dos alunos de 3ª série do Ensino Médio.

### **QUAL A MAGNITUDE DO IMPACTO DO PROJETO JOVEM DE FUTURO?**

Saber que o projeto gerou impacto é bom, mas talvez não o suficiente, pois a expectativa dos gestores é que esse impacto seja de magnitude relevante.

Para saber a magnitude do impacto do Jovem de Futuro, o procedimento utilizado é muito similar ao que vimos na seção

anterior. Até então, rejeitamos a hipótese de que o projeto não tem impacto, considerando o número de grupos em que se observou que as escolas beneficiadas melhoraram mais do que as de controle. Nesta seção, qualificamos a análise, considerando a informação sobre quantos pontos a mais na escala Saeb<sup>3</sup> as escolas beneficiadas melhoram comparadas com as de controle. A precisão dessas estimativas pode ser obtida a partir da variabilidade dos diferenciais na melhoria em proficiência entre grupos, conforme indicado nas tabelas a seguir.

IMPACTO DO JOVEM DE FUTURO: PROFICIÊNCIA EM LÍNGUA PORTUGUESA						
GERAÇÃO	ÁREA	Impacto sobre o aprendizado na 3ª série, três anos após o início do projeto				
		Média (pontos na escala Saeb)	Número de agrupamentos	Erro padrão (pontos na escala Saeb)	Limite inferior (95% de confiança)	Estatística para detectar heterogeneidade
	Média simples dos experimentos	4	141	1,4	2,1	-
	Média ponderada das áreas	5	-	1,7	1,7	
1ª	Minas Gerais (RM* de Belo Horizonte)	10	4	4,7	2,0	1,1
	São Paulo (RM* de São Paulo)	5	17	4,2	-1,5	0,2
	São Paulo (Vale do Paraíba)	3	20	2,9	-1,3	-0,4
	Rio de Janeiro (RM* do Rio de Janeiro)	6	12	6,8	-5,6	0,1
2ª	Pará	2	15	4,6	-5,6	-0,6
	Ceará	5	25	2,7	0,5	0,1
	Goiás	5	25	4,0	-1,4	0,1
	Mato Grosso do Sul	3	23	3,4	-2,5	-0,4

\*Região metropolitana.

Fonte: elaboração dos autores.

3. A escala Saeb é a escala brasileira na qual expressamos a proficiência em língua portuguesa e matemática, construída pelo Ministério da Educação (MEC).

IMPACTO DO JOVEM DE FUTURO: PROFICIÊNCIA EM MATEMÁTICA						
GERAÇÃO	ÁREA	Impacto sobre o aprendizado na 3ª série, três anos após o início do projeto				
		Média (pontos na escala Saeb)	Número de agrupamentos	Erro padrão (pontos na escala Saeb)	Limite inferior (95% de confiança)	Estatística para detectar heterogeneidade
	Média simples dos experimentos	6	141	1,2	3,8	-
	Média ponderada das áreas	5	-	1,0	3,0	7,9
1ª	Minas Gerais (RM* de Belo Horizonte)	12	4	6,0	1,9	1,2
	São Paulo (RM* de São Paulo)	1	17	2,1	-3,0	-1,9
	São Paulo (Vale do Paraíba)	7	20	2,9	2,8	1,0
	Rio de Janeiro (RM* do Rio de Janeiro)	7	12	5,0	-1,4	0,4
2ª	Pará	10	15	4,7	2,4	1,2
	Ceará	5	25	2,1	1,8	0,3
	Goiás	6	25	3,8	-0,5	0,3
	Mato Grosso do Sul	4	23	2,1	0,6	-0,3

\*Região metropolitana.

Fonte: elaboração dos autores.

Uma primeira estimativa simples da magnitude do impacto do projeto sobre a proficiência média dos alunos ao final da 3ª série do Ensino Médio é de 4 pontos na escala Saeb em língua portuguesa e de 6 pontos na escala Saeb em matemática.

No entanto, como o número de escolas em cada grupo é distinto e a variabilidade do desempenho entre escolas pode mudar entre os estados, a forma estatisticamente mais eficiente de agregar as estimativas de impacto dos 141 grupos não é necessariamente igual à média aritmética simples adotada acima. Em princípio, cada grupo oferece um estimador da magnitude do impacto do projeto com precisão diferenciada, dependendo do número de escolas presentes no grupo, do balanceamento no número de tratamentos e de controles e também da variabilidade

de desempenho entre escolas no estado. Idealmente, a precisão do estimador de impacto será maximizada se as informações dos grupos forem ponderadas por sua precisão, entendida como o inverso da variância.

Quando esse segundo procedimento é utilizado, o impacto sobre a proficiência em língua portuguesa aumenta, passando de 4 para 5 pontos na escala Saeb. No caso da matemática ocorre o contrário. A estimativa de impacto cai, passando de 6 para 5 pontos na escala Saeb.

Em resumo, nossas estimativas pontuais apontam para um impacto do projeto de magnitude entre 4 e 6 pontos na escala Saeb, tendendo a ser um pouco maior em matemática do que em língua portuguesa.

## *HETEROGENEIDADE DO IMPACTO*

Nas tabelas anteriores (pp. 325 e 326) apresentamos tanto estimativas agregadas do impacto do projeto Jovem de Futuro para o conjunto dos estados atendidos, como também estimativas individuais para cada estado. Existe considerável variabilidade entre os impactos aferidos em cada estado. Em língua portuguesa, por exemplo, o impacto nos estados vai de 2 a 10 pontos na escala Saeb. Em matemática variam em um intervalo ainda maior: entre 1 e 12 pontos na escala Saeb.

Existem duas possíveis explicações para essa variabilidade: (a) o projeto realmente tem impacto diferenciado a depender do estado em que é implantado e, nesse caso, dizemos que o impacto do projeto é heterogêneo; ou (b) trata-se apenas de flutuação estatística resultante do erro de medição. Como conhecemos a precisão das estimativas de impacto para cada estado, é perfeitamente possível avaliar a legitimidade ou a importância dessas duas possibilidades.

Uma estatística comumente utilizada para avaliar se as diferenças de impacto entre os estados é real (o que tornaria o impacto heterogêneo) consiste em expressar a diferença entre a magnitude do impacto em um estado e a média para o conjunto deles em múltiplos do erro padrão do estimador de impacto



*O impacto sobre a proficiência é de 5 pontos na escala Saeb [...] Resta, agora, verificar qual a relevância dessa magnitude de impacto*

para o estado. Por exemplo, como vimos no caso de língua portuguesa, o impacto estimado para o conjunto dos estados foi de 5 pontos na escala Saeb. Como revela a tabela da p. 325, em Minas Gerais a magnitude estimada para o impacto do projeto sobre a proficiência em língua portuguesa foi de 10 pontos na escala Saeb, com um erro padrão associado de 4,7. Assim, a diferença em relação ao impacto agregado expressa em múltiplos desse erro padrão seria de 1,1 ( $\frac{9,7-4,6}{4,7} = 1,1$ ). O valor dessa razão em módulo deve ser próximo à unidade ou menor para afirmarmos que se tratam puramente de erros de medida. Já valores bem acima da unidade apontam para alguma dose de heterogeneidade do impacto. A evidência disponível é perfeitamente compatível com a argumentação de que a variabilidade observada é fruto exclusivamente de erros aleatórios, naturalmente presentes na mensuração do impacto em cada estado. Se existe alguma evidência de heterogeneidade, ela se restringe a São Paulo – região metropolitana. Nesse caso, a estatística detectora da heterogeneidade é -1,9, que tomada em módulo nos dá um valor bem superior à unidade.

## **INTERPRETANDO O IMPACTO ESTIMADO**

### **RELEVÂNCIA DA MAGNITUDE DO IMPACTO ESTIMADO**

Já sabemos que é possível afirmar, de forma generalizada, que o impacto do projeto Jovem de Futuro é de 5 pontos na escala Saeb, tanto sobre a proficiência em língua portuguesa quanto em matemática. Resta, agora, verificar qual a relevância dessa magnitude de impacto.

O impacto de uma intervenção é comumente expresso como uma porcentagem do desvio padrão do resultado sobre o qual se está medindo o impacto. Com vistas a facilitar a comparabilidade, tomemos como parâmetro o desvio padrão em proficiência tanto em matemática como em língua portuguesa no Brasil, que é da ordem de 33 pontos na escala Saeb. Assim, um impacto de 5 pontos na escala Saeb corresponde a 15% de um desvio padrão.

Considerando os achados de uma importante meta-análise internacional (Hill, Bloom, Black & Lipsey (2007)<sup>4</sup>), que reúne avaliações de intervenções educacionais voltadas à melhoria do aprendizado de jovens na Educação Média, o impacto médio encontrado em 76 avaliações foi de 24% de um desvio padrão. Se restringirmos o universo aos experimentos, temos 43 avaliações que indicam um impacto médio de 27% de um desvio padrão. Um pouco mais de um terço das avaliações experimentais incluídas nessa meta-análise apresentaram impactos inferiores ao do Jovem de Futuro.

Ainda recorrendo às meta-análises realizadas no Brasil e no mundo sobre avaliação de impacto<sup>5</sup>, sabemos que 5 pontos na escala Saeb (ou 15% de um desvio padrão) é um efeito maior do que o observado em intervenções que fazem os estudantes passarem a ter aulas com um professor mais experiente. O impacto de ter aula com professores mais experientes é 10% de um desvio padrão. Também é maior do que o efeito produzido pela redução do tamanho da sala de aula de 22 para 15 alunos, que também está na ordem dos 10% de um desvio padrão. Por outro lado, considerando especificamente a qualidade do professor, a literatura indica que o impacto de se estudar com um dos melhores professores da escola (aqueles que mais conseguem bons resultados de aprendizagem com seus estudantes) ao invés de estudar com um dos mais fracos é da ordem de 30% de um desvio padrão, portanto o dobro do impacto do Jovem de Futuro.

---

4. Hill, Carolyn J.; Bloom, Howard S.; Black, Alison Rebeck; Lipsey, Mark W. (2007). *Empirical Benchmarks for Interpreting Effect Sizes in Research. MDRC Working Papers on Research Methodology*.

5. Consultar o site Caminhos para Melhorar o Aprendizado ([www.paramelhoraroaprendizado.org.br](http://www.paramelhoraroaprendizado.org.br)).

Uma terceira alternativa para se avaliar a relevância de 5 pontos a mais na escala Saeb é comparar esse efeito com o quanto um aluno aprende ao longo de todo o Ensino Médio. A diferença entre a proficiência média em matemática e língua portuguesa dos alunos ao final do Ensino Médio e a correspondente proficiência ao final da Educação Fundamental foi de 18 pontos na escala Saeb em 2013. Se considerarmos essa diferença como uma boa medida do que tipicamente um aluno aprende ao longo do Ensino Médio e se, a cada série, os alunos adquirem mais ou menos a mesma quantidade nova de conhecimento, o aprendizado por série seria de 6 pontos na escala Saeb. Assim, o impacto do Jovem de Futuro seria mais ou menos 80% do que se aprende em uma série do Ensino Médio.

Por fim, cabe analisar o quanto os estados brasileiros foram capazes de avançar ao longo da última década em proficiência. A evidência disponível revela que, entre 2003 e 2013, apenas quatro unidades da federação conseguiram melhorar a proficiência no Ensino Médio público em 5 pontos na escala Saeb ou mais. Assim, o impacto do projeto é maior do que o progresso alcançado por 85% das unidades da federação ao longo de uma década.

## **IMPACTO SOBRE A ESCOLA VERSUS IMPACTO SOBRE O ALUNO**

Caso não houvesse reprovação, abandono ou mesmo transferência de estudantes entre escolas, *todos* os concluintes do Ensino Médio teriam cursado as três séries na mesma escola. Esse é o caso que poderia nos dar o maior impacto pleno possível e nos permitiria afirmar que o impacto do projeto sobre o desempenho da escola é igual ao impacto sobre a proficiência dos estudantes que ingressaram na 1ª série do Ensino Médio no primeiro ano de implantação e cursaram todas as três séries na escola beneficiada.

No entanto, essa não é a realidade, e há muitos entraves no fluxo. Muitos dos estudantes que participaram da avaliação de impacto do Jovem de Futuro ao final da 3ª série do Ensino Médio foram, na realidade, repetentes na 2ª série ou ainda são repetentes (por duas vezes) da 3ª série. Ao menos, nesse caso, todos

esses estudantes permaneceram por três anos na escola beneficiada, o que nos permitiria afirmar que o impacto sobre a escola é igual ao impacto sobre os alunos que passaram três anos em uma escola beneficiada, independentemente da série em que se encontravam no primeiro ano de implantação da intervenção.

Contudo, sabemos que há alunos que vieram transferidos de outras escolas e que, portanto, não ficaram expostos por três anos ao projeto Jovem de Futuro, e isso faz com que o impacto sobre a escola, que viveu três anos de projeto, não seja o mesmo que o impacto sobre os alunos que cursavam a 3ª série do Ensino Médio no momento final da avaliação.

Ocorre também que o projeto pode mudar o próprio fluxo escolar. Suponhamos que ele aumente a taxa de aprovação e reduza a de abandono. Ao fazer isso, os estudantes que acabam frequentando a 3ª série do Ensino Médio não são os mesmos que a frequentariam na ausência do projeto. Possivelmente seriam estudantes com proficiência média menor do que aqueles que chegariam à 3ª série se a escola não estivesse no projeto. Afinal, na ausência do projeto, as taxas de reprovação e abandono seriam mais altas, permitindo chegar à 3ª série apenas os “sobreviventes”. Mais uma vez, portanto, o impacto sobre o desempenho da escola ao final da 3ª série não é igual ao impacto sobre o desempenho dos alunos que frequentavam essa escola ao final da 3ª série, pois o projeto provocaria mudança na composição do grupo que chega à 3ª série. É possível ainda que bons alunos de outras escolas sejam atraídos para a escola beneficiada e essa experimente uma melhora de resultado sem ter feito, na realidade, qualquer progresso pedagógico com os alunos que lá estavam anteriormente.

*Se todos os concluintes tivessem cursado as três séries na mesma escola, teríamos o maior impacto pleno possível [...]  
No entanto, há muitos entraves no fluxo*

Em suma, ao avaliarmos o impacto sobre o desempenho da escola, o que estamos capturando é o quanto uma escola beneficiada evolui a mais do que evoluiria caso o projeto não existisse. E isso inclui tanto as mudanças na composição de perfil dos estudantes que frequentaram a escola beneficiada em qualquer um dos três anos de projeto, como também o quanto realmente a aprendizagem teria sido influenciada pelo projeto.

Para calcular o impacto do Jovem de Futuro sobre o real desempenho dos estudantes e não apenas sobre o desempenho da escola, seria necessário neutralizar o efeito da mudança de composição no perfil dos estudantes. Para isso, bastaria restringir a análise a um universo fixo de alunos, por exemplo, àqueles matriculados no momento em que o projeto foi inicialmente implantado. Tal estratégia foi seguida em versões anteriores da avaliação de impacto do Jovem de Futuro, revelando efeitos de magnitude similar ao impacto sobre a escola. Além disso, no passado também foram investigados os impactos sobre o fluxo escolar, mas não foram encontrados efeitos significativos, apontando para baixa modificação no perfil dos estudantes que chegam à 3ª série comparado ao que se observaria caso o projeto não existisse. Há alguma evidência de que o impacto sobre a escola pode ser similar ao impacto sobre os estudantes.

## **PRINCIPAIS CONCLUSÕES E RECOMENDAÇÕES**

Os resultados apontam inequivocamente para um impacto positivo do projeto sobre a aprendizagem em língua portuguesa e matemática nas escolas públicas. A magnitude do impacto em ambas as disciplinas é de 5 pontos na escala Saeb ou 15% de um desvio padrão. Embora o tamanho do impacto varie dependendo do estado em que o projeto foi implantado, esse efeito é reflexo de flutuações estatísticas provenientes de erros de medida.

Cinco pontos a mais na aprendizagem em escolas beneficiadas pelo Jovem de Futuro não é pouco, principalmente se consideramos o quanto um aluno brasileiro tem aprendido durante toda a etapa do Ensino Médio. Cinco pontos na aprendizagem é 80% do que se aprende tipicamente em uma série do Ensino

Médio, e, portanto, o efeito do programa é parecido com o que aconteceria caso o Ensino Médio passasse a ter uma série a mais. Ou, ainda, 5 pontos em 3 anos é mais do que 23 dos 27 estados brasileiros melhoraram em proficiência durante toda a última década.

Segundo a literatura internacional, impactos medidos por meio de experimentos podem chegar a 27% de um desvio padrão. É fato que as intervenções com maior capacidade de afetar a aprendizagem estão ligadas ao processo pedagógico e incidem sobre os professores. Mais de um terço das avaliações internacionais mostraram efeitos inferiores aos do Jovem de Futuro. Há espaço para ajustes no desenho e na implementação do projeto visando a elevação do nível atual de impacto.

Caso pretenda-se tomar como meta alcançar níveis próximos aos 25% de um desvio padrão, seria necessário aumentar o nível atual de impacto em cerca de 10 pontos percentuais, o que soa extremamente desafiador.

Perante tal desafio e dado o limite da potência das transformações em gestão para alterar a aprendizagem na escola, muito provavelmente será necessário incorporar no desenho atual componentes que atuem diretamente no processo de ensino-aprendizagem. Percebemos a preocupação do Instituto Unibanco em discutir e caminhar nessa direção.

© 2017 Fundação Santillana e Instituto Unibanco.

**Organização**

Instituto Unibanco

**Superintendente Executivo**

Ricardo Henriques

**Gerente de Gestão do Conhecimento**

Mirela de Carvalho

**Coordenação**

Alan Ary Meguerditchian

**Tradução**

Sibelle Pedral

**Produção Editorial**

Fundação Santillana

**Diretoria**

André Luiz de Figueiredo Lázaro

Luciano Monteiro

Karyne Arruda de Alencar Castro

**Edição**

Ana Luisa Astiz | AA Studio

**Revisão**

Juliana Caldas, Lessandra Carvalho, Marcia Menin e Maria A. Medeiros | AA Studio

**Assistente Editorial**

Lígia Arata Barros

**Projeto Gráfico**

Paula Astiz

**Editoração Eletrônica e Gráficos**

Paula Astiz Design

Dados Internacionais de Catalogação na Publicação (CIP)  
(Câmara Brasileira do Livro, SP, Brasil)

---

Caminhos para a qualidade da educação pública : impactos e evidências / [organização] Instituto Unibanco. — São Paulo : Fundação Santillana, 2017.

Vários autores.  
Bibliografia.  
ISBN: 978-85-63489-38-8

1. Aprendizagem 2. Educação pública 3. Escolas — Administração 4. Gestão educacional 5. Política e educação I. Instituto Unibanco.

17-08108

CDD-371.2

---

Índices para catálogo sistemático:

1. Escolas públicas : Gestão escolar : Educação 371.2

## **Instituto Unibanco**

### **CONSELHO DE ADMINISTRAÇÃO**

#### **Presidente**

Pedro Moreira Salles

#### **Vice-Presidente**

Pedro Sampaio Malan

#### **Conselheiros**

Antonio Matias

Claudia Costin

Cláudio de Moura Castro

Cláudio Luiz da Silva Haddad

Marcos de Barros Lisboa

Ricardo Paes de Barros

Rodolfo Villela Marino

Tomas Tomislav Antonin Zinner

#### **DIRETORIA**

Andréa Pinotti

Claudio José Coutinho Arromatte

Fernando Marsella Chacon Ruiz

Jânio Gomes

Leila Cristiane B. B. de Melo

Marcelo Luis Orticelli

Ricardo Lazcano

Sergio Miron

#### **EQUIPE TÉCNICA**

##### **Superintendente Executivo**

Ricardo Henriques

##### **Gerentes**

**Administração, Finanças  
e Tecnologia da Informação**

Fábio Santiago

**Desenvolvimento de Soluções**

Alexsandro Santos

**Implementação de Projetos**

Maria Julia Azevedo

**Gestão do Conhecimento**

Mirela de Carvalho

**Planejamento, Articulação e Comunicação**

Tiago Borba

## **Fundação Santillana**

### **DIRETORIA**

André Luiz de Figueiredo Lázaro

Luciano Monteiro

Karyne Arruda de Alencar Castro



Este livro foi composto nas fontes Trasandina e Milo Serif  
e impresso em setembro de 2017.

O diálogo virtuoso entre os gestores educacionais e a ciência, motivado pela busca das melhores alternativas para garantir a todos o direito à aprendizagem, é o tema deste segundo volume da coleção *Caminhos para a qualidade da Educação*, que resulta de uma série de seminários de mesmo título realizada desde 2015 pelo Instituto Unibanco. Organizado em torno do tema “impactos e evidências”, traz exemplos práticos e discute a produção científica e o uso de evidências para tomada de decisão na Educação: o projeto Jovem de Futuro, implementado pelo Instituto Unibanco em parceria com redes públicas estaduais, e experiências internacionais; a avaliação de impacto com desenho experimental, a utilização de meta-análises pelos gestores e a combinação de métodos. O primeiro volume trata de gestão escolar.

