


COLEÇÃO
Diálogos sobre Avaliação

POLÍTICAS E AVALIAÇÕES EXTERNAS

Sistemas, Indicadores e Accountability

VOLUME 1



Wagner dos Santos
Denilson Junio Marques Soares
ORGANIZADORES

encontrografia

COLEÇÃO
Diálogos sobre Avaliação

POLÍTICAS E AVALIAÇÕES EXTERNAS

Sistemas, Indicadores e Accountability

VOLUME 1



Wagner dos Santos
Denilson Junio Marques Soares
ORGANIZADORES

encontrografia

Copyright © 2026 Encontrografia Editora.

Todos os direitos reservados. É proibida a reprodução parcial ou total desta obra sem a expressa autorização da editora.

EDITOR CIENTÍFICO

Décio Nascimento Guimarães

EDITORA ADJUNTA

Carolina Gonçalves Caldas

DESIGN

Diagramação: Nadini Mádhava

Design de capa: Nadini Mádhava

Foto de capa: Freepik.com

REVISÃO

Paula Vigneron

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Políticas e avaliações externas : sistemas, indicadores e accountability / organizadores Wagner dos Santos, Denilson Junio Marques Soares. -- Campos dos Goytacazes, RJ : Encontrografia Editora, 2026. -- (Diálogos sobre avaliação ; 1)

Vários autores.
Bibliografia.
ISBN 978-65-5456-163-1

1. Avaliação educacional 2. Educação - Pesquisa
3. Políticas educacionais I. Santos, Wagner dos.
II. Soares, Denilson Junio Marques. III. Série.

26-339820.0

CDD-379

Índices para catálogo sistemático:

1. Políticas educacionais : Educação 379

Aline Grazielle Benitez - Bibliotecária - CRB-1/3129

DOI: 10.52695/978-65-5456-163-1

encontrografia

ENCONTROGRAFIA EDITORA COMUNICAÇÃO E ACESSIBILIDADE LTDA

Av. Alberto Torres, 371 - Sala 1101, Centro

Campos dos Goytacazes, RJ, 28035-581 | Tel: (22) 2030-7746

www.encontrografia.com | editora@encontrografia.com

Comitê científico/editorial

- Prof. Dr. Antonio Hernández Fernández – UNIVERSIDAD DE JAÉN (ESPANHA)
- Prof. Dr. Carlos Henrique Medeiros de Souza – UENF (BRASIL)
- Prof. Dr. Casimiro M. Marques Balsa – UNIVERSIDADE NOVA DE LISBOA (PORTUGAL)
- Prof. Dr. Cássius Guimarães Chai – MPMA (BRASIL)
- Prof. Dr. Daniel González – UNIVERSIDAD DE GRANADA (ESPANHA)
- Prof. Dr. Décio Nascimento Guimarães - IFF (BRASIL)
- Prof. Dr. Douglas Christian Ferrari de Melo – UFES (BRASIL)
- Prof. Dr. Eduardo Shimoda – UCAM (BRASIL)
- Prof.^a Dr.^a Emilene Coco dos Santos – IFES (BRASIL)
- Prof.^a Dr.^a Fabiana Alvarenga Rangel – UFES (BRASIL)
- Prof. Dr. Fabrício Moraes de Almeida – UNIR (BRASIL)
- Prof.^a Dr.^a Fernanda Luísa de Miranda Cardoso – IFRS (BRASIL)
- Prof. Dr. Francisco Antonio Pereira Fialho – UFSC (BRASIL)
- Prof. Dr. Francisco Elias Simão Merçon – FAFIA (BRASIL)
- Prof.^a Dr.^a Gisele Pessin - UENF (BRASIL)
- Prof. Dr. Iêdo de Oliveira Paes – UFRPE (BRASIL)
- Prof. Dr. Javier Vergara Núñez – UNIVERSIDAD DE PLAYA ANCHA (CHILE)
- Prof. Dr. José Antonio Torres González – UNIVERSIDAD DE JAÉN (ESPANHA)
- Prof.^a Dr.^a Margareth Vetis Zaganelli – UFES (BRASIL)
- Prof.^a Dr.^a Martha Vergara Fregoso – UNIVERSIDAD DE GUADALAJARA (MÉXICO)
- Prof.^a Dr.^a Patrícia Teles Alvaro – IFRJ (BRASIL)
- Prof.^a Dr.^a Rita de Cássia Barbosa Paiva Magalhães – UFRN (BRASIL)
- Prof. Dr. Rogério Drago – UFES (BRASIL)
- Prof.^a Dr.^a Shirlena Campos de Souza Amaral – UENF (BRASIL)
- Prof. Dr. Wilson Madeira Filho – UFF (BRASIL)

Este livro passou por avaliação e aprovação às cegas de dois ou mais pareceristas *ad hoc*.

SUMÁRIO

PREFÁCIO	9
APRESENTAÇÃO	19
1. REFORMA CURRICULAR E POLÍTICAS DE AVALIAÇÃO: ESTAMOS NO CAMINHO CERTO?	23
Adriana Bauer	
2. POLÍTICAS DE EVALUACIÓN ESCOLAR Y RENDICIÓN DE CUENTAS: SU IMPACTO EDUCATIVO EN UN SISTEMA DE BAJAS CONSECUENCIAS	53
Mauro C. Moschetti Gerard Ferrer-Esteban Edgar Quilabert Antoni Verger	
3. AVALIAÇÃO EM LARGA ESCALA E ACCOUNTABILITY EM EDUCAÇÃO	105
Luís Antônio Fajardo Pontes Tatiane Gonçalves Moraes	
4. CONCEPÇÕES EDUCACIONAIS DA OCDE E A SUA INFLUÊNCIA POR MEIO DO PISA NA AMÉRICA DO SUL	123
Talita Emidio Andrade Soares Denilson Junio Marques Soares Wagner dos Santos	
5. EVALUACIÓN Y DINÁMICAS DE PRIVATIZACIÓN EDUCATIVA EN ARGENTINA: ENTRE LA MEDICIÓN DE APRENDIZAJES Y LA INTERVENCIÓN DE ORGANIZACIONES NO GUBERNAMENTALES	155
Lucrecia Rodrigo	

**6. POLÍTICAS DE BONIFICAÇÃO E GESTÃO EDUCACIONAL:
CENÁRIOS 187**

Cristiane Machado
Paulo Henrique Arcas

**7. ANÁLISE DO IDEB COMO INDICADOR DE MONITORAMENTO DA
QUALIDADE DA EDUCAÇÃO BÁSICA 215**

Girlene Ribeiro de Jesus
Francisco Augusto da Costa Garcia
Gizelle Alves Xavier dos Santos

**8. AVALIAÇÕES EXTERNAS E IDEB: PERCEPÇÕES SOBRE USOS E
IMPLICAÇÕES NO TRABALHO ESCOLAR 237**

Daiane Ansaloni Mendes
Daniel Abud Seabra Matos
Breyenner Ricardo de Oliveira

**9. LACUNAS NA APRENDIZAGEM E DESIGUALDADES
EDUCACIONAIS NA PANDEMIA: EVIDÊNCIAS DA REDE ESTADUAL
DE MINAS GERAIS 269**

Joyce Soares Rodrigues Petrus
Flavia Pereira Xavier

**10. DIF EM TESTES EDUCACIONAIS: COMPARAÇÃO DE APLICAÇÕES
EM PAPEL E COMPUTADOR 301**

Alexandre Jaloto
Gustavo Henrique Martins

**11. AVALIAÇÃO EM LARGA ESCALA E DESEMPENHO ESTUDANTIL:
UMA ANÁLISE COMPARATIVA ENTRE SRES CAPIXABAS NO PAEBES
(2021-2024) 327**

Lara Nunes Barbosa
Ronildo Stieg
Wagner dos Santos

MINICURRÍCULO DOS AUTORES E DAS AUTORAS 355

PREFÁCIO

1. Introdução e contexto histórico

Nas últimas décadas, os movimentos de reforma do Estado e de modernização da gestão pública, guiados pelos princípios de *eficiência*, *transparência* e *responsabilização*, estimularam a criação de diversas iniciativas voltadas à avaliação de políticas e programas públicos. No campo da educação, esse movimento ganhou especial relevância. Consolidou-se a convicção de que a escola deve assegurar, a cada estudante, os conhecimentos e habilidades necessários para sua participação autônoma e bem-sucedida na vida em sociedade.

Essa concepção, que vincula o desenvolvimento da aprendizagem ao direito à educação, está consagrada no artigo 205 da Constituição Federal e reafirmada em documentos da UNESCO, que definem a aprendizagem como expressão concreta desse direito. O Decreto nº 6.094/2007, ao instituir o *Plano de Metas Compromisso Todos pela Educação*, incorporou esse princípio à política nacional ao estabelecer como primeira diretriz “colocar a aprendizagem como foco e apontar resultados concretos a atingir”.

A partir desse marco, a noção de resultado ganhou centralidade e normatividade no debate educacional. A ideia de garantir o direito à educação

por meio de resultados mensuráveis foi materializada em políticas públicas — inicialmente, com o Sistema de Avaliação da Educação Básica (Saeb) e, a partir de 2007, com o Índice de Desenvolvimento da Educação Básica (Ideb).

Ao longo de quase duas décadas, o teste e o indicador mobilizaram gestores, redes de ensino e a sociedade, trazendo a aprendizagem escolar para o centro das políticas públicas. A importância adquirida por esses instrumentos, naturalmente, tornou-os objeto de críticas e revisões constantes, próprias de políticas que atingem maturidade. As normativas vigentes estabeleceram o encerramento do primeiro ciclo do Ideb em 2022, ano simbólico em que o Brasil celebrou os 200 anos de sua independência. Esse marco abriu uma oportunidade para revisar o sistema de avaliação, incorporando novas funcionalidades e perspectivas e corrigindo limitações identificadas.

2. Contribuição desta coletânea de artigos

É nesse contexto que os artigos desta coletânea foram elaborados. Todas as reflexões sobre o Saeb e o Ideb aqui reunidas partem do reconhecimento de sua relevância histórica e influência profunda na gestão educacional brasileira. No entanto, mais do que discutir sua trajetória, é necessário, neste momento, pensar seu futuro — ou seja, como esses instrumentos devem evoluir para continuar pertinentes e socialmente significativos.

Os textos apontam para uma transição: da ênfase exclusiva nos indicadores para uma visão mais interpretativa e pedagógica da avaliação. A coletânea é plural e multifocal, reunindo análises teóricas e empíricas que discutem a avaliação como fenômeno técnico, político e ético. Produzida por autores de diferentes gerações e instituições, a obra evidencia o vínculo crescente entre as políticas de avaliação e a pesquisa universitária, sinal de amadurecimento do campo.

Os diferentes artigos podem ser lidos como respostas e reformulações para um conjunto de perguntas fundamentais:

- O que deve ser ensinado e, portanto, avaliado?
- Como representar, medir e interpretar os aprendizados esperados?

- Como explicar os resultados à luz das desigualdades sociais e das escolhas pedagógicas de sistemas e escolas?
- Como sintetizar as medidas de resultado obtidas com os testes?
- É possível evitar usos indevidos dos dados obtidos nas avaliações externas?

Essas questões são aprofundadas e reformuladas ao longo da obra, que adquire importância estratégica no momento atual do debate educacional — quando o novo ciclo de avaliação está em fase de concepção e reorientação.

3. Quais aprendizados?

A definição de quais aprendizagens devem ser consideradas essenciais e, portanto, avaliadas permanece um desafio sem solução consensual. Os resultados medidos pelo Saeb têm sido alvo de críticas por induzirem um estreitamento curricular, segundo a percepção de muitos analistas e educadores. Argumenta-se que a concentração dos testes nas áreas de Língua Portuguesa e Matemática teria tornado mais superficial a experiência educativa.

Na base dessa crítica, está a constatação de que a perspectiva consagrada nos quatro pilares do aprender — conhecer, fazer, conviver e ser — introduzidos pelo *Relatório Delors* (UNESCO, 1996), embora seja uma referência fundamental para a formação integral, não encontra expressão nas matrizes de avaliação. Essa distância entre a visão formativa e os instrumentos de mensuração revela uma tensão permanente entre currículo e avaliação, tema central da política educacional contemporânea.

É amplamente reconhecido que essas duas dimensões devem dialogar de forma mais profunda, pois a aprendizagem escolar envolve aspectos cognitivos, sociais, éticos e afetivos, nem todos mensuráveis por testes padronizados. Por outro lado, é preciso reconhecer que a avaliação constitui parte essencial do projeto pedagógico da escola, e suas limitações técnicas não anulam sua relevância como instrumento de reflexão e decisão pedagógica.

Há, contudo, espaços inexplorados para tornar os processos avaliativos mais significativos. O primeiro é reconhecer que os testes de Leitura e Matemática tratam de linguagens estruturantes, indispensáveis para o desenvolvimento de qualquer conhecimento ou competência. Assim, esses componentes deveriam ser tratados não como fins em si mesmos, mas como meios de expressão e integração das aprendizagens. Essa concepção permitiria a criação de instrumentos que, ao mesmo tempo que medem proficiências, favorecem a aprendizagem interdisciplinar.

Os instrumentos atuais, centrados quase exclusivamente em itens de múltipla escolha e de baixa complexidade cognitiva, são suficientes para classificar estudantes em níveis de desempenho, mas incapazes de inspirar e apoiar práticas pedagógicas mais adequadas à promoção do aprendizado. É especialmente significativo que o Saeb e grande parte das avaliações estaduais ainda não tenham incorporado plenamente a Base Nacional Comum Curricular (BNCC)..

De fato, a noção de competência, entendida como a capacidade de mobilizar diferentes habilidades para resolver problemas complexos e contextualizados, é o pilar conceitual da BNCC. No entanto, os testes permanecem baseados em itens isolados, avaliando desempenhos fragmentados. Assim, a importante ideia de que a vida exige a integração harmônica de múltiplas habilidades está ausente dos instrumentos de avaliação, o que os torna ferramentas de medida, mas não de aprendizagem.

Um sistema de avaliação da Educação Básica comprometido com a aprendizagem deve ter como componente estruturante a avaliação formativa, realizada na e pela escola. Essa seria, essencialmente, uma *avaliação para a aprendizagem*, na qual os instrumentos são tarefas significativas e não apenas itens isolados.

Em um modelo assim, as avaliações de larga escala assumiriam função complementar, produzindo medidas que orientem a gestão e o monitoramento das políticas educacionais, sem substituir o protagonismo do professor e da escola no acompanhamento da aprendizagem. Avaliar, nesse sentido, deixa de ser apenas mensurar resultados; passa a ser promover o desenvolvimento humano e o aprimoramento pedagógico.

4. Como representar, medir e interpretar os aprendizados esperados?

A avaliação educacional é um processo inferencial fundamentado em três elementos inter-relacionados: cognição, observação e interpretação. O primeiro passo em qualquer esforço avaliativo é definir claramente o que será medido — etapa denominada *cognição*, na qual se estabelece o modelo de como os estudantes representam o conhecimento e desenvolvem competência no domínio avaliado. A *observação* refere-se às tarefas e às situações apresentadas aos estudantes, que produzem evidências de seu conhecimento; e a *interpretação* diz respeito ao raciocínio utilizado para inferir, a partir dos dados obtidos, o que os estudantes sabem e são capazes de fazer.

Essas três dimensões formam o chamado “triângulo da avaliação”, no qual a cognição ocupa o núcleo teórico, definindo o que significa saber algo e como a aprendizagem se desenvolve. Essa dimensão é operacionalizada por meio da análise de domínio, que identifica os conhecimentos, processos cognitivos e progressões de aprendizagem característicos de determinado campo. Essas análises sustentam a construção do modelo do estudante — isto é, o que se deseja inferir sobre a aprendizagem a partir das evidências — e orientam o desenho das tarefas e critérios de interpretação dos resultados.

Essa abordagem, já adotada em testes internacionais como o Programa Internacional de Avaliação de Estudantes (PISA) e o Estudo Internacional de Progresso em Leitura (PIRLS), também serve de referência para os sistemas de avaliação de diversos países. No caso brasileiro, uma das prioridades na revisão do Saeb deve ser a redefinição das análises de domínio, pois as matrizes atualmente em uso foram elaboradas há muitos anos, antes da consolidação de novas demandas sobre as aprendizagens essenciais para o século XXI.

Além disso, o processo de avaliação precisa ir além da mensuração. Compreender os resultados requer mais do que descrevê-los: exige explicá-los. Há, nesse sentido, duas dimensões complementares da explicação. A primeira é pedagógica, centrada nas *devolutivas pedagógicas* que, ao

identificar, a partir das respostas dos estudantes, quais apoios e intervenções podem favorecer sua aprendizagem, indicam as ações pedagógicas necessárias. A segunda é sistêmica e pública, pois busca compreender os fatores e processos que influenciam o desempenho, permitindo aprimorar políticas e práticas educacionais.

A contribuição de Amartya Sen é especialmente relevante: sua noção de resultados abrangentes destaca que não basta conhecer os resultados obtidos; é necessário entender como foram alcançados e em que condições. Essa perspectiva é particularmente fecunda na educação, na qual muitos determinantes do aprendizado — como as condições familiares, culturais e socioeconômicas — estão além do controle direto das escolas.

Uma política de avaliação verdadeiramente relevante deve, portanto, combinar evidência e interpretação, articulando dados quantitativos e análises qualitativas. Somente assim, a avaliação poderá cumprir sua dupla função: *informar decisões e promover a aprendizagem*.

5. Como sintetizar as medidas de resultado obtidas com os testes?

Uma vez obtidos os resultados das avaliações, surge naturalmente a questão de como sintetizá-los em indicadores que orientem decisões e políticas educacionais. Embora reduzir a complexidade da aprendizagem a um único indicador seja impossível, o sucesso do Ideb demonstrou a importância de dispor de uma medida sintética que, da melhor forma possível, indique em que medida escolas e sistemas asseguram o direito à educação.

As experiências acumuladas ao longo de duas décadas revelam que o Ideb, apesar de seu papel histórico, apresenta limitações que precisam ser enfrentadas no novo ciclo.

Três dimensões centrais podem orientar seu aperfeiçoamento:

1. Interpretação pedagógica das proficiências, com base em níveis de desempenho que expressem o que os estudantes sabem e são capazes de fazer, conferindo sentido formativo ao indicador.

2. Consideração das desigualdades educacionais, reconhecendo que um indicador que ignora as variações de desempenho entre grupos sociais, territoriais, raciais ou de gênero perde validade como medida de equidade.
3. Inclusão de informações sobre evasão e ausência, pois o não comparecimento às avaliações também expressa vulnerabilidade e desigualdade.

A incorporação dessas dimensões é necessária, mas complexa. Alguns propõem a criação de uma cesta de indicadores complementares; outros defendem a existência de dois principais — qualidade e equidade — a serem ou não combinados em uma única medida.

O desafio é reformular a síntese dos resultados sem perder de vista sua função pedagógica e social. Um indicador justo e informativo deve combinar rigor técnico, sensibilidade às desigualdades e relevância educacional para que a avaliação continue sendo instrumento de aprendizagem, justiça e aprimoramento do sistema educacional.

6. Usos polêmicos dos resultados das avaliações

O primeiro ciclo do Ideb, apesar de seus méritos, foi fortemente marcado por um viés regulatório, baseado na crença de que o estabelecimento de metas e o oferecimento de incentivos financeiros seriam suficientes para elevar o desempenho dos estudantes. Essa concepção gerou uma tensão recorrente entre, de um lado, os gestores educacionais, cada vez mais pressionados pelo cumprimento de metas, e, de outro, as equipes pedagógicas, que defendem uma avaliação voltada à compreensão e promoção da aprendizagem.

Muitos dos artigos desta coletânea analisam usos dos resultados das avaliações que, sob uma perspectiva pedagógica orientada para a aprendizagem e o desenvolvimento dos estudantes, são considerados inadequados ou contraproducentes. Entre os usos mais controversos destacam-se a premiação de escolas e a bonificação de docentes com base nos resultados das avaliações externas.

A política de premiação difundiu-se por todo o país, frequentemente apoiada em critérios que desconsideram fatores extraescolares, como as condições socioeconômicas e culturais das famílias. Já a bonificação parte da suposição de que recompensas financeiras estimulariam mais dedicação dos professores, ignorando a complexidade das motivações docentes e o caráter coletivo do trabalho escolar.

Essas práticas assumem que as equipes escolares podem compensar, com seu esforço, dificuldades que são estruturais e sistêmicas. Além disso, não há evidências consistentes de que tais políticas aumentem o engajamento dos estudantes, elemento decisivo para qualquer melhoria efetiva da aprendizagem.

Esses mecanismos também dificultam a identificação de boas práticas educacionais, pois os critérios de reconhecimento baseiam-se, em geral, no desempenho em testes de baixa demanda cognitiva. Práticas que promovem aprendizagens profundas, criativas e transformadoras tendem, assim, a serem invisibilizadas ou subvalorizadas.

Em certo sentido, os testes do Saeb acabaram cristalizando expectativas modestas de aprendizagem, contribuindo para a previsibilidade dos resultados e favorecendo a adoção de estratégias de “treinamento para o teste”. O foco deslocou-se da formação integral dos estudantes para o aperfeiçoamento de técnicas de resposta, esvaziando o potencial educativo da avaliação.

Essas questões, amplamente debatidas nesta coletânea, exigem análise crítica e cuidadosa neste momento em que o país redesenha seu sistema de avaliação. Refletir sobre os usos e limites dos resultados é condição indispensável para que a avaliação cumpra seu propósito essencial: *sustentar a aprendizagem, promover a equidade e fortalecer o compromisso público com uma educação de qualidade para todos.*

7. Encerramento

Destinada a educadores, gestores, pesquisadores e formuladores de políticas, esta coletânea não oferece prescrições, mas convites à reflexão. Ao percorrer suas páginas, o leitor encontrará não apenas resultados

de pesquisa, mas também um chamado à responsabilidade pública pela qualidade e pela justiça educacional.

Esta coletânea reafirma que a avaliação educacional é muito mais que um instrumento técnico; é uma linguagem pela qual a sociedade expressa o valor que atribui à aprendizagem e à justiça.

Ao revisitar conceitos, práticas e resultados, os textos aqui reunidos convidam o leitor a compreender que avaliar é também um ato de escolha — de quais saberes reconhecemos, de quais oportunidades ampliamos e de que futuro educacional desejamos construir.

Que estas reflexões inspirem um novo ciclo de avaliação no Brasil, mais atento às pessoas do que às estatísticas, mais orientado à aprendizagem do que ao controle e mais comprometido com o direito de todos a uma educação de qualidade, equitativa e transformadora.

Chico Soares

Professor Emérito da Universidade Federal de Minas Gerais

APRESENTAÇÃO

O debate sobre avaliação educacional ocupa, nas últimas décadas, posição estratégica na agenda das políticas públicas brasileiras. A expansão do Saeb, a consolidação do Ideb, o fortalecimento das avaliações estaduais e a crescente influência de organismos internacionais têm reconfigurado o modo como sistemas, escolas e profissionais compreendem a qualidade e tomam decisões pedagógicas e institucionais. Este volume, que inaugura a coleção *Diálogos sobre Avaliação*, nasce do propósito de reunir pesquisadores do Brasil e do exterior para discutirem esse cenário, articulando rigor conceitual, análise empírica e reflexão política.

A coleção constitui um desdobramento das ações desenvolvidas no âmbito da pesquisa “Dos exames padronizados ao direito à aprendizagem: experiências para a melhoria da qualidade do Ensino Médio na Rede Estadual do Espírito Santo”, financiada pelo Edital Universal nº 28/2022 da Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES), protocolo nº 980/2023, SIAFEM nº 2023-99KC.

Especificamente, este livro oferece ao/à leitor/a uma análise ampla e diversificada das políticas e sistemas de avaliação externa, situando seus usos, tensões e implicações para a gestão educacional e para o direito à educação. Parte-se do entendimento de que as avaliações não são instrumentos neutros: elas expressam escolhas, prioridades, disputas e

valores que atravessam a construção de agendas públicas e influenciam diretamente escolas, redes e sujeitos. Ao longo dos capítulos, os/as autores/as examinam de que maneira noções como qualidade, equidade, governança e *accountability* são construídas e mobilizadas em diferentes contextos históricos e institucionais.

O capítulo 1, de Adriana Bauer, discute a relação entre políticas curriculares e avaliações externas, tomando a BNCC como eixo organizador de reformas que buscam alinhar currículos, instrumentos e *accountability*. A autora revisita bases teóricas, evidencia tensões entre prescrição e prática e problematiza a crença de que a padronização curricular garantiria, por si só, mais qualidade e equidade. Sua análise articula literatura nacional e internacional, além de destacar os riscos de leituras tecnocráticas que ignoram contextos locais e dinâmicas de implementação.

O capítulo 2, de Mauro C. Moschetti, Gerard Ferrer-Esteban, Edgar Quilabert e Antoni Verger, insere o leitor no debate internacional ao examinar a política de avaliação e prestação de contas da Catalunha/Espanha. A partir de análises empíricas, realizadas mediante questionários aplicados a dezenas de centros educativos, e entrevistas com docentes, diretores e técnicos, os autores analisam como medidas avaliativas, introduzidas sob o discurso da melhoria, produzem tensões entre autonomia escolar, competição institucional e usos regulatórios da informação.

O capítulo 3, de Luís Antônio Fajardo Pontes e Tatiane Gonçalves Moraes, aprofunda o estudo sobre *accountability*, reconstruindo sua trajetória no Brasil e situando as justificativas que legitimaram sua adoção. O texto examina aspectos conceituais, usos políticos, controvérsias e desafios técnicos da responsabilização baseada em resultados. Com rigor analítico, os autores mostram como, em diferentes governos e contextos federativos, o discurso da eficácia e da meritocracia moldou instrumentos, induziu práticas de gestão e, muitas vezes, tensionou princípios de equidade e justiça educacional.

No capítulo 4, Talita Emidio Andrade Soares, Denilson Junio Marques Soares e Wagner dos Santos examinam a influência da OCDE e do PISA nas agendas educacionais da América do Sul. O capítulo discute como modelos globais de avaliação e governança se difundem, são reinterpretados e produzem efeitos concretos sobre políticas nacionais, currículos, indicadores e

reformas pedagógicas. A análise destaca os mecanismos de circulação de ideias, os limites da comparação internacional e as tensões entre agendas globais e projetos educativos locais.

O capítulo 5, de Lucrecia Rodrigo, traz a experiência argentina para o centro da discussão sobre privatização educacional. A autora reconstrói historicamente a atuação crescente de organizações não governamentais, fundações empresariais e redes híbridas que influenciam políticas avaliativas e curriculares. O capítulo mostra como a avaliação se converteu em dispositivo de regulação e poder no país, impactando o trabalho docente, a gestão escolar e o próprio sentido público da educação.

No capítulo 6, Cristiane Machado e Paulo Henrique Arcas analisam políticas de bonificação que premiam resultados obtidos em avaliações externas. O texto dissecar arranjos institucionais, incentivos, indicadores e efeitos sobre o trabalho docente, demonstrando como essas políticas podem produzir tanto engajamento quanto distorções e pressões indevidas. A contribuição é particularmente relevante num cenário em que bonificações e incentivos de mérito se expandem nos estados brasileiros.

O capítulo 7, de Gírlene Ribeiro de Jesus, Francisco Augusto da Costa Garcia e Gizelle Alves Xavier dos Santos, revisita o Ideb como principal indicador de qualidade educacional no Brasil. Os autores analisam bases conceituais, virtudes, limites e incoerências do indicador, discutindo como ele se consolidou como referência nacional e como seus usos nem sempre dialogam com o caráter complexo da aprendizagem. O capítulo contribui com o debate ao propor leituras mais críticas e contextualizadas sobre o Ideb, evitando reducionismos que confundem indicador e qualidade.

O capítulo 8, de Daiane Ansaloni Mendes, Daniel Abud Seabra Matos e Breyner Ricardo de Oliveira, apresenta uma revisão sistemática e uma análise temática das percepções de profissionais da educação sobre avaliações externas e o Ideb. Com base em um *corpus* robusto de pesquisas, os autores identificam regularidades, contradições, tensões e interpretações que influenciam práticas pedagógicas, culturas escolares e processos de gestão. O capítulo ilumina a dimensão subjetiva e vivida da avaliação, muitas vezes invisibilizada nos debates técnicos.

No capítulo 9, Joyce Soares Rodrigues Petrus e Flavia Pereira Xavier examinam as desigualdades educacionais acirradas pela pandemia de Covid-19. O capítulo analisa o ensino remoto emergencial, as lacunas de aprendizagem, os impactos diferenciados entre grupos sociais e as evidências provenientes das avaliações do Sistema Mineiro de Avaliação e Equidade da Educação Pública (SIMAVE). A discussão articula dados, marcos legais e literatura recente, evidenciando como a crise sanitária ampliou vulnerabilidades e impôs novos desafios à política educacional.

O capítulo 10, de Alexandre Jaloto e Gustavo Henrique Martins, aprofunda a discussão metodológica ao apresentar uma análise de Funcionamento Diferencial de Itens (DIF) entre aplicações em papel e computador de exames educacionais. O capítulo combina avanços psicométricos, Teoria de Resposta ao Item (TRI), simulações e o método RMSD, oferecendo um estudo robusto sobre validade, invariância e justiça avaliativa. É uma contribuição essencial para pesquisadores interessados na transição para avaliações digitais.

Por fim, o capítulo 11, de Lara Nunes Barbosa, Ronildo Stieg e Wagner dos Santos, analisa comparativamente o desempenho das 11 SREs capixabas no Paebes entre 2021 e 2024. O texto discute variações regionais, desigualdades territoriais e potenciais implicações para a gestão educacional, integrando análise documental, dados do Centro de Políticas Públicas e Avaliação da Educação (CAEd) e abordagens comparativas.

Coerente com a proposta da coleção *Diálogos sobre Avaliação*, este volume reafirma a necessidade de compreender a avaliação educacional para além de seus instrumentos, cifras e métricas, assumindo-a como norteadora de políticas públicas comprometidas com a garantia do direito à educação com aprendizagem. Esperamos que este livro contribua para qualificar o debate público, fortalecer a produção de conhecimento e inspirar políticas educacionais comprometidas com equidade, justiça social e o efetivo direito à aprendizagem.

Wagner dos Santos
Universidade Federal do Espírito Santo

Denilson Junio Marques Soares
Instituto Federal de Minas Gerais

CAPÍTULO 1

REFORMA CURRICULAR E POLÍTICAS DE AVALIAÇÃO: ESTAMOS NO CAMINHO CERTO?

Adriana Bauer¹

DOI: 10.52695/978-65-5456-163-1.1

Este capítulo trata da aposta das políticas educacionais propostas desde a última década do século XX de relacionar as proposições curriculares e os resultados das avaliações externas, questionando se essa aposta tem tido os efeitos desejados para a melhoria da qualidade da educação no Brasil, tendo em vista a recente introdução da Base Nacional Comum Curricular (BNCC) no cenário das reformas curriculares. O texto inicia contextualizando as relações entre currículo e avaliação e o cenário de implementação da BNCC. Posteriormente, apresenta uma revisão de literatura sobre a sua implementação, discutindo as poucas ações e pesquisas que têm sido propostas nesse sentido. Finaliza argumentando sobre a necessidade de iniciativas de avaliação da implementação que estejam concatenadas entre si e que considerem as dinâmicas estabelecidas nos diferentes contextos de implementação, a fim de poder compreender o que tem sido concretizado nas redes de ensino, visando à melhoria da qualidade da educação.

1 Fundação Carlos Chagas. Faculdade de Educação da USP. E-mail: adbauer@fcc.org.br; dri.bauer@usp.br.

Avaliação e currículo: contextualizando a relação

No Brasil, as relações entre avaliação e currículo têm sido alvo de reflexão em diversos trabalhos acadêmicos, principalmente após a consolidação de sistemas de avaliação em larga escala que começaram a se desenvolver a partir da década de 1990. No entanto, de acordo com Bauer (2020), as conexões entre avaliação e currículo existem pelo menos desde o século XIX, com mudanças em suas características e propósitos ao longo do tempo. Tais relações são indissociáveis e fundamentais para o desenvolvimento educacional.

No Brasil, a União passou a demandar avaliações em larga escala de sistemas de ensino no final da década de 1980, a partir de uma preocupação com os índices de fluxo e com a qualidade dos serviços educacionais e também impelida pela necessidade de um gerenciamento eficaz dos recursos disponíveis para a educação. Rapidamente, os sistemas estaduais de ensino também passaram a valorizar as informações obtidas por meio da aplicação de testes aos alunos como suporte à tomada de decisões na área educacional (Lopes, 2007; Brooke, 2011) e, mais recentemente, observa-se o crescimento da adoção de sistemas de avaliação também por municípios, com a mesma lógica do sistema estadual e federal.

Desde a implementação das avaliações em larga escala em contexto nacional, diversos são os autores que têm criticado as relações que se estabelecem entre currículo e avaliação, apontando que tais avaliações geram indução e reducionismo curricular, controle externo sobre os currículos praticados, falta de autonomia dos professores e padronização do trabalho pedagógico, com mudanças nas formas de avaliar a aprendizagem dos estudantes e organizar os planos de ensino (Menegao, 2015). São poucos os estudos que apontam aspectos positivos na conexão currículo-avaliação, como a organização do trabalho escolar, a indução ao ensino de conteúdos relevantes frequentemente negligenciados na sala de aula e o apoio à formação continuada de professores.

A passagem dos desenhos avaliativos de uma lógica amostral para uma lógica censitária impulsionou uma série de práticas de gestão, muitas delas equivocadas, a partir dos resultados obtidos como, por exemplo,

a bonificação de docentes a partir dos resultados dos estudantes (Bauer, 2012), o afastamento de diretores de escolas com rendimento considerado baixo, dentre outras. Além disso, a adequação do currículo às matrizes avaliativas tem se estendido amplamente entre os sistemas de ensino, quer por meio da elaboração de materiais estruturados ou orientações mais direcionadas sobre o que ensinar, quer pela adoção de materiais apostilados produzidos por empresas.

Por exemplo, a pesquisa de Batista *et al.* (2015), que procurou analisar os currículos propostos nos estados brasileiros e no Distrito Federal para o segundo segmento do nível fundamental, apontou um movimento de produção e revisão de documentos curriculares pelos estados marcado por uma maior padronização, atribuída pelos autores a um alinhamento entre políticas curriculares e políticas de avaliação.

Essa tendência de maior direcionamento dos conteúdos e habilidades a serem ensinados culminou na proposição, em 2017, de uma Base Nacional Comum Curricular (BNCC), a partir da qual estados e municípios revisaram seus currículos.

BNCC: da proposição à implementação

De acordo com o Ministério da Educação (MEC), órgão responsável pela proposição do documento curricular, a BNCC é essencial para balizar a qualidade da educação no país, já que, ao estabelecer o que todos os estudantes têm o direito de aprender, ela indica um patamar mínimo de aprendizagens a serem consolidadas pelos estudantes brasileiros.

Promulgado o documento, que se tornou um marco referencial normativo, o MEC, em colaboração com o Conselho Nacional de Secretários de Educação (Consed) e a União Nacional dos Dirigentes Municipais de Educação (Undime), empreendeu um esforço sistemático de estruturação de diretrizes orientadoras para subsidiar estados e municípios no complexo processo de (re)elaboração e implementação de seus documentos curriculares.

Tal iniciativa buscou articular as diferentes esferas federativas em torno de consensos mínimos sobre o que deve constituir o núcleo comum da

formação dos estudantes da Educação Básica por meio de um movimento que envolve dois processos: o de padronização curricular e o de valorização de escolhas curriculares locais. Enfatizando que a BNCC é uma *base curricular* e não um *currículo* a ser implementado tal e qual está proposto, a aposta subjacente a essa iniciativa é a de que, com as aprendizagens básicas garantidas a todos os estudantes, a reforma curricular propicia melhoria da qualidade da educação, observada por meio da melhoria dos resultados nas avaliações em larga escala. Em outras palavras, observa-se um movimento que procura equilibrar padronização e diversidade, universalidade e particularidade em um país de dimensões continentais e marcado por profundas desigualdades regionais.²

O MEC reconhece a necessidade de se construírem pontes entre as orientações nacionais e as especificidades locais e de contextualização curricular sem, contudo, abrir mão da garantia de direitos de aprendizagem comuns a todos os estudantes brasileiros. A fim de fomentar esse processo, realizou uma série de indicações sobre a implementação da BNCC em estados e municípios.

No site oficial da BNCC, para além do próprio texto da base, estão disponíveis textos e materiais complementares; dentre esses, o *Guia de implementação da Base Nacional Comum Curricular* (Brasil, 2018). O documento se apresenta como um material de apoio para gestores estaduais e municipais e também escolas que estão em processo de (re)elaboração³ curricular. O guia é direcionado às etapas de Educação Infantil e Ensino Fundamental e destaca sete dimensões orientadoras do processo de implementação:

-
- 2 Em uma perspectiva crítica, alguns estudiosos questionam se, de fato, haverá espaço para as escolhas curriculares locais, tendo em vista que, como as avaliações incidem sobre a parte comum do currículo, a parte diversificada será relegada a um segundo plano. Apontamentos como esse são fundamentais e necessitam de pesquisas empíricas que produzam informações relevantes para subsidiar processos de avaliação da implementação da BNCC (Rocha; Pereira, 2016).
 - 3 O guia usa esse termo sob o pretexto de que ele contempla tanto as redes que estão reformulando seus documentos curriculares quanto aquelas que farão sua primeira elaboração.

1. Estruturação da governança da implementação;
2. Estudos das referências curriculares;
3. (Re)elaboração curricular;
4. Formação continuada para os novos currículos;
5. Revisão dos projetos pedagógicos;
6. Materiais didáticos; e
7. Avaliação e acompanhamento de aprendizagem.

Apesar de organizadas como um percurso linear, podem ser desenvolvidas de forma concomitante.

O documento apresenta, ainda, mais cinco ações transversais que seriam essenciais para a execução de cada uma das etapas anteriores: Planejamento e monitoramento; Comunicação e engajamento; Processos formativos; Apoio técnico e financeiro; e Fortalecimento da gestão pedagógica. Também foi criado o Programa de Apoio à Implementação da Base Nacional Comum Curricular (Pro-BNCC), que oferece suporte técnico e financeiro aos estados e municípios na elaboração e implementação de seus currículos alinhados à BNCC.

Além dos documentos oficiais supracitados, instituições da sociedade civil, tais como o Movimento Pela Base e o Instituto Ayrton Senna, oferecem algum tipo de suporte à implementação da BNCC, fato amplamente criticado por autores que discutem a relação público-privado na educação e a atuação de organizações não governamentais na implementação de políticas públicas educacionais.

Passados mais de sete anos da homologação da BNCC, cabe verificar como anda seu processo de implementação e o que estudos têm apontado em relação aos resultados do processo. Esse é o movimento realizado neste texto.

Mas por que avaliar a implementação de uma política pública?

No Brasil, a análise de políticas públicas, em especial no que se refere à implementação, embora tenha se desenvolvido no último quarto de século, ainda é área de pesquisa incipiente e “fica restrita a analisar como as ações postas em curso são distintas daquelas planejadas, focando, portanto, os ‘erros’ que daí derivam” (Lotta, 2012, p. 21). Tal realidade não é diferente no campo da educação que, em geral, baseia-se pouco em aportes da análise de políticas e de sua implementação.

Reforçando a necessidade de mais estudos sobre análise de implementação de políticas públicas, Lotta (2014), apoiada em autores como Hill e Lipsky, reafirma a importância de se atentar para as diferentes compreensões que os implementadores têm acerca das políticas, programas e ações que implementam, visto que, em geral, a falta de direcionamento e de materiais de apoio à ação dos executores, aliada aos conflitos que emergem durante o processo de implementação, modifica a própria proposta em jogo.

Ademais, aportes da análise de políticas sugerem a necessidade de conhecer e compreender os fatores contextuais e suas influências nos processos de implementação. Lotta (2014) defende a incorporação de dois focos de atenção às análises no campo: o olhar sobre a cadeia de agentes envolvidos no processo de implementação (compreendida como interação) e sobre os valores desses atores, que influenciam a construção de suas práticas e o exercício de sua discricionariedade.

A autora afirma que parte dos estudos sobre implementação baseia-se no princípio de que os atores (tanto os políticos quanto os implementadores) compartilham o significado das políticas. No entanto, não raramente esses sentidos são vagos ou se encontram em disputa, o que remete à importância de focalizar também a compreensão dos implementadores quanto às políticas e à forma de colocá-las em prática. Esse olhar sobre a cadeia de agentes envolvidos e seus valores é fundamental para que se compreendam as práticas concretas que caracterizam a própria implementação.

O reconhecimento de que a implementação da política é sujeita a uma série de disputas, negociações e decisões que influenciam sua consecução e de que cabe compreender esses movimentos para que se possa contribuir para a melhoria desse processo também motiva a análise dos elementos que influenciam sua implementação.

Tais aspectos discutidos no campo teórico são importantes quando o objeto de análise é uma política pública como a BNCC. Sua proposição e formulação foram eivadas de disputas, controvérsias e resistências.

O fundamento legal da BNCC foi instituído pelo artigo 210 da Constituição Federal de 1988, que estabeleceu a necessidade de conteúdos mínimos para o Ensino Fundamental, visando a uma formação básica comum e o respeito à cultura nacional e regional. Tal fundamento foi reiterado, anos depois, pela Lei de Diretrizes e Bases da Educação Nacional (Lei nº 9.394/96), que determinou, em seu artigo 26, que os currículos do Ensino Fundamental e Médio deveriam ter uma base nacional comum a ser complementada por uma parte diversificada.

Mesmo assim, o formato e o conteúdo do documento de referência curricular nacional nunca constituíram um ponto pacífico nas discussões sobre a matéria, nem após as indicações feitas no Plano Nacional de Educação (2014–2024), que postulou a necessidade de lhe dar tratamento.

Em suma, mesmo após a homologação do documento curricular, as disputas em torno da BNCC permanecem e são deslocadas para a etapa da implementação. Devido às dimensões continentais do país e às realidades socioeconômicas e culturais diversas, *intra* e *entre* as diferentes regiões, os desafios colocados à implementação da BNCC são muitos. Nakad e Skaf (2016), por exemplo, ao estudarem esses desafios, elencam questões como a qualificação de professores e do pessoal técnico das secretarias de educação, a diversidade de tipos de escolas existentes, a precariedade da infraestrutura de muitas escolas, as dificuldades de estabelecimento de relações realmente colaborativas no federalismo brasileiro, dentre outros aspectos.

Tais dificuldades não são desconhecidas dos propositores da política curricular. O Movimento pela Base Nacional Curricular Comum (Mello,

2014) elenca algumas adversidades à sua implementação, com qualidade, no cotidiano das escolas. Além disso, observa-se que, mesmo que se estabeleçam diretrizes para o trabalho das escolas, a BNCC não consegue garantir que tais diretrizes sejam implementadas da forma como foram formuladas.

Decorre do exposto que, embora a necessidade de acompanhar a implementação da BNCC, avaliando-a e monitorando-a, seja reconhecida por propositores da política e pelo MEC, até o momento, poucas são as iniciativas que transcorrem nesse sentido.

Este trabalho busca sistematizar os achados de estudos de avaliação da implementação e de implementação da BNCC, pesquisados nas bases de dados do Scielo, Educa e *Google Scholar*.

Implementação da BNCC: o que dizem os estudos?

Em termos metodológicos, procedeu-se à busca de estudos, nas bases de dados do Scielo Brasil, Educa e *Google Scholar*, por meio de dois descritores de pesquisa: “Avaliação da Implementação da BNCC” e “Implementação da BNCC”. Buscaram-se estudos que pudessem elucidar as seguintes questões de pesquisa: têm sido verificadas mudanças nos currículos locais a partir de sua adequação ao documento curricular oficial? Como tem sido a implementação dos currículos nas escolas, construídos a partir da BNCC? Que dificuldades são enfrentadas nesse processo? Como ele é avaliado?

Desde o início da busca, estabeleceram-se, como recorte, estudos que tratassem da implementação da BNCC na Educação Infantil e no Ensino Fundamental ou da avaliação dessa implementação, tendo em vista que as polêmicas em torno do Novo Ensino Médio e da BNCC dessa etapa de ensino, bem como as recentes mudanças propostas para o currículo, permitem supor que o processo de implementação ainda está se consolidando. Definiu-se também que seriam contemplados apenas estudos de caráter geral, ou seja, os que tratassem da implementação da BNCC ou de avaliação dessa implementação em componentes curriculares específicos (Matemática, Ciências, etc.) seriam descartados. Finalmente, optou-se por

considerar apenas artigos em revistas acadêmicas, pospondo os trabalhos de conclusão de curso (TCCs), as teses e dissertações, os capítulos de livros e os artigos apresentados em congressos. Também foram excluídos estudos repetidos nas diferentes bases de dados.

Os resultados da busca, após a aplicação dos critérios de inclusão e exclusão, são organizados no quadro 1.

Quadro 1 – Síntese dos resultados obtidos na busca de trabalhos sobre avaliação da implementação da BNCC

Descritor de busca	Scielo Brasil	Educa	<i>Google Scholar</i>	Total
“Avaliação da Implementação da BNCC”	0	3	3	6
“Implementação da BNCC”	2	17	18	37
Total	2	20	21	43

Fonte: Elaboração da autora

Primeiramente, cabe observar o baixíssimo número de trabalhos que se propuseram a realizar uma avaliação da implementação da BNCC. Mesmo os trabalhos mostrados como resultados da busca com esse descritor não tratam, de fato, de uma avaliação da implementação em alguma região ou localidade. Apenas um trabalho acadêmico, de natureza teórica, trata, especificamente, da avaliação da implementação da BNCC, propondo alguns princípios orientadores para realizar essa avaliação. O que se tem são, em geral, trabalhos que se debruçam sobre o processo de implementação da política e buscam, de forma descolada dos aportes teóricos da avaliação de programas, analisar características e entraves desse processo.

Os resumos desses textos foram lidos para subsidiar uma classificação inicial dos temas em quatro categorias: estudos que discutem a implementação da BNCC em diversas localidades; estudos que discutem ferramentas e processos gerais de implementação; estudos que criticam a influência de agentes externos nas redes na implementação; e estudos que discutem as relações entre BNCC, formação e atuação docente. Além

dessas categorias, observou-se que dois estudos focalizam aspectos teóricos para subsidiar a avaliação e implementação da BNCC, tratando especificamente de pressupostos a serem considerados para se fazer uma análise da política curricular (Meira; Bonamino, 2021) ou a avaliação da sua implementação (Galian; Silva, 2019).

Verificou-se, por meio da análise dos resumos, que, dos 43 estudos encontrados, 18 discutem o processo de implementação curricular em contextos específicos (estados, municípios ou escolas). Nove estudos tratam, de alguma forma, da influência de agentes externos nas redes na implementação da BNCC, discutindo as relações entre o documento aprovado e as diretrizes de instituições multilaterais e do terceiro setor (OCDE, Banco Mundial, Unesco, Todos pela Educação) ou empresariais (Fundação Itaú, Fundação Lemann, Fundação Getúlio Vargas, Fundação Maria Cecília Vidigal, por exemplo). Seis estudos discutem as implicações da BNCC para o trabalho e a formação docente.

Além dos estudos quantificados no quadro 1, pesquisa no *site* oficial do MEC mostrou que um estudo, realizado em duas etapas, procurou analisar o processo de implementação da BNCC, realizado pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd) da Universidade de Juiz de Fora. Optou-se, assim, por incorporar esses estudos aos resultados apresentados na próxima seção em que, primeiramente, são discutidos os resultados do estudo oficial e, em um segundo momento, são apresentados os resultados da revisão da literatura acadêmica realizada a partir da busca dos artigos.

Vale salientar que os estudos selecionados para análise foram sistematizados com o auxílio de ferramentas de Inteligência Artificial (IA). Para cada estudo, foi solicitada uma síntese dos argumentos principais dos autores e a resposta às questões de pesquisa definidas anteriormente: têm sido verificadas mudanças nos currículos locais a partir de sua adequação ao documento curricular oficial? Como tem sido a implementação dos currículos nas escolas, construídos a partir da BNCC? Que dificuldades são enfrentadas nesse processo?

Resultados

O estudo oficial

Em 2021, o CAEd realizou, a pedido do Ministério da Educação, a *Pesquisa de Avaliação e Monitoramento da Implementação da BNCC*, com o objetivo de monitorar o processo de implementação da BNCC nas redes de Educação Infantil, Ensino Fundamental e Ensino Médio e identificar como a Base estava impactando a percepção dos profissionais sobre a gestão educacional e o lugar do currículo na prática pedagógica, com ênfase na compreensão do estágio de implementação, nos desafios enfrentados e nas condições que favorecem ou dificultam o processo.

A primeira etapa da pesquisa buscou mapear percepções de profissionais da educação e identificar elementos que permitissem subsidiar políticas públicas voltadas à consolidação da reforma curricular. Os objetivos específicos incluíram mapear o estágio de implementação, identificar práticas e lacunas na formação de professores, avaliar a disponibilidade e adequação de materiais e recursos didáticos, examinar os mecanismos de gestão e monitoramento, bem como levantar desafios e condições que favorecem a consolidação das reformas curriculares. Metodologicamente, o estudo adotou uma abordagem quantitativa e qualitativa, valendo-se de um *survey* nacional. Os dados foram coletados por meio de questionários *online* aplicados a diferentes segmentos da comunidade educacional, incluindo profissionais de secretarias estaduais e municipais, equipes técnicas, gestores escolares e professores. A análise integrou tratamento estatístico das respostas fechadas e interpretação das respostas abertas.

Os resultados indicaram que a maioria das redes já havia iniciado processos de adequação curricular à BNCC, embora em estágios distintos de consolidação. As redes estaduais, em geral, apresentaram avanço maior na produção de documentos curriculares do que as redes municipais. No que se refere à formação de professores, embora diversas iniciativas tenham sido implementadas, observaram-se lacunas quanto ao alcance, à frequência e à correspondência entre o conteúdo da formação e a prática docente efetiva. Em relação aos materiais e recursos

didáticos, foram identificados problemas de insuficiência e inadequação, com desigualdades regionais na distribuição e nem sempre consonância com as diretrizes da BNCC.

Quanto à gestão e ao monitoramento, verificou-se que, na maioria das redes, os mecanismos eram incipientes ou inexistentes. Os principais desafios elencados pelos participantes incluíram a falta de clareza sobre as diretrizes curriculares, a resistência por parte dos profissionais, a sobrecarga de trabalho e a escassez de recursos financeiros e tecnológicos. Por outro lado, foram apontadas como condições facilitadoras para a implementação: o engajamento das equipes gestoras, a presença de equipes técnicas capacitadas e a oferta de formação continuada de caráter contextualizado. Tais achados evidenciaram que, embora a implementação da BNCC tivesse avançado de forma significativa no território nacional, persistiam desafios estruturais e operacionais que demandavam ações articuladas entre União, estados e municípios, especialmente para o fortalecimento de processos formativos, a ampliação do acesso a recursos adequados e o estabelecimento de mecanismos consistentes de acompanhamento e avaliação.

A segunda etapa da pesquisa também foi conduzida pelo CAEd/UFJF por solicitação do Ministério da Educação (MEC). Para essa etapa da pesquisa, houve uma alteração do instrumento de coleta de dados, com a retirada de questões sobre participação na construção do currículo e a inclusão de perguntas para os professores sobre formação específica em Educação Infantil. As perguntas abertas foram eliminadas, substituídas por novos itens fechados. Manteve-se o mesmo universo da aplicação anterior, o que significa que a pesquisa foi aplicada em 222 municípios, ainda que tenha abrangido um maior número de escolas (n=1490, sendo 521 estaduais e 969 municipais).

Ao todo, segundo o relatório da pesquisa, participaram dessa etapa 23.528 profissionais, sendo 17.658 docentes das escolas (72% dos docentes das escolas participantes), 2.191 coordenadores pedagógicos, 1.288 diretores (86,5% dos diretores das escolas), 1.310 técnicos de secretarias municipais, 227 técnicos de secretarias estaduais e 854 técnicos de regionais.

Tais números representam um aumento significativo da participação dos profissionais da primeira para a segunda aplicação da pesquisa.

Os resultados obtidos na segunda aplicação do estudo de avaliação e monitoramento da BNCC aprofundaram as observações do relatório inicial e, em relação à participação no processo de implementação, mostraram que houve um aumento no envolvimento de todos os profissionais, principalmente dos profissionais das secretarias de educação e escolas. Segundo o relatório, mesmo em um contexto de desafios decorrentes da pandemia, o ritmo e a intensidade da implementação foram mantidos.

Da mesma forma do estudo anterior, a visão dos profissionais acerca dos efeitos da BNCC sobre a gestão educacional parece ter sido positiva para o grupo de técnicos das secretarias municipais e coordenadores pedagógicos. No entanto, essa percepção positiva decaiu para os técnicos das secretarias estaduais (de 89% para 54% de profissionais com percepção positiva) e discretamente para os professores (de 72% para 69%). Uma das conclusões do estudo é que, à medida que o tempo avança, a percepção positiva desses profissionais se torna cada vez mais dependente de seu engajamento direto e contínuo nas atividades de implementação da BNCC.

Os estudos acadêmicos

O Quadro 2 apresenta a seleção de 18 artigos considerados mais relevantes para a discussão sobre a implementação da BNCC e avaliação da implementação da BNCC, utilizados como base para as discussões que seguem.

Primeiramente, cabe destacar que os estudos sobre implementação da BNCC são relativamente recentes, abrangendo principalmente os anos de 2022 e 2024. A pandemia de Covid-19, que fechou as escolas por um longo período no final da segunda década do século, provavelmente foi a responsável por esse fato, tendo em vista a dificuldade de se avaliar a implementação de uma política quando ela fica em suspenso.

Quadro 2 – Estudos sobre avaliação da implementação e implementação da BNCC selecionados (continua)

Autores	Título	Publicação	Ano	Contexto
ALARCÃO, Luciana Sousa Teixeira; SILVA, Mille Anne Ribeiro da; OLIVEIRA, Teresa Helena Batelli de.	A Base Nacional Comum Curricular (BNCC) na prática: experiências e resultados em sala de aula.	Revista Ilustração	2024	Revisão bibliográfica, não identifica os contextos
ANJOS, Ricardo Eleutério; COELHO, Lurdinei de Souza Lines; CRESPI, Maria Clara Favarão	O processo de implementação da BNCC no âmbito da gestão e da docência: uma análise a partir da pedagogia histórico-crítica	Atos de Pesquisa em Educação	2024	Diretoria de Ensino localizada no interior do estado de São Paulo
DESTRO, Denise de Souza.	A ressignificação da Base Nacional Comum Curricular na rede municipal de Juiz de Fora/MG.	Roteiro	2021	Juiz de Fora
FALEIRO, Betina Wiebusch; WANDERER, Fernanda.	BNCC e suas formas de conduzir a docência nos anos iniciais do Ensino Fundamental.	Série-Estudos	2023	Município gaúcho
JESUS, Fernanda; BITTENCOURT, Ricardo Luiz de. I.	Implementação da BNCC na perspectiva dos professores do ensino fundamental	Revista Saberes Pedagógicos	2023	Escola da rede estadual de Criciúma/SC
LIMA, Elmo Souza.	A implementação da BNCC nas escolas famílias agrícolas do Piauí: os impactos na proposta pedagógica e curricular	Linguagens, Educação e Sociedade	2022	Piauí
LIMA, Maria Aires de, RODRIGUES, Bruno Alysson Soares and COSTA, Frederico Jorge Ferreira	Os Desafios de Pensar um Currículo Contra-Hegemônico na Escola do Campo no Contexto da Implementação da BNCC	e-Curriculum	2023	Assentamento no Ceará

Quadro 2 – Estudos sobre avaliação da implementação
 e implementação da BNCC selecionados (continua)

LIMA, Michelle Fernandes; BIER-NASKI, Caroline	Implementação da Base Nacional Comum Curricular (BNCC) nas redes municipais de ensino da microrregião de Irati/Paraná	Jornal de Políticas Educacionais	2022	Irati, Rebouças, Rio Azul e Mallet
MACEDO, Maria de Lourdes Leoncio; SANTOS, Jocyleia Santana dos	BNCC do ensino fundamental: implantação e implementação no Tocantins	Revista Foco	2022	Tocantins
MAGALHÃES, Ramon Mendes da Costa <i>et al.</i>	Análise dos documentos orientadores para a implementação da BNCC e o currículo de referência de Minas Gerais	Revista Interinstitucional Artes de Educar	2023	Minas Gerais
MAGDALENO, Beatriz Vieira; FARIA, Rejane Waiandt S. de C.	Construção colaborativa da BNCC: vozes dissonantes	Educar em Revista	2024	Viçosa/MG
MARCHELLI, Paulo Sérgio	A BNCC e o currículo da educação infantil e do ensino fundamental em Sergipe	Roteiro	2021	Aracajú e SEE Sergipe
OLIVEIRA, Marcia Betania de.	Regime de colaboração e formação docente: práticas discursivas na produção política Pró-BNCC/RN	Roteiro	2021	12ª (DIREC) e (SEME/Mossoró)
OLIVEIRA, Breynner Ricardo de; ALVES, Maria Michelle Fernandes	Capacidades estatais e implementação de políticas educacionais no nível local: a trajetória da Base Nacional Comum Curricular em um município mineiro.	Cadernos de Gestão Pública e Cidadania	2024	Contagem (MG)
PINTO, Jacyguara Costa <i>et al.</i>	Desafios na implementação da BNCC em Macapá: infraestrutura e formação docente	Revista Brasileira de Ensino e Aprendizagem	2025	Macapá

Quadro 2 – Estudos sobre avaliação da implementação e implementação da BNCC selecionados (conclusão)

SANTOS, Adriana Ramos; SILVA, Maria Elisângela M. M.	Formação continuada de professores: uma análise das ações formativas destinadas a implementação da BNCC e do novo currículo do Acre	Debates em Educação	2023	Acre
SANTOS, Geniana	Processos de resignificação após a BNCC: aspectos da produção curricular em Mato Grosso.	Roteiro	2021	Mato Grosso
TEDESCO, Anderson Luiz <i>et al.</i>	A Base Nacional Comum Curricular (BNCC) e o currículo catarinense: travessias perigosas na luta contra as desigualdades sociais.	Práxis Educativa	2023	Estado de Santa Catarina

Fonte: Elaboração da autora (2025).

A partir da análise dos dezoito artigos que compõem o grupo de estudos que discutem a implementação da BNCC a partir de estudos empíricos, é possível identificar que diversas mudanças têm ocorrido nos currículos locais, decorrentes da adequação à BNCC, embora com variações significativas e críticas substanciais ao processo.

De modo geral, os estudos apontam, em consonância com a pesquisa MEC/CAEd, que os currículos locais têm sido remodelados para se alinhar à estrutura da base curricular comum, com foco em competências e habilidades, ainda que, em algumas localidades, prevaleça a valorização dos documentos nelas construídos. No entanto, os agentes implementadores do currículo têm enfrentado desafios na sua adaptação às realidades regionais, ou seja, concatenar o currículo comum e a parte diversificada está longe de ser um processo simples, de acordo com os estudos.

Por exemplo, a análise dos artigos que abordam a implementação do currículo em Minas Gerais — *Análise dos Documentos Orientadores para a Implementação da BNCC e o Currículo de Referência de Minas Gerais*

(Magalhães, 2023), *A resignificação da Base Nacional Comum Curricular na rede municipal de Juiz de Fora/MG* (Destro, 2021) e *Construção colaborativa da BNCC: vozes dissonantes* (Magdaleno; Faria, 2024) — revela que existem pontos em comum quando se trata de como os currículos locais têm sido adaptados à BNCC e os desafios enfrentados nesse processo. Embora com perspectivas metodológicas distintas (análise documental, entrevistas e revisão bibliográfica), as pesquisas convergem em críticas sobre o caráter centralizador da política, a metodologia de implementação e os desafios práticos vivenciados pelos profissionais da educação.

Magalhães (2023) mostra que o conteúdo da BNCC tem sido incorporado, de forma integral, em um processo que difere dos postulados dos guias e materiais de orientação à implementação, isto é, observa-se a discussão em torno do currículo ocorrendo de maneira centralizada e verticalizada, com a opção do estado por elaborar um único documento curricular para todos os 853 municípios, em vez de apoiar grupos de municípios na reelaboração de seus próprios documentos, o que desobriga as instituições de ensino, em última instância, a considerarem as especificidades regionais e locais.

Em Juiz de Fora, a Secretaria de Educação optou por não aderir ao Currículo de Referência de Minas Gerais, preservando suas propostas curriculares locais, construídas desde 2012 com a participação de profissionais da rede. No entanto, Destro (2021) mostra que a BNCC, por seu caráter normativo, impulsionou a “rediscussão” e “ressignificação” dessas propostas locais para se alinharem às diretrizes nacionais. Esse movimento de “tradução” curricular permitiu, segundo a autora, que os profissionais buscassem pontos de convergência e mantivessem aspectos que consideravam importantes, como a organização do currículo por componentes disciplinares e não por áreas. Houve também a manutenção de conhecimentos caros à área de História, como a história afro-brasileira e pré-colombiana, que foram apagados no documento federal.

Já em Viçosa, os professores, para cumprir as indicações da secretaria de educação local, acabaram alinhando seus planejamentos à BNCC e ao currículo estadual, no movimento esperado pelos propositores da nova base curricular.

Na contramão dos estudos que reforçam o caráter indutor da política curricular nacional, Oliveira e Alves (2024), ao estudarem as capacidades estatais do município de Contagem e sua relação com a implementação da BNCC, concluíram que essa implementação não foi mera reprodução das diretrizes federais. Pelo contrário, afirmam que a SME se apropriou da política e a traduziu em um processo próprio, apoiado em sua capacidade técnica (estrutura, burocracia qualificada, diretoria de formação), mas com limites em relação à sua capacidade político-relacional (baixa articulação externa e cooperação intergovernamental, com movimentos de resistência, disputas ideológicas e limitações).

É possível perceber, a partir dos estudos, que os desafios para a implementação dos currículos em Minas Gerais são múltiplos e recorrentes. A crítica mais evidente é à natureza *top-down* da elaboração dos documentos, embora os autores afirmem que ocorreram consultas públicas com um grande número de participantes. As coordenadoras pedagógicas de Viçosa relatam que, apesar dos convites, a participação dos professores foi limitada.

Outro desafio se refere à insuficiência da formação continuada para apoiar os professores no processo de implementação. Os estudos apontam que há ações de formação, mas que essas têm carga horária insuficiente. Esse aspecto pode ter sido esvaziado nas discussões, apartando o currículo de seu caráter político.

Por fim, os estudos ressaltam a falta de apoio e infraestrutura como um obstáculo à implementação da Base. Embora os documentos da BNCC e do ProBNCC orientem sobre a necessidade de recursos e apoio técnico, o estudo em Juiz de Fora destaca a falta de material didático adequado para o currículo local.

Esse aspecto não é apenas uma realidade em Minas Gerais. Em Mossoró, o estudo de Oliveira (2021) mostra que, embora o ProBNCC busque alinhar currículos locais à BNCC de modo centralizado, sua implementação não assume essa perspectiva, visto que a autora argumenta que a implementação é mediada por negociações, resistências e adaptações nas redes municipais e escolas. Ao contrário do que a literatura tem apontado, em relação

a um engessamento curricular que seria causado pela implementação do novo currículo, a autora observou que há espaço para a reinterpretação e ressignificação dos documentos em um movimento que leva a política curricular a se concretizar de forma plural.

Macedo e Santos (2022), ao tratarem do processo de construção do currículo em Tocantins, afirmam que o estado cumpriu as formalidades legais na implantação da BNCC e que o Documento Curricular do Tocantins (DCT) foi aprovado e adotado pelos 139 municípios. No entanto, a pesquisa destacou a falta de material de apoio pedagógico e de uma oferta adequada de formação continuada para os professores, visando a seu preparo para compreender e implementar o novo currículo nas salas de aula, corroborando as análises realizadas pelo CAEd e os estudos já citados sobre a situação da implementação da BNCC em Minas Gerais.

As autoras reportam que a maioria dos professores entrevistados (51%) acredita que o DCT melhorou o planejamento pedagógico, mas uma porcentagem significativa (49%) considera que essa melhora foi apenas parcial. Além disso, para 57% dos professores, não há material didático específico que atenda ao DCT, sendo necessária a busca por materiais complementares, o que é difícil para muitos professores que ainda não possuem acesso à *internet* ou ao computador. O estudo também aponta que metade dos professores ainda não possui formação adequada na área em que leciona, o que agrava a dificuldade de lidar com a nova abordagem curricular.

Aspectos parecidos são apontados por estudo que tem como foco o município de Macapá. A pesquisa qualitativa destaca a precariedade da infraestrutura escolar e a inadequada formação docente como os principais desafios para a implementação da BNCC. A falta de recursos básicos, como energia elétrica e saneamento, e a ausência de conectividade à *internet* tornam inviável a aplicação de metodologias ativas e o uso de tecnologias, aspectos preconizados pelo documento.

A formação continuada dos professores na região é considerada escassa pelos entrevistados, e muitos docentes não têm acesso a programas de atualização que os auxiliem a adaptar suas práticas às novas diretrizes. A

alta rotatividade de professores na rede pública também é um fator crítico que impede a continuidade de projetos pedagógicos. O estudo conclui que a BNCC não contempla as especificidades socioculturais de Macapá, como as comunidades indígenas e ribeirinhas, o que gera um distanciamento entre o currículo e a realidade dos estudantes.

As dificuldades apresentadas nos estudos supracitados não são realidade apenas nas regiões mais pobres do Brasil. Tedesco *et al.* (2023), ao estudarem a realidade catarinense, reforçam achados de outros estudos aqui discutidos: embora o Currículo Catarinense represente uma adequação formal à BNCC, sua implementação concreta se choca com as desigualdades estruturais do território; em consequência, a promessa de enfrentamento das desigualdades sociais, anunciada como horizonte da BNCC, não se concretiza.

Os demais estudos analisados apenas repetem, com nuances, as realidades mencionadas pelos estudos anteriores, sugerindo que o processo de implementação dos currículos locais precisa ser refletido e apoiado por políticas públicas de formação e de reestruturação das escolas.

No que se refere à avaliação das aprendizagens, uma das etapas propostas pelo *Guia de Implementação da BNCC*, pode-se dizer que os textos selecionados para a realização da pesquisa ora reportada silenciam sobre esse aspecto. No máximo, há comentários sobre a estrutura de competências e habilidades proposta pela BNCC, com vistas à adequação às avaliações em larga escala. Na análise do processo de implementação, a avaliação da aprendizagem, cuja melhoria é uma das justificativas da proposição do novo currículo, não entra, ainda, em discussão/reflexão.

Considerações finais

O primeiro aspecto a salientar, em relação ao presente estudo, foi o relativo baixo número de estudos que se debruçaram sobre a implementação da BNCC ou se propuseram a fazer alguma avaliação dessa implementação. Os estudos de natureza empírica tendem a ser pontuais, analisando a realidade de uma escola ou uma diretoria de ensino, deixando uma lacuna para pesquisas de levantamento como as realizadas pelo CAEd. Estudos de

painel ou levantamentos regionais, com diferentes tipos de escolas e etapas e modalidades de ensino, contribuiriam para o melhor entendimento acerca do processo de implementação da BNCC nas localidades.

Outro aspecto a ser mencionado é o pouco diálogo dos autores com aportes teóricos da implementação de políticas públicas. As lentes teóricas mais comumente adotadas referem-se aos autores críticos de políticas neoliberais ou ainda teóricos afiliados às correntes pós-críticas do currículo. Sem dúvidas, são abordagens interessantes, mas que informam pouco sobre aspectos que podem ser aprimorados no processo de implementação, que poderiam ser utilizados para a melhoria do processo.

Feitas essas ponderações iniciais, pode-se dizer que o levantamento bibliográfico realizado para a produção deste texto revelou que, apesar de a BNCC ter sido aprovada há quase uma década, boa parte dos estudos recentes assume um tom crítico à própria política ou aos mecanismos existentes para sua implementação. Tais estudos, apesar de imprescindíveis para a compreensão do contexto, das ferramentas e estratégias existentes para a implementação da política e das reações que fervilham no chão da escola a partir das mudanças curriculares, não foram foco de análise neste trabalho, tendo em vista que se buscou compreender o processo de incorporação da BNCC aos diferentes contextos.

Em seu conjunto, os estudos que tratam diretamente do processo de implementação nas secretarias de educação e escolas iluminaram como esse tem se dado em diferentes localidades, quer seja em municípios de diferentes portes, quer em redes estaduais maiores. As considerações desses estudos abrangeram a realidade das cinco regiões brasileiras e permitem afirmar que o processo de implementação é bastante desigual e pode acirrar ainda mais as desigualdades de acesso a conhecimento já existentes.

A maioria dos estudos aponta que a formação docente se mostra insuficiente para subsidiar o processo de implementação, concluindo que é necessário que o poder público ofereça mais atividades de formação continuada aos professores, a fim de subsidiar sua prática docente. No entanto, em algumas regiões, as dificuldades de implementação vão para além da formação docente, envolvendo a falta de estrutura básica nas

escolas. Ou seja, o currículo não chega a ser implementado em sua totalidade devido à falta de recursos e materiais de apoio para concretizar as propostas da BNCC. Isso significa que há um movimento de alteração da política curricular que não é, necessariamente, acompanhado da garantia das condições necessárias à sua implementação, o que acaba influenciando o processo e deixando a cargo das escolas realizar o que é possível, dentro das condições existentes. Tal situação permite afirmar que apenas a política curricular não é suficiente para melhorar a qualidade de ensino.

Os estudos apontam que, apesar de as redes estarem envidando esforços para adequar os currículos locais às diretrizes, habilidades e conteúdos curriculares propostos pela BNCC, sua implementação, de fato, no cotidiano escolar ainda enfrenta vários desafios, o que permite questionar se a base curricular tem se convertido em ferramenta de melhoria da qualidade do ensino.

Referências

ALARCÃO, Luciana Sousa Teixeira; SILVA, Mille Anne Ribeiro da; OLIVEIRA, Teresa Helena Batelli de. A Base Nacional Comum Curricular (BNCC) na prática: experiências e resultados em sala de aula. **Revista Ilustração**, Cruz Alta, v. 5, n. 9, p. 231–244, 2024. Disponível em: <https://journal.editorailustracao.com.br/index.php/ilustracao/article/view/397>. Acesso em: 20 jul. 2025.

ANJOS, Ricardo Eleutério; COELHO, Lurdinei de Souza Lines; CRESPI, Maria Clara Favarão. O processo de implementação da BNCC no âmbito da gestão e da docência: uma análise a partir da pedagogia histórico-crítica. **Atos de Pesquisa em Educação**, Blumenau, v. 19, p. 1–23, 2024. Disponível em: <https://ojsrevista.furb.br/ojs/index.php/atosdepesquisa/article/view/10941>. Acesso em: 20 jul. 2025.

ARAÚJO, Enith Romão *et al.* A Base Nacional Comum Curricular (BNCC) e os desafios para a sua implementação. **Lumen et Virtus**, [S. l.], v. 16, n. 46, p. 2103–2116, 2025. Disponível em: <https://periodicos.newsciencepubl.com/LEV/article/view/3760>. Acesso em: 20 jul. 2025.

ARAUJO, Hellen Gregol; LOPES, Alice Casimiro. Redes políticas de currículo: a atuação da Fundação Getulio Vargas. **Práxis Educativa**, [S. l.], v. 16, p. 1809–4309, 2021. Disponível em: <https://revistas.uepg.br/index.php/praxiseducativa/article/view/18297>. Acesso em: 20 jul. 2025.

BAUER, Adriana. É possível relacionar avaliação discente e formação de professores? A experiência de São Paulo. **Educação em Revista**, [S. l.], v. 28, n. 2, p. 61–82, 2012. Disponível em: <https://www.scielo.br/j/edur/a/TPPxVXCHHdM-W6MG96twZJyG/abstract/?lang=pt>. Acesso em: 20 jul. 2025.

BAUER, Adriana. “Novas” relações entre currículo e avaliação? Recolocando e redirecionando o debate. **Educação em Revista**, Belo Horizonte, v. 36, p. 1–29, 2020. Disponível em: <https://www.scielo.br/j/edur/a/LPNd9sV7qbfkGBjv8zcS-vWK/?lang=pt>. Acesso em: 20 jul. 2025.

BATISTA, Antônio Augusto G. *et al.* Centralização e padronização dos currículos: posições e tomadas de posição. *In*: 37ª REUNIÃO NACIONAL DA ANPED, 37., 2015, Florianópolis. **Anais** [...]. Florianópolis: UFSC, 2015.

BITTENCOURT, Jane. Educação integral no contexto da BNCC. **e-Curriculum**, São Paulo, v. 17, n. 4, p. 1759–1780, out. 2019. Disponível em: <https://share.google/bVwqGF0hWepGGGfmo>. Acesso em: 20 jul. 2025.

BORGES, Wendla Mendes Silva; NASCIMENTO, Ilma Vieira do. O Blog como produto técnico tecnológico que auxilia na implementação da BNCC. **Revista de Estudos Aplicados em Educação**, Campinas, v. 6, n. 12, 2021. Disponível em: https://seer.uscs.edu.br/index.php/revista_estudos_aplicados/pt_BR/article/view/8167. Acesso em: 20 jul. 2025.

BORTOT, Camila Maria *et al.* Agenda supranacional e currículo da educação infantil no Brasil: convergências entre proposições do Banco Mundial e a BNCC. **Journal of Supranational Policies of Education (JOSPOE)**, [S. l.], n. 21, p. 07–27, 2025. Disponível em: <https://revistas.uam.es/jospoe/article/view/19565>. Acesso em: 20 jul. 2025.

BRASIL. **Guia de implementação da Base Nacional Comum Curricular**: Orientações para o processo de implementação da BNCC. Brasília, DF: MEC, 2018. Disponível em: https://www.gov.br/mec/pt-br/assuntos/eb/guia_BNCC_2018_atualizacao_2020_cap_1_ao_6_interativo_28.pdf. Acesso em: 1 jul. 2025.

BROOKE, Nigel; CUNHA, Maria Amália. A avaliação externa como instrumento da gestão educacional nos estados. *In*: FUNDAÇÃO VITOR CIVITA. **Estudos & Pesquisas**, São Paulo, 2011. p. 17–79.

COSTA, Raquel; MOLINA, Adão Aparecido. Elaboração e implementação da BNCC (2015/2017) na educação pública brasileira: aproximações com o ideário político-ideológico da Unesco. **Revista Cocar**, [S. l.], v. 14, n. 29, p. 477–497, 2020. Disponível em: <https://periodicos.uepa.br/index.php/cocar/article/view/3389>. Acesso em: 20 jul. 2025.

DESTRO, Denise de Souza. A resignificação da Base Nacional Comum Curricular na rede municipal de Juiz de Fora/MG. **Roteiro**, [S. l.], v. 46, jan. 2021. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S2177-60592021000102013. Acesso em: 20 jul. 2025.

DIAS, Maria Angélica Dornelles *et al.* Implementação da base nacional comum curricular: impactos e perspectivas. **Aracê**, [S. l.], v. 6, n. 2, p. 1324-1338, 2024. Disponível em: <https://periodicos.newsciencepubl.com/arace/article/view/722>. Acesso em: 20 jul. 2025.

DIÓGENES, Camila Gomes; VALOYES, Angie Yirlesa Valoyes; EUZEBIO, Umberto. Implementación de la competencia 10 de la Base Nacional Común Curricular en Brasil: un análisis desde el concepto de Ciudadanía Global de la Agenda 2030. **Revista Brasileira de Estudos Pedagógicos**, [S. l.], v. 101, n. 259, p. 583-606, set. 2020. Disponível em: <https://www.scielo.br/j/rbeped/a/cVNmyjW9Y-GbDhHnKgPfrgFx/abstract/?lang=es>. Acesso em: 20 jul. 2025.

FALEIRO, Betina Wiebusch; WANDERER, Fernanda. BNCC e suas formas de conduzir a docência nos anos iniciais do Ensino Fundamental. **Série-Estudos**, Campo Grande, v. 28, n. 63, p. 91-110, 2023. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S2318-19822023000200091. Acesso em: 20 jul. 2025.

FARIA, Carlos A. P. de. **Implementação de políticas públicas**: teoria e prática. Belo Horizonte: PUC Minas, 2012.

FREITAS, Henrique Campos; NOVODVORSKI, Ariel. O programa UAM: *corpus tool* aplicado à análise do engajamento em corpus sobre a implementação da BNCC. **Cadernos da FUCAMP**, Campinas, v. 22, n. 58, p. 85-107, 2023. Disponível em: <https://revistas.fucamp.edu.br/index.php/cadernos/article/view/3029>. Acesso em: 20 jul. 2025.

GALIAN, Cláudia Valentina Assumpção; SILVA, Roberto Rafael Dias da. Apontamentos para uma avaliação de currículos no Brasil: a BNCC em questão. **Estudos em Avaliação Educacional**, São Paulo, v. 30, n. 74, p. 508-535, mai./ago. 2019. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/5693>. Acesso em: 20 jul. 2025.

GALIAN, Cláudia Valentina Assumpção; PIETRI, Émerson de; SASSERON, Lúcia Helena. Vozes e mensagens nos materiais de apoio à implementação da BNCC. **Currículo sem fronteiras**, [S. l.], v. 21, n. 3, p. 1106-1127, 2021. Disponível em: <https://www.curriculosemfronteiras.org/vol21iss3articles/galian-pietri-sasseron.html>. Acesso em: 20/07/2025 Acesso em: 20 jul. 2025.

JESUS, Fernanda; BITTENCOURT, Ricardo Luiz de. Implementação da BNCC na perspectiva dos professores do ensino fundamental I. **Revista Saberes Pedagógicos**, [S. l.], v. 7, n. 2, p. 40–58, 2023. Disponível em: <https://share.google/ULSdbSeJOVdTOPajl>. Acesso em: 20 jul. 2025.

LIMA, Elmo Souza. A implementação da BNCC nas escolas famílias agrícolas do Piauí: os impactos na proposta pedagógica e curricular. **Linguagens, Educação e Sociedade**, [S. l.], v. 26, n. 52, p. 219–248, 2022. Disponível em: <https://share.google/q2B4NmX5QIGLJw6FJ>. Acesso em: 20 jul. 2025.

LIMA, Maria Aires de, RODRIGUES, Bruno Alysso Soares and COSTA, Frederico Jorge Ferreira. Os Desafios de Pensar um Currículo Contra-Hegemônico na Escola do Campo no Contexto da Implementação da BNCC. **e-Curriculum**, [S. l.], v. 21, 2023. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S1809-38762023000100318. Acesso em: 20 jul. 2025.

LIMA, Michelle Fernandes; BIERNASKI, Caroline. Implementação da Base Nacional Comum Curricular (BNCC) nas redes municipais de ensino da microrregião de Irati/Paraná. **Jornal de Políticas Educacionais**, [S. l.], v. 16, 2022. Disponível em: <https://revistas.ufpr.br/jpe/article/view/86391/47313>. Acesso em: 20 jul. 2025.

LOPES, Valéria Virgínia. **Cartografia da Avaliação Educacional no Brasil**. 2007. 396 f. Tese (Doutorado em Educação) — Faculdade de Educação, Universidade de São Paulo, São Paulo, 2007.

LOTTA, Gabriela S. **Implementação de políticas públicas**: o impacto dos fatores relacionais e organizacionais sobre a atuação dos Burocratas de Nível de Rua no Programa Saúde da Família. 2010. 295 f. Tese (Doutorado em Ciência Política) — Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2010.

LOTTA, Gabriela S. O papel das burocracias do nível de rua na implementação de políticas públicas: entre o controle e a discricionariedade. *In*: FARIA, Carlos A. P. de. **Implementação de políticas públicas**: teoria e prática. Belo Horizonte: PUC Minas, 2012, p. 20–49.

LOTTA, G. S. Agentes de implementação: uma forma de análise de políticas públicas. **Cadernos Gestão Pública e Cidadania**, São Paulo, v. 19, n. 65, p. 186–206, jul./dez. 2014. Disponível em: <https://share.google/oFpzovtqRiBmx29iU>. Acesso em: 20 jul. 2025.

MELLO, Guiomar Namó de. Currículo da educação básica no Brasil: concepções e políticas. **Movimento pela Base**, 1 setembro de 2014, Disponível em: http://movimentopelabase.org.br/wp-content/uploads/2015/09/guiomar_pesquisa.pdf. Acesso em: 20 jul. 2025.

MACEDO, Maria de Lourdes L.; SANTOS, Jocyleia Santana dos. BNCC do ensino fundamental: implantação e implementação no Tocantins. **Revista Foco**, Curitiba, v. 15, n. 7, p. 1-18, 2022. Disponível em: <https://ojs.focopublicacoes.com.br/foco/article/view/611>. Acesso em: 20 jul. 2025.

MAGALHÃES, Ramon Mendes da Costa *et al.* Análise dos documentos orientadores para a implementação da BNCC e o currículo de referência de Minas Gerais. **Revista Interinstitucional Artes de Educar**, Rio de Janeiro, v. 9, n. 3, p. 11-31, 2023. Disponível em: <https://www.e-publicacoes.uerj.br/riae/article/view/74962>. Acesso em: 20 jul. 2025.

MAGDALENO, Beatriz Vieira; FARIA, Rejane Waiandt Schuwartz de Carvalho. Construção colaborativa da BNCC: vozes dissonantes. **Educar em Revista**, Curitiba, v. 40, 2024. Disponível em: <https://www.scielo.br/jj/er/a/fbd6VJrQCRbz9JpLKFVJ-dDS/?format=html&lang=pt>. Acesso em: 20 jul. 2025.

MARCHELLI, Paulo Sérgio. A BNCC e o currículo da educação infantil e do ensino fundamental em Sergipe. **Roteiro**, Joaçaba, v. 46, jan./dez. 2021. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S2177-60592021000102012. Acesso em: 20 jul. 2025.

MEIRA, Marina; BONAMINO, Alicia Maria Catalano de. Contribuições dos estudos de implementação para a análise de políticas educacionais: uma breve discussão do contexto de implementação da BNCC. **Educação em Revista**, Belo Horizonte, v. 37, e78979, 2021. Disponível em: <https://www.scielo.br/jj/er/a/Q96pqGjpFmbYfbzcfwz6bxM/?lang=pt>. Acesso em: 20 jul. 2025.

MELLO, Adriene Ferreira; CAETANO, Joane Marieli Pereira. Gestão democrática e participativa na implementação da BNCC: análise do guia para gestores escolares. **Ensino em Perspectivas**, [S. l.], v. 2, n. 2, p. 1-19, 2021. Disponível em: <https://share.google/X6AFjCyZMK1LYTPh7>. Acesso em: 20 jul. 2025.

MENEGAO, Rita de Cássia Silva Godoi. **Impactos das avaliações externas nos currículos escolares**: percepção de professores e gestores. 2015. 200 f. Tese (Doutorado em Educação) – Universidade Estadual de Campinas, Campinas, 2015.

MORGAN, Karine; MOCARZEL, Marcelo. Formação para a conformação? Uma análise das redes da Fundação Itaú Social na implementação da BNCC. **Interfaces da Educação**, [S. l.], v. 12, n. 35, p. 1039-1064, 2021. Disponível em: <https://share.google/zHcv8Gli7GOuWikJE>. Acesso em: 20/07/2025.

NAKAD, Fabricio Abdo; SKAF, Gabriel Junqueira Pamplona. **Desafios para a implementação da Base Nacional Comum Curricular**. 2017. 84 f. Dissertação (Mestrado Profissional em Gestão e Políticas Públicas – MPGPP) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas (FGV), São Paulo, 2017. Disponível em: <https://repositorio.fgv.br/server/api/core/bitstreams/952272b5-3105-4632-bb7c-987dad31319b/content>. Acesso em: 20 jul. 2025.

OLIVEIRA, Breyner Ricardo de; ALVES, Maria Michelle Fernandes. Capacidades estatais e implementação de políticas educacionais no nível local: a trajetória da base nacional comum curricular em um município mineiro. **Cadernos Gestão Pública e Cidadania**, [S. l.], v. 29, p. 1-18, 2024. Disponível em: <https://www.scielo.br/j/cgpc/a/7bNwMD4GRk6rB885Pn5JhrC/?format=html&lang=pt>. Acesso em: 20 jul. 2025.

OLIVEIRA, Marcia Betania de; LEMOS, Anaylla da Silva; CANUTO, Mônica Barbosa. O Global Education Reform Movement (GERM) e a BNCC: performatividade para formação e práticas docentes. **Revista Teias**, [S. l.], v. 24, n. 74, p. 10-22, jul. 2023. Disponível em: <https://www.e-publicacoes.uerj.br/revistateias/article/view/76015>. Acesso em: 20 jul. 2025.

OLIVEIRA, Marcia Betania de. Regime de colaboração e formação docente: práticas discursivas na produção política Pró-BNCC/RN. **Roteiro**, Joaçaba, v. 46, p. 1-26, jan./dez. 2021. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S2177-60592021000102010. Acesso em: 20 jul. 2025.

PINTO, Jacyguara Costa *et al.* Desafios na implementação da BNCC em Macapá: infraestrutura e formação docente. **Revista Brasileira de Ensino e Aprendizagem**, [S. l.], v. 10, p. 180-190, 2025. Disponível em: <https://rebena.emnuvens.com.br/revista/article/view/330>. Acesso em: 20 jul. 2025.

PONCE, Branca Jurema; ARAÚJO, Wesley. A justiça curricular em tempos de implementação da BNCC e de desprezo pelo PNE (2014-2024). **e-Curriculum**, [S. l.], v. 17, n. 3, p. 1045-1074, jul. 2019. Disponível em: <https://share.google/u9K-NetqIml3Z3fLEsv>. Acesso em: 20 jul. 2025.

ROCHA, Nathália F. E.; PEREIRA, Maria Z. da C. O que dizem sobre a BNCC? Produções sobre a Base Nacional Comum Curricular no período de 2010 a 2015. **Espaço do Currículo**, [S. l.], v. 9, n. 2, p. 215-236, 2016. Disponível em: <https://mail.sumarios.org/artigo/o-que-dizem-sobre-bncc-produ%C3%A7%C3%B5es-sobre-base-nacional-comum-curricular-bncc-no-per%C3%ADodo-de-2010>. Acesso em: 20 jul. 2025.

SANTOS, Adriana Ramos dos; MENDONÇA, Maria Elisangela Martins da Silva. Formação continuada de professores: uma análise das ações formativas destinadas à implementação da BNCC e do novo currículo do Acre. **Debates em Educação**, [S. l.], v. 15, n. 37, p. 1-19, 2023. Disponível em: <https://ufal.emnuvens.com.br/debateseducacao/article/view/14978>. Acesso em: 20 jul. 2025.

SANTOS, Camila Rodrigues dos; MELO, Elda Silva do Nascimento; MORAIS, Erivania Melo de. Base nacional: Uma reflexão a partir da representação social de professores (as) sobre a BNCC. **Revista Diálogo Educacional**, [S. l.], v. 21, n. 70, p. 1241-1269, 2021. Disponível em: <https://periodicos.pucpr.br/dialogoeducacional/article/view/27015>. Acesso em: 20 jul. 2025.

SANTOS, Geniana. Processos de ressignificação após a BNCC: aspectos da produção curricular em Mato Grosso. **Roteiro**, [S. l.], v. 46, p. 1-22, jan. 2021. Disponível em: <https://periodicos.unoesc.edu.br/roteiro/article/view/23803>. Acesso em: 20 jul. 2025.

SCHNEIDER, Marilda Pasqual; ROSTIROLA, Camila Regina. Atuação dos aparelhos privados de hegemonia na implementação da BNCC nos estados da Região Sul. **Revista Educação em Questão**, Natal, v. 59, n. 60, p. 1-23, abr. 2021. Disponível em: <https://periodicos.ufrn.br/educacaoemquestao/article/view/25485>. Acesso em: 20 jul. 2025.

SENA, Ivânia Paula Freitas de Souza; SANTOS, Valéria Prazeres dos; RIOS, Roseane Oliveira. A atuação dos aparelhos privados de hegemonia para controle da educação. **Práxis Educacional**, [S. l.], v. 21, n. 52, p. 1-24, 2025. Disponível em: <https://periodicos2.uesb.br/praxis/article/view/16851/10191>. Acesso em: 20 jul. 2025.

SOUZA, Celina. Capacidade Burocrática para a Implementação de Políticas: Brasil e Argentina em Perspectiva Comparada. In: 10º ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE CIÊNCIA POLÍTICA (ABCP), 2016, Belo Horizonte. **Anais [...]**. Belo Horizonte: ABCP, 2016.

TEDESCO, Anderson Luiz *et al.* A Base Nacional Comum Curricular (BNCC) e o Currículo Catarinense: travessias perigosas na luta contra as desigualdades sociais. **Práxis Educativa**, Ponta Grossa, v. 18, p. 1-23, 2023. Disponível em: <https://revistas.uepg.br/index.php/praxiseducativa/article/view/21450>. Acesso em: 20 jul. 2025.

TOMÉ, Cláudia; SANTOS, Maria. Formação em deslocamentos: ficção e contrariedade em torno na BNCC. **Roteiro**, Joaçaba, v. 46, p. 1-27, jan./dez. 2021. Disponível em: <https://periodicos.unoesc.edu.br/roteiro/article/view/23829>. Acesso em: 20 jul. 2025.

VIEIRA, Kelmara Mendes *et al.* Os temas transversais na Base Nacional Comum Curricular: da legislação à prática. **Educação**: Teoria e Prática, Rio Claro, v. 32, n. 65, p. 1-21, 2022. Disponível em: <https://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/15719>. Acesso em: 20 jul. 2025.

VIEIRA FILHO, Vanderlei José Valim; GONÇALVES, Fábio Peres. Gerencialismo na formação continuada de professores no Brasil: uma análise de documentos propostos pelo Conselho Nacional de Secretários de Educação. **Educação em Revista**, Belo Horizonte, v. 39, p. 1-23, 2023. Disponível em: <https://www.scielo.br/j/er/a/x754kfnZNPnGF55NDVTchs/?format=html&lang=pt>. Acesso em: 20 jul. 2025.

WANDERER, Fernanda; FREITAS, Juliana Veiga de; PRATES, Maria Eduarda Leidens. A terceirização da formação de professores para a implementação da BNCC: os caminhos para o endividamento da docência. **Cadernos do Aplicação**, [S. l.], v. 35, p. 1-13, 2022. Disponível em: <https://seer.ufrgs.br/index.php/CadernosdoAplicacao/article/view/124617>. Acesso em: 20 jul. 2025.

WEBER, Karine *et al.* A BNCC e seus múltiplos sentidos: desafios na (re)criação de práticas curriculares de uma unidade federal de Educação Infantil. **Revista Educação e Políticas em Debate**, [S. l.], v. 13, n. 2, p. 1-17, 2024. Disponível em: <https://seer.ufu.br/index.php/revistaeducaopoliticas/article/view/69680>. Acesso em: 20 jul. 2025.

CAPÍTULO 2

POLÍTICAS DE EVALUACIÓN ESCOLAR Y RENDICIÓN DE CUENTAS: SU IMPACTO EDUCATIVO EN UN SISTEMA DE BAJAS CONSECUENCIAS

Mauro C. Moschetti¹

Gerard Ferrer-Esteban²

Edgar Quilabert³

Antoni Verger⁴

DOI: 10.52695/978-65-5456-163-1.2

Introducción

En las últimas décadas, numerosos países han puesto en marcha reformas para modernizar sus sistemas educativos y mejorar simultáneamente su efectividad y equidad. La centralidad creciente de la educación en la llamada sociedad del conocimiento, los acelerados cambios tecnológicos y la globalización, así como el impacto de evaluaciones internacionales como PISA o PIRLS en el debate público, han consolidado un énfasis de la

1 Universitat Autònoma de Barcelona. E-mail: mauro.moschetti@uab.cat.

2 Universitat Oberta de Catalunya.

3 Universitat Autònoma de Barcelona.

4 Universitat Autònoma de Barcelona.

política educativa en elevar el rendimiento académico del alumnado. Bajo esta lógica, y con la intención de superar las rigideces asociadas a los modelos burocráticos del siglo XX, se han impulsado políticas que promueven la autonomía de los centros, introducen evaluaciones basadas en resultados académicos y refuerzan los mecanismos de rendición de cuentas. Estas estrategias buscan que las escuelas amplíen sus márgenes de decisión sobre aspectos organizativos y pedagógicos, desarrollen proyectos de mejora educativa propios y asuman una responsabilidad más directa sobre su desempeño.

Actualmente, la autonomía escolar, la evaluación y la rendición de cuentas (RdC) se articulan como un bloque de políticas interdependientes. A cambio de una mayor libertad organizativa, presupuestaria o curricular, las administraciones educativas exigen mayores niveles de supervisión sustentada en evidencias — como pruebas externas estandarizadas o procesos de inspección. Esta lógica permite que docentes y equipos directivos adapten sus proyectos educativos a las características del entorno y accedan a un abanico más amplio de recursos formativos y herramientas orientadas a la mejora escolar. Sin embargo, esta mayor autonomía implica también una mayor responsabilidad sobre los resultados obtenidos.

Aunque esta tendencia de reforma educativa se ha generalizado en países muy diversos — en lo que algunos autores describen como un proceso de convergencia internacional (Breakspear, 2012; Teltemann; Jude, 2019) —, la evidencia sobre su eficacia es mixta (Torres, 2021). En algunos contextos, la tríada autonomía, evaluación y RdC fomenta la innovación, refuerza la enseñanza en el aula y eleva el rendimiento escolar (De Grauwe, 2005; Greany, 2018); en otros, favorece la competencia entre escuelas, la segmentación de la oferta y prácticas centradas en *teaching to the test* (una enseñanza centrada en la preparación para las pruebas), la reducción del currículo a los contenidos estrictamente evaluados o la separación del alumnado por niveles de rendimiento, con efectos negativos sobre la equidad (Ferrer-Esteban, 2015; Pagès *et al.*, 2023; Werler; Færevaaag, 2017).

Por último, estas políticas persiguen objetivos distintos según el contexto: mientras que en algunos países se orientan a mejorar la calidad y la equidad educativas, en otros buscan fomentar la elección escolar y la

competencia entre centros (Verger; Fontdevila; Parcerisa, 2019). Su impacto, evolución y configuración dependen en gran medida de las características de los sistemas educativos en los que se implementan, así como del diseño específico de las propias políticas. La mayoría de los estudios sobre estas políticas se han centrado en países anglosajones, donde han predominado modelos de evaluación con altas consecuencias. En estos contextos, se ha observado una intensificación de comportamientos oportunistas y la aparición de efectos no deseados (Mittleman; Jennings, 2018).

En Cataluña (España), foco de este capítulo, el sistema de evaluación y RdC se caracteriza por ser eminentemente *low-stakes*, es decir, de bajas consecuencias, en el sentido de que los resultados no se hacen públicos ni determinan la carrera profesional del profesorado ni la financiación de los centros educativos. La institucionalización de estas políticas se produjo con la *Llei d'Educació de Catalunya* (2009) y sus decretos reglamentarios (2010), que reforzaron la autonomía de los centros, reglamentaron las evaluaciones externas y profesionalizaron el liderazgo y la gestión escolar. Dicha ley aspiraba a fomentar la innovación pedagógica y una gestión más flexible en un sistema históricamente muy centralizado.

En este contexto, las pruebas de Competencias Básicas (CB), iniciadas en el curso 2000–2001, han ido adquiriendo un peso creciente como instrumento de diagnóstico y mejora. Sin embargo, aún existe un número limitado de estudios que analicen cómo se implementan estas políticas en la práctica en contextos *low-stakes*, qué significados adquieren para los distintos actores escolares y qué usos efectivos se hacen de los datos generados por las pruebas en el ámbito de los centros educativos.

Este capítulo tiene por objetivo analizar los efectos y significados de las políticas de evaluación externa y rendición de cuentas a partir de un estudio de caso en Cataluña. Lejos de tratarse de un análisis meramente local, el caso catalán permite observar las tensiones, mediaciones y resignificaciones que emergen entre los marcos normativos de la política educativa y su implementación cotidiana en los centros escolares. Cataluña ha desarrollado un modelo que combina autonomía escolar con evaluaciones de carácter fundamentalmente formativo, en un entorno cada vez más competitivo — marcado por una notable presencia del sector privado,

el descenso demográfico y la ampliación de la elección escolar. En este contexto, el estudio ofrece claves analíticas útiles para comprender cómo estas políticas operan realmente en la práctica y en qué condiciones generan determinados efectos, tanto esperados como no esperados.

El capítulo se organiza del siguiente modo. Primero presentamos el marco conceptual del estudio, centrado en la relación entre autonomía escolar y evaluación externa, y exponemos el enfoque teórico que orienta nuestro análisis, basado en las lógicas de apropiación de las políticas educativas por parte del profesorado. A continuación, describimos el contexto normativo catalán en el que se insertan estas políticas. En el apartado siguiente detallamos el enfoque metodológico, que combina un cuestionario administrado a una muestra representativa de 78 centros de enseñanza obligatoria con entrevistas semiestructuradas a docentes y equipos directivos de 20 centros.

Posteriormente, presentamos los resultados empíricos, organizados en tres apartados: en el primero analizamos las fuentes de información que utilizan los centros para orientar sus procesos de mejora; en el segundo, examinamos el uso que se hace de los datos generados por las evaluaciones externas, en particular por las pruebas de competencias básicas; en el tercero, abordamos las experiencias de presión asociadas a dichas pruebas y sus efectos en la cultura escolar; finalmente, cerramos el capítulo con una discusión de los hallazgos y algunas recomendaciones de política derivadas del análisis. El caso de Cataluña se propone como un ejemplo paradigmático de sistema *low-stakes*, cuya trayectoria ofrece claves valiosas para repensar las políticas de evaluación en otros contextos con características similares.

Diseño y efectos de las políticas de evaluación y rendición de cuentas escolar

Las políticas de evaluación y rendición de cuentas han ido habitualmente de la mano de procesos de descentralización y autonomía escolar, pues exigen que un amplio abanico de agentes — desde la administración hasta las familias y los centros — asuman responsabilidad sobre los resultados educativos. Estas políticas se apoyan en estándares de aprendizaje

y marcos de competencias, sistemas de monitorización, planes de mejora e incentivos vinculados a los logros académicos. En la práctica, son sobre todo los docentes y los equipos directivos quienes rinden cuentas de los avances de sus estudiantes, medidos por evaluaciones externas que funcionan como indicadores de calidad a nivel de aula, escuela, zona e incluso del sistema en su conjunto (Nichols; Harris, 2016).

En las dos últimas décadas ha crecido exponencialmente el número de países que aplican evaluaciones o pruebas externas censales, sobre todo en lengua, ciencias y matemáticas (UNESCO, 2019; Verger *et al.*, 2019). Aunque puedan diferir en nivel educativo, frecuencia, materias y modalidad de administración, las pruebas persiguen, en mayor o menor medida, dos objetivos: por un lado, ofrecer información diagnóstica y formativa que oriente al profesorado, ajuste el currículo y ayude a planificar el desarrollo profesional; por otro lado, reforzar la rendición de cuentas mediante la publicación de resultados y la vinculación de incentivos o sanciones a los logros alcanzados (Clarke *et al.*, 2021).

Diversos estudios han identificado también beneficios potenciales asociados a estos sistemas. Se ha señalado, por ejemplo, que las pruebas externas pueden elevar las expectativas de rendimiento de estudiantes y centros (Dee y Jacob, 2011), facilitar la identificación de buenas prácticas y orientar los esfuerzos hacia estrategias respaldadas por evidencia (Datnow *et al.*, 2007; Vogler, 2017), dirigir recursos adicionales a los grupos con mayores dificultades y promover una cultura de transparencia y mejora continua entre el profesorado. Según la OCDE, los efectos más positivos de la rendición de cuentas tienden a observarse cuando la publicación de resultados se combina con altos niveles de autonomía en la gestión de recursos humanos y económicos (OCDE, 2011; Torres, 2021). Sin embargo, en el contexto europeo esta relación aparece empíricamente más débil y heterogénea (European Commission, 2018), lo que sugiere que los resultados dependen de cómo se articulan estos instrumentos en cada sistema educativo.

Una variable central para comprender los efectos diferenciales de la rendición de cuentas es el tipo y nivel de consecuencias que se derivan de las evaluaciones externas. Cuando las pruebas están asociadas a incentivos o sanciones significativas para estudiantes, docentes o centros

— lo que caracteriza a los sistemas *high-stakes* —, abundan las evidencias de efectos no deseados: la enseñanza se concentra en los contenidos evaluados, se estrecha el currículo y aumenta el estrés entre docentes y alumnado; además, se incrementan las prácticas de selección y agrupación que comprometen la equidad interna de los centros (Jerrim y Sims, 2021; Mittleman y Jennings, 2018). A ello se suma que la combinación de autonomía escolar y publicación de resultados puede potenciar la competencia entre escuelas: las que obtienen mejores puntuaciones las utilizan como recurso para atraer alumnado con mayor capital económico y cultural, mientras las familias más informadas eligen a partir de estas referencias, contribuyendo así a una mayor segmentación del sistema (Hamilton *et al.*, 2002; Allen y Vignoles, 2007).

En contraste, los sistemas de bajas consecuencias conciben las evaluaciones externas como herramientas principalmente formativas. En estos casos, los resultados sólo generan implicaciones para los centros con rendimientos persistentemente bajos y, cuando esto ocurre, suelen traducirse en mayor acompañamiento institucional y oportunidades de desarrollo profesional, en lugar de sanciones. El uso de los datos se orienta a promover la reflexión pedagógica y la identificación de áreas de mejora. Aun así, diversos estudios muestran que, incluso en estos entornos, el profesorado puede cuestionar o resistir estas políticas cuando las percibe como mecanismos de control externo que interfieren en su autonomía profesional (Maroy y Voisin, 2017; Barbana *et al.*, 2020; Thiel *et al.*, 2017).

Asimismo, distintos autores han subrayado que la efectividad de la rendición de cuentas no depende únicamente de los instrumentos utilizados, sino del uso que se haga de los datos generados. Para que esta información contribuya realmente a la mejora, resulta imprescindible que los actores implicados la perciban como válida y dispongan de la formación, el tiempo y los recursos necesarios para analizarla e interpretarla de manera rigurosa. De lo contrario, la infrautilización o el uso indiscriminado de los resultados puede convertir la inversión en evaluaciones externas en una estrategia ineficaz e incluso costosa (Terhart, 2013; UNESCO, 2019).

En cualquier caso, los estudios empíricos sobre los efectos de las políticas de evaluación y rendición de cuentas *low-stakes* siguen siendo

mucho más limitados que aquellos centrados en contextos *high-stakes* (Levatino *et al.*, 2024).

Profesorado y apropiación de políticas educativas

El análisis de la implementación de las políticas educativas de gobernanza escolar plantea importantes retos. Parte de su complejidad radica en que se trata de políticas en continua evolución, replanteamiento y ajuste. Además, su impacto sobre el comportamiento y las prácticas del profesorado suele ser indirecto y está fuertemente mediado por una multiplicidad de factores, como la cultura escolar, las dinámicas organizativas internas, las trayectorias profesionales o las condiciones materiales de los centros.

En este estudio, centrado en el caso de Cataluña, analizamos la implementación de las políticas de evaluación y RdC en el ámbito escolar. También exploramos su impacto en la práctica educativa y en la organización interna de los centros educativos. De forma general, el estudio busca examinar hasta qué punto se materializan en la práctica los principales supuestos y efectos esperados de las políticas de autonomía, evaluación y RdC, especialmente en lo que respecta a los procesos de mejora e innovación educativa y a la conformación de equipos docentes cohesionados. Por lo tanto, centramos nuestro análisis en la asociación de estas políticas con las prácticas pedagógicas y organizativas, así como al uso que hacen los centros escolares de las evaluaciones externas, entre otras fuentes de información, para producir cambios y mejoras escolares.

En este estudio entendemos que la asociación de las políticas de autonomía y evaluación, a nivel de sistema, con las prácticas escolares concretas, a nivel de centro, no depende únicamente de cómo la administración educativa regula e implementa las políticas. Es una asociación que también está mediada por una serie de variables relevantes de contexto, como la composición social y el origen del alumnado, el carácter público o privado de los centros, sus resultados en las evaluaciones de las competencias básicas o las percepciones y opiniones que los docentes y equipos directivos tienen tanto de estas políticas como de su propia práctica profesional. Por tanto, los análisis de este estudio se llevan a cabo considerando la diversidad de los centros en titularidad, los niveles

de rendimiento, la composición y vulnerabilidad social del alumnado y las etapas educativas.

Para profundizar en la importancia de las variables de tipo contextual, en los análisis también consideramos la posición de las escuelas en su mercado educativo local (MEL) que, según la definimos en este estudio, viene determinada por tres factores: la reputación de los centros en su entorno más inmediato, su rendimiento académico medio en las pruebas de CB y el nivel de demanda que reciben. Se trata de una variable que, a pesar de que en investigación educativa se utiliza menos que otras variables de contexto educativo — como la composición social del alumnado o el estilo de liderazgo de la escuela —, nos ayuda a entender por qué los centros escolares experimentan de forma diferente y responden con distintas actuaciones ante el marco regulativo y la presión competitiva (Maroy y van Zanten, 2009).

En relación con esto último, creemos que la adhesión del personal docente a este tipo de políticas — cómo las viven y las interpretan — es muy variable y tiene implicaciones directas en la implementación de la política educativa (Malen, 2006). Una de las principales premisas teóricas de este estudio es que los procesos de significación de las políticas por parte del profesorado — lo que en la literatura internacional se conoce como *'enactment'* — es clave para entender cómo éstas se traducen en diferentes decisiones y prácticas en los centros educativos (Ball *et al.*, 2011; Spillane *et al.*, 2002). En esta línea, el concepto de *'apropiación'* que desarrollan Levinson *et al.* (2009) es particularmente útil para entender cómo una política es interpretada de forma activa — y a menudo creativa — por una determinada comunidad de práctica, responsable de su implementación en función de las condiciones institucionales y las normas culturales y sociales que la definen.

La adopción de políticas de autonomía, evaluación y RdC en el sistema educativo catalán

La *Llei d'Educació de Catalunya* (LEC) de 2009 marcó un hito incorporando, por un lado, disposiciones de la Ley Orgánica de Educación española

(LOE) (2006) y trazando, por otro, las bases de un modelo educativo propio en la comunidad autónoma (Farré, 2009). Inspirada en los preceptos de la Nueva Gestión Pública, favoreció la descentralización, la planificación estratégica y la gestión por resultados, lineamientos plasmados en 2010 en los decretos que reglamentarían la Autonomía de Centro, la Dirección Escolar y la Evaluación Educativa. En materia de autonomía, las escuelas ampliaron sus competencias de gestión económica y organizativa, asumieron la definición de estructuras y proyectos y profesionalizaron la figura del director, tanto mediante formación y procesos de selección específicos como a través de una mayor capacidad para gestionar las plantillas docentes.

A diferencia de los decretos de autonomía y dirección, el decreto de evaluación tuvo un recorrido limitado y, debido a la crisis financiera, en 2011 se suspendió la creación de una agencia de evaluación profesionalizada e independiente. En el momento de realizar este estudio, el Consejo Superior de Evaluación del Sistema Educativo (CSASE) — dependiente del Departamento de Educación y con recursos económicos y humanos restringidos — era el organismo encargado de administrar las pruebas de Competencias Básicas (CB) y de coordinar la participación de Cataluña en evaluaciones internacionales. Según la Resolución EDU/3.614/2022, las pruebas de CB constituyen “una de las líneas de mejora de la educación en Cataluña”, ya que proporcionan información global sobre el sistema y permiten que cada centro valore sus propios resultados y planifique acciones de mejora. Aunque la LEC no permite la difusión de los resultados con fines promocionales o comparativos, los datos generados alimentan otros instrumentos de planificación escolar y de supervisión por parte de la inspección educativa.

Por su parte, la profesionalización de las direcciones de centro ha liberado a la Inspección Educativa de tareas de gestión para centrarse en la evaluación y el acompañamiento. Hoy, la normativa vigente refuerza la autonomía de gestión de los equipos directivos y la responsabilidad por los resultados de aprendizaje ante familias, inspección y administración. Se espera que la devolución de datos genere debate en los equipos docentes y directivos e impulse planes de mejora, ya sea de forma autónoma o con la ayuda de agentes como la Inspección, los llamados Centros de Recursos Pedagógicos o la llamada Red de Competencias Básicas.

Este modelo de gobernanza escolar sitúa la autonomía escolar en el núcleo del sistema — y como condición necesaria para la articulación de proyectos educativos innovadores y contextualizados. Las evaluaciones externas, por su parte, constituyen un componente clave para diagnosticar el rendimiento del alumnado y orientar la gestión escolar. También se contemplan como uno de los instrumentos para la promoción docente por méritos, aunque esta vía ha tenido hasta ahora un desarrollo limitado en comparación con otras modalidades de promoción de carácter más burocrático.

En este estudio, analizamos estos supuestos, así como la percepción y el grado de penetración de estas políticas en los centros, y su alineación con los objetivos normativos que las sustentan.

Métodos y datos

Nuestro estudio se basa en un enfoque de métodos mixtos que combina un cuestionario administrado a una muestra representativa de 78 centros de enseñanza obligatoria de Cataluña (155 directoras/es y 852 docentes) con entrevistas semiestructuradas a directores/as y profesorado de 20 centros seleccionados por heterogeneidad (titularidad, complejidad y composición socioeconómica). De forma complementaria, se emplearon fuentes secundarias del Departamento de Educación de la Generalitat de Catalunya para caracterizar variables contextuales (categoría de complejidad, porcentaje de alumnado extranjero, tamaño del claustro y número de estudiantes), así como entrevistas a informadores clave del sistema educativo (representantes del Departamento, del CSASE y de la Inspección Educativa) para triangular resultados.

En la fase cuantitativa, se utilizó un muestreo estratificado (por área urbana, titularidad y nivel educativo), combinado con un muestreo aleatorio proporcional al tamaño (PPS) y la selección aleatoria de 10 docentes por centro, considerando criterios de carga lectiva y asignatura. Se realizaron análisis descriptivos (distribuciones, correlaciones, comparaciones de medias) y modelos de regresión lineal y logística. Las variables de interés incluyeron índices de autonomía educativa y de gestión, presión hacia el rendimiento, cultura de pruebas y posición en el mercado educativo local,

mientras que los controles abarcaron características de escuela (titularidad, composición socioeconómica, rendimiento medio, reputación y tamaño) y de profesor (edad, género, experiencia, contrato, afiliación sindical, uso de datos), con efectos fijos de provincia. La imputación de las variables con valores perdidos se hizo con *missing dummies* (Cohen; Cohen, 1983).

La fase cualitativa, basada en estudios de caso, incluyó entrevistas semiestructuradas de 45–60 minutos de duración a 20 directoras/es y 48 profesoras/es, complementadas con análisis documental y notas de campo. El guion abordó tanto aspectos biográficos como el contexto escolar, las percepciones sobre evaluación y rendición de cuentas, las respuestas organizativas y pedagógicas, la autonomía en el proyecto educativo y las estrategias frente a la competición escolar. Todas las entrevistas se transcribieron íntegramente y se codificaron mediante una estrategia flexible que combinó códigos deductivos y emergentes, agrupándolos temáticamente y analizando sus coocurrencias para identificar patrones. El análisis documental de proyectos institucionales, planes de mejora y webs de los centros reforzó la interpretación de los hallazgos.

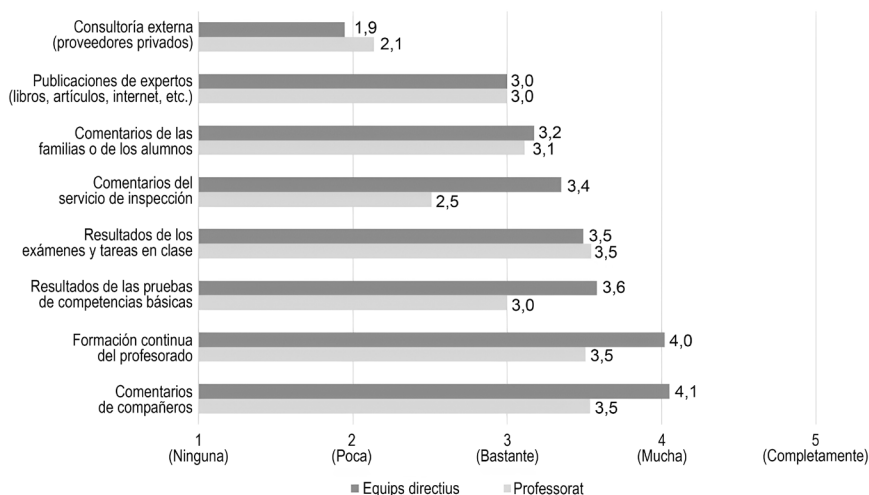
Resultados

Organizamos los resultados del estudio en torno a los siguientes ejes: en primer lugar, analizamos las fuentes de información y evaluación utilizadas por el profesorado y los equipos directivos, así como el uso y la importancia que les otorgan, con especial énfasis en las pruebas externas. A continuación, examinamos la presión que el sistema de evaluación y rendición de cuentas ejerce sobre los centros educativos catalanes, y cómo ésta influye en sus decisiones organizativas y prácticas pedagógicas. Aunque el diseño de la política catalana responde a un modelo formalmente *low-stakes*, los resultados muestran que su implementación no siempre se traduce en experiencias de baja presión percibida, ni en usos de los datos predominantemente formativos. Los resultados permiten observar cómo los centros utilizan las evaluaciones externas junto con otras fuentes de información, qué factores influyen en su uso y qué prácticas emergen en contextos con mayor nivel de presión o con mayores limitaciones para trabajar con los datos.

Fuentes de datos para la mejora

Los datos de las pruebas externas constituyen solo una de las varias fuentes a las que recurren los actores escolares para detectar áreas de mejora. Los docentes y directivos valoran especialmente los comentarios entre compañeros, la formación continua y el análisis de los resultados de exámenes y tareas en el aula, ya que les aportan información práctica y orientadora (gráfico 1). Las direcciones conceden además una alta utilidad a los resultados de las pruebas de competencias básicas y a los informes de la inspección; en cambio, el profesorado tiende a valorar menos estas fuentes externas, salvo cuando se traducen en discusiones y reflexiones concretas sobre los resultados de su propia práctica evaluativa.

Gráfico 1 – Utilidad de las fuentes de información para mejorar la calidad de la enseñanza del centro, equipos directivos y docentes



Fuente: Proyecto Reformed (Cataluña).

El hecho de que los equipos directivos otorguen mayor relevancia a los insumos de la inspección educativa responde a su canal de interlocución casi exclusiva con este servicio. Sin embargo, este vínculo varía notablemente según el inspector/a asignado al centro: en unos centros se percibe un estilo distante y centrado en la supervisión de documentos e indicadores, mientras que en otros predomina un acompañamiento cercano, orientado

a asesorar en los procesos de mejora pedagógica. Esta diversidad de enfoques modula la utilidad real que cada escuela extrae de los informes y recomendaciones de la inspección. Así lo ilustra una directora:

Lo cierto es que yo, por ejemplo, he visto tres inspecciones diferentes en esta escuela en cinco años: la primera era una inspección muy cercana, que valoraba muchísimo todo el trabajo que hacíamos, nos ayudaba a recabar recursos, incluso nos acompañaba en el aula y nos observaba. O sea, una inspección espectacular. Luego tuvimos otra inspección que estuvo un año, que hubo como una transición. Ahora tenemos una inspección que ha aparecido en dos ocasiones durante todo el curso. Una de ellas era por los indicadores del proyecto de dirección, que deben evaluarse y no sé qué. Mucho desde la burocracia, la verdad. Y es un acompañamiento muy técnico, con una mirada menos comprensiva, podríamos decir, de las dificultades del día a día (Directora 20, escuela pública 20).

Sin embargo, también hay quien considera que, de forma progresiva, se está transitando de un modelo de supervisión más distante a un modelo de acompañamiento más cercano, tal y como se ilustra en la siguiente entrevista:

La imagen histórica que se tiene de la Inspección de que vienen aquí a joderte, esto ha cambiado, ellos también tienen otro rol, ahora el inspector viene a aquí a ayudar, a informarse, viene a trasladar información a gente que está por encima de él, no es lo mismo. Debo decir también que cuando nos hicieron la auditoría pedagógica, ellos fueron también muy cercanos, muy humanos, no teníamos la sensación de estar haciendo [una rendición] de cuentas, sino que era más como un acompañamiento, un asesoramiento, [...] debemos decir que estamos contentos (Directora 7, escuela pública 7).

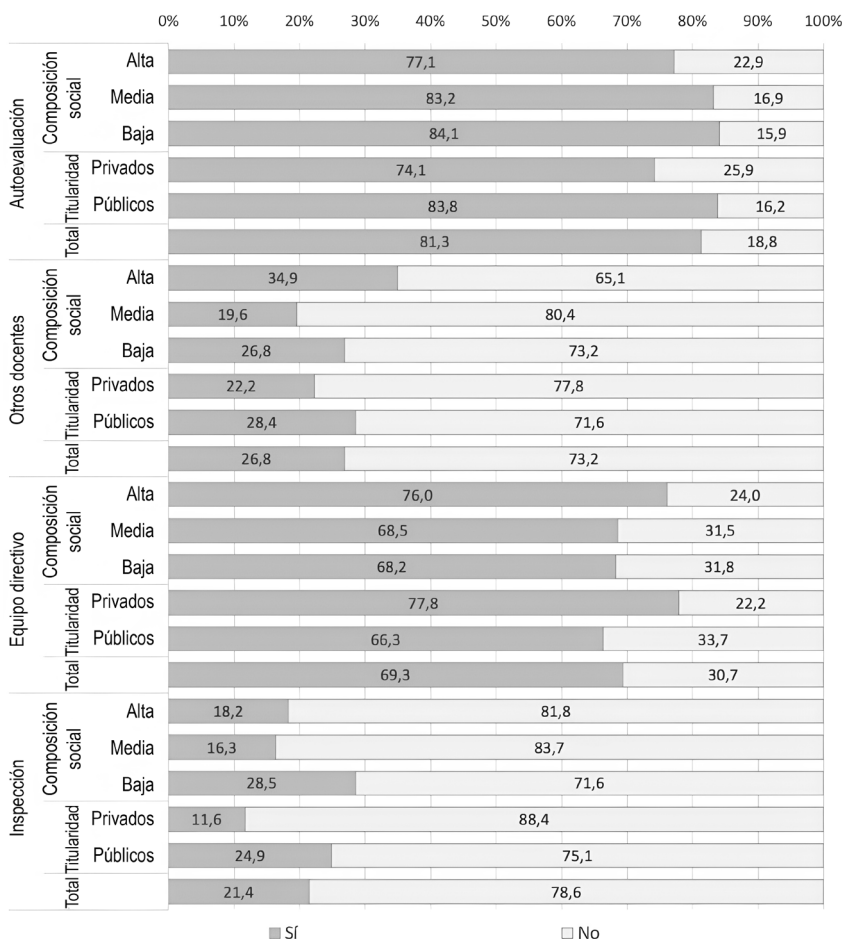
Por lo general, la inspección educativa no actúa de forma intervencionista en lo pedagógico: rara vez entra en el aula u observa directamente la implementación del proyecto educativo. Su relación con el profesorado suele ser indirecta, a través del equipo directivo, y se centra en revisar la documentación de planificación y organización del centro, así como en analizar indicadores de resultados y mejora generados internamente y por la administración. Esta dinámica quedó registrada en las entrevistas realizadas a directoras y maestros:

P: ¿Qué rol tiene la inspección en este caso? R: Ella cada cierto tiempo se reúne con el equipo directivo y debe supervisar el trabajo que se está realizando y, a final de curso, el equipo directivo debe exponer, con el departamento, inspección, etc., los resultados alcanzados. Hay unos indicadores muy claros con los niños, hay que recabar información, pruebas, para ser un poco más objetivo, a veces cuestan ponderar estas cosas [...] Yo creo que la inspección delega más en el equipo directivo, no sé si quizás porque asumen muchas escuelas, no sé, es el equipo directivo quien hace esta función (Docente 39).

Junto a las fuentes de información para la mejora educativa, también exploramos qué agentes participan en la evaluación de la labor docente (gráfico 2). La modalidad más habitual es la autoevaluación, utilizada por el 81% del profesorado, que resulta especialmente frecuente en la escuela pública, mientras que en los centros concertados (de titularidad privada y con financiamiento público) y aquellos de menor complejidad su práctica disminuye. En segundo lugar, los equipos directivos desempeñan un rol relevante para casi el 70% de los docentes, con un peso mayor en la escuela concertada (78%) que en la pública (66%). Por el contrario, la evaluación entre iguales y la realizada por la inspección son poco habituales: solo el 27% identifica a compañeros como evaluadores y el 21% recurre a la inspección, pese a su función oficial de supervisión. Aunque la inspección evalúa algo más en la pública (25%) que en la concertada (11%), su presencia es limitada y tiende a concentrarse en los centros de menor composición social.

En esta primera sección observamos que la autonomía en el uso de información no conduce necesariamente a usos profundos o formativos de los datos. La evaluación externa constituye una referencia más dentro de un ecosistema de fuentes, no siempre la más influyente, lo que anticipa las tensiones entre autonomía, valoración de las pruebas y uso real para la mejora que emergen en los apartados posteriores.

Gráfico 2 – Fuentes de evaluación docente: ¿quién evalúa tu trabajo?



Nota: los porcentajes se han calculado considerando el profesorado que identificaba al menos una fuente de evaluación

Fuente: Proyecto Reformed (Cataluña).

Uso de los datos de las evaluaciones externas

Los centros pueden utilizar los datos de las pruebas estandarizadas para múltiples fines: evaluar el trabajo docente, detectar estudiantes con necesidades educativas, ajustar el currículo, informar a familias y profesorado, agrupar alumnos por nivel, compararse con otros centros o reforzar su reputación. Su empleo más habitual consiste, sin embargo, en definir y monitorizar el plan de mejora escolar. De hecho, los resultados de las pruebas de competencias básicas suelen servir de referencia en planificaciones anuales, memorias, proyectos de dirección y acuerdos de corresponsabilidad (gráfico 3).

Este uso genérico de los datos como herramienta de ‘mejora escolar’ explica que se declare con más frecuencia que otros usos más específicos, e incluso controvertidos. Las entrevistas muestran que, en la práctica, los datos de las pruebas suelen generar debate y reflexividad en el claustro docente y actúan como motor de los procesos de mejora. Esto se expresa en el siguiente testimonio de una directora:

P: ¿Las pruebas de competencias han contribuido a mejorar la calidad educativa del centro y el rendimiento de los estudiantes? R: A ver, te hace reflexionar para realizar mejoras. El hecho de que como equipo directivo te propongas llegar a la media o intentar superarla, evidentemente hace poner en marcha un engranaje de reflexión, que puedes estar más de acuerdo o menos de acuerdo con las pruebas en sí, pero sí es verdad que dices: “como esto es lo que tenemos” debemos intentar sacarle el máximo provecho. Dices: “tienes esto, diseña unos objetivos hacia la mejora” y al final es a partir del resultado de competencias básicas (Directora 1, escuela pública 1).

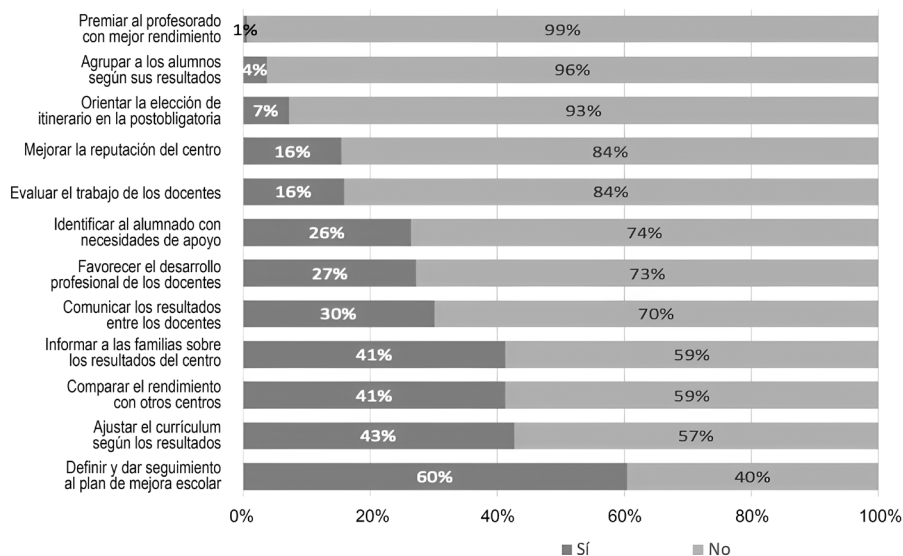
El resto de usos de los datos tiene un papel más secundario en el contexto educativo catalán. Poco más del 40% del profesorado recurre a ellos para ajustar el currículo, compararse con otras escuelas o informar a las

familias sobre el nivel de consecución de aprendizajes (gráfico 3). Entre el 25% y el 30% los emplea para identificar alumnado con necesidades de apoyo o impulsar el desarrollo profesional.

Los usos más instrumentales, como agrupar a los estudiantes según su rendimiento, premiar a los de mejor rendimiento o realzar la reputación del centro, son muy poco frecuentes en Cataluña, como también lo es el uso de estos datos para detectar a quienes requieren refuerzo. Sin embargo, el bajo aprovechamiento formativo puede estar relacionado con los plazos de administración de las pruebas de competencias básicas, que se aplican al final de cada etapa (primaria y secundaria) y cuyos resultados se retornan a los centros educativos cuando el alumnado evaluado ya se ha graduado o está a punto de hacerlo. Dos directoras describen esta des-coordinación entre el calendario administrativo y el escolar:

Siempre se hace [comentar los resultados de las pruebas en el claustro], lo que ocurre es que llegan muy tarde. Llegaron, por ejemplo, la semana pasada y sí que hicimos una reunión y entonces [fue] la única reflexión que nos dio tiempo de hacer... pero después sí que cuando encaramos el trabajo con los grupos entonces hacemos una más específica. Pero al final es el tema de reflexionar sobre estas formas de trabajar en el aula, qué resultados nos han ayudado a obtener, hacemos esta reflexión (Docente 2, escuela pública 1)

Gráfico 3 – Uso de los datos de las pruebas de competencias básicas (docentes)



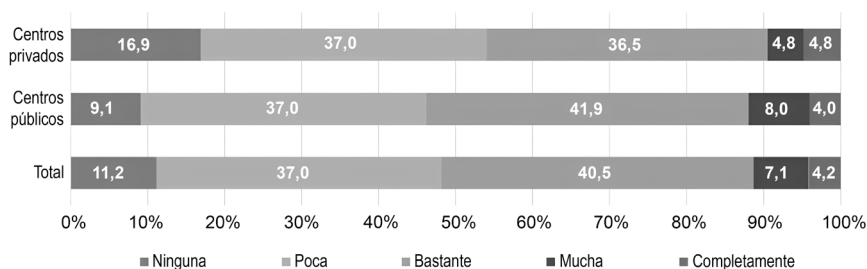
Fuente: Proyecto Reformed (Cataluña).

El limitado uso de los resultados de las pruebas de competencias básicas también obedece a que dichos datos no siempre se consideran indicadores fiables de calidad educativa (cuestión que abordamos más adelante) y a las dificultades de los centros para convertirlos en propuestas concretas. De hecho, poco más de la mitad del profesorado catalán declara tener al menos cierto grado de dificultad para transformar estos datos en medidas específicas de mejora (gráfico 4). Esta percepción varía según la titularidad: mientras el 54% de los docentes de centros concertados señalan poca o ninguna dificultad, en la pública este porcentaje desciende al 48%.

Asimismo, la complejidad social del centro influye en estas dificultades (gráfico 5): en los centros de composición alta, seis de cada diez docentes encuentran poca o ninguna barrera para traducir los datos en acciones; en los de composición media y baja, este porcentaje cae al 47% y al 40%, respectivamente. Esto puede deberse tanto a la mayor presión por mejorar resultados en entornos desfavorecidos, que expone

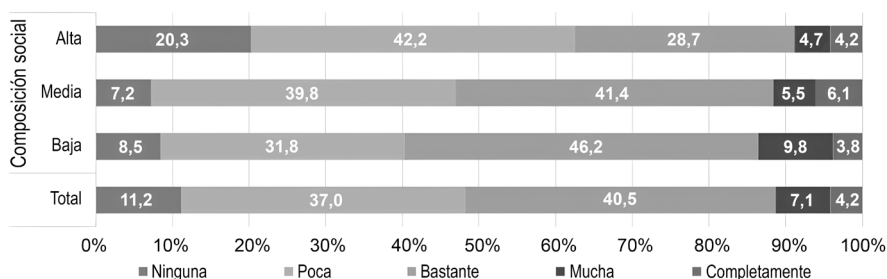
las limitaciones a la hora de incorporar los datos en planes de mejora, como a diferencias en prioridades y disponibilidad de tiempo para trabajar aspectos concretos de la enseñanza.

Gráfico 4 – Dificultad para transformar los datos provenientes de las pruebas de CB en medidas y acciones para mejorar la enseñanza, por titularidad (docentes)



Fuente: Proyecto Reformed (Cataluña).

Gráfico 5 – Dificultad para transformar los datos provenientes de las pruebas de CB en medidas y acciones para mejorar la enseñanza, por niveles de composición social (docentes)



Fuente: Proyecto Reformed (Cataluña).

El 40% del profesorado señala la falta de tiempo como el principal obstáculo para traducir los datos en acciones, cifra que desciende al 33% en centros de composición social aventajada y asciende al 42% en los de composición social media y desaventajada (gráfico 6). Casi la mitad

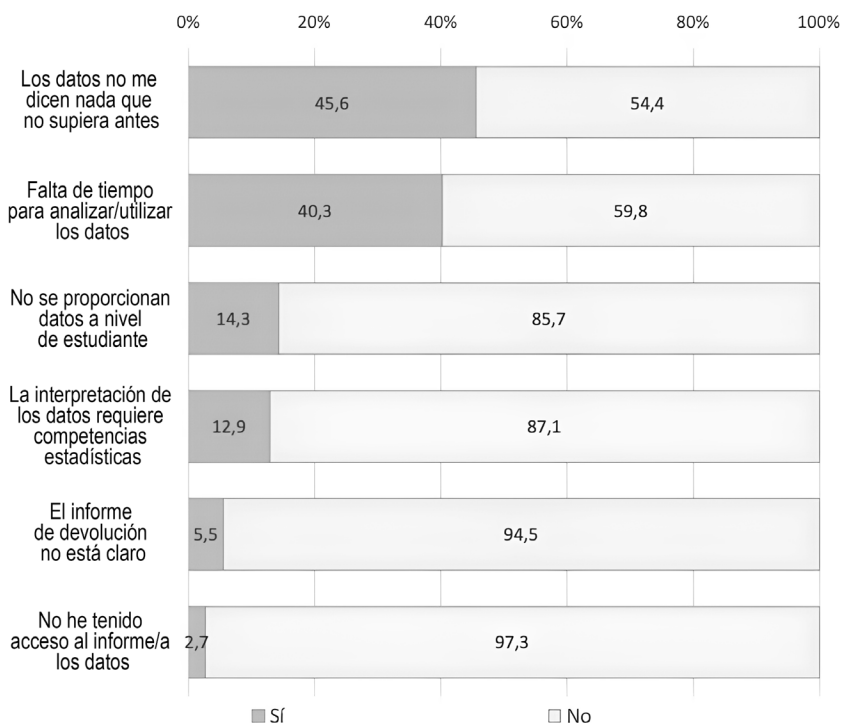
de los docentes añade que los resultados de las pruebas externas no les aportan información nueva. Esta percepción se refleja en las siguientes declaraciones de dos maestras, una de una escuela con población más desaventajada y una de una concertada:

P: ¿Le ayudan [los resultados de las pruebas] a identificar carencias de la escuela o algunos puntos donde se deba mejorar algo? R: Es que las carencias de esta escuela ya las sabemos por el tipo de escuela que somos y las intentamos mejorar cada día... (Docente 27, escuela pública 9).

Los resultados de las competencias es, más o menos, lo que sabíamos y la información que nos da no es, 'ostras, qué sorpresa'. Puede pasar, puede variar, pero no nos da ningún dato nuevo, la verdad es ésta, no nos sorprende. Es verdad que lo que hacemos es buscar unos indicadores de éxito para superar aquellos vacíos que vemos de estas pruebas (Docente 36, escuela concertada 13).

El 14 % del profesorado señala la falta de datos a nivel de estudiante y el 13 % apunta que la interpretación exige competencias estadísticas. Solo un 6 % encuentra poco claro el informe de devolución y un 3 % cita problemas de accesibilidad, lo que sugiere que la documentación del CSASE es, en general, clara.

Gráfico 6 – Factores que explican las dificultades percibidas por transformar los datos en medidas y acciones concretas (docentes)



Fuente: Proyecto Reformed (Cataluña).

Factores asociados con el uso de los datos

Los datos de las evaluaciones externas pueden emplearse con fines formativos y expresivos. Por ejemplo, para evaluar el trabajo docente, detectar estudiantes con necesidades de apoyo, planificar el desarrollo profesional o ajustar el currículo. También pueden utilizarse con propósitos sumativos e instrumentales, como informar a las familias, separar al alumnado por nivel, comparar rendimientos con otras escuelas o mejorar la reputación del centro. En este apartado, analizamos qué factores organizativos (estructura, recursos, liderazgo) e individuales (perfil y competen-

cias del profesorado) están asociados a estos distintos usos de los datos dentro de los centros educativos.

Uso de los datos de carácter instrumental

Nuestro análisis revela que la presión percibida para lograr buenos resultados en las pruebas de competencias básicas se asocia con un uso de los datos de corte competitivo: comparar rendimientos con otros centros y reforzar la reputación del centro (tabla 1). Esta presión también vincula los datos con la evaluación del trabajo docente y, en menor medida, con la agrupación de alumnado por niveles, lo que sugiere la existencia de una cultura escolar que instrumentaliza la información. Este patrón es especialmente marcado en la red concertada: tras ajustar por presión, cultura de evaluación y composición escolar, el profesorado de centros concertados duplica las probabilidades de emplear los datos para compararse con otras escuelas o mejorar su reputación, y los utiliza en mayor medida para informar a las familias.

En cambio, en la pública es más frecuente el uso de datos, de cualquier origen, no solo de CB, para agrupar a los estudiantes con miras a optimizar sus resultados. Además, la confianza en la validez de las pruebas se correlaciona de forma notable con estos usos instrumentales y competitivos. Por nivel educativo, el profesorado de secundaria recurre con más frecuencia a los datos para realzar la imagen del centro y evaluar la acción docente, mientras que en primaria se emplean con mayor intensidad para fundamentar decisiones de formación y desarrollo profesional.

Uso de los datos de carácter formativo

Nos centramos ahora en los factores que promueven o dificultan un uso formativo de los datos orientado a la mejora pedagógica. El principal factor identificado que obstaculiza su uso son las dificultades percibidas por el profesorado para transformar los datos en medidas y acciones concretas de mejora. En este sentido, el profesorado que tiene dificultades para traducir los datos en acciones de mejora, tiene casi la mitad de probabilidad de utilizar los datos para ajustar el currículum o tomar

decisiones sobre el desarrollo profesional. Además, cabe destacar que la dificultad percibida no está asociada a ninguna otra finalidad del uso de los datos. Sin embargo, como vemos en el gráfico 6, las dificultades no son de naturaleza técnica, sino que se deben tanto a la falta de utilidad percibida de los datos para iniciar procesos de mejora, como a las dificultades asociadas a la falta de tiempo y la carga de trabajo.

Por otra parte, el uso de los datos con fines formativos también está asociado con dinámicas de colaboración del profesorado en materia didáctica y pedagógica. Así, observamos que la frecuencia de estas dinámicas de colaboración está relacionada con la probabilidad de hacer un uso de los datos tanto para tomar decisiones en materia de formación como para informar a los demás profesores y profesoras sobre los resultados de las pruebas.

Tabla 1 – Presión por obtener buenos resultados y uso de los datos de las pruebas de competencia (Continúa)

	(1)	(2)	(3)	(4)
	Uso de los datos para evaluar el trabajo docente	Uso de los datos para identificar a estudiantes con necesidad de apoyo	Uso de los datos para tomar decisiones de desarrollo profesional	Uso de los datos para ajustar el currículum
Presión para obtener buenos resultados	1.46** (0.24)	0.98 (0.12)	1.02 (0.14)	1.20* (0.13)
Cultura favorable a las pruebas	1.09 (0.17)	1.15* (0.09)	0.98 (0.13)	0.94 (0.10)
Dificultad para transformar los datos	0.85 (0.25)	0.65 (0.22)	0.60* (0.18)	0.58** (0.16)
Edad	0.97*** (0.01)	0.99 (0.01)	0.99 (0.01)	1.03*** (0.01)
Género (ser mujer vs. Hombre)	0.70 (0.19)	0.64** (0.12)	0.74* (0.12)	1.29 (0.31)

Tabla 1 – Presión por obtener buenos resultados y uso de los datos de las pruebas de competencia (continúa)

Nivel educativo (docentes secundaria vs. Primaria)	1.79** (0.51)	0.92 (0.29)	0.47*** (0.12)	1.04 (0.20)
Titularidad (escuela pública vs. Privada)	0.97 (0.22)	1.22 (0.28)	0.89 (0.17)	1.11 (0.24)
Competencia performativa entre profesores	1.28*** (0.12)	1.06 (0.08)	1.07 (0.11)	0.98 (0.08)
Colaboración con los compañeros	0.93 (0.14)	0.83 (0.14)	1.51* (0.35)	0.74 (0.15)
Observaciones	652	652	655	655
Pseudo R 2	0.051	0.022	0.037	0.039

	(5)	(6)	(7)	(8)	(9)
	Uso de los datos para informar a las familias	Uso de los datos para agrupar a los estudiantes	Uso de los datos para comparar el rendimiento con otras escuelas	Uso de los datos para favorecer la reputación de la escuela	Uso de los datos para informar a los profesores
Presión para obtener buenos resultados	0.87 (0.09)	1.71* (0.54)	1.33** (0.19)	1.58** (0.33)	1.21 (0.16)
Cultura favorable a las pruebas	1.21* (0.14)	1.42* (0.27)	1.13 (0.10)	1.41*** (0.17)	1.09 (0.10)
Dificultad para transformar los datos	0.91 (0.21)	1.15 (0.64)	1.41 (0.34)	1.25 (0.35)	1.25 (0.35)
Edad	1.02*** (0.01)	1.04* (0.02)	0.99 (0.01)	0.97*** (0.01)	0.99* (0.01)
Género (ser mujer vs. Hombre)	1.26 (0.18)	1.25 (0.59)	0.99 (0.17)	0.81 (0.27)	0.88 (0.16)
Nivel educativo (docentes secundaria vs. Primaria)	0.81 (0.19)	1.55 (1.19)	0.89 (0.23)	2.54*** (0.56)	0.85 (0.22)

Tabla 1 – Presión por obtener buenos resultados y uso de los datos de las pruebas de competencia (conclusión)

Titularidad (escuela pública vs. Privada)	0.42*** (0.10)	3.60** (2.27)	0.58* (0.18)	0.53* (0.19)	0.66 (0.20)
Competencia performativa entre profesores	1.09 (0.09)	1.16 (0.15)	1.19** (0.09)	1.16 (0.13)	0.95 (0.08)
Colaboración con los compañeros	1.09 (0.25)	0.90 (0.35)	0.87 (0.16)	0.94 (0.24)	1.33* (0.23)
Observaciones	655	627	655	655	655
Pseudo R 2	0.057	0.082	0.051	0.106	0.029

Nota: Los valores corresponden a los *odds ratio* de uso de los datos de las pruebas. Cada columna refleja un modelo de regresión logística con efectos fijos, en el que las distintas variables de uso de los datos son función de una batería de variables individuales y del centro educativo. Además de las variables mostradas, se incluyeron como controles: proporción de estudiantes extranjeros (composición social), proporción de estudiantes repetidores (desventaja educativa), carga de trabajo (horas), efecto fijo de la provincia de Barcelona y *missing dummies* para variables con valores perdidos (Cohen; Cohen, 1983; OCDE, 2007). Los coeficientes están exponentiados (*odds-ratio*); los errores estándar aparecen entre paréntesis.

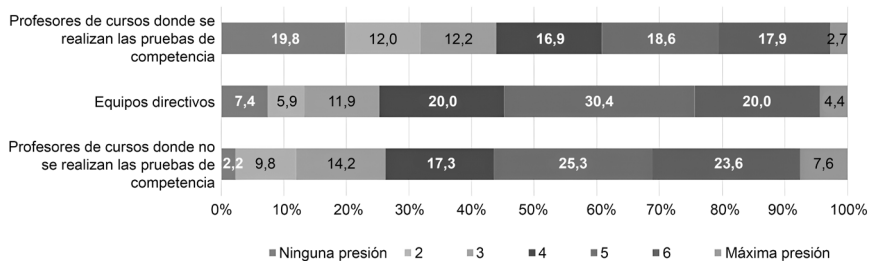
Presión por rendir en las pruebas de competencias básicas

Una vez descrito el uso limitado de los datos de evaluación externa, así como los usos formativos e instrumentales identificados y los factores que restringen su aprovechamiento pedagógico, este apartado aborda el segundo eje de resultados: la presión que experimentan los centros en un sistema formalmente *low-stakes*. Como se ha presentado, en el sistema de evaluación escolar catalán, las pruebas de competencias básicas no llevan aparejadas sanciones económicas ni la publicación de resultados escolares, y su influencia en los procesos de promoción docente es mínima. Se utilizan principalmente como un indicador de referencia en los planes de mejora escolar, y sirven como uno de los elementos para valorar la labor del equipo directivo, aunque siempre en el marco de un conjunto de indicadores mucho más amplio. Por ello, se considera que es un sistema *low-stakes*.

Sin embargo, que el sistema sea formalmente *low-stakes* no impide que docentes y directivos experimenten presiones derivadas de factores más sutiles: la reputación del centro ante las familias, la valoración interna del cuerpo de profesores/as, y los procesos de mejora escolar y evaluación directiva, en el marco de los cuales la inspección y la administración educativa hacen cada vez más hincapié en los resultados de CB para clasificar y exigir planes de mejora a los centros. Por lo tanto, aun sin consecuencias formales, la percepción de que 'la imagen del centro' o 'el prestigio profesional' depende de esos datos genera un ambiente de exigencia. El progresivo descenso en los resultados de las pruebas de CB, junto con el bajo rendimiento del sistema educativo catalán en evaluaciones internacionales, especialmente en comparación con la media europea y con países considerados de referencia, ha contribuido a reforzar esta tendencia (Quilabert *et al.*, 2024).

Los datos recogidos revelan que esta presión percibida es considerable. Según se presenta en el gráfico 7, aproximadamente el 40% del profesorado cuyos alumnos afrontan las pruebas sitúa su nivel de presión en las categorías más altas (5–7 en una escala de 1 a 7), mientras que únicamente un 20 % declara no percibir ninguna presión. Entre los equipos directivos, la tensión se acentúa: más de la mitad señala niveles elevados de presión, y sólo alrededor del 7% afirma no sentirse presionado. Este desajuste entre el carácter *low-stakes* formal y la presión real sugiere que los mecanismos informales de responsabilidad (reputación, presión de pares, creciente interés en los medios por los resultados de las CB, etc.) pueden ejercer un efecto performativo similar al de las sanciones directas, lo cual acaba repercutiendo en la cultura escolar y en las prácticas educativas que se llevan a cabo en los centros.

Gráfico 7 – Nivel de presión para obtener buenos resultados en las pruebas (profesorado y equipos directivos)

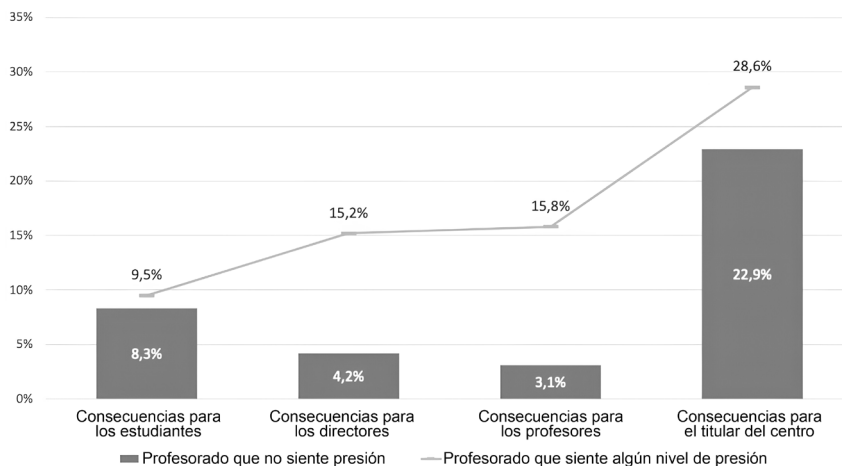


Fuente: Proyecto Reformed (Cataluña).

En el gráfico 7 también observamos que, de forma contraintuitiva, la presión por rendir no se concentra necesariamente en quienes trabajan en cursos evaluados mediante las pruebas de competencias básicas, sino que resulta más intensa entre quienes no participan directamente en ellas. Mientras que un 40% de los docentes de cursos evaluados reporta niveles altos o muy altos de presión, esta proporción asciende a casi el 60% entre quienes se desempeñan en cursos no evaluados. Igualmente llamativa es la diferencia en la ausencia total de presión: prácticamente ningún maestro de un curso no evaluado declara no sentir exigencia alguna, frente a dos de cada diez profesores de cursos evaluados que sí perciben un entorno libre de exigencias.

Este último grupo atribuye su sensación de ‘cero presión’ a la convicción de que no existen riesgos, ni reputacionales ni profesionales, asociados a los resultados: solo un 3 % cree que puedan derivarse consecuencias negativas para el profesorado, frente al 16% de quienes admiten al menos una presión mínima (Gráfico 8). Estos datos sugieren que, en contextos *low-stakes*, la presión puede desplazarse hacia docentes de cursos no evaluados, posiblemente por dinámicas internas de centro, expectativas implícitas, responsabilidades indirectas en la preparación del alumnado evaluado o simplemente falta de información sobre las consecuencias de las pruebas. No obstante, este patrón contraintuitivo plantea interrogantes que requieren una exploración más profunda para comprender sus causas y mecanismos.

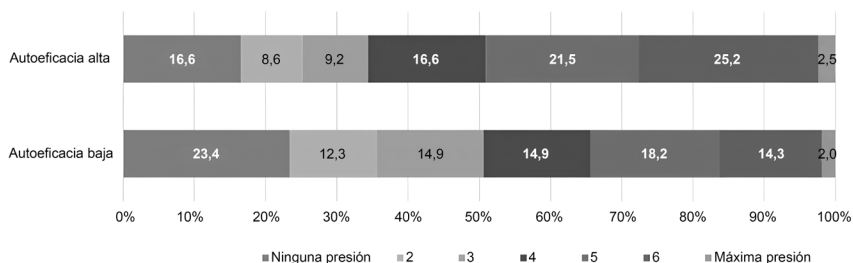
Gráfico 8 – Profesores que consideran que las pruebas de competencias tienen algún tipo de consecuencia (económicas, relativas al trabajo, reputacional, etc.), según nivel de presión percibida



Fuente: Proyecto Reformed (Cataluña).

La presión percibida también guarda relación con la autoeficacia del profesorado, esto es, la confianza en su capacidad para gestionar eficazmente el aula, mejorar el contenido de sus lecciones y mantener relaciones comunicativas positivas con el alumnado. Quienes creen firmemente que pueden impulsar cambios en su centro son, además, quienes más sienten la presión por los resultados, probablemente porque consideran que estos dependen en buena medida de su propio desempeño y no solo de factores externos. Así, cerca del 50% de los docentes con alto nivel de autoeficacia reporta sentir niveles elevados de presión, frente al 34% de aquellos que no se consideran capaces de gestionar la clase con eficacia (gráfico 9).

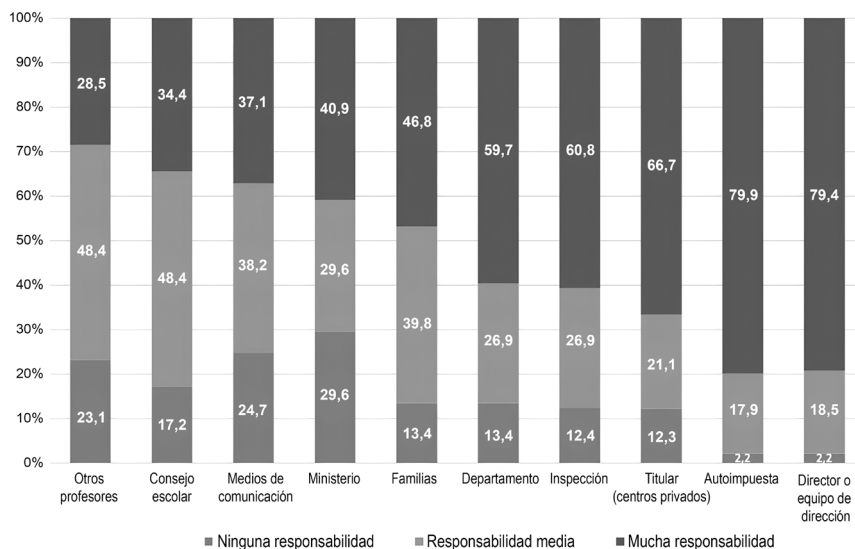
Gráfico 9 – Presión por obtener buenos resultados
y autoeficacia percibida por el profesorado



Fuente: Proyecto Reformed (Cataluña).

Una vez establecido el nivel de presión performativa, analizamos sus fuentes entre el profesorado que declara estar en las categorías más altas (5–7 del gráfico 9). Como muestra el Gráfico 10, tras la presión autoimpuesta, los principales responsables percibidos son los equipos directos y, en los centros privados, los titulares de la escuela. Así, el 79% de quienes sienten mucha presión atribuye parte de ella al equipo directivo, y el 67% a la titularidad del centro. Además, seis de cada diez docentes apuntan al Departamento de Educación y a la Inspección como fuentes significativas de esa misma presión para lograr buenos resultados en las pruebas de CB. Por lo general, los agentes a los que se les atribuye menor responsabilidad como fuentes de presión son el consejo escolar y los/as compañeros/as.

Gráfico 10 – Nivel de responsabilidad de los agentes por la presión por resultados percibida por el profesorado (docentes que perciben un alto nivel de presión)



Fuente: Proyecto Reformed (Cataluña).

Aunque el equipo directivo constituye la principal fuente de presión percibida, tanto los datos del cuestionario como las entrevistas indican que el profesorado no experimenta una presión explícita por parte de la dirección para orientar la enseñanza hacia los resultados. Más bien, esta exigencia se manifiesta como una expectativa compartida que emerge del clima de opinión existente entre el profesorado del centro y en el propio equipo directivo, pero que raramente se traduce en directrices concretas o indicaciones formales. Así lo plantean dos docentes de una escuela pública, entrevistadas conjuntamente:

R2: Yo pienso que partimos de un equipo directivo que, por ejemplo, ya sabe cómo trabajamos las dos tutoras que hay en sexto; por tanto, confía en nosotras, ni nos presionan ni nos dicen 'es muy importante'.

R1: Es que también nosotras sabemos lo que es importante, pero yo en ningún momento, ni tú tampoco, hemos sido presionadas. Ellos saben que sabemos que es importante; si suspendiéramos las pruebas, supongo que el equipo directivo diría ‘escucha, que estas pruebas son importantes, que estamos a sexto’, pero yo pienso que no es el caso.

R2: Yo pienso que hay dos personas que saben cómo somos y confían en nosotras; si no, no nos hubieran puesto a sexto, habrían puesto a otros maestros. Es lo que a mí me transmiten: que confían plenamente, saben cómo somos y cómo trabajamos, y que saben que haremos lo que podamos y que ya lo estamos haciendo, de hecho. Por tanto, este aspecto es muy importante: que un equipo directivo te transmita esto (Docentes 26 y 27, escuela pública 9).

El profesorado experimenta un alto nivel de presión autoimpuesta: el 80% de quienes sienten un alto nivel de exigencia identifica este tipo de presión. Sin embargo, la presión autoimpuesta no se da en el vacío. Podría reflejar, en gran medida, la internalización de demandas y expectativas externas. Como se presenta en la tabla 2, el 38% de la variabilidad de la presión autoimpuesta coincide con la presión percibida por parte del equipo directivo y, en los centros privados, un 20% con la presión del titular del centro, lo que sugiere que cuanto más presión ejerzan estos agentes, mayor es el grado en que los docentes “se exigen” a sí mismos.

Otras fuentes vinculadas a la presión autoimpuesta son las familias (27%), otros compañeros (23%), el consejo escolar (20%) y la inspección (17%). En conjunto, estos datos indican que la presión que los docentes sienten como propia se alimenta de una mezcla de expectativas profesionales internas y de las demandas que perciben en su entorno educativo, que va desde la dirección y la titularidad hasta el nivel de implicación de la comunidad escolar.

Tabla 2 – Relación entre la presión autoimpuesta y la presión que proviene de otras fuentes (docentes)

Otras fuentes de presión:	Presión autoimpuesta	
	Coefficiente de correlación	Variabilidad compartida
Director o equipo de dirección	0.615	38%
Familias	0.516	27%
Otros profesores	0.479	23%
Titular (centros privados)	0.452	20%
Consejo escolar	0.451	20%
Inspección	0.415	17%
Departamento	0.378	14%
Medios de comunicación	0.335	11%
Ministerio	0.298	8%

Nota: Todas las correlaciones son significativas al 1%.

Fuente: Proyecto Reformed (Cataluña).

Complementariamente, de las entrevistas se desprende que la presión autoimpuesta deriva de una combinación de autoexigencia, interés por el aprendizaje del alumnado y mantenimiento de cierta reputación profesional frente a los iguales. Una docente de una escuela de composición social media-baja lo plantea con estas palabras:

P: ¿De dónde dirías que viene esa presión que describes?

R: Bueno, de mí, ¿no? Porque yo soy exigente y quiero que los niños escriban bien y dices: 'vaya, si he estado trabajando los plurales, pues que me lo hagan bien', porque muchas veces lo has trabajado mucho y ese día no lo aplican, y entonces parece que no lo hayas trabajado, ¿sabes? [...]

P: ¿Dirías que es una sensación similar para el resto de docentes o piensas que hay quien no la siente?

R: Yo creo que sí, porque tengo conocidos que dedican tiempo a preparar las pruebas.

P: ¿Y en este centro?

R: Claro, es que el resto no, porque solo es el de 6.º... claro, quizá sea una sensación mía, ¿eh?, pero cuando salen los resultados siempre miran al maestro de sexto; pero claro, al resto no creo que les preocupe...

P: Tú vas haciendo un año 5.º y un año 6.º, ¿te sientes diferente un año respecto al otro?

R: Por eso te decía que no hay presión porque no te dicen nada, pero sabes que si salen mal, te mirarán a ti (Docente 45, escuela pública 15).

Presión de las evaluaciones externas: impacto en prácticas escolares y efectos colaterales

En este apartado examinamos cómo la presión derivada de las evaluaciones externas se relaciona con diversas prácticas pedagógicas: por un lado, con aquellas habitualmente señaladas como efectos colaterales de las políticas de rendición de cuentas (*teaching to the test* – TTT, reducción del currículo y agrupación del alumnado por niveles de desempeño), por otro, con el uso de datos con fines de mejora escolar. No obstante, mientras que las agrupaciones homogéneas de rendimiento preocupan por sus impactos negativos en equidad y cohesión social, el TTT solo se considera problemático cuando monopoliza el tiempo lectivo o se aplica de forma mecánica y sin *feedback* adecuado.

Organizamos este apartado en tres bloques. Primero, exploramos la asociación entre niveles de presión percibida y determinadas prácticas escolares. A continuación, profundizamos en cómo distintas fuentes de presión influyen en la práctica específica de *teaching to the test*. Por último, analizamos otros factores que, más allá de la presión percibida, también se vinculan con determinadas prácticas y efectos colaterales.

Fuentes de presión por los resultados y prácticas educativas: una exploración

El profesorado que siente mayor presión para obtener buenos resultados en las pruebas de CB tiende a practicar más *teaching to the test* y a alinear el currículo escolar con los contenidos de las pruebas (véase la tabla 3). Esta asociación es especialmente fuerte cuando la presión proviene de otros maestros, de las familias y, sobre todo, del equipo directivo (y, en los centros concertados, también del titular). Aunque menos intensa, la relación persiste cuando la fuente de presión es externa al centro (Departamento o Inspección) y cuando es autoimpuesta, lo que sugiere una cultura de responsabilidad compartida. Por el contrario, la vinculación entre presión percibida y uso de los datos de las pruebas de CB (por ejemplo, para diseñar planes de mejora) es más débil, y destaca únicamente un uso ligeramente mayor cuando la presión procede del equipo directivo (y del titular en la concertada).

Conviene señalar, además, que la simple conciencia de posibles consecuencias derivadas de los resultados no se asocia estrechamente con prácticas de TTT o reducción curricular, salvo cuando el profesorado percibe que el equipo directivo ejerce esa presión. Algunas entrevistas apuntan a que la incertidumbre sobre el alcance real de esas consecuencias alimenta la preparación intensiva para las pruebas en un sistema formalmente *low-stakes*, pero percibido como exigente. Esto se ilustra en la siguiente cita:

P: ¿Sientes presión por parte de las pruebas de competencias básicas?

R: Nosotras sí. Tenemos presión porque son unos números: si salen bajos, nos preguntamos qué consecuencias tendrá esto... (Docente 21, escuela pública 7).

Por último, a diferencia de lo que sucede en otros países, la presión atribuida a los medios de comunicación no muestra ninguna asociación con prácticas tácticas de optimización de resultados.

Tabla 3 – Correlación entre el grado de responsabilidad atribuida a los agentes de la presión por resultados y las prácticas escolares (docentes) (continúa)

	Ministerio	Departamento	Inspección	Titular (privados)	Director y equipo directivo	Consejo escolar	Compañeros/as	Familias	Autoimpuesta	Medios comunicación	
Uso de los resultados de las pruebas	Ajustar el currículum	0.07	0.10	0.09	0.10	0.10	0.10	0.09	0.08	0.11	0.05
	Informar a los padres sobre el rendimiento escolar	-0.03	-0.01	-0.03	0.26	0.12	0.05	0.09	0.11	0.14	0.05
	Agrupar a los estudiantes	0.06	0.07	0.06	0.04	0.10	0.08	0.14	0.11	0.07	0.08
	Comparar el rendimiento con el de otras escuelas	0.01	0.02	0.02	0.25	0.16	0.04	0.08	0.04	0.09	0.05
	Favorecer la reputación de la escuela	-0.07	-0.07	-0.04	0.19	0.21	0.08	0.11	0.09	0.08	-0.02
	Informar al profesorado	-0.01	0.00	0.01	0.14	0.08	0.00	0.05	0.03	0.06	0.01
Reducción / alineación del currículum	Las pruebas de CB llevan a una redistribución de los recursos	0.19	0.20	0.20	0.32	0.33	0.26	0.28	0.24	0.20	0.11
	Los estándares de aprendizaje influyen en el enfoque pedagógico	0.14	0.21	0.21	0.33	0.33	0.21	0.24	0.23	0.22	0.11
	Se consideran las pruebas de CB por las decisiones curriculares	0.12	0.19	0.23	0.31	0.38	0.27	0.29	0.26	0.28	0.10

Tabla 3 – Correlación entre el grado de responsabilidad atribuida a los agentes de la presión por resultados y las prácticas escolares (docentes) (conclusión)

Teaching to the test	Frecuencia de actividades para preparar a los estudiantes	0.05	0.11	0.16	0.18	0.16	0.09	0.06	0.13	0.26	0.04
	Actividades para preparar a los estudiantes durante todo el año	0.13	0.16	0.17	0.11	0.20	0.09	0.10	0.06	0.22	0.14
	Actividades para preparar a los estudiantes el mes anterior de la prueba	0.09	0.14	0.15	0.25	0.23	0.10	0.09	0.12	0.31	0.07
Conocimiento de las consecuencias de las pruebas	Consecuencias de los resultados de las pruebas para el director	0.08	0.15	0.17	0.15	0.20	0.12	0.15	0.13	0.17	0.15
	Consecuencias de los resultados de las pruebas para el profesorado	0.02	0.08	0.12	0.24	0.19	0.06	0.10	0.10	0.18	0.07
	Consecuencias de los resultados de las pruebas para el titular de la escuela	-0.01	0.07	0.13	0.11	0.15	0.06	0.06	0.10	0.16	0.07

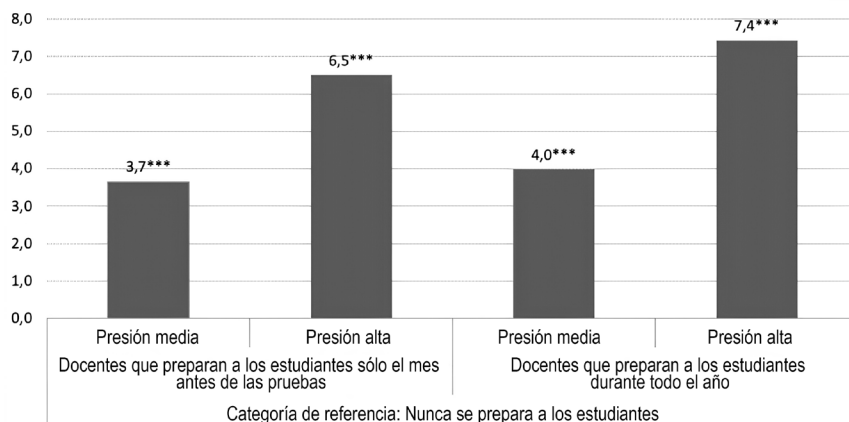
Fuente: Proyecto Reformed (Cataluña).

Preparación de los estudiantes para las pruebas según niveles de presión percibida

En este subapartado, relacionamos la presión percibida para obtener buenos resultados en las pruebas de competencias básicas con la práctica del *teaching to the test*, es decir, la preparación específica del alumnado para afrontar con éxito los tests estandarizados. Para ello, comparamos la probabilidad (riesgo) de que el profesorado que siente presión, en

comparación con el que no la percibe, haya dedicado tiempo a preparar a sus estudiantes durante el mes previo a las pruebas o a lo largo de todo el curso. Los resultados de los modelos, ilustrados en el Gráfico 11, apuntan que quienes reportan niveles medios o altos de presión performativa adoptan con mayor frecuencia prácticas de TTT. Concretamente, un docente que percibe alta presión tiene hasta 7,4 veces más probabilidad de realizar TTT durante todo el año que uno que no siente presión. Esta relación se mantiene, aunque en menor medida (odds ratio = 6,5), cuando se analiza únicamente la preparación en el mes anterior a las pruebas.

Gráfico 11 – Teaching to the test y presión para obtener buenos resultados en las pruebas de competencias básicas (odds ratio)



Modelos de regresión logística multinomial. Variable respuesta: preparar a los estudiantes para las pruebas de competencias básicas (nunca / mes antes / durante todo el año). Los valores ilustrados corresponden a los odds ratio de preparar a los estudiantes para las pruebas, según la frecuencia y la presión percibida para obtener buenos resultados. Variables de control a nivel individual: experiencia, tipo de contrato, afiliación sindical, nivel educativo, género, satisfacción con el trabajo y con la escuela, enfoque tradicional / innovador, preferencia por metodologías metacognitivas. Control a nivel de centro: efecto fijo de centro educativo. Signif. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Pseudo R^2 : 0.32.

Fuente: Proyecto Reformed (Cataluña).

Además de la presión, exploramos si la posición de la escuela en su mercado educativo local (MEL; determinado por su reputación, rendimiento

medio y nivel de demanda) también influye en la probabilidad de practicar *teaching to the test*. El gráfico 12 revela que, junto a la presión, la posición en el MEL se asocia con prácticas de TTT: los centros con mala posición que además perciben alta presión muestran una mayor predisposición a preparar a sus estudiantes que aquellos con igual posición pero menor presión. De igual modo, entre los centros bien situados en el MEL, la frecuencia de TTT aumenta o disminuye según el nivel de presión percibida, lo cual confirma que ambos factores operan de forma conjunta en la adopción de prácticas orientadas a optimizar los resultados de las pruebas.

Gráfico 12 – *Teaching to the test*, según niveles de presión y posición en el MEL (continúa)

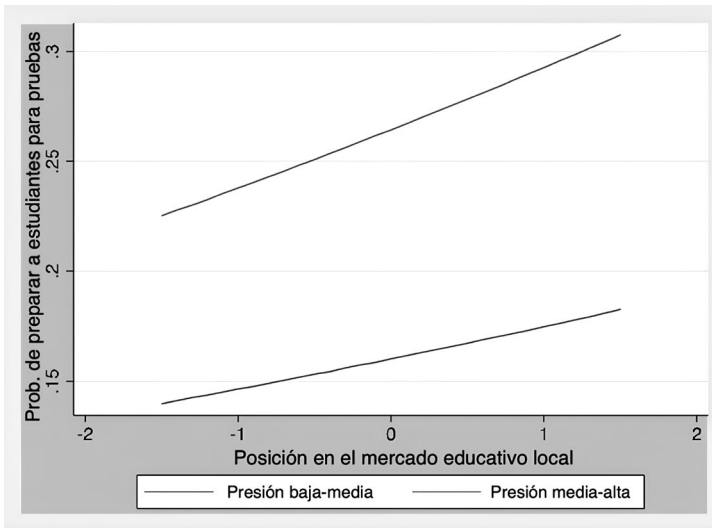
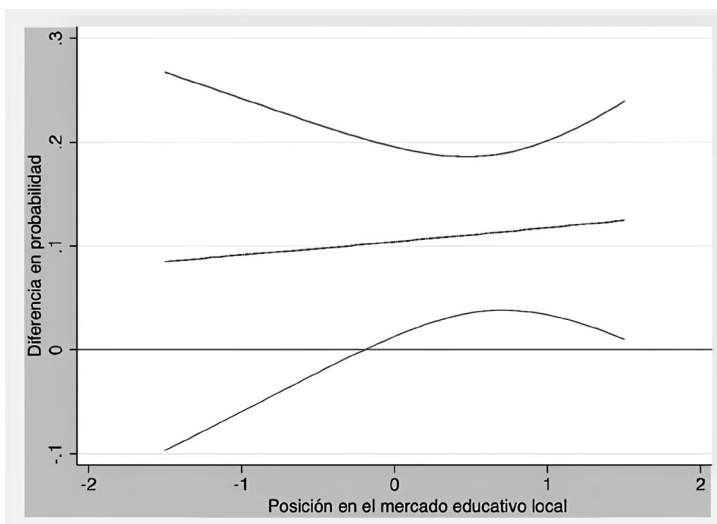


Gráfico 12 – *Teaching to the test*, según niveles de presión y posición en el MEL (conclusión)



Nota: El gráfico muestra la interacción entre la posición en el MEL (variable continua) y el nivel de presión (variable categórica) en un modelo de regresión logística donde la probabilidad de realizar *teaching to the test* (0/1) depende de una batería de variables individuales y de centro: género, media jornada, afiliación sindical, edad, nivel educativo, titularidad, autonomía de gestión, índice de autonomía educativa, satisfacción con el trabajo, satisfacción con la escuela e índice de cultura escolar favorable a las pruebas. Las variables con valores perdidos se imputaron mediante *missing dummies* (Cohen; Cohen, 1983; OCDE, 2007). Signif. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Fuente: Proyecto Reformed (Cataluña).

Cabe matizar que el *teaching to the test* no se percibe únicamente como una estrategia orientada a inflar los resultados. En algunos centros, estas pruebas se valoran positivamente por su potencial educativo: familiarizan al alumnado con formatos estandarizados, facilitan el entrenamiento en habilidades como la comprensión auditiva y ayudan a gestionar el tiempo de examen, especialmente en escuelas donde los exámenes escritos no son habituales. Desde esta perspectiva, la preparación para las pruebas puede entenderse como un modo de promover el bienestar del estudiante y reducir su ansiedad, más que como una práctica instrumental. En las entrevistas, varios docentes subrayan este enfoque formativo:

P: ¿Qué haces para intentar mejorar los resultados de las competencias básicas?

R: En sexto utilizamos algunos ejercicios de años anteriores: no hemos hecho muchas, pero las que sí practicamos ayudan por su tipología de actividades. Así evitamos que les sorprendan 'los listenings' o el tiempo limitado; en clase, a veces de manera inconsciente, pautamos las tareas, explicamos paso a paso, sin interrupciones, y luego hacemos prácticas en el tiempo estipulado (Docente 33, escuela pública 11).

Respecto a la carga de trabajo, corregimos y preparamos más a fondo: ya no es solo ensayar, sino explicar cada parte y calcular el tiempo disponible. Dedicamos muchas horas y sesiones a estas competencias porque los alumnos no suelen enfrentarse a un dossier tan extenso, de muchas páginas (Docente 30, escuela pública 12).

Rendición de cuentas y respuestas pedagógicas colaterales: otros factores explicativos

Más allá de la presión, otros factores inciden en prácticas como el *teaching to the test*, la reducción del currículo y la agrupación por niveles. En primer lugar, como muestra el modelo de la tabla 4, una cultura escolar que valora positivamente las pruebas de competencia se vincula con la redistribución de recursos hacia las materias evaluadas, la reducción curricular, el alineamiento con estándares de aprendizaje y la agrupación homogénea del alumnado. En segundo lugar, la presión por obtener buenos resultados se asocia con todas estas respuestas, excepto la agrupación. Las entrevistas confirmaron que, en escuelas donde se confía en la validez de las pruebas, estas no se usan solo de forma instrumental sino también como herramientas pedagógicas legítimas:

Lo que sí decimos es que realmente hay algunas pruebas que están bien hechas y entonces, si

esta forma de trabajar consideramos que es buena, pues vayamos por ahí, independientemente de que los resultados sean mejores o peores. Por ejemplo, la prueba de expresión escrita, lo que evalúa las rúbricas que tienen y la manera de poder desarrollarlo a nosotros nos parece muy bien y entonces nosotros lo hemos aplicado (Directora 14, escuela pública 14).

En tercer lugar, la composición social se asocia con una mayor intensidad de *teaching to the test*. El profesorado de estos centros recurre a simulacros regulares tanto para compensar las condiciones de partida del alumnado como para demostrar que, pese a las dificultades sociales, pueden ofrecer una educación de calidad.

P: ¿Hacéis ensayos de las pruebas?

R2: Sí, les llaman simulacros; vamos muy en serio.

R1: Hacen bastantes.

R2: E incluso utilizamos modelos de otros años, que se realizaron pruebas en cursos anteriores. De hecho, a partir del próximo año vuelve a cambiar...

R1: Sí, pero practicamos esas pruebas con antelación; no se las damos sin avisar, que sería un tema complicado de sexto. Les hablamos de ellas, les explicamos el formato.

P: Entiendo que tomáis elementos característicos de las competencias básicas e intentáis irlos introduciendo poco a poco en las clases, para que los alumnos se familiaricen.

R1: Sí, y después hacemos simulacros tal cual: tomamos una prueba antigua, la realizamos y revisamos los resultados (Directora 4 y Docente 12, escuela pública 4).

Tenemos descargadas pruebas de años anteriores, clasificadas en carpetas, y van practicando.

Una vez a la semana hacen un simulacro (Docente 26, escuela pública 9).

En cuarto lugar, un planteamiento pedagógico más centrado en el docente se asocia con una mayor frecuencia de preparación de los alumnos para las pruebas de CB y con la adaptación del currículo a sus contenidos. De manera más moderada, los enfoques tradicionales también muestran vínculos positivos con el alineamiento a los estándares de aprendizaje y la formación de agrupamientos homogéneos. Por ejemplo, en una escuela con un planteamiento educativo tradicional predominante, la directora y una docente consideran que los resultados de la evaluación externa son un indicador anual fundamental para la toma de decisiones en el centro:

P: ¿La prueba de competencias básicas es un instrumento que os sirve también a vosotros?

R1: Sí. Incluso ahora, cuando hacemos la memoria de final de curso, hay indicadores que nos sirven para entender cómo ha ido la expresión escrita de los niños; sacamos esta información de las pruebas de competencias.

R2: Lo que hacemos es comparar las notas que ponemos en la escuela, por ejemplo de expresión escrita, con la nota que obtienen los alumnos de sexto en expresión escrita, y luego comparamos: no es que estemos evaluando super alto, de qué van super bien y después llegan a la prueba y les sale muy bajo. En otras pruebas más estrictas ponían notas bajas, pero con las competencias dices escucha, pues lo hacen bastante bien; esto también es un buen análisis.

R1: La utilizamos y la contabilizamos como resultado.

R2: Y en los institutos también lo piden y lo utilizan de referencia (Directora 4 y Docente 12, escuela pública 4).

Finalmente, en quinto lugar, según la titularidad, las escuelas públicas tienden a concentrar recursos en las áreas evaluadas por las pruebas de CB, mientras que las concertadas responden a la presión por resultados intensificando la preparación específica para las pruebas. En las entrevistas en centros concertados, el profesorado describe cómo incorporan ejercicios estándar dentro del programa habitual y dedican tiempo a ensayos de exámenes anteriores:

P: ¿Hacéis alguna especie de ensayo de pruebas de años anteriores o dedicáis más horas, separáis en grupos o algo así?

R: Sí, te aseguro que sí. Ahora hablo de matemáticas: el programa que seguimos incluye día sí y día también ejercicios que podrían pertenecer a cualquier prueba de competencias básicas, pero lo preparamos con pruebas de años anteriores (Docente 35, escuela concertada 12).

Tabla 4 – Autonomía, niveles de presión y respuestas no esperadas (continúa)

	(1)	(2)	(3)	(4)	(5)
	Teaching to the test: actividades para preparar a los estudiantes (último mes antes de las pruebas)	Redistribución de los recursos en favor de las áreas evaluadas	El contenido de las pruebas se tiene en cuenta para tomar decisiones curriculares	Los estándares de aprendizaje influyen el enfoque pedagógico del centro	Agrupar a los estudiantes de forma homogénea
Autonomía educativa (currículo, métodos, material y formación)	0.10*	-0.08	-0.04	-0.04	-0.05
	(0.06)	(0.06)	(0.05)	(0.04)	(0.07)
Autonomía de gestión (salarios, selección docentes y directores, presupuestos)	-0.06	-0.01	-0.02	0.02	0.09
	(0.06)	(0.05)	(0.05)	(0.04)	(0.06)

Tabla 4 – Autonomía, niveles de presión y respuestas no esperadas (conclusión)

Posición en el MEL	0.06	0.11*	0.10	0.07	-0.06
	(0.08)	(0.06)	(0.07)	(0.05)	(0.07)
Planteamiento pedagógico tradicional	0.18***	-0.01	0.08**	0.06	0.01
	(0.07)	(0.05)	(0.04)	(0.06)	(0.06)
Escuela pública (ref. privada)	-0.27**	0.28**	0.10	0.07	0.12
	(0.12)	(0.11)	(0.13)	(0.10)	(0.15)
Proporción de estudiantes extranjeros	1.29***	0.25	0.02	-0.22	-0.62
	(0.43)	(0.24)	(0.22)	(0.16)	(0.49)
Cultura escolar favorable a las pruebas	0.03	0.21***	0.09**	0.10**	0.17***
	(0.06)	(0.04)	(0.04)	(0.05)	(0.05)
Presión para obtener buenos resultados	0.25***	0.29***	0.36***	0.20***	0.10
	(0.08)	(0.06)	(0.05)	(0.05)	(0.07)
Índice de autoeficacia	0.06	0.06**	0.03	0.03	0.07**
	(0.04)	(0.03)	(0.02)	(0.03)	(0.03)
Observaciones	426	622	628	624	636
R2	0.111	0.129	0.169	0.079	0.048

Nota: Cada columna corresponde a un modelo OLS con efectos fijos en el que las diferentes prácticas o respuestas escolares son función de una batería de variables individuales y del centro educativo. Además de las variables mostradas, se incluyeron como controles: tipo de contrato laboral, afiliación sindical, nivel educativo, edad y efecto fijo de la provincia de Barcelona. Las variables con valores perdidos se imputaron mediante *missing dummy variables* (Cohen; Cohen, 1983; OCDE, 2007). Los errores estándar aparecen entre paréntesis. Signif. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. En la tabla, los niveles de significatividad se indican con colores claros para * $p < 0.10$ y con colores oscuros para ** $p < 0.05$ y *** $p < 0.01$.

Fuente: Proyecto Reformed (Cataluña).

Conclusiones

En este estudio hemos analizado cómo los centros educativos implementan, perciben y resignifican las políticas de evaluación externa y rendición de cuentas en el contexto de un sistema formalmente de bajas consecuencias (*low-stakes*) que, al mismo tiempo, sitúa la autonomía escolar como pieza central de la gobernanza educativa. A partir de un

enfoque mixto —que combina cuestionarios, entrevistas y análisis documental— mostramos que el carácter *low-stakes* de la política no garantiza, por sí solo, una experiencia de baja presión en los centros. De hecho, una proporción significativa de docentes y equipos directivos declara sentir una fuerte presión por obtener buenos resultados en las pruebas de competencias básicas, aunque esta presión no derive de incentivos o sanciones formales, sino de mecanismos informales vinculados a la reputación institucional, las expectativas de las familias, la autoexigencia profesional y el clima de opinión dentro de las propias comunidades docentes.

Este hallazgo cuestiona el supuesto de que sistemas *low-stakes* conducen automáticamente a una apropiación formativa de la evaluación. Nuestros resultados muestran que la presión performativa se asocia con prácticas más instrumentales —como el *teaching to the test* y el estrechamiento curricular—, que desdibujan el propósito diagnóstico y formativo de la evaluación y tienden a limitar la enseñanza a los contenidos evaluados. Estos efectos aparecen con mayor intensidad en centros que concentran alumnado vulnerable y se encuentran en posiciones más desfavorecidas en el mercado educativo local, lo que evidencia que la competencia entre escuelas puede erosionar el carácter *low-stakes* del sistema y reforzar desigualdades ya existentes.

Frente a visiones que reducen los efectos no deseados a la presencia de incentivos materiales, nuestros hallazgos subrayan que, incluso en ausencia de sanciones explícitas, los mecanismos simbólicos de responsabilización y rendición de cuentas pueden producir dinámicas similares a las observadas en contextos *high-stakes*. Además, identificamos que la presión percibida no se concentra únicamente en quienes enseñan en cursos evaluados, sino que también se desplaza hacia docentes de cursos no sometidos a pruebas, un patrón contraintuitivo que sugiere la existencia de expectativas implícitas y responsabilidades compartidas dentro de los centros.

A su vez, el estudio destaca el papel clave de la cultura escolar y de la legitimidad de las pruebas en la configuración de los usos de los datos. Cuando el profesorado percibe las evaluaciones externas como válidas, aumenta la probabilidad de que los datos se utilicen de manera formativa.

Sin embargo, la utilidad real de los datos suele verse limitada por la falta de tiempo, recursos y espacios institucionales para traducir los resultados en ajustes curriculares y pedagógicos. En entornos con mayor complejidad social esta brecha es más aguda, lo que revela un sesgo estructural en la capacidad para aprovechar el potencial de mejora asociado a la evaluación externa.

Asimismo, la implementación efectiva de estas políticas está fuertemente mediada por actores intermedios. La inspección educativa, en particular, desempeña un rol decisivo: cuando ejerce un acompañamiento pedagógico cercano, los centros reportan más reflexión y uso significativo de los datos; cuando la relación se reduce a supervisión burocrática, la evaluación tiende a ser percibida como un mecanismo de control y no como un recurso para la mejora.

En conjunto, estos resultados revelan que el impacto de las políticas de autonomía, evaluación y RdC depende menos de su diseño formal y más de su apropiación por parte de los actores escolares, así como de las condiciones institucionales que habilitan o constriñen su implementación. Incluso cuando se conciben explícitamente como herramientas de mejora, pueden generar efectos colaterales y resistencias si no se acompañan de una infraestructura sólida de apoyo, una comunicación clara sobre sus objetivos y una estrategia sostenida de desarrollo de capacidades docentes.

Desde una perspectiva de política educativa, este estudio pone de relieve la necesidad de reforzar el carácter formativo del sistema de evaluación en Cataluña. Ello implica avanzar hacia un marco normativo coherente, estable y comprensible; garantizar tiempos, recursos y apoyos adecuados para el análisis pedagógico de los datos; y consolidar estructuras intermedias capaces de articular la autonomía con una responsabilidad compartida orientada al aprendizaje de todo el alumnado. No se trata de fomentar la competencia entre centros ni de poner el foco en la rendición de cuentas individual, sino de promover una cultura de responsabilidad profesional colectiva que reconozca las desigualdades de partida y priorice el valor pedagógico de la evaluación sobre su función clasificatoria.

En definitiva, el caso catalán muestra que un sistema *low-stakes* puede operar de facto como un sistema de alta presión ambiental si no se acompaña de las condiciones necesarias para que los datos sean útiles, legítimos y transformadores. La evaluación externa, entendida como herramienta de conocimiento y justicia educativa, solo cumplirá su potencial cuando vaya unida a una apuesta clara por el apoyo institucional, la equidad y la mejora escolar en todas las escuelas, especialmente en aquellas que más lo necesitan.

Referências

ALLEN, R.; VIGNOLES, A. What should an index of school segregation measure? **Oxford Review of Education**, [S. l.], v. 33, n. 5, p. 643–668, 2007. DOI: <https://doi.org/10.1080/03054980701366306>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/03054980701366306>. Acesso em: 5 jul. 2025.

BALL, S. J.; MAGUIRE, M.; BRAUN, A. **How schools do policy**: Policy enactments in secondary schools. Abingdon: Routledge. 2011.

BARBANA, S.; DUMAY, X.; DUPRIEZ, V. Local implementation of accountability instruments in the French-speaking community of Belgium. **European Educational Research Journal**, [S. l.], v. 19, n. 2, p. 94–108. 2020. DOI: <https://doi.org/10.1177/1474904119850964>. Disponível em: <https://journals.sagepub.com/doi/10.1177/1474904119850964>. Acesso em: 12 jul. 2025

BREAKSPEAR, S. **The policy impact of PISA**: An exploration of the normative effects of international benchmarking in school system performance. Paris: OECD Education Working Papers, 2012. DOI: <https://doi.org/10.1787/5k9fdfqffr28-en>. Disponível em: https://www.oecd.org/en/publications/the-policy-impact-of-pisa_5k9fdfqffr28-en.html. Acesso em: 3 jul. 2025

CLARKE, M.; LUNA-BAZALDUA, D. **Primer on large-scale assessments of educational achievement**. Washington, D.C.: World Bank. 2021. DOI: <https://doi.org/10.1596/978-1-4648-1659-8>. Disponível em: <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/827991619500628075>. Acesso em: 14 de julho de 2025

COHEN, J.; COHEN, P. **Applied multiple regression/correlation analysis for the behavioural sciences**. Mahwah: Erlbaum, 1983.

COMISSION EUROPEA. **Education and training monitor 2018**. Luxemburgo: Publications Office of the European Union, 2018.

DATNOW, A.; PARK, V.; WOHLSTETTER, P. **Achieving with data**: How high-performing school systems use data to improve instruction for elementary students. Los Angeles: Center on Educational Governance; USC Rossier School of Education, 2007.

DE GRAUWE, A. Improving the quality of education through school-based management: Learning from international experiences. **International Review of Education**, [S. l.], v. 51, n. 4, p. 269–287, 2005. DOI: <https://doi.org/10.1007/s11159-005-7733-1>. Disponível em: <https://link.springer.com/article/10.1007/s11159-005-7733-1>. Acesso em: 2 jul. 2025

DEE, T. S.; JACOB, B. The impact of No Child Left Behind on student achievement. **Journal of Policy Analysis and Management**, [S. l.], v. 30, n. 3, p. 418–446, 2011. <https://doi.org/10.1002/pam.20586>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/pam.20586>. Acesso em: 9 jul. 2025.

FARRE, R. **Crònica de la Llei d'educació de Catalunya**. Catalunya: Fundació Bofill, 2009.

FERRER-ESTEBAN, G. **Autonomia de centre, segregació escolar i mercats en educació**: Una anàlisi internacional dels efectes de l'autonomia en l'equitat i l'eficàcia dels sistemes educatius. 2015. 269 f. Tese (Doutorado em Sociologia) – Faculdade de Sociologia e Políticas Públicas, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, 2015.

GREANY, T. Innovation is possible, it's just not easy: Improvement, innovation and legitimacy in England's autonomous and accountable school system. **Educational Management Administration & Leadership**, [S. l.], v. 46, n. 1, p. 65–85, 2018. DOI: <https://doi.org/10.1177/1741143216659297>. Disponível em: <https://journals.sagepub.com/doi/10.1177/1741143216659297>. Acesso em: 11 jul. 2025.

HAMILTON, L. S.; STECHER, B. M.; KLEIN, S. P. **Making sense of test-based accountability in education**. Santa Monica: Rand Corporation, 2002.

JERRIM, J.; SIMS, S. When is high workload bad for teacher wellbeing? Accounting for the non-linear contribution of specific teaching tasks. **Teaching and Teacher Education**, [S. l.], v. 105, 2021. DOI: <https://doi.org/10.1016/j.tate.2021.103395>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0742051X21001190?via%3Dihub>. Acesso em: 4 jul. 2025.

LEVATINO, A.; PARCERISA, L.; VERGER, A. Understanding the stakes: The influence of accountability policy options on teachers' responses. **Educational Policy**, [S. l.], v. 38, n. 1, p. 31–60. 2024. DOI: <https://doi.org/10.1177/08959048221142048>. Disponível em: <https://journals.sagepub.com/doi/10.1177/08959048221142048>. Acesso em: 14 jul. 2025

LEVINSON, B. A.; SUTTON, M.; WINSTEAD, T. Education policy as a practice of power: Theoretical tools, ethnographic methods, democratic options. **Educational Policy**, [S. l.], v. 23, n. 6, p. 767–795, 2009. DOI: <https://doi.org/10.1177/0895904808320676>. Disponível em: <https://journals.sagepub.com/doi/10.1177/0895904808320676>. Acesso em: 13 jul. 2025.

MALEN, B. Revisiting policy implementation as a political phenomenon. In: HONIG, M. I. (ed.). **New directions in education policy implementation: Confronting complexity**. Albany: State University of New York Press, 2006. p. 83–104.

MAROY, C.; VAN ZANTEN, A. Regulation and competition among schools in six European localities. **Sociologie du travail**, [S. l.], v. 51, p. 67–79, 2009. DOI: <https://doi.org/10.1016/j.soctra.2009.01.005>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0038029609000065?via%3Dihub>. Acesso em: 7 jul. 2025.

MAROY, C.; VOISIN, A. **Think piece on accountability**. Paris: UNESCO, 2017.

MITTMAN, J.; JENNINGS, J. L. Accountability, achievement, and inequality in American public schools: A review of the literature. In: SCHNEIDER, B. (ed.). **Handbook of the sociology of education in the 21st century**. Cham: Springer, 2018. p. 475–492.

NICHOLS, S. L.; HARRIS, L. R. Accountability assessment's effects on teachers and schools. In: BROWN, G. T. L.; HARRIS, L. R. (eds.). **Handbook of human and social conditions in assessment**. Abingdon: Routledge, 2016. p. 40–56.

OECD. **PISA 2006: Science competencies for tomorrow's world. Analysis** (vol. 1). Paris: OECD Publishing, 2007.

OECD. **PISA in focus**. n° 9. School autonomy and accountability: Are they related to student performance? Paris: OECD Publishing, 2011.

PAGES, M.; FERRER-ESTEBAN, G.; VERGER, A.; PRIETO, M. At the crossroad of performativity and the market: Schools' logics of action under a hybrid accountability regime. **Compare: A Journal of Comparative and International Education**, [S. l.], p. 1–21, 2023. DOI: <https://doi.org/10.1080/03057925.2023.2175316>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/03057925.2023.2175316>. Acesso em: 10 jul. 2025.

QUILABERT, E.; VERGER, A.; MOSCHETTI, M. C.; FERRER-ESTEBAN, G.; PAGÈS, M. The obstacle race to educational improvement: governance, policies, and practices in disadvantaged schools. **British Journal of Sociology of Education**, [S. l.], v. 45, n. 6, p. 934–956, 2024. DOI: <https://doi.org/10.1080/01425692.2024.2376596>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01425692.2024.2376596>. Acesso em: 6 jul. 2025.

SPILLANE, J. P.; DIAMOND, J. B.; BURCH, P.; HALLETT, T.; JITA, L.; ZOLTNERS, J. Managing in the middle: School leaders and the enactment of accountability policy. **Educational Policy**, [S. l.], v. 16, n. 5, p. 731–762, 2002. DOI: <https://doi.org/10.1177/089590402237311>. Disponível em: <https://journals.sagepub.com/doi/10.1177/089590402237311>. Acesso em: 1 jul. 2025.

TELTEMANN, J.; JUDE, N. Assessments and accountability in secondary education: International trends. **Research in Comparative and International Education**, [S. l.], v. 14, n. 2, p. 249–271, 2019. DOI: <https://doi.org/10.1177/1745499919846174>. Disponível em: <https://journals.sagepub.com/doi/10.1177/1745499919846174>. Acesso em: 8 jul. 2025.

TERHART, E. Teacher resistance against school reform: Reflecting an inconvenient truth. **School Leadership & Management**, [S. l.], v. 33, n. 5, p. 486–500, 2013. DOI: <https://doi.org/10.1080/13632434.2013.793494>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/13632434.2013.793494>. Acesso em: 15 jul. 2025.

THIEL, C.; SCHWEIZER, S.; BELLMANN, J. Rethinking side effects of accountability in education: Insights from a multiple methods study in four German school systems. **Education Policy Analysis Archives**, [S. l.], v. 25, n. 93, 2017. DOI: <https://doi.org/10.14507/epaa.25.2662>. Disponível em: <https://epaa.asu.edu/index.php/epaa/article/view/2662>. Acesso em: 2 jul. 2025.

TORRES, R. Does test-based school accountability have an impact on student achievement and equity in education? A panel approach using PISA. **OECD Education Working Papers**, 250. OECD Publishing, 2021. DOI: <https://doi.org/10.1787/0798600f-en>. Disponível em: https://www.oecd.org/en/publications/does-test-based-school-accountability-have-an-impact-on-student-achievement-and-equity-in-education_0798600f-en.html. Acesso em: 5 jul. 2025.

UNESCO. **The promise of large-scale learning assessments**: Acknowledging limits to unlock opportunities. Paris: UNESCO, 2019. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000369947>. Acesso em: 14 jul. 2025.

VERGER, A.; FONTDEVILA, C.; PARCERISA, L. Reforming governance through policy instruments: How and to what extent standards, tests and accountability in education spread worldwide. **Discourse: Studies in the Cultural Politics of Education**, [S. l.], v. 40, n. 2, p. 248–270, 2019. DOI: <https://doi.org/10.1080/01596306.2019.1569882>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01596306.2019.1569882>. Acesso em: 9 de julho de 2025.

VOGLER, K. E. Comparing the impact of accountability examinations on Mississippi and Tennessee social studies teachers' instructional practices. **Educational Assessment**, [S. l.], v. 13, n. 1, p. 1–32. 2008. DOI: <https://doi.org/10.1080/10627190801968158>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/10627190801968158>. Acesso em: 7 jul. 2025.

WERLER, T.; FÆREVAAG, M. K. National testing data in Norwegian classrooms: A tool to improve pupil performance? **Nordic Journal of Studies in Educational Policy**, [S. l.], v. 3, n. 1, p. 67–81, 2017. DOI: <https://doi.org/10.1080/20020317.2017.1320188>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/20020317.2017.1320188>. Acesso em: 3 jul. 2025.

CAPÍTULO 3

AVALIAÇÃO EM LARGA ESCALA E ACCOUNTABILITY EM EDUCAÇÃO

Luís Antônio Fajardo Pontes¹

Tatiane Gonçalves Moraes²

DOI: 10.52695/978-65-5456-163-1.3

Introdução

O propósito deste capítulo é abordar a questão da *accountability* e sua relação com a avaliação em larga escala dos sistemas educacionais. Trata-se, naturalmente, de um tema de grande atualidade e relevância, tanto no campo educacional como também no político e acadêmico, entre outros. Mas, além disso, nele também costumam surgir controvérsias e debates em torno da divulgação dos resultados das avaliações em larga escala e do impacto que isso pode gerar para as escolas públicas e para os profissionais que lá trabalham. Esse é o caso, por exemplo, das medidas do Índice de Desenvolvimento da Educação Básica — o Ideb — ou dos resultados das avaliações realizadas pelos diferentes sistemas estaduais de aferição da qualidade educacional.

1 Centro de Políticas Públicas e Avaliação da Educação – CAEd. E-mail: luis@caed.ufjf.br

2 Universidade Federal de Juiz de Fora

Contudo, nosso objetivo no presente texto não é o de abordar os embates ideológicos por trás dessas questões, mas, sim, o de esboçar uma linha histórica através da qual o movimento da *accountability* tem se desenvolvido e ganhado força no Brasil, ao mesmo tempo que também pretendemos abordar algumas das justificativas por trás dessa tendência, sem esquecermos de mencionar algumas das principais dificuldades técnicas – e políticas – que colocam desafios consideráveis a esse processo.

Quando se trata da questão da *accountability*, uma questão que surge de imediato é uma certa pluralidade de vocábulos e expressões que costumam ser utilizados de modo meio indistinto entre si. Dessa forma, temos, por exemplo, a tradução desse termo em português para “responsabilização”, que costuma também ser empregada nesse mesmo sentido. E há também a expressão “prestação de contas”, que igualmente costuma se fazer presente em textos sobre esse mesmo tema.

Em relação a esse problema de nomenclatura, Brooke e Cunha (2008, p. 21) observam o seguinte:

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

Apesar dessa ressalva, cabe dizer que o termo *accountability*, de modo mais ou menos indistinto, refere-se a ambos os tipos de utilização mencionados por Brooke, que também serão abordados nesta exposição. Mas, antes disso, faz-se necessário contextualizar mais exatamente o surgimento desse movimento no âmbito mais geral da avaliação das organizações públicas, algo que será realizado na próxima seção.

A *accountability* no contexto das avaliações de desempenho das organizações públicas

A avaliação de desempenho no setor público tornou-se uma prática amplamente difundida em diversos países, impulsionada pelas exigências contemporâneas de transparência, eficiência na aplicação dos recursos públicos e responsabilização dos gestores. A crescente defasagem entre as demandas sociais e os recursos disponíveis intensificou a necessidade de que governos demonstrem capacidade de resposta eficaz, o que inclui a geração de dados objetivos e comparáveis sobre os resultados alcançados por suas instituições (Arellano *et al.*, 2012).

No entanto, ao tratar-se de organizações públicas, esse processo assume características particulares. A avaliação de desempenho ocorre em um ambiente marcadamente político, em que diversos atores, pertencentes a esferas administrativas, sociais e ideológicas distintas, influenciam e são influenciados pelas decisões avaliativas. Como destaca Shaw (2006), a construção de sistemas de avaliação de desempenho (aqui chamado de SAD) exige não apenas critérios técnicos robustos, mas também um esforço de negociação, articulação e consenso político em torno dos objetivos e critérios a serem utilizados.

O campo educacional ilustra com nitidez essas tensões. As escolas operam em contextos multifacetados; os resultados não podem ser compreendidos apenas a partir dos processos internos, como práticas pedagógicas e gestão escolar. Fatores externos, sobretudo o nível socioeconômico dos alunos e suas famílias, exercem influência significativa sobre o desempenho acadêmico. O Relatório Coleman (1966), pioneiro na análise da influência do contexto socioeconômico na educação, demonstrou que

esse fator explicava grande parte da variabilidade nos resultados dos estudantes. Ainda que relevante, esse achado foi posteriormente criticado por não considerar suficientemente o papel das interações escolares e da qualidade pedagógica (Brooke; Cunha, 2008).

Frente a esse cenário, a construção de matrizes de marco lógico (MML) ou teorias de programa constitui ferramenta essencial para a organização dos processos avaliativos. Segundo Weiss (1997), essas teorias descrevem os mecanismos causais que ligam os insumos (*inputs*), as ações executadas e os resultados esperados de um programa ou política pública. Além de tornarem explícitas as hipóteses subjacentes, essas estruturas ajudam a identificar os pontos estratégicos em que os indicadores de desempenho devem ser posicionados, viabilizando o monitoramento e a retroalimentação das decisões de gestão.

Como um exemplo concreto no âmbito educacional, essas cadeias causais poderiam ser ilustradas por políticas que visam, digamos, ao aumento salarial dos professores como estratégia para melhorar os resultados de aprendizagem. A eficácia dessa medida depende de diversas mediações: a atração de profissionais mais qualificados, a redução do acúmulo de funções ou cargos em diferentes escolas por parte dos docentes, a maior dedicação ao planejamento e à formação continuada, entre outras (Reimers, 2008). A avaliação, nesse caso, deve buscar indicadores intermediários e finais que possam comprovar ou refutar essas hipóteses.

Porém, uma forma alternativa — e, na verdade, bastante utilizada — de se promover a avaliação educacional é concentrar-se nos *outputs*, o que, nesse caso, poderia corresponder, ao menos em parte, ao desempenho dos estudantes nos testes padronizados, a fim de se tentar aferir a qualidade do seu aprendizado. Indo por essa linha, nas últimas décadas, os testes padronizados em larga escala consolidaram-se como principal instrumento para a mensuração de resultados educacionais. Eles oferecem dados amplos, comparáveis e de aplicação periódica, o que facilita análises longitudinais e decisões baseadas em evidências. No Brasil, os testes de Língua Portuguesa e Matemática, como os aplicados pelo Sistema de Avaliação da Educação Básica (SAEB), Prova Brasil e sistemas estaduais

como o SIMAVE, tornaram-se centrais para políticas de avaliação e responsabilização (Mello, 2013).

Apesar de sua importância, esses instrumentos enfrentam críticas consideráveis. Muitos autores alertam para o risco de um estreitamento curricular, no qual os conteúdos avaliados se sobrepõem às diretrizes curriculares, mais amplas, reduzindo o escopo da educação escolar. Há também preocupações sobre a desvalorização de habilidades socioemocionais, do pensamento crítico e da participação cidadã — aspectos essenciais, mas dificilmente mensuráveis por testes de múltipla escolha (Hilton; Pellegrino, 2012).

Outro debate relevante refere-se à validade e à fidedignidade das avaliações. Muitas vezes, os resultados são tratados como precisos e inequívocos, ignorando possíveis margens de erro ou limitações metodológicas. Essa prática pode levar a decisões injustas, como premiar ou punir escolas com base em dados imprecisos e comprometer políticas educacionais baseadas em evidências (Koretz, 2008).

Nesse contexto, o conceito de “efeito-escola” torna-se central. Trata-se da capacidade de uma instituição escolar de influenciar o desempenho de seus alunos, independentemente do contexto externo. Teddlie, Reynolds e Sammons (2000) propõem sete dimensões para análise desse efeito: (1) existência; (2) magnitude; (3) estabilidade temporal; (4) consistência entre áreas de conhecimento; (5) variação por contexto; (6) efeitos diferenciais entre grupos de alunos; e (7) continuidade dos efeitos ao longo do tempo.

Complementarmente, Raudenbush e Willms (1995) propõem uma tipologia dos efeitos-escola. Os de tipo A referem-se aos ganhos absolutos de desempenho dos alunos — mais valorizados por famílias e comunidade. Já os efeitos de tipo B consideram o desempenho obtido à luz dos recursos e condições disponíveis, sendo de maior interesse para gestores preocupados com eficiência e equidade. Essa distinção é fundamental para entender o lugar de cada tipo de informação na tomada de decisão educacional.

Outra possibilidade de se classificarem as avaliações é quanto às suas respectivas consequências, que podem ser baixas ou altas para os atores

e instituições envolvidos, como alunos, professores, diretores e escolas, por exemplo. Em alguns casos, as consequências dessas avaliações se limitam à informação dos resultados para as partes envolvidas, sendo, por isso, chamadas de baixa consequência, ou *low-stakes* na literatura internacional. Por outro lado, há outras avaliações que resultam em consequências capazes de afetar mais intensamente os atores e instituições avaliados, tanto no sentido positivo quanto negativo. Exemplo do primeiro caso é a premiação pelo resultado, ou a aprovação em um processo seletivo, enquanto, no segundo caso, poderia ser a interferência externa numa escola com baixo desempenho nas provas padronizadas ou a substituição dos profissionais da escola; e, em alguns casos extremos, até o fechamento de determinados programas ou instituições educacionais. Nesses últimos casos, tem-se a avaliação de altas consequências, ou *high-stakes*.

A construção de sistemas de avaliação que equilibrem precisão técnica, legitimidade política e sensibilidade ao contexto é, portanto, um dos principais desafios da gestão educacional contemporânea. No caso brasileiro, os dados oriundos das avaliações nacionais e estaduais dos sistemas educacionais representam oportunidades valiosas, mas requerem análises mais refinadas, que levem em conta a instabilidade dos resultados ao longo do tempo, a composição do corpo discente e os múltiplos fatores que interferem na aprendizagem.

Porém, igualmente, é preciso observar que as avaliações educacionais não devem ser vistas apenas como instrumentos de controle, mas também como ferramentas para promover o autoconhecimento institucional, o aprimoramento pedagógico e a democratização da gestão. A construção coletiva de indicadores, a participação dos diferentes atores escolares no processo avaliativo e o uso formativo dos resultados podem transformar os SADs em importantes aliados da qualidade e da justiça educacional.

Outro ponto que cabe observar é o caráter marcadamente internacional dessa tendência de avaliação educacional em larga escala, segundo um fenômeno que já vem se intensificando desde as últimas décadas do século 20. E, nesse sentido, os Estados Unidos foi um país que, conforme adiante se verá, influenciou bastante as políticas públicas brasileiras

nessa mesma direção, principalmente as que se elaboraram e implementaram nacionalmente no Brasil a partir da década de 1990 e ao longo das primeiras décadas do século 20. Por essa razão, serão apresentadas, na sequência, algumas das principais características do sistema norte-americano de avaliação e responsabilização educacional.

O movimento norte-americano de reforma educacional: evolução, desafios e sistemas de avaliação em larga escala

O cenário da avaliação educacional e *accountability* nos EUA pode ser mais bem compreendido quando observado através de uma síntese de seus principais aspectos históricos, políticos e metodológicos, com ênfase na articulação entre tais reformas e os sistemas de avaliação educacional em larga escala. Esse panorama se revela especialmente relevante na medida em que a experiência norte-americana exerce influência direta sobre outros contextos educacionais, notadamente o brasileiro, que vem paulatinamente adotando elementos similares em suas práticas avaliativas.

Historicamente, a educação pública dos Estados Unidos apresenta um desenvolvimento precoce associado à formação das colônias da Nova Inglaterra, onde já se institucionalizava, desde o século XVII, a obrigatoriedade da criação de escolas em cidades com mais de cinquenta famílias. A legislação educacional do estado de Massachusetts de 1647 já previa tais obrigações. Essa tradição consolidou-se ao longo do tempo, promovendo tanto a ampliação da cobertura escolar quanto o debate sobre igualdade de oportunidades, que viria a ganhar especial relevo com o movimento dos direitos civis nos anos 1960. Apesar do acesso relativamente precoce à educação formal, o sistema escolar norte-americano foi marcado por uma história de segregação racial institucionalizada, cujas últimas manifestações formais somente foram abolidas no final da década de 1970 (Pontes, 2015).

Ao longo dos séculos 19 e 20, o sistema educacional dos Estados Unidos passou por diversas críticas e reformas, com vistas a melhorar não apenas o acesso, mas também a qualidade da educação. Os primeiros movimentos reformistas advogavam a necessidade de um currículo mais amplo, integrando disciplinas humanísticas às chamadas habilidades básicas, como

leitura, escrita e matemática. Posteriormente, com a consolidação da sociedade industrial, passaram a ser exigidas também competências técnicas e industriais, dando origem a uma concepção de educação voltada ao mercado de trabalho (Chung; Reimers, 2014).

Na segunda metade do século 20, ganhou destaque o relatório *A Nation at Risk* (1983), que denunciava uma suposta crise na educação norte-americana, relacionando baixos indicadores educacionais ao declínio da competitividade econômica dos Estados Unidos. O documento gerou amplo impacto e levou à proposição de uma agenda de reformas pautadas na responsabilização de escolas e professores. Nesse contexto, fortaleceu-se o movimento por *accountability*, que culminou na aprovação da legislação *No Child Left Behind* (NCLB, literalmente Nenhuma Criança Deixada para Trás), em 2001. Essa legislação previa a aplicação de testes padronizados para medir o desempenho dos alunos em disciplinas como leitura e matemática, condicionando o repasse de recursos às escolas às metas estabelecidas pelo governo federal.

A perspectiva de responsabilização baseava-se na premissa de que o controle rigoroso por meio de testes geraria um incentivo à melhoria da qualidade do ensino. No entanto, ao longo dos anos, diversas críticas emergiram, apontando falhas conceituais e metodológicas no modelo proposto. Dentre elas, destacam-se a imprecisão das medidas adotadas, o uso de modelos agregados de análise que desconsideravam as especificidades contextuais, além de efeitos perversos sobre o currículo escolar, como o “estreitamento curricular”, ou seja, a tendência à priorização de conteúdos diretamente cobrados nas avaliações em detrimento de outras dimensões da formação dos estudantes. Outra crítica nesse sentido é a que aponta a possibilidade de ocorrência do chamado *gaming*, pelo qual a avaliação passa a ser vista como uma espécie de esporte em que o aumento do desempenho pode ocorrer de forma um tanto quanto artificial, sem refletir uma melhoria qualitativa do sistema em questão, de que é exemplo a prática de “ensinar para o teste”.

As críticas mais severas à abordagem tecnocrática da reforma recaem sobre a incapacidade do sistema de promover igualdade de condições

de aprendizagem, visto que o desempenho escolar é influenciado por fatores socioeconômicos que escapam à esfera de controle da escola. Assim, a adoção de sanções baseadas exclusivamente em resultados de testes ignora a complexidade das condições escolares, gerando penalizações indevidas sobretudo para escolas localizadas em contextos mais vulneráveis.

Ainda assim, há defensores do modelo de avaliação em larga escala proposto nesses moldes. E a argumentação, nesse sentido, aponta como avanço a padronização de indicadores, a ampliação da transparência na gestão escolar e a produção de dados comparáveis sobre o desempenho dos estudantes. Para esses autores, a existência de metas claras e instrumentos de monitoramento permite à gestão educacional atuar com mais racionalidade e eficácia.

Nos anos mais recentes, o debate sobre a reforma educacional nos Estados Unidos passou a incorporar as chamadas habilidades do século 21, como pensamento crítico, comunicação, colaboração e criatividade, além de competências digitais. Essas habilidades, amplamente demandadas pelo mercado de trabalho contemporâneo, são vistas como essenciais para a formação de cidadãos autônomos e adaptáveis. No entanto, surge, então, o desafio de mensurá-las de forma confiável e consistente, dado seu caráter subjetivo e contextualizado.

A complexidade dessas novas competências questiona a centralidade dos testes padronizados na avaliação da aprendizagem, demandando abordagens mais qualitativas e formativas que considerem o desenvolvimento integral dos estudantes. Nesse sentido, o movimento atual tende a equilibrar os modelos tradicionais de responsabilização com práticas pedagógicas mais reflexivas, inclusivas e contextualizadas.

Quanto às avaliações tradicionais do aprendizado estudantil, o caso do *National Assessment of Educational Progress* (NAEP), instrumento nacional norte-americano criado em 1964, merece destaque por adotar um modelo amostral de aplicação de testes, permitindo a observação de tendências e padrões de desempenho educacional sem a pressão de responsabilização direta de escolas e professores. Tal abordagem se contrapõe ao modelo

do NCLB e inspirou iniciativas semelhantes em outros países, incluindo o Sistema de Avaliação da Educação Básica (SAEB) no Brasil.

Em resumo, a experiência estadunidense com reformas educacionais e sistemas de avaliação em larga escala revela tanto os avanços quanto os limites de uma abordagem centrada em resultados mensuráveis. Ao mesmo tempo que contribui para o desenvolvimento de uma cultura de avaliação e de *accountability*, também expõe os riscos de uma racionalidade tecnocrática que reduz a complexidade da educação a indicadores quantitativos. A análise dessa experiência oferece subsídios importantes para a formulação de políticas educacionais em outros contextos, especialmente em países como o Brasil, que buscam aprimorar seus sistemas avaliativos respeitando suas especificidades sociais, culturais e institucionais.

Um caso brasileiro: a avaliação externa como instrumento da gestão educacional nos estados

Conforme se disse anteriormente, o Brasil tendeu a se espelhar no modelo de um sistema nacional e amostral de avaliação do desempenho discente, de tal forma que o SAEB, em boa medida, inspirou-se no NAEP. Mas, além disso, ao longo do tempo, os estados brasileiros também tenderam a criar seus próprios sistemas de avaliação educacional que fossem mais sensíveis às suas próprias demandas e especificidades. Nesse sentido, segundo Brooke e Cunha (2011), as avaliações externas em larga escala têm evoluído nos estados brasileiros, passando de instrumentos técnicos pouco utilizados para ferramentas integradas às políticas de gestão educacional. O estudo desses dois autores mostra que, diferentemente do cenário descrito por Sousa e Oliveira (2007), em que os resultados das avaliações educacionais anteriormente pouco informavam sobre as decisões gerenciais, os sistemas atuais revelam um uso mais variado e sofisticado das informações coletadas.

Indo por essa linha, os autores observam que as formas tradicionais de classificar as avaliações (formativa *versus* somativa ou diagnóstica *versus* certificatória) não explicam a multiplicidade de usos na gestão. Eles, então, propõem um modelo baseado na utilidade concreta das avaliações

nas políticas públicas, inspirado na matriz de Shepard, que categoriza os usos em sete funções: diagnóstico, certificação, avaliação de pessoal, avaliação de programas, alocação de recursos, responsabilização (*accountability*) e pesquisa. Contudo, Brooke e Cunha (2011) ajustam essa proposta, fundindo e expandindo categorias para refletir melhor o contexto brasileiro, incluindo o planejamento e o monitoramento.

Em linhas gerais, Brooke e Cunha concluem que o uso dos resultados de avaliações externas como ferramenta de gestão educacional tornou-se mais presente e estruturado nas políticas estaduais. As avaliações passaram a influenciar decisões sobre currículo, formação, planejamento, alocação de recursos e incentivos. Ainda que persistam desafios, como a apropriação pedagógica dos dados e a resistência de alguns setores, observa-se um movimento em direção à institucionalização da avaliação como parte integrante da gestão educacional pública.

Avaliação educacional e *accountability* no Brasil: um panorama crítico

Acerca da associação entre a avaliação educacional e a *accountability* no caso brasileiro, Reinaldo Fernandes e Amaury Gremaud (2009) produziram uma análise do papel das avaliações externas e dos indicadores de qualidade no contexto da Educação Básica no Brasil, com ênfase na incorporação de mecanismos de prestação de contas. Segundo esses autores, as avaliações ganharam uma função central na política educacional após a adoção do Plano de Desenvolvimento da Educação (PDE), sendo articuladas a metas e índices como o Ideb.

Numa das seções de seu texto, os autores discutem os potenciais e também os riscos dos sistemas de avaliação. Afirmam que, embora as avaliações possam ser feitas com base em insumos, processos ou resultados, é a última abordagem que ganhou destaque com os sistemas de *accountability*. A preferência por indicadores de resultados baseia-se na fraca correlação observada entre insumos/processos e o desempenho dos alunos. Contudo, reconhece-se que os resultados de exames são influenciados por fatores externos à escola, como as condições socioeco-

nômicas. Por outro lado, as tentativas de isolar o “efeito escola”, como os Modelos de Valor Adicionado, enfrentam limitações metodológicas.

Na terceira parte desse trabalho, seus autores discutem a incorporação dos objetivos de *accountability* no Brasil a partir da Prova Brasil e da divulgação dos resultados do ENEM por escola. Observam que, com a Prova Brasil, as avaliações passam a ser censitárias e permitem mais transparência, visto que, através delas, todas as escolas passam a ter seu respectivo desempenho mensurado. Os autores apontam que, embora o sistema brasileiro seja recente, ele avança no sentido de oferecer informações comparáveis por escola, o que caracteriza um modelo de “*accountability* fraca”. Por sua vez, o ENEM, mesmo sendo de participação voluntária por parte dos estudantes, passou a integrar esse sistema ao divulgar resultados por escola durante certa fase de seu respectivo ciclo.

Fernandes e Gremaud abordam também a questão do Ideb, que busca integrar informações de fluxo escolar e desempenho em exames. Os autores alertam para o risco de que os sistemas baseados apenas na proficiência podem aumentar a reprovação e evasão. E a lógica por trás disso é o surgimento de uma pressão maior sobre os estudantes para obterem melhores resultados acadêmicos, o que pode levá-los a terem maiores taxas de repetência, desistência e abandono escolar. Por sua vez, o Ideb, ao combinar uma nota média (de português e matemática) com uma taxa de aprovação, pretende equilibrar esses efeitos.

Ainda em relação ao Ideb, analisam as metas a ele atreladas. Apontam que, para promover o comprometimento das redes de ensino, foi estipulado um sistema de metas progressivas até 2021, com base em uma função logística e com o objetivo de convergência entre regiões, ou seja, as metas das diferentes regiões convergem para um mesmo ponto num futuro “longínquo” (no caso, o ano de 2095, muito além, portanto, da faixa “real” do Ideb original, que ia até o ano do cumprimento de suas respectivas metas, em 2021/22). Cada escola e rede tem metas diferenciadas, de modo a harmonizar melhor o “esforço” de cada uma delas para cumprir sua respectiva meta (INEP, [s. d.]a).

A análise dos referidos autores termina com a conclusão de que o sistema de avaliação com *accountability* adotado no Brasil visa induzir melhorias educacionais respeitando o federalismo. Eles reconhecem o potencial desses mecanismos para elevar o desempenho escolar, mas também os riscos envolvidos, como a exclusão de alunos ou a distorção de incentivos. A combinação de exames e indicadores de fluxo no Ideb, associada a metas pactuadas, configura uma estratégia que busca alinhar qualidade e equidade na Educação Básica.

Foi com esse pensamento que o Brasil, a partir da primeira década dos anos 2000, viu ampliar sua rede de avaliação educacional em larga escala, ancorada numa política federal de *accountability* de relativamente baixas consequências.

E, nesse ponto, cabe comentar como se deu mais exatamente a evolução desse movimento no Brasil. Em linhas gerais, poder-se-ia dizer que, do seu “nascimento” até a sua “maturidade”, o processo brasileiro de avaliação educacional conheceu três diferentes fases, que serão tratadas na próxima seção.

Avaliação educacional em larga escala no Brasil: uma visão de conjunto

A avaliação em larga escala da educação no Brasil passou a desempenhar um papel central nas políticas públicas a partir da década de 1990. Essa transformação ocorreu paralelamente à ampliação do acesso à Educação Básica e à crescente valorização da mensuração do desempenho escolar como estratégia de gestão e responsabilização dos sistemas de ensino.

Mas, ao mesmo tempo, também se podem perceber diferentes fases nesse processo até se chegar a uma espécie de “consolidação” ao final da primeira década do século 21 e ao longo das seguintes. Desse modo, serão a seguir apresentados seus três períodos distintos, destacando-se os avanços, as principais preocupações e as inovações metodológicas.

Fase fundacional (1990 – 1997): estruturação dos sistemas de avaliação

O processo de democratização da Educação Básica no Brasil implicou o ingresso de parcelas historicamente excluídas do sistema escolar. Essa expansão gerou pressões por mais eficiência e qualidade, provocando o surgimento de instrumentos formais de avaliação institucional baseados em testes padronizados e comparáveis entre unidades escolares, municípios e estados. O uso desses sistemas passou a ser visto como um mecanismo para promover transparência, identificar deficiências e orientar políticas públicas.

Nessa primeira fase, o principal marco foi a criação do Sistema de Avaliação da Educação Básica (Saeb), que, inicialmente, era aplicado de forma amostral e focalizado nas séries finais de cada ciclo da Educação Básica. Com apoio do Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), o Saeb introduziu o uso de escalas de proficiência e, mais tarde, incorporou a Teoria da Resposta ao Item (TRI), permitindo a construção de instrumentos com poder maior de comparação longitudinal. Essa etapa marcou a consolidação de uma agenda técnica e política voltada para a mensuração do rendimento dos estudantes em língua portuguesa e matemática.

Fase de expansão (1997 – 2005): ampliação e consolidação das avaliações

Durante esse período, houve uma significativa ampliação do escopo das avaliações. A Prova Brasil foi instituída como um desdobramento do Saeb, com aplicação censitária nas escolas públicas urbanas e foco em turmas de 4º e 8º anos do Ensino Fundamental. A Prova Brasil forneceu dados mais precisos por escola, permitindo sua utilização para formulação de políticas em nível local. Além disso, reforçou-se a associação entre avaliação externa e prestação de contas, criando-se um ambiente em que gestores e professores passaram a ser responsabilizados pelos resultados obtidos.

Outra realização relevante dessa fase foi que, ao seu final, a avaliação do SARESP, o Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo, passou a adotar a mesma métrica do SAEB em suas respectivas

avaliações do Ensino Básico, algo que também acabou se verificando em praticamente a totalidade de sistemas estaduais brasileiros. E isso também coincidiu com que, em 1997, fossem criadas as escalas unificadas do SAEB, uma para língua portuguesa e outra para matemática, que vêm vigorando desde então.

Fase de responsabilização (2005 – presente): Ideb e suas implicações

O ano de 2005 marca a criação do Índice de Desenvolvimento da Educação Básica (Ideb), que passou a agregar os dados da Prova Brasil e do Censo Escolar, combinando rendimento (fluxo escolar) e desempenho (proficiência) (INEP, [s. d.]). O Ideb tornou-se um indicador-síntese de qualidade da Educação Básica, amplamente utilizado para definir metas, distribuir recursos e comparar o desempenho. Seu caráter público e sua ampla disseminação intensificaram os efeitos simbólicos e práticos das avaliações, reforçando a cultura da mensuração e do resultado.

Contudo, a vinculação entre as metas e o desempenho no Ideb tem suas limitações e desafios, como os já comentados anteriormente acerca das avaliações em geral, do que são exemplos a indução ao ensino “para o teste”, o estreitamento curricular e a desigualdade de incentivos entre redes e escolas com contextos socioeconômicos distintos. Tais críticas não anularam a importância do Ideb, mas evidenciaram a necessidade de aperfeiçoamentos metodológicos e cuidados na sua interpretação e uso político.

Em paralelo à consolidação do Saeb e do Ideb, nessa terceira fase, os estados brasileiros começaram a implantar seus próprios sistemas de avaliação. Inspirados no modelo federal, esses sistemas – como o Sistema Mineiro de Avaliação da Educação Básica (SIMAVE) – ampliaram a capilaridade e a capacidade de diagnóstico regional, promovendo mais autonomia técnica e política das secretarias estaduais de educação em relação às instituições federais.

Os sistemas estaduais, muitas vezes, utilizam a mesma metodologia do Saeb, incluindo o uso da TRI, o que permite o alinhamento entre os indicadores locais e nacionais. No caso mineiro, por exemplo, o SIMAVE

destaca-se por sua longevidade e consistência metodológica, oferecendo uma base importante para análises longitudinais do desempenho escolar no estado.

Conclusões

Com base no que se viu no presente texto, podem-se citar as seguintes conclusões principais:

As avaliações das organizações públicas surgem como resposta às demandas por transparência e eficiência, sendo aplicadas em um ambiente político marcado por múltiplos interesses. No campo educacional, o desafio é ainda maior, dado que o desempenho dos alunos depende de fatores tanto internos quanto externos à escola, como o nível socioeconômico.

Nesse contexto, assume relevância o papel dos testes padronizados como forma prática, mas nem sempre suficiente, de aferição de resultados.

Um emblemático caso internacional de avaliação educacional em larga escala associada à *accountability* é o norte-americano. Nesse sentido, um de seus marcos foi a lei do *No Child Left Behind*, de 2001, marcada por uma forte ênfase na responsabilização com base em testes padronizados.

Porém, embora modelos como esse tenham ampliado a transparência e a padronização das avaliações, também podem gerar efeitos colaterais negativos, como o estreitamento curricular e a prática do “ensinar para o teste”.

No caso brasileiro, os estados da Federação, além do governo federal, desenvolveram seus próprios sistemas de avaliação externa em larga escala. E, a partir da década de 1990, esses sistemas evoluíram de instrumentos meramente técnicos para mecanismos efetivos de gestão educacional.

Outro marco da avaliação educacional brasileira foi o surgimento do Ideb, sendo caracterizado por uma “*accountability* fraca”, de consequências moderadas. Esse índice, por sua vez, tenta conciliar desempenho e fluxo escolar por meio de metas diferenciadas por escola e por rede.

O movimento da avaliação educacional e *accountability* no Brasil também pode ser visto como decomposto em três fases: fundacional (1990

– 1997), de expansão (1997 – 2005) e de responsabilização (2005 – presente). Inicialmente, o destaque foi a criação do Saeb e a incorporação de métricas como a Teoria da Resposta ao Item. Depois, com a Prova Brasil, houve a ampliação das aplicações e mais detalhamento dos dados por escola. A terceira fase é marcada pela criação do Ideb e pelo fortalecimento dos sistemas estaduais, que passaram a adotar metodologias semelhantes às federais. Consolidou-se, assim, um modelo nacional de avaliação baseado em evidências, mas que ainda enfrenta desafios relacionados à equidade e ao uso pedagógico dos resultados.

Referências

- ARELLANO, D. *et al.* **Sistemas de evaluación del desempeño para organizaciones públicas: ¿Cómo construirlos efectivamente?** México: CIDE, 2012.
- BROOKE, N.; CUNHA, M. A avaliação externa como instrumento de regulação da educação básica no Brasil. **Cadernos de Pesquisa**, São Paulo, v. 38, n. 133, p. 573–599, 2008.
- COLEMAN, J. *et al.* **Equality of Educational Opportunity**. Washington: U.S. Government Printing Office, 1966.
- FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. Qualidade da educação: avaliação, indicadores e metas. *In*: VELOSO, Fernando; PESSÔA, Samuel; HENRIQUES, Ricardo; GIAMBIAGI, Fábio (orgs.). **Educação básica no Brasil: construindo o país do futuro**. Rio de Janeiro: Elsevier, 2009.
- HILTON, M.; PELLEGRINO, J. **Education for life and work: developing transferable knowledge and skills in the 21st century**. Washington: National Academies Press, 2012.
- INEP. **Nota técnica: metodologia utilizada para o estabelecimento das metas intermediárias para a trajetória do Ideb no Brasil, Estados, Municípios e Escolas**. Brasília, DF: INEP, [s. d.].a. Disponível em: http://portalideb.inep.gov.br/index.php?option=com_content&view=article&id=40&Itemid=58. Acesso em: 10 abr. 2025.
- INEP. **Índice de Desenvolvimento da Educação Básica – Ideb: nota técnica**. Brasília, DF: INEP, [s. d.].b. Disponível em: https://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf. Acesso em: 9 abr. 2025.

KORETZ, D. **Measuring Up**: what educational testing really tells us. Cambridge: Harvard University Press, 2008.

MELLO, G. N. **Avaliação educacional**: indicadores de rendimento escolar. São Paulo: Moderna, 2013.

PONTES, L. A. F. **Medidas de eficácia escolar no contexto das políticas brasileiras de responsabilização educacional**: o caso do Índice de Desenvolvimento da Educação Básica, o Ideb, em Minas Gerais. 2015. 243 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal de Juiz de Fora, Juiz de Fora, 2015.

RAUDENBUSH, S.; WILLMS, J. D. The estimation of school effects. **Journal of Educational and Behavioral Statistics**, [S. l.], v. 20, n. 4, p. 307–335, 1995.

REIMERS, F. **Assessing education reform**: a framework for research. Washington: IADB, 2008.

SHAW, I. Evaluation and the postmodern dilemma. **Evaluation**, [S. l.], v. 12, n. 2, p. 203–214, 2006.

SOUSA, S. M. Z. L.; OLIVEIRA, R. P. D. **Sistemas de avaliação educacional no Brasil**: características, tendências e uso dos resultados. São Paulo: Universidade de São Paulo, Faculdade de Educação; Centro de Estudos e Pesquisas em Políticas Públicas de Educação (CEPPPE), 2007.

TEDDLIE, C.; REYNOLDS, D.; SAMMONS, P. **The international handbook of school effectiveness research**. London: Falmer Press, 2000.

WEISS, C. **Evaluation**: methods for studying programs and policies. 2. ed. Upper Saddle River: Prentice Hall, 1997.

CAPÍTULO 4

CONCEPÇÕES EDUCACIONAIS DA OCDE E A SUA INFLUÊNCIA POR MEIO DO PISA NA AMÉRICA DO SUL

Talita Emidio Andrade Soares¹

Denilson Junio Marques Soares²

Wagner dos Santos³

DOI: 10.52695/978-65-5456-163-1.4

Introdução

Em 14 de dezembro de 1960, ocorreu a assinatura da convenção que estabeleceu a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), uma organização que rapidamente se consolidaria como uma força de grande influência internacional, inclusive no campo educacional. Desde sua criação, a OCDE declara como objetivo principal a promoção do bem-estar mundial, buscando “orientar” os governos por meio de análises e recomendações políticas fundamentadas em evidências, padrões e redes de políticas globais (OCDE, 2024).

1 Instituto Federal de Minas Gerais. E-mail: talita.soares@ifmg.edu.br.

2 Instituto Federal de Minas Gerais.

3 Universidade Federal do Espírito Santo.

Nesse contexto, a educação é instrumentalizada como um meio para atender aos objetivos estratégicos da organização. Com a crescente demanda por dados internacionais comparáveis sobre desempenho escolar, a OCDE lançou o Programa Internacional de Avaliação de Alunos (Pisa), conhecido em inglês como *Programme for International Student Assessment*.

Realizado a cada três anos, o objetivo do Pisa é avaliar o conhecimento e as habilidades dos estudantes nas áreas de literatura, matemática e ciências por meio de testes cognitivos que utilizam uma métrica padronizada, acordada internacionalmente pelos países membros da OCDE, permitindo a produção de dados comparáveis entre diferentes sistemas educacionais. De acordo com o *Pisa 2000 Technical Report* (OCDE, 2002), as análises comparativas internacionais têm o potencial de “auxiliar” na interpretação dos resultados nacionais, oferecendo subsídios para a formulação de políticas educacionais, como aquelas voltadas ao currículo escolar e às práticas de ensino e aprendizagem (OCDE, 2002).

No contexto sul-americano, os sistemas de avaliação educacional emergiram na década de 1970, em meio à ascensão do neoliberalismo e a uma crise econômica global. Esse período foi marcado pela consolidação de uma cultura política orientada para resultados quantitativos. Desde então, impulsionada pela OCDE e outras organizações multilaterais, a avaliação consolidou-se como um elemento central das políticas educacionais implementadas no continente.

Assim, além das avaliações nacionais, diversos países do continente também aderiram a iniciativas internacionais, como o Pisa, cujos objetivos incluem:

[...] a) Avaliar conhecimentos e habilidades que são necessários em situações da vida real; [...] b) Relacionar diretamente o desempenho dos alunos a temas de políticas públicas; [...] c) Permitir o monitoramento regular dos padrões de desempenho (OCDE, 2001a).

No contexto apresentado, questionamos: qual é o discurso central que permeia essa participação? Quais concepções educacionais são defendidas promovidas pela OCDE por meio do Pisa?

Considerando o exposto, o objetivo deste artigo é analisar a participação dos países sul-americanos no Pisa e as concepções educacionais da OCDE, buscando compreender suas influências na formação de uma *cultura política* (Berstein, 1998) centrada em avaliações externas aplicadas em larga escala. Para isso, abrangemos o período de 2000 a 2024, durante o qual ocorreram os processos avaliativos do programa. Inicialmente, apresentamos uma seção dedicada a descrever, de forma geral, a metodologia empregada. Em seguida, exploramos as concepções educacionais defendidas pela OCDE por meio do Pisa. O capítulo é concluído com as considerações finais.

Metodologia

O presente trabalho adota uma metodologia de natureza qualitativa e do tipo exploratória, assumindo a pesquisa bibliográfica e o paradigma indiciário (Ginzburg, 2002) como abordagens teórico-metodológicas. Para atingir os objetivos propostos, este estudo também é conduzido com base nos conceitos de *usos e consumo* (Certeau, 2014) e *cultura política* (Berstein, 1998). Utiliza como fontes as publicações acadêmicas relacionadas ao Pisa e os documentos oficiais produzidos pela OCDE, disponibilizados, na íntegra, na *internet*.

Desse modo, efetua um levantamento das concepções educacionais defendidas pela OCDE, destacando os princípios propagados por meio dos relatórios do Pisa. Para isso, considera o período a partir do final da década de 1990, momento em que ocorreram os primeiros processos avaliativos do programa. Essas fontes são consideradas privilegiadas para fornecer informações sobre o desenvolvimento de uma *cultura política* (Berstein, 1998) avaliativa nos países sul-americanos e sobre como ela é perpetuada no continente. Os documentos estão descritos a seguir no quadro 1.

Quadro 1 – Documentos Normativos

Ano	Documento
1998	<i>Education at a Glance - OECD Indicators</i>
1999	Política y Promoción de la Inversión Extranjera Directa en América Latina
	<i>Measuring Student Knowledge and Skills: A New Framework for Assessment</i>
2000	<i>Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy</i>
2001	Pisa 2000 Relatório Nacional
2004	<i>What Makes School Systems Perform?</i>
	<i>Pisa 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills</i>
2008	<i>Latin American Economic Outlook 2008</i>
2009	<i>Iberoamérica en Pisa 2006 Informe regional</i>
	<i>Latin American Economic Outlook 2009</i>
2010	Latin American Economic Outlook 2011: How Middle-Class Is Latin American?
2014	<i>Latin American Economic Outlook 2015: Education, Skills and Innovation for Development</i>
2015	<i>The Experience of Middle-Income Countries Participating in PISA 2000-2015</i>
2016	Fomentando o Crescimento Inclusivo da Produtividade na América Latina
2017	<i>Activos con América Latina y el Caribe</i>
2018	<i>Competencias en Iberoamérica: Análisis de PISA 2015</i>
2021	How's Life in Latin America? Measuring Well-being for Policy Making
2023	<i>Latin American Economic Outlook 2023: Investing in Sustainable Development</i>

Fonte: Elaboração dos autores.

Pretende-se, assim, buscar “pistas”, “indícios”, “vestígios” (Ginzburg, 2002) que ajudem a compreender os interesses defendidos pela OCDE por meio do Pisa e suas possíveis implicações para o desenvolvimento da educação nos países participantes. Também se esforça em entender como as condutas não ditas oficialmente podem revelar a influência e/ou imposição de organismos multilaterais e outras instâncias no seu processo de origem.

Concepções educacionais da OCDE e a sua influência por meio do Pisa

Ao analisar os relatórios selecionados como fontes, identificamos *indícios* (Ginzburg, 2002) de que os princípios educacionais propagados pela OCDE estão fundamentados na responsabilização (do inglês, *accountability*), inclusive na América do Sul. Observamos que a organização defende o estabelecimento de normas internacionais por um grupo multilateral, a privatização, a prática de responsabilização e a comparação entre sistemas — características que remontam ao modelo *accountability*. Diante disso, apresentaremos alguns trechos extraídos dos relatórios originais para dar suporte a nossa análise.

Aderindo a um grupo multilateral

No primeiro relatório encontrado sobre a América Latina, intitulado *Política y Promoción de la Inversión Extranjera Directa en América Latina* (OCDE, 1999a) — visto que não foram encontrados relatórios específicos sobre a América do Sul —, a OCDE defende a adesão de um grupo multilateral para estabelecer normas internacionais em escala global. Conforme a OCDE (1999), a proposta é que esse grupo aborde questões de relevância mundial, que influenciam o investimento estrangeiro direto (IED), abrangendo leis, regulamentações e práticas administrativas em todos os níveis governamentais.

Si se concibe sobre esta base, un marco equilibrado de reglas internacionales trataría de las cuestiones clave relativas a la mundialización y garantizaría el respeto de las prerrogativas de política gubernamentales. Juntamente con otros instrumentos de cooperación internacional en ámbitos como la competencia, la tributación, el buen gobierno de los corporativos, las normas laborales y ambientales y las medidas contra la corrupción, ofrecería una asociación mundial para los países desarrollados y en desarrollo y complementaría

los acuerdos bilaterales, regionales y sectoriales (OCDE, 1999a, p. 36).

Embora o relatório destaque algumas implicações negativas, como a possibilidade de investidores estrangeiros usarem seu poder de negociação para contornar as exigências dos países anfitriões ou implementarem estratégias que possam prejudicar a capacidade dos governos, a adesão ainda é incentivada pela OCDE.

Vale destacar que a abordagem multilateral defendida pela OCDE não se restringe ao campo econômico. O discurso aplicado aos mercados é análogo ao que está sendo utilizado na educação, sustentando que, ao comparar as experiências nacionais com as de outros países do grupo, compartilham-se políticas e práticas consideradas “bem-sucedidas”. Essas comparações são feitas com base no desempenho escolar dos países em avaliações externas realizadas em grande escala, como o Pisa.

The assessment being carried out in 2000 [...] represents the fruition of a long collaborative process. PISA is a highly ambitious venture. Its aim is to make an authentic assessment of how well students are equipped for the future, in a way that is valid across many different cultures and languages. It has brought together scientific expertise from a wide range of participating countries, and has worked in association with governments to produce a survey that will be relevant for policy making (OCDE, 2000, p. 16).

Os primeiros documentos da OCDE sobre o programa, elaborados com o objetivo de obter informações educacionais de maneira rápida sobre os países (OCDE, 2004b), detalham a metodologia e as estruturas de análise, além de apresentarem argumentos em defesa da consolidação desse grupo. O principal argumento está fundamentado no próprio ideal do Pisa, que requer um “esforço colaborativo” (OCDE, 2001b), reunindo a *expertise* dos países participantes e sendo conduzido de maneira “conjunta” pelos governos, com base em interesses compartilhados (OCDE, 1999b).

No entanto, somos levados à seguinte reflexão: como essa colaboração acontece na prática? De que forma os países se organizam para participar do Pisa? Qual é o papel de cada um deles? Há diferenças significativas nesse processo?

O relatório *The Experience of Middle-Income Countries Participating in PISA 2000–2015* (Lockheed; Prokic–Bruer; Shadrova, 2015) aponta que a estratégia de parceria adotada no Pisa está fundamentada na adesão a um grupo multilateral no qual alguns países assumem a liderança sobre os demais. Com esse propósito, foi instituído um conselho composto por representantes de todos os países participantes, responsável por definir as prioridades em políticas educacionais do Pisa, alinhadas aos objetivos da OCDE (OCDE, 2004b). Além disso, cabe a esse conselho monitorar e garantir o cumprimento dessas prioridades ao longo do programa.

This includes setting priorities for the development of indicators, for the establishment of the assessment instruments and for the reporting of the results. Experts from participating countries also serve on working groups charged with linking the OECD/PISA policy objectives with the best internationally available technical expertise in the different assessment domains. By participating in these expert groups, countries ensure that the instruments are internationally valid and take into account the cultural and educational contexts in OECD Member countries. They also ensure that the assessment materials have strong measurement properties, and that the instruments emphasise authenticity and educational validity (OCDE, 2004b, p. 19).

Nesse contexto, é importante ressaltar a relevância do papel desempenhado pelos especialistas em avaliação educacional no desenvolvimento do Pisa, sendo responsáveis pela elaboração dos instrumentos e metodologias de pesquisa, além de conduzir os testes de campo e os estudos necessários para validá-los (Lockheed; Prokic–Bruer; Shadrova, 2015). Eles

também são responsáveis por definir o escopo e a natureza das avaliações (OCDE, 2004b), bem como por coletar informações de apoio e aprimorar os indicadores comparativos que permitem avaliar o desempenho dos sistemas educacionais (OCDE, 2001b).

[...] These indicators provide information on the human and financial resources invested in education, on how education and learning systems operate and evolve, and on the individual, social and economic returns from educational investment. In the past, the absence of regular and reliable indicators of education *outcomes* across countries, especially indicators of knowledge and skills, has been a significant gap in the available data. Without such indicators, policy-makers, taxpayers, educators and parents lack a means of judging the comparative effectiveness of their education systems (OCDE, 2001b, p. 17).

Nesse sentido, o relatório *Iberoamérica en PISA 2006: Informe regional* (OCDE, 2009a) enfatiza a experiência limitada dos países ibero-americanos⁴ no campo da avaliação educacional, apontando o “atraso” em relação a alguns países da OCDE, que já desenvolviam essa *expertise* desde meados do século XX. Como resultado dessa “defasagem”, os países ibero-americanos tiveram uma participação restrita nas duas primeiras edições do Pisa, não contribuindo para a elaboração dos estudos e instrumentos.

Además de muchos rasgos similares debidos a su herencia cultural, algunos de esos países tenían en común una reducida experiencia en el campo de la evaluación de aprendizajes mediante instrumentos estandarizados de aplicación a gran escala, que desde mediados del siglo xx eran ya usuales en algunos países de la OCDE, en especial del ámbito

4 Na ocasião participaram dois países europeus (Espanha e Portugal) e seis da América Latina (Argentina, Brasil, Chile, Colômbia, México e Uruguai).

anglosajón. Por esta razón, los países que habían participado en 2000 y 2003 lo habían hecho de forma limitada, sin intervenir en la preparación de los estudios ni en la elaboración de ítems o preguntas para la conformación de las pruebas y los cuestionarios que se aplican a alumnos y directores de escuela (OCDE, 2009a, p. 7).

O relatório também ressalta as dificuldades enfrentadas pelos países para atenderem às exigências da OCDE para a aplicação do Pisa.

[...] Las personas responsables de Pisa en cada uno de nuestros países, especialmente quienes participaban por vez primera en una aplicación, se enfrentaban además a dificultades considerables para atender adecuadamente a las precisas y múltiples exigencias prácticas que implica la aplicación de este tipo de instrumentos en muestras formadas por miles de alumnos, para que los resultados sean confiables y válidos (OCDE, 2009a, p. 7).

Por essa razão, em 2005, foi criado o Grupo Ibero-americano do Pisa (GIP), com o objetivo declarado de fomentar a cooperação, a reflexão e a assistência mútua entre os países ibero-americanos participantes do Pisa diante dos desafios científicos e técnicos envolvidos na avaliação. Mais uma vez, a iniciativa da OCDE busca promover um “trabalho colaborativo” em que os países com mais *expertise* apoiem aqueles com menos experiência (OCDE, 2009a).

Dessa maneira, conforme a OCDE (2009a), essa abordagem permitiria que todos os países “atendessem” aos padrões de qualidade do Pisa, enquanto desenvolviam sua própria capacidade técnica de forma mais ágil e com menor esforço do que seria necessário caso cada grupo nacional tentasse alcançar esses objetivos de maneira isolada. Com a criação do GIP, os países ibero-americanos inauguraram uma nova etapa de colaboração e, a partir de 2007, passaram a contribuir para a elaboração de itens do programa. Ou seja, esses países só puderam efetivamente “integrar-se”

ao processo quando foram capazes de atender plenamente às normativas estabelecidas pela OCDE.

A participação desse grupo, embora regional, reflete a abordagem multilateral promovida pela OCDE em escala global. Ademais, a própria organização reconhece a importância de períodos de transição, permitindo que esses países possam “ajustar” suas legislações nacionais para se alinharem às regras do quadro (OCDE, 1999a, 2009), tanto no âmbito econômico quanto no educacional.

Carvalho (2016) analisa as estratégias utilizadas pelas organizações internacionais para expandir seu alcance, não apenas difundindo ideias e normas sobre estruturas, processos e práticas educativas, mas também criando novos espaços e instrumentos para a governança da educação. Assim, conforme discutido pelo autor, o estabelecimento de um grupo multilateral para definir regras reflete uma das estratégias utilizadas pela OCDE para construir agendas direcionadas aos governos, legitimando-as em nome de um conhecimento especializado produzido a partir do Pisa.

Desse modo, as informações fornecidas pelo programa — que vão além do desempenho escolar nas áreas avaliadas e incluem dados contextuais — constituem o que Carvalho (2016) denomina de monopólio de conhecimento da OCDE. Para ter acesso a esses dados e às experiências compartilhadas pelos países do grupo, especialmente por aqueles que apresentaram maior crescimento nos índices, é necessário que os países sigam as normas estabelecidas e utilizem o Pisa como instrumento, conforme as diretrizes recomendadas pela organização. Esse cenário evidencia um primeiro *indício* (Ginzburg, 2002) de regulação exercida pela OCDE.

Promoção da responsabilização na educação

Embora não utilize explicitamente o termo, a OCDE desenvolve cuidadosamente o conceito de responsabilização em seus relatórios. Buscando alinhar as práticas educacionais com seus padrões e recomendações, a organização procura influenciar os países participantes a adotarem abordagens que promovam a responsabilização das escolas, transferindo responsabilidades anteriormente atribuídas ao Estado para essas instituições.

Isso está em consonância com o propósito declarado na primeira edição do relatório *Education at a Glance* (OCDE, 1998), que apresenta indicadores educacionais com o intuito de “orientar” o desenvolvimento de políticas e reforçar a responsabilidade pública nos sistemas educacionais.

Dessa forma, é possível afirmar que o interesse da OCDE por trás de seus indicadores vai além da simples mensuração, buscando promover a adoção da cultura política que ela defende, incluindo normas, valores e dinâmicas político-econômicas específicas, como discutido no item anterior. Para além das diretrizes obrigatórias para a participação no Pisa, a organização difunde essa *cultura política* (Berstein, 1998) por meio de relatórios amplamente divulgados, fundamentados em argumentos quantitativos e assinados por especialistas, conferindo-lhes uma aparência de incontestabilidade.

Em 2004, por exemplo, a OCDE publicou o relatório *What Makes School Systems Perform? Seeing School Systems through the Prism of Pisa* (OCDE, 2004a), que, embora não fosse especificamente voltado para América Latina ou América do Sul, relaciona os resultados do Pisa 2000 (primeira edição) a evidências qualitativas sobre políticas públicas implementadas em seis países com bom desempenho. Vale destacar que nenhum dos países mencionados era latino-americano.⁵

Uma das questões discutidas no relatório diz respeito à ampliação da autonomia das escolas, destacando que essa medida visa elevar os níveis de desempenho dos estudantes e atender às necessidades locais. Os resultados apresentados parecem reforçar essa perspectiva, mostrando que, na maioria dos países com bom desempenho no Pisa 2000, as escolas e autoridades locais dispõem de significativa liberdade para adaptar e implementar conteúdos educacionais, além de gerenciar recursos de maneira independente (OCDE, 2004a).

Ampliar a autonomia das escolas significa transferir para as próprias instituições educacionais parte da responsabilidade que antes era atribuí-

5 Os países abordados foram Canadá, Inglaterra, Finlândia, França, Holanda e Suécia.

da ao Estado. Essa reforma envolve estratégias que incluem mudanças na governança e na alocação de recursos, o estabelecimento de abordagens nacionais para definir padrões, avaliar e monitorar tanto os sistemas educacionais quanto o desempenho dos professores, além da implementação de medidas destinadas a reduzir as desigualdades socioeconômicas entre os alunos (OCDE, 2004a).

Nesse sentido, Carvalho (2016) discute a intencionalidade presente nos relatórios da OCDE e o modo como a organização “autoriza a si mesma”, já que as análises e publicações são elaboradas por equipes e entidades que as validam internamente e de forma coletiva. Nota-se, entretanto, que esses relatórios não passam por avaliações de qualidade do conteúdo em revistas especializadas nem por um processo de revisão por pares. Isso nos leva a pensar: até que ponto o conhecimento produzido pela OCDE realmente reflete interesses globais e não apenas os de um grupo restrito?

O autor caracteriza esse processo de produção de conhecimento do Pisa como uma “regulação transnacional”, uma vez que tanto o ponto de partida quanto o ponto de chegada são parcialmente definidos pelos representantes nacionais, incluindo atores do setor privado, e não exclusivamente por pesquisadores da área. Dessa forma, essa regulação influencia as relações entre os “produtores de conhecimento” e as políticas educacionais, moldando o direcionamento e aplicação desse conhecimento, como nas práticas de responsabilização das escolas.

[...] por um lado, as autoridades públicas outorgam o estatuto de *experts* a certos atores sociais, bem como selecionam e determinam que conhecimento é válido para ser usado no campo político; por outro lado, as autoridades aceitam que esses putativos especialistas sejam parceiros-chave no processo de construção de representações e de regras para coordenação, orientação e controle do sector educativo; finalmente, as relações entre as autoridades públicas e os peritos nacionais e internacionais são geradas e geridas sob a égide da OCDE, a qual, depois de reunir essa variedade

de atores em torno do estudo de um objeto singular, cria condições para que esses interajam regularmente, fixa e arbitra as relações entre esses diferentes mundos, de acordo com valores e regras por si determinadas (Carvalho, 2016, p. 673).

Em resumo, ao confiarem no conhecimento de especialistas renomados, frequentemente vinculados a institutos de pesquisa de prestígio, os governos partem do pressuposto de que esses especialistas têm autoridade para estabelecer prioridades educacionais e o conhecimento técnico para implementar políticas públicas. Contudo, na produção desse conhecimento, os especialistas baseiam-se nos dados do Pisa, aceitos sem questionamentos e conformados às regras impostas pelo grupo para participação. Assim, ao final do processo, os governos acabam sendo orientados pela OCDE.

Ademais, Carvalho (2016) chama atenção para uma iniciativa do programa que exemplifica essa orientação direcionada: o *PISA-based Test for Schools* (Pisa para a Escola). Conforme observado pelo autor, essa iniciativa estabelece uma conexão direta com as instituições de ensino, dispensando a mediação das autoridades nacionais. Assim, podemos afirmar que essa prática representa, portanto, um segundo *indício* (Ginzburg, 2002) de regulação exercida pela OCDE.

Apoiando a privatização

Quando se trata de governança e alocação de recursos, a educação é vista como um “investimento” no desenvolvimento das habilidades humanas, com potencial para promover o crescimento econômico, aumentar a produtividade, favorecer o desenvolvimento pessoal e social, além de reduzir as desigualdades sociais (OCDE, 1998, 2014, 2016). Tal como qualquer outro investimento, envolve tanto “custos” quanto “retornos”. Nesse cenário, grupos interessados em informações educacionais, como pais/responsáveis, autoridades administrativas e a sociedade em geral, passam a ser vistos como “clientes”.

With increased participation drawing from new groups of clients and a wider range of choices on what, when, how and where to learn, governments are forging new partnerships to mobilise the necessary resources, to encourage efficiency and to introduce flexibility to permit everyone to pursue the pathways and learning opportunities which best meet their requirements. New policies are designed to allow the different actors and stakeholders to participate more fully and to share the costs and benefits of education more equitably. New financing strategies aim also at influencing student behaviour in ways that make education more cost-effective (OCDE, 1998, p. 68).

O relatório *Education at a Glance* (OCDE, 1998) destaca a necessidade de estabelecer novas parcerias e políticas, visando não apenas à mobilização de recursos pelo setor público, mas também ao envolvimento de diversos atores e partes interessadas, promovendo o “compartilhamento” dos custos. O documento enfatiza que, embora o financiamento público responda por uma parcela significativa do “investimento” em educação, as fontes privadas têm assumido um papel cada vez mais relevante nesse contexto.

Nesse sentido, o relatório *Política y Promoción de la Inversión Extranjera Directa en América Latina* (OCDE, 1999a) oferece *indícios* (Ginzburg, 2002) da postura pró-privatização defendida pela organização. A privatização é apresentada como um “estímulo” ao crescimento econômico dos países, especialmente por meio do aumento do IED. O relatório aponta para a expectativa de continuidade dessa tendência, proporcionando às empresas acesso a novos mercados, recursos estratégicos, redução de custos e maior eficiência, entre outros benefícios (OCDE, 1999a).⁶

6 Por exemplo, no âmbito econômico, o relatório menciona a privatização da empresa brasileira Telebrás, que foi vendida por um total de 19 bilhões de dólares, sendo 12 bilhões provenientes de investidores estrangeiros.

No relatório *Activos con América Latina y el Caribe* (OCDE, 2017), também encontramos *pistas* (Ginzburg, 2002) do apoio da OCDE a essa abordagem conhecida como “Estado aberto”. A organização defende que, em um cenário de recursos públicos limitados, é necessário estabelecer uma governança eficiente, com regulamentações eficazes, voltadas para estimular o crescimento, o investimento e a inovação. De acordo com a OCDE (2017), essa estratégia visa atingir metas sociais, econômicas e ambientais na região da América Latina e do Caribe (ALC).

ALC es una de las regiones del mundo más urbanizada; sin embargo, es necesario mejorar la calidad de la urbanización. Promover unas ciudades y regiones más competitivas, sostenibles e integradoras exigirá una formulación de políticas efectiva a todas las escalas de la Administración, con el fin de garantizar que todas esas ciudades y regiones puedan movilizar su potencial de crecimiento. Mediante el Programa regional de la OCDE para ALC, la Organización trabaja para consolidar y expandir su asociación con las autoridades a todas las escalas de la Administración en la región de ALC, con el fin de asistirles en la ejecución de las reformas de descentralización, y de reforzar las capacidades de gobernanza nacional y subnacional para diseñar estrategias de desarrollo más efectivas (OCDE, 2017, p. 64).

O relatório *Perspectivas Econômicas da América Latina 2009* (OCDE, 2009b) analisa o “gasto” público em educação, apontando que, em comparação com padrões internacionais, o continente apresenta um nível considerado de investimento.⁷ No entanto, ao reconhecer a demografia como um desafio relevante para esses países, o documento chama atenção para a

7 O relatório da OCDE (2008) cita o Brasil como exemplo, indicando que o país arrecada e gasta consideravelmente, mas essa abundância não implica em melhores resultados. O Brasil não se destaca positivamente entre os países da América Latina e, de acordo com os padrões da organização, é avaliado de forma desfavorável.

disparidade observada no gasto por aluno. Esses argumentos são utilizados para reforçar a perspectiva de que os recursos públicos alocados à educação podem ser insuficientes para atender às necessidades específicas da América Latina.

Nessa perspectiva, o relatório *Fomentando o Crescimento Inclusivo da Produtividade na América Latina* (OCDE, 2016) delinea as prioridades de reforma para o continente, enfatizando a necessidade de criar ambientes de negócios sólidos, abertos e competitivos. A OCDE argumenta que ainda existe espaço para a implementação de políticas de liberalização em determinados mercados e, para estimular a concorrência, recomenda “[...] a remoção ou reformulação das regulações ou outras formas de intervenção estatal que restringem desnecessariamente a concorrência, especialmente as que dificultam a entrada dos recém-chegados” (OCDE, 2016, p. 10). O relatório também aponta que as restrições nos mercados de produtos na América Latina são, em geral, mais rígidas do que nos países membros da OCDE.

Em relação ao Pisa, o relatório *The Experience of Middle-Income Countries Participating in PISA 2000–2015* (Lockheed; Prokic–Bruer; Shadrova, 2015) destaca que os países participantes do programa podem contar com o “apoio” de agências de cooperação internacional, como o Banco Mundial, com o qual as autoras possuem vínculos. Essas agências fornecem suporte financeiro e técnico para o desenvolvimento das pesquisas, mas condicionam esse apoio ao cumprimento de determinadas exigências estabelecidas para a sua realização. Como resultado, a educação acaba sendo orientada pelas prescrições e normativas do economicismo (Ball, 2001).

Within each partner country, bilateral and multilateral development co-operation agencies will provide both financial and technical support for the initiative, including assistance in developing the survey instruments and methodologies. The results of the survey can be used to inform their own policies and practices (Lockheed; Prokic–Bruer; Shadrova, 2015, p. 4).

Lockheed, Prokic-Bruer e Shadrova (2015) ressaltam que o financiamento externo é “facilitador” para a participação de países de renda média ou baixa em avaliações internacionais. As autoras apontam que esse financiamento pode ocorrer por meio de doações diretas ou concessão de empréstimos voltados para projetos educacionais. Além disso, destacam que organizações internacionais, como o Banco Mundial, utilizam os dados do Pisa para monitorar os resultados desses projetos.⁸ Vale mencionar que quase metade dos países de renda baixa e média que participaram do Pisa 2015 contou com financiamento externo.

A descentralização do Estado é defendida pela OCDE com o argumento de que, ao distribuir os “custos” da educação por meio da participação de instituições privadas, seria possível promover mais transparência, prestação de contas e tomada de decisões sobre políticas educacionais (OCDE, 2008, 2017).

Ball (2001, p. 111) argumenta que, nessa nova configuração, o Estado deixa de atuar como provedor para assumir papel de regulador, “[...] aquele que estabelece as condições através das quais se autoriza os vários mercados internos a operar”. O autor destaca que isso configura “[...] processos de re-regulação; representam não propriamente o abandono por parte do Estado dos seus mecanismos de controle, mas sim o estabelecimento de nova forma de controle” (Ball, 2001, p. 104).

Para Ball (2001), esses processos de reforma não se limitam à introdução de novas estruturas e incentivos, mas também implicam na criação de novas relações, culturas e valores. Assim, para alcançar os objetivos alinhados às estratégias de reforma da OCDE, são promovidas abordagens nacionais que abrangem a definição de padrões, a avaliação e o mo-

8 Conforme Lockheed, Prokic-Bruer e Shadrova (2015), entre 2003 e 2012, as atividades de avaliação internacional em projetos do Banco Mundial foram realizadas em 11 países que participaram do Pisa: Azerbaijão, Brasil, Bulgária, Indonésia, Jordânia, Quirguistão, Moldávia, Montenegro, Peru, Tunísia e Turquia. Além disso, o Banco Mundial também apoiou avaliações internacionais no mesmo período em outros 11 países que não eram participantes do Pisa: Bangladesh, El Salvador, Gâmbia, Honduras, Líbano, Mongólia, Nepal, Nicarágua, Santa Lúcia, Senegal e Ucrânia.

nitoramento tanto do sistema educacional (OCDE, 2008) quanto do desempenho dos professores (OCDE, 2016). Nesse cenário, segundo o autor, consolida-se o Estado auditor, responsável por avaliar os resultados educacionais, utilizando a prática da comparação como ferramenta central.

Prática de comparação

Como está a preparação dos jovens para enfrentar os desafios do futuro? Eles são capazes de analisar, raciocinar e comunicar suas ideias de forma eficaz? Eles possuem a capacidade de continuar aprendendo ao longo da vida? Conforme descrito no relatório *Measuring Student Knowledge and Skills: A New Framework for Assessment* (OCDE, 1999b), essas informações são consideradas essenciais para pais/responsáveis, estudantes, o público em geral e gestores educacionais. A busca por respostas a essas questões é vista como crucial, e a realização de análises internacionais comparativas é apresentada como uma estratégia para ampliar e enriquecer a compreensão nacional, oferecendo um contexto mais abrangente para interpretar os resultados educacionais.

[...] Las evaluaciones internacionales pueden mostrar qué niveles de calidad y equidad son alcanzados por los sistemas educativos más avanzados. Pueden servir de pauta del ritmo de progreso educativo y ayudar a valorar, desde las posiciones más avanzadas, la realidad de la educación ofrecida. Por último, aunque no menos importante, las evaluaciones internacionales permiten que los sistemas educativos se valoren a sí mismos desde la óptica de las políticas planificadas, implementadas y realmente conseguidas en otros lugares, y compararlas con los esfuerzos nacionales que se realizan para ayudar a los alumnos a que aprendan mejor, a los docentes a que enseñen mejor y a las escuelas a que sean más efectivas (OCDE, 2009a, p. 11).

Acredita-se que uma análise quantitativa do desempenho dos sistemas educacionais permite que os países sejam avaliados comparativamente

com outras nações, identificando suas “fraquezas” e reconhecendo seus pontos fortes para possíveis melhorias (OCDE, 1998). Conforme argumentado por Ball (2011), a aprendizagem é redefinida como o “produto final de políticas custo-efetivas”, enquanto as realizações educacionais passam a ser interpretadas como um conjunto de “metas de produtividade”.

A quantitative description of the functioning of education systems allows countries to see themselves in the light of other countries' performance. Through international comparisons, countries may come to recognise weaknesses in their own education systems, while also identifying strengths that can otherwise be ignored in the heat of domestic debate. The OECD education indicators show whether variations in educational experiences within a country are unique or if they mirror differences observed elsewhere (OCDE, 1998, p. 5).

Defende-se, portanto, a utilização do Pisa como uma ferramenta capaz de monitorar o progresso do desempenho escolar em diferentes países (OCDE, 1999b). O relatório *The Experience of Middle-Income Countries Participating in PISA 2000-2015* (Lockheed; Prokic-Bruer; Shadrova, 2015) ressalta o interesse da OCDE em consolidar o Pisa como uma referência global de sucesso educacional. Assim, a organização não apenas conduz as pesquisas do programa, mas também oferece “recomendações” para a implementação de políticas públicas nos países.

The OECD remains committed to maintain and develop PISA as a global yardstick for educational success. It will continue to contribute its expertise and its platforms to support collaboration for developing and conducting PISA surveys and to assist policy-makers and practitioners throughout the world to use them most productively (Lockheed; Prokic-Bruer; Shadrova, 2015, p. 4).

No entanto, é essencial cautela ao interpretar e comparar os resultados. O relatório *Knowledge and Skills for Life: First Results from PISA*

2000 (OCDE, 2001b) alerta que diferenças nos resultados entre países não indicam, necessariamente, mais eficácia das escolas. A aprendizagem ocorre em diversos contextos, incluindo ambientes institucionais e extra-curriculares, e tem início antes mesmo da entrada do estudante na escola. Contudo, o mesmo relatório indica que o impacto cumulativo de todas as experiências de aprendizagem em determinado país pode ser associado a desempenhos mais favoráveis nos aspectos avaliados pelo programa. Dessa forma, sugere-se que os demais países possam “aprender” com aqueles que alcançaram resultados estatisticamente significativos, aproveitando essas experiências para aprimorar suas próprias práticas.

Com base na discussão levantada, vale refletir: será que comparar sistemas educacionais tão distintos realmente nos oferece *insights* relevantes ou estamos simplificando questões complexas demais? Existe alguma evidência sólida de que as reformas propostas pela OCDE tenham melhorado os sistemas educacionais nos países participantes? E, além disso, como essas comparações estão sendo feitas e quem, de fato, é beneficiado por elas?

Nesse contexto, a OCDE, mais uma vez, promove comparações entre a América Latina e seus países membros, destacando as disparidades existentes. O relatório *Latin American Economic Outlook 2015: Education, Skills and Innovation for Development* (OCDE, 2014) enfatiza que um dos principais desafios ainda pendentes no continente é ampliar o impacto dos investimentos em educação na melhoria da qualidade e do desempenho dos sistemas educacionais.

Brazil and Mexico sit alongside Tunisia and Turkey as the countries that have made the most progress in secondary-education performance (15-16 year-olds), increasing their scores in the OECD's PISA tests (Programme for International Student Assessment) by between three and four points a year since 2003. Nevertheless, there is still a huge gulf between the performances of Latin American students and those in OECD countries, amounting to almost two years of schooling according to the 2012 PISA tests (OCDE, 2014, p. 23).

No relatório, embora se reconheça o progresso no desempenho escolar dos estudantes latino-americanos no Pisa, ainda é apontado um “atraso” significativo em relação aos países da OCDE, equivalente a quase 2 anos de escolaridade na edição do Pisa de 2012. Na edição mais recente, intitulada *Latin American Economic Outlook 2023: Investing in Sustainable Development* (OCDE, 2023), destaca-se que as condições socioeconômicas permanecem desafiadoras na região da ALC.

Esse “atraso” dos países latino-americanos também se reflete em relatórios cujo foco principal não é a educação. No *Latin American Economic Outlook 2011: How Middle-Class Is Latin America?* (OCDE, 2010), embora a análise econômica esteja centrada na classe média da região, a educação é abordada em um capítulo como um fator capaz de promover a mobilidade social,⁹ ressaltando a relevância do Pisa na compreensão dos contextos socioeconômicos. O relatório aponta que “[...] levels of educational attainment are converging over time, but the middle sector remains closer to the disadvantaged than the affluent” (OCDE, 2010, p. 122).

No documento *Fomentando o Crescimento Inclusivo da Produtividade na América Latina* (OCDE, 2016), também são apresentados argumentos que ressaltam as fragilidades do continente. O texto observa que, “[...] apesar das melhorias nos resultados da educação, os países latino-americanos estão atrasados em comparação à média da OCDE” (OCDE, 2016, p. 5). Além das questões educacionais, aborda o atraso em áreas como mercados, pesquisa e inovação, destacando, ainda, que a América Latina apresenta a maior desigualdade de renda do mundo. Em resposta a esses desafios, o documento afirma que “[...] a OCDE está pronta para trabalhar com os governos da América Latina, para desenvolver e oferecer melhores políticas para uma vida melhor em toda a região” (OCDE, 2016, p. 5).

As conclusões para a Iberoamérica seguem uma linha semelhante às observadas para a América Latina. O relatório *Competencias en Iberoamérica: Análisis de PISA 2015* (OCDE, 2018) inicia descrevendo a região

9 De acordo com a OCDE (2010), mobilidade social é a capacidade dos indivíduos e suas famílias de melhorar seu *status* econômico e social ao longo do tempo.

como uma terra de “potencial desperdiçado”. O objetivo do documento é destacar sinais “encorajadores” de transformação no campo educacional, fundamentados nos resultados obtidos pelo Pisa 2015.

Los resultados del Programa para la Evaluación Internacional de Alumnos (PISA) de 2015 plantean cuestiones de calidad y equidad en la educación en Iberoamérica. Además de los miembros de la OCDE: Chile, México, Portugal y España, siete países iberoamericanos asociados participaron en PISA 2015: Argentina, Brasil, Colombia, Costa Rica, República Dominicana, Perú y Uruguay. Mientras que España y Portugal obtuvieron resultados cercanos al promedio de la OCDE, todos los países latinoamericanos rindieron sistemáticamente por debajo de ese promedio en las tres materias: lectura, matemáticas y ciencias (OCDE, 2018, p. 27).

De forma semelhante, embora reconheça avanços como o aumento no número de matrículas, o relatório aponta que “[...] la región se mantiene rezagada en el desempeño educativo general” (OCDE, 2018, p. 33) e que “[...] el desempeño de la región ha sido insatisfactorio en lo que atañe a la mejora de los resultados educativos y a la consecución del mayor grado de competencia de su población” (OCDE, 2018, p. 34). Esses trechos reforçam as dificuldades persistentes na região em termos de desempenho e qualidade educacional.

Entre as recomendações para a Iberoamérica, destacam-se a necessidade de garantir o acesso ao e a permanência dos estudantes no sistema educacional por períodos mais longos, bem como a melhoria da qualidade do ensino e da gestão escolar. Para atingir esses objetivos, o relatório enfatiza a importância de “assegurar a consistência dos indicadores” ao longo do tempo, permitindo “comparações confiáveis” entre os países participantes do Pisa.

Dessa forma, compreendemos que a intenção da OCDE é estimular que os países considerados menos desenvolvidos, como os latino-americanos, ibero-americanos e sul-americanos, façam uso (Certeau, 2014)

do Pisa, apropriando-se de seus indicadores, diretrizes e recursos. Em outras palavras, conforme discutido por Ball (2001), busca-se a disseminação de uma cultura de desempenho competitivo, baseada na combinação de devolução, metas e incentivos, com o objetivo de promover uma reconfiguração institucional.

[...] Mercados e sistemas de competição e escolha através dos quais eles operam, re-processam os seus próprios atores-chave – no nosso caso, famílias, crianças e professores/as – e exigem que as escolas assumam novos tipos de preocupações extrínsecas e, por consequência, re-configurem e re-valorizem o significado da educação (Ball, 2001, p. 107).

De acordo com o autor, o estudante é cada vez mais mercantilizado, sendo posicionado conforme os resultados obtidos em diferentes avaliações, como o Pisa. O desempenho, tanto individual quanto das escolas, passa a ser tratado como uma medida de produtividade ou resultado, servindo como uma exposição de “qualidade”. Por isso, o autor destaca a importância de refletir sobre quem exerce o controle no campo da avaliação.

Portanto, a atenção direcionada às informações divulgadas pelo programa — tanto no que diz respeito ao contexto quanto ao desempenho escolar — e à posição ocupada nos *rankings* promovidos pela OCDE, configura-se como uma forma de *consumo* (Certeau, 2014). Esse consumo reflete a produção resultante, ou seja, as ações e iniciativas realizadas por esses governos/países/escolas a partir do processo de apropriação.

Dale (2004) descreve esse processo de apropriação como uma “Agenda Globalmente Estruturada para a Educação” e examina como essa nova forma de influência supranacional molda os sistemas educacionais nacionais, identificando o capitalismo, impulsionado pela busca do lucro, como a força motriz por trás desse fenômeno. O autor acrescenta que a relação com essas organizações deixa de ser uma “emulação voluntária” para um “ajustamento coercivo”.

Destarte, concordamos com Rodrigues e Ferrarotto (2024) ao observarem que, para a OCDE, a qualidade de um sistema educacional é percebida como proporcional ao seu nível de alinhamento, monitoramento e subordinação aos objetivos da organização, geralmente associados aos interesses dos reformadores empresariais.

Considerações finais

Este capítulo se propôs a analisar as concepções educacionais defendidas pela OCDE, destacando os princípios propagados por meio dos relatórios do Pisa. Para isso, foi realizado um levantamento das publicações acadêmicas relacionadas ao programa e dos documentos oficiais produzidos pela OCDE, disponibilizados, na íntegra, na *internet*.

Observamos que a *cultura política* (Berstein, 1998) centrada em avaliações externas nos países sul-americanos é amplamente influenciada pela OCDE e outras organizações multilaterais, como o Banco Mundial, por meio do Pisa. Essas avaliações são interpretadas pelos Estados como ferramentas de “prestação de contas” à sociedade, justificadas pelo argumento de promover mais transparência na alocação dos recursos públicos destinados à educação. Em teoria, espera-se que os resultados dessas avaliações contribuam para a autoavaliação das escolas e dos sistemas de ensino, o aperfeiçoamento de projetos educacionais, a realização de diagnósticos sobre a qualidade do ensino oferecido e a formulação de políticas educacionais, entre outros.

Notamos, ainda, que muitos relatórios internacionais da OCDE estão disponíveis exclusivamente em inglês e francês, mesmo quando abordam temas relacionados à América Latina ou à Iberoamérica. Quando traduzidos para outros idiomas, como o espanhol, essas publicações frequentemente ocorrem de forma tardia, com atrasos que podem se estender por anos. Ademais, ao se analisarem as assinaturas dos especialistas envolvidos nesses documentos, destaca-se a ausência de profissionais originários desses continentes.

Além disso, não encontramos relatórios específicos voltados exclusivamente para a América do Sul, onde os contextos dos países são mais

semelhantes. Isso nos leva a refletir sobre a influência que esses agentes podem exercer na interpretação dos resultados, tornando os países da região “reféns” da OCDE e de outras organizações. Ressaltamos, assim, a necessidade de estudos que abordem os contextos particulares da América do Sul de maneira aprofundada e contextualizada.

Também percebemos que a concepção educacional promovida pela OCDE é cuidadosamente construída em seus relatórios, defendendo o estabelecimento de normas internacionais por meio de um grupo multilateral, a privatização e a prática de responsabilização e a comparação entre sistemas — elementos fundamentais para a implementação de um sistema de *accountability*. Auld e Morris (2014) analisam como essa abordagem é justificada pela busca de comprovações sobre o que “funciona” em países ou sistemas educacionais de alto desempenho, com base em avaliações como o Pisa. Os autores denominam essa prática de “Novo Paradigma”, cujo propósito é identificar as estratégias de sucesso dessas nações para “aconselhar” outros países, auxiliando-os a melhorar seus resultados e, por extensão, a assegurar seu futuro econômico.

Auld e Morris (2014) também analisam como essas “evidências” comparativas têm sido usadas para iniciar ou legitimar padrões de empréstimo entre sistemas educacionais. No Pisa 2015, Lockheed, Prokic-Bruer e Shadrova (2015) constataram que quase metade dos países de renda baixa e média só conseguiu participar com o auxílio de financiamento externo. Esse fato ressalta os altos custos envolvidos na avaliação e também pode fornecer uma *pista* (Ginzburg, 2002) para justificar as dificuldades para cumprir os requisitos técnicos, como o tamanho da amostra e a *expertise* na área de avaliação educacional.

Além disso, a comparação entre países revela-se intencional, pois opta por comparar o desempenho de países membros com não membros, bem como entre os de diferentes continentes, sem considerar as especificidades do contexto socioeconômico de cada nação. Esse cenário reforça as conclusões de Pereira (2019), que, ao examinar as proposições educacionais da OCDE para a América Latina, identificou, que, mesmo antes de lançar um programa específico para a região, a organização já delineava uma política clara para esses países. No momento do lançamento, a OCDE já

apresentou um documento com recomendações detalhadas para melhorias no campo educacional. Isso nos leva a questionar: quem se beneficia com os resultados dessas comparações internacionais? Quais interesses estão em jogo? E como essas práticas influenciam as dinâmicas de poder e as relações de influência global?

Dessa forma, entendemos que as concepções educacionais defendidas pela OCDE, fundamentadas no *accountability*, constituem uma estratégia para ampliar e consolidar sua influência sobre as políticas educacionais globais. A análise revela que, ao estabelecer normas e padrões internacionais por meio do Pisa, a organização não apenas promove a comparação entre os sistemas educacionais, mas também estabelece uma lógica de dependência técnica e normativa, especialmente para os países de menor experiência no campo avaliativo, refletindo o que Carvalho (2016) denomina de “regulação transnacional”, em que o controle sobre as políticas educacionais é exercido por meio de normas e indicadores.

Embora apresentada como uma oportunidade de colaboração, essa forma de “governo a partir dos números” (Silva; Silva; Ferreira, 2022) reforça o papel da OCDE como um *centro* (Ginzburg, 1989) regulador, que dita diretrizes educacionais globais em nome de um conhecimento especializado. Nesse contexto, a produção de conhecimento, sustentada por dados aceitos sem contestação, reforça a dependência técnica dos governos em relação à organização, muitas vezes ignorando os contextos específicos de cada país, como os da América do Sul.

Para terminar, constatamos que o Pisa se configura como uma poderosa ferramenta para a OCDE consolidar o seu papel como agente regulador em um sistema educacional global cada vez mais pautado pelo economicismo, uma vez que o programa, ao mesmo tempo que fomenta a cooperação, perpetua relações de assimetria e de subordinação no campo educacional.

Por fim, esperamos que este artigo esclareça alguns aspectos das concepções educacionais da OCDE e sua influência por meio do Pisa na América do Sul, contribuindo como um instrumento para a disseminação de conhecimento sobre o programa para a promoção de discussões dessa avaliação no ambiente escolar e acadêmico.

Referências

AULD, Euan; MORRIS, Paul. Comparative education, the 'New Paradigm' and policy borrowing: Constructing knowledge for educational reform. **Comparative Education**, [S. l.], v. 50, n. 2, p. 129–155, 2014. DOI: <https://doi.org/10.1080/03050068.2013.826497>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/03050068.2013.826497>. Acesso em: 10 mai. 2024.

Ball, Stephen J. Diretrizes políticas globais e relações políticas locais em educação. **Currículo sem fronteiras**, [S. l.], v. 1, n. 2, p. 99–116, jul./dez. 2001. Disponível em: <https://www.curriculosemfronteiras.org/volliss2articles/ball.htm>. Acesso em: 15 dez. 2024.

BERSTEIN, Serge. A cultura política. In: RIOUX, Jean-Pierre; SIRINELLI, Jean-François (orgs.). **Para uma história cultural**. Lisboa: Editorial Estampa, 1998.

CARVALHO, Luís Miguel. Intensificação e sofisticação dos processos da regulação transnacional em educação: o caso do Programa Internacional de Avaliação de Estudantes. **Educação & Sociedade**, v. 37, n. 136, p. 669–683, 2016. DOI: <https://doi.org/10.1590/ES0101-73302016166669>. Disponível em: <https://www.scielo.br/j/es/a/ms5Rh69wtSB5k6m3pxP9hxp/?lang=pt>. Acesso em: 10 mai. 2024.

CERTEAU, Michel de. **A invenção do cotidiano**: 1. Artes de Fazer/Michel de Certeau; Tradução de Ephraim Ferreira Alves. 22. ed. Petrópolis: Vozes, 2014.

DALE, Roger. Globalização e educação: demonstrando a existência de uma "Cultura Educacional Mundial Comum" ou localizando uma "Agenda Globalmente Estruturada para a Educação"? **Educação & Sociedade**, v. 25, n. 87, p. 423–460, 2004. DOI: <https://doi.org/10.1590/S0101-73302004000200007>. Disponível em: <https://www.scielo.br/j/es/a/bJbBCJS5DvngSvwz9hngDXK/?lang=pt>. Acesso em: 10 mai. 2024.

GINZBURG, Carlo. **A micro-história e outros ensaios**. Rio de Janeiro: Difel: Bertrand Brasil, 1989.

GINZBURG, Carlo. **Mitos, emblemas, sinais**: morfologia e história. São Paulo: Companhia das Letras, 2002.

LOCKHEED, M.; PROKIC-BRUER, T.; SHADROVA, A. **The Experience of Middle-Income Countries Participating in PISA 2000–2015**. Washington: Banco Mundial; Paris: OECD Publishing, 2015. DOI: doi.org/10.1787/9789264246195-en. Disponível em: https://www.oecd.org/content/dam/oecd/en/publications/reports/2015/11/the-experience-of-middle-income-countries-participating-in-pisa-2000-2015_g1g5d6b5/9789264246195-en.pdf. Acesso em: 10 mai. 2024.

MORRIS, Paul. Política educacional, exames internacionais de desempenho e a busca da escolarização de classe mundial: uma análise crítica. **Estudos em Avaliação educacional**, São Paulo, v. 28, n. 68, p. 302–342, 2017. DOI: <https://doi.org/10.18222/eae.v28i68.4696>. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/4696>. Acesso em: 12 abr. 2024.

NATIONAL RESEARCH COUNCIL. **International Comparative Studies in Education**: Descriptions of Selected Large-Scale Assessments and Case Studies. Washington, DC: The National Academies Press, 1995. DOI: <https://doi.org/10.17226/9174>. Disponível em: <https://nap.nationalacademies.org/catalog/9174/international-comparative-studies-in-education-descriptions-of-selected-large-scale>. Acesso em: 12 abr. 2024

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO OCDE. **Education at a Glance**: OCDE Indicators. Paris: OCDE, 1998. DOI: <https://doi.org/10.1787/eag-1998-en>. Disponível em: https://www.oecd.org/en/publications/education-at-a-glance-1998_eag-1998-en.html. Acesso em: 20 abr. 2024

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Política y Promoción de la Inversión Extranjera Directa en América Latina**. Paris: OCDE, 1999a. DOI: <https://doi.org/10.1787/9789264080072-es>. Disponível em: https://www.oecd.org/es/publications/politica-y-promocion-de-la-inversion-extranjera-directa-en-america-latina_9789264080072-es.html. Acesso em: 22 abr. 2024

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Measuring Student Knowledge and Skills**: A New Framework for Assessment. Paris: OCDE, 1999b. Disponível em: <https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33693997.pdf>. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Measuring Student Knowledge and Skills**: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy. Paris: OCDE, 2000. DOI: <https://doi.org/10.1787/9789264181564-en>. Disponível em: https://www.oecd.org/en/publications/measuring-student-knowledge-and-skills_9789264181564-en.html. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **PISA 2000 Relatório Nacional**. Paris: OCDE, 2001a. Disponível em: <https://download.inep.gov.br/download/internacional/pisa/PISA2000.pdf>. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Knowledge and Skills for Life: First Results from PISA 2000.** Paris: OCDE, 2001b. DOI: <https://doi.org/10.1787/9789264195905-en>. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **PISA 2000 Technical Report.** Paris: OCDE, 2002. Disponível em: https://www.oecd.org/content/dam/oecd/en/publications/reports/2002/12/programme-for-international-student-assessment-pisa_g1gh2e07/9789264199521-en.pdf. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **What Makes School Systems Perform?** Seeing School Systems through the Prism of PISA. Paris: OCDE, 2004a. DOI: <https://doi.org/10.1787/9789264017726-en>. Disponível em: https://www.oecd.org/en/publications/what-makes-school-systems-perform_9789264017726-en.html. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Pisa 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills.** Paris: OCDE, 2004b. DOI: <https://doi.org/10.1787/9789264101739-en> Disponível em: https://www.oecd.org/en/publications/the-pisa-2003-assessment-framework_9789264101739-en.html. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Latin American Economic Outlook 2008.** Paris, OCDE, 2008. DOI: <https://doi.org/10.1787/leo-2008-en>. Disponível em: https://www.oecd.org/en/publications/latin-american-economic-outlook-2008_leo-2008-en.html. Acesso em: 10 abr. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Iberoamérica en PISA 2006 Informe regional.** Paris: OCDE, 2009a. Disponível em: https://download.inep.gov.br/download/internacional/pisa/ibero-america_en_Pisa_2006.pdf. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Perspectivas Econômicas da América Latina 2009.** Paris: OCDE, 2009b. DOI: <https://doi.org/10.1787/leo-2009-en>. Disponível em: https://www.oecd.org/en/publications/latin-american-economic-outlook-2009_leo-2009-en.html. Acesso em: 10 abr. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **American Economic Outlook 2011: How Middle-Class Is Latin American?** Paris: OCDE, 2010. DOI: <https://doi.org/10.1787/leo-2011-en>. Disponível em: https://www.oecd.org/en/publications/latin-american-economic-outlook-2011_leo-2011-en.html. Acesso em: 12 abr. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Latin American Economic Outlook 2015: Education, Skills and Innovation for Development.** Paris, OCDE, 2014. DOI: <https://doi.org/10.1787/leo-2015-en>. Disponível em: https://www.oecd.org/en/publications/latin-american-economic-outlook-2015_leo-2015-en.html. Acesso em: 20 jun. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Fomentando o Crescimento Inclusivo da Produtividade na América Latina.** Paris: OCDE, 2016. DOI: <https://doi.org/10.1787/9789264258426-pt>. Disponível em: https://www.oecd.org/pt/publications/fomentando-o-crescimento-inclusivo-da-productividade-na-america-latina_9789264258426-pt.html. Acesso em: 20 jun. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Activos con América Latina y el Caribe.** Paris: OCDE, 2017. Disponível em: <https://www.oecd.org/latin-america/Activos-con-America-Latina-y-el-Caribe.pdf>. Acesso em: 10 mai. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Competencias en Iberoamérica: Análisis de PISA 2015.** Paris: OCDE, 2018. Disponível em: <https://www.oecd.org/skills/piaac/Competencias-en-Iberoamerica-Analisis-de-PISA-2015.pdf>. Acesso em: 15 jun. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **How's Life in Latin America? Measuring Well-being for Policy Making.** Paris: OCDE, 2021. DOI: <https://doi.org/10.1787/2965f4fe-en>. Disponível em: https://www.oecd.org/en/publications/how-s-life-in-latin-america_2965f4fe-en.html. Acesso em: 10 jun. 2024.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Latin American Economic Outlook 2023: Investing in Sustainable Development.** Paris: OCDE, 2023. DOI: <https://doi.org/10.1787/8c93ff6e-en>. Disponível em: https://www.oecd.org/en/publications/latin-american-economic-outlook-2023_8c93ff6e-en.html. Acesso em: 20 jun. 2024.

PEREIRA, Rodrigo da Silva. Proposições da OCDE para América Latina: o Pisa como instrumento de padronização da educação. **Revista Ibero-americana de Estudos em Educação**, v. 14, n. esp. 3, p. 1717-1732. DOI: <https://doi.org/10.21723/riaee.v14iesp.3.12756>. Disponível em: <https://periodicos.fclar.unesp.br/iberoamericana/article/view/12756/8513>. Acesso em: 18 dez. 2024.

RODRIGUES, Jean Douglas Zeferino; FERRAROTTO, Luana. Alinhamento e controle: as recomendações da OCDE ao Saeb. **Horizontes**, v. 42, n. 1, p. 1-25, 2024. DOI: <https://doi.org/10.24933/horizontes.v42i1.1796>. Disponível em: <https://revistahorizontes.usf.edu.br/horizontes/article/view/1796>. Acesso em: 10 dez. 2024.

SILVA, Maria Abádia da; SILVA, Maicon Donizete Andrade; FERREIRA, Neusa Souza Rêgo. Governar por números: política da Organização para a Cooperação e Desenvolvimento Econômico para a educação básica. **Revista Brasileira de Educação**, [S. l.], v. 27. p. 1-20, 2022. DOI: <https://doi.org/10.1590/S1413-24782022270120>. Disponível em: <https://www.scielo.br/j/rbedu/a/qdw-qlQwJJgsGsQyMt8qN8Rs/?lang=pt>. Acesso em: 23 abr. 2024.

WALKER, Maurice. **PISA 2009 Plus Results: Performance of 15-year-olds in reading, mathematics and science for 10 additional participants**. Melbourne: ACER Press, 2011. Disponível em: <https://research.acer.edu.au/pisa/1/>. Acesso em: 4 jun. 2024.

CAPÍTULO 5

EVALUACIÓN Y DINÁMICAS DE PRIVATIZACIÓN EDUCATIVA EN ARGENTINA: ENTRE LA MEDICIÓN DE APRENDIZAJES Y LA INTERVENCIÓN DE ORGANIZACIONES NO GUBERNAMENTALES¹

Lucrecia Rodrigo²

DOI: 10.52695/978-65-5456-163-1.5

-
- 1 Este capítulo se enmarca en las actividades de investigación de los siguientes proyectos: PID-AC 06/23 "Evaluación y gubernamentalidad del sistema educativo argentino desde una perspectiva federal", radicado en el Instituto de Educación, Justicia y Trabajo de la Universidad Nacional del Oeste (UNO), Argentina; UBACyT "Políticas públicas y derecho a la educación en la Argentina del siglo XXI. Análisis sobre las trayectorias de las políticas educativas en un país federal", radicado en el Instituto de Ciencias de la Educación, Facultad de Filosofía y Letras de la Universidad de Buenos Aires (UBA).
 - 2 Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Universidad Nacional del Oeste. Instituto de Investigaciones en Ciencias de la Educación, Universidad de Buenos Aires. E-mail: lucrecia.rodrigo@gmail.com

Introducción

Las evaluaciones estandarizadas de aprendizaje a gran escala han ganado un protagonismo creciente en la política educativa argentina, consolidándose como un pilar del actual sistema nacional de evaluación.³ Estas pruebas suelen presentarse como instrumentos capaces de promover la calidad, optimizar la eficiencia y garantizar la transparencia en la gestión escolar. Su centralidad se sostiene en una narrativa dominante que interpreta sus resultados como evidencia incuestionable del supuesto “fracaso” del sistema educativo, atribuyendo directamente a docentes y escuelas la responsabilidad por los desempeños observados. Sin embargo, este diagnóstico desatiende las condiciones estructurales que configuran las trayectorias escolares y desplaza el foco hacia indicadores individuales, reforzando una lectura reduccionista y fragmentada del complejo proceso de enseñanza y aprendizaje.

Esta orientación, cuyos cimientos se remontan a la década de 1990 con la institucionalización de un dispositivo nacional de evaluación, se profundizó entre 2015 y 2025 con la expansión de políticas guiadas por los principios de la rendición de cuentas. En este período, se profundizaron métricas cuantificables, comparables y orientadas al control de resultados. Entre estas iniciativas, las pruebas de aprendizaje ocuparon un lugar central, al priorizar la producción de datos para justificar decisiones gubernamentales, legitimar intervenciones focalizadas y fiscalizar el desempeño de estudiantes, docentes e instituciones. Esta configuración se vio reforzada por la intensificación de los procesos de privatización educativa (Feldfeber; Caride; Duhalde, 2020). El avance de actores no estatales — agencias multilaterales, organizaciones de la sociedad civil, fundaciones empresariales y consultoras especializadas — habilitó nuevas formas de gobernanza caracterizadas por articulaciones flexibles entre el sector público y el privado.

3 Por estas evaluaciones entendemos a aquellas pruebas aplicadas en el sistema educativo por organismos nacionales, jurisdiccionales o multilaterales, que miden aprendizajes en muestras o poblaciones completas bajo condiciones uniformes (por ejemplo, homogeneidad en los instrumentos, tiempos y criterios de evaluación) para garantizar la comparabilidad de resultados.

En sintonía con las tendencias globales, estas configuraciones promovieron el desplazamiento del tradicional modelo jerárquico de gobierno, dando lugar a la emergencia de redes de colaboración y coproducción de políticas, sustentadas en formas de autoridad informal y en mecanismos de gestión basados en la eficiencia, la competencia y la rendición de cuentas (Ball; Junemann, 2012; Ball; Youdell, 2007).

Respaldadas por un entramado normativo que posibilitó su institucionalización y por redes privadas que impulsaron activamente su expansión, las evaluaciones estandarizadas se asentaron en el país, desplegándose tanto a nivel nacional como en las jurisdicciones subnacionales. A este proceso se suma la participación en programas internacionales, que profundiza la inserción del sistema educativo argentino en la agenda global y refuerza la lógica de comparación y competencia entre sistemas. En este contexto, el capítulo analiza la relación entre la expansión reciente de las pruebas estandarizadas y las recomendaciones promovidas por el sector privado para consolidar reformas guiadas por los principios del *accountability*. Sostenemos que estas iniciativas constituyen una expresión paradigmática de los procesos de privatización de y en la educación (Ball; Youdell, 2007), cuyo despliegue funciona como un mecanismo de regulación que restringe el margen de acción estatal, limitándolo a aquello que las mediciones logran visibilizar y jerarquizar como problemático. Esta configuración plantea interrogantes sobre las redefiniciones que atraviesa el Estado y sobre la emergencia de nuevas formas de articulación entre lo público y lo privado en el diseño e implementación de las políticas educativas en el país.

El capítulo se estructura en tres secciones. La primera presenta una breve discusión conceptual sobre las propuestas de evaluación en el marco de las políticas de *accountability* y de los procesos de privatización. La segunda examina las políticas educativas implementadas en Argentina entre 2015 y 2025, las cuales han buscado consolidar estrategias de evaluación basadas en los principios de la rendición de cuentas y la responsabilidad por resultados. La tercera se centra en el análisis de la intervención del sector privado en la definición e implementación de estas medidas, a partir del estudio de una organización no gubernamental: Argentinos por la Educación. La elección de esta organización como caso

de estudio se sostiene en tres razones. En primer lugar, desde su creación en 2017 ha elaborado diagnósticos basados casi exclusivamente en datos provenientes de pruebas estandarizadas, reafirmando la premisa de que solo “lo medible” define la calidad educativa. En segundo término, su estructura multisectorial — que reúne actores del ámbito público y privado, entre ellos empresarios, investigadores, funcionarios y referentes educativos — le otorga una capacidad singular para producir y legitimar sentidos comunes sobre el desempeño escolar e incidir en la agenda pública. Finalmente, su creciente protagonismo en la política nacional de alfabetización confirma el alcance de su influencia: el programa actualmente vigente retoma de manera directa los diagnósticos y enfoques impulsados por la organización, consolidando un modelo de gobernanza donde la evaluación estandarizada se erige como dispositivo central para definir problemas, orientar decisiones y monitorear resultados. En este contexto, constituye un caso paradigmático para examinar cómo la evaluación se transforma en un mecanismo de intervención y direccionamiento de la política educativa.

Los hallazgos presentados se basan en el análisis documental de fuentes normativas, informes institucionales, comunicados oficiales y convenios interinstitucionales producidos por organismos públicos y privados entre 2015 y 2025. Esta revisión permite reconstruir el entramado político e institucional en el que se inscriben las estrategias de evaluación en Argentina, con un enfoque particular en el rol desempeñado por el sector privado.

Evaluaciones estandarizadas de aprendizaje, políticas de *accountability* y procesos de privatización

Las evaluaciones estandarizadas a gran escala se han consolidado como uno de los pilares de las políticas de *accountability* impulsadas globalmente en el marco de las reformas educativas asociadas a la Nueva Gestión Pública (Verges; Normand, 2015). Estas políticas incorporan principios del sector privado y buscan fortalecer la competencia, la eficiencia y la rendición de cuentas, promoviendo la comparación entre escuelas, estudiantes y docentes a partir de indicadores cuantitativos. Su propósito es incidir en las prácticas escolares mediante mecanismos de incentivo o sanción, reconfigurando tanto la autonomía institucional como el trabajo

docente para alinearlos con estándares y metas de rendimiento previamente establecidos. En este escenario, la evaluación se convierte en un dispositivo regulador que no solo redefine el rol del Estado — cada vez más orientado a la supervisión de resultados —, sino que también moldea los comportamientos dentro del sistema educativo. Así, las pruebas operan como “tecnologías de gubernamentalidad” que, mediante la articulación de saberes técnicos y estrategias de poder, moldean la percepción de la calidad educativa y jerarquizan a estudiantes e instituciones según estándares preestablecidos (Ramos-Zincke, 2018).

Aunque su diseño y aplicación responden a decisiones políticas que expresan relaciones de poder entre distintos actores del sistema, suelen presentarse como herramientas científicas, objetivas y de aplicación universal, capaces de diagnosticar con precisión el desempeño educativo (Aboites, 2012). Esta aparente neutralidad técnica y legitimidad científica otorgan autoridad a los resultados, que se posicionan como insumos “confiables” para la toma de decisiones sustentadas en criterios externos de control, eficiencia y productividad.

Desde una lógica funcionalista, estas evaluaciones se han enfocado inicialmente en medir aprendizajes individuales en áreas consideradas estratégicas para el desarrollo económico y social, como lengua y matemática. Sin embargo, en los últimos años, su alcance se ha ampliado hacia otros ámbitos, entre ellos las habilidades socioemocionales, digitales y financieras. Esta expansión ha sido impulsada por organismos internacionales que han desempeñado un rol central en la difusión y legitimación global de estos enfoques. Un caso emblemático es la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y su Programa para la Evaluación Internacional de Alumnos (PISA), cuya influencia ha sido decisiva en la consolidación de un modelo de gobernanza educativa sustentado en indicadores cuantitativos (Ozga, 2008; Popkewitz, 2013).

Este esquema redefine las formas de intervención estatal al priorizar el uso intensivo de evidencia empírica, indicadores estandarizados y rankings internacionales elaborados, en gran medida, por agencias no estatales que orientan crecientemente las decisiones de política educativa. En este contexto, los criterios de calidad se reconfiguran globalmente bajo

parámetros comparativos y descontextualizados, alineados con una lógica tecnocrática que desplaza enfoques situados, participativos y sensibles a las particularidades locales.

Así, en el marco de una gobernanza post-burocrática, flexible y en red, sustentada en los números, los datos y lo mensurable, se ha consolidado un modelo híbrido en el que funciones estratégicas del Estado comienzan a ser progresivamente delegadas a actores no estatales, como organismos internacionales y organizaciones de la sociedad civil (Ball, 2012; Saura, 2021). Estos actores se posicionan como “expertos” en el diseño, la implementación y el monitoreo de evaluaciones, participando directamente en la definición y ejecución de las políticas públicas. Esta transformación no solo favorece la incorporación de principios de mercado en las estructuras estatales, sino que también redefine las fronteras entre lo público y lo privado (Verger; Moschetti; Fontdevila, 2017). El resultado es una reconfiguración de la función estatal, cada vez más orientada a la regulación por resultados y a la rendición de cuentas, en detrimento de su rol como garante del derecho social a la educación.

No obstante, la adopción de políticas de *accountability*, como los programas de evaluación del desempeño, no se desarrolla de manera lineal ni homogénea en todos los contextos. Su implementación se ve atravesada por complejas dinámicas de negociación, disputa y adaptación entre actores estatales y privados, dando lugar a configuraciones híbridas en las que confluyen elementos globales con tradiciones y prácticas locales (Peck; Theodore, 2010). En consecuencia, estas políticas se inscriben en realidades multifacéticas que adquieren características específicas según el contexto en que se despliegan, moldeadas tanto por condicionantes históricos, políticos y culturales como por las tensiones entre agendas internacionales y prioridades locales. Estas diferencias se reflejan en los marcos normativos e institucionales que estructuran los sistemas de evaluación, en los instrumentos empleados, en las áreas de conocimiento priorizadas y en el tipo de impacto que los resultados ejercen sobre los distintos actores. En este sentido, los sistemas de evaluación se distinguen por el peso que sus resultados tienen en las trayectorias escolares, las condiciones laborales docentes o las decisiones institucionales (Ravela, 2006). Se denomina evaluaciones de

“alta implicancia” (*high-stakes*) a aquellas cuyos resultados generan consecuencias directas para estudiantes, docentes o escuelas, mientras que las de “baja implicancia” (*low-stakes*) no producen efectos inmediatos sobre los actores involucrados.

Partiendo de estas perspectivas conceptuales, los apartados siguientes analizan cómo estas dinámicas se han expresado en el contexto argentino, atendiendo a las especificidades históricas, institucionales y políticas que han moldeado su desarrollo a lo largo de la última década.

La expansión de las evaluaciones estandarizadas en el sistema educativo argentino

En Argentina, la implementación de evaluaciones estandarizadas a gran escala encontró un fuerte respaldo en las gestiones políticas de corte neoliberal y neoconservador que, desde 2015, profundizaron la adopción de estos mecanismos de control de la educación.⁴ A partir de ese momento, estas medidas ampliaron su alcance y consolidaron su lugar estratégico en la agenda nacional, considerándose esenciales tanto para la formulación de diagnósticos como para la toma de decisiones.

La centralidad otorgada a las pruebas de desempeño y a sus resultados se apoyó en discursos que alertaban sobre una supuesta crisis del sistema educativo argentino, entendida en términos de su incapacidad para formar sujetos competitivos, adaptables y flexibles, acordes con las demandas del mercado (Feldfeber, 2024). Esta narrativa se nutrió de la lectura de datos cuantitativos provenientes de evaluaciones nacionales e internacionales que mostraban bajos niveles de aprendizaje. Bajo esta perspectiva, el sistema educativo argentino fue retratado como obsoleto, ineficiente y carente de calidad, cuestionando de manera particular el rol de la escuela pública y la labor docente, a quienes se atribuyeron presun-

4 Nos referimos a la presidencia de Mauricio Macri, quien asumió el gobierno en diciembre de 2015 y gobernó hasta 2019 al frente de la alianza Cambiemos, integrada por el PRO (Propuesta Republicana), la UCR (Unión Cívica Radical) y la Coalición Cívica (CC).

tas deficiencias en su formación, compromiso y capacidad profesional. Este discurso encontró respaldo en los diagnósticos producidos por organismos internacionales, como el Banco Mundial (2018), que alertaban sobre una “crisis global del aprendizaje”, especialmente aguda en países de ingresos bajos y medios, entre ellos Argentina. Desde esta perspectiva, una proporción significativa de estudiantes asiste a la escuela pero no adquiere habilidades básicas de lectura, escritura y matemática, lo que pone en duda la eficacia de los sistemas educativos. A partir de la difusión de este enfoque, la noción de “crisis de aprendizaje” se consolidó como una categoría central en los discursos y en las políticas de reforma educativa promovidas por agencias multilaterales en América Latina.

Asimismo, esta situación se inscribe en un proceso más amplio de profundización de la privatización, que favoreció la incorporación de principios de gestión empresarial en el sistema educativo y fortaleció el protagonismo de actores privados en la definición de la política nacional (Feldfeber; Caride; Duhalde, 2020). Durante el período 2015–2019, se consolidó la presencia en el Estado de funcionarios provenientes de organizaciones del sector privado, fundaciones y grupos empresariales de alcance nacional, regional y multinacional. Este fenómeno, denominado “CEOcracia”, alude al desembarco de altos ejecutivos de corporaciones económicas en puestos clave de la administración pública, con el riesgo de que intereses corporativos avancen sobre sectores estratégicos del aparato estatal e incidan de manera directa en la orientación de la política pública (Castellani, 2019).

En este contexto, la política educativa argentina reorientó sus prioridades alrededor de los principios de eficiencia, competencia y rendición de cuentas. Se enfatizó la medición de aprendizajes y se promovió la renovación de los dispositivos de evaluación como estrategia para fortalecer los diagnósticos y legitimar la toma de decisiones (Rodríguez Moyano; Rodrigo, 2022). En este marco, se relanzaron los operativos bajo la denominación Aprender, con el objetivo de actualizar y robustecer técnica y metodológicamente los programas preexistentes, que habían sido cuestionados por su baja rigurosidad metodológica, la discontinuidad en su aplicación y la limitada comparabilidad de los datos a lo largo del tiempo. Según el diagnóstico oficial, estas

deficiencias afectaban la confiabilidad y legitimidad del sistema nacional de evaluación (Rodríguez; Vior; Más Rocha, 2018). En línea con estos objetivos, las pruebas comenzaron a aplicarse de forma censal y anual a estudiantes del último año de los niveles primario y secundario.

Se impulsaron también iniciativas destinadas a evaluar el sistema de formación docente. Al respecto, el programa Enseñar representó una experiencia inédita en el país, al relevar por primera vez, a nivel nacional, información sobre las capacidades de los estudiantes que se encontraban próximos a egresar de los profesorados (Rodrigo; Rodríguez Moyano, 2023). Estas iniciativas se alinearon con las propuestas del Banco Mundial para América Latina, en las que la mejora educativa se sustenta en la aplicación de pruebas censales a los estudiantes, con una difusión transparente de sus resultados; la asignación de bonificaciones a las escuelas en función del progreso observado en los aprendizajes; y la evaluación individual y voluntaria de los docentes, acompañada de incentivos económicos para quienes acepten ser evaluados y demuestren un desempeño destacado (Banco Mundial, 2014).

En su conjunto, estos programas fueron impulsados por la Secretaría de Evaluación Educativa (SEE), creada en 2016 como parte del organigrama del Ministerio de Educación de la Nación, con el objetivo de jerarquizar institucionalmente el área dentro del Estado. La iniciativa buscó fortalecer las capacidades técnicas estatales en el diseño, implementación y análisis de los instrumentos de evaluación. Entre sus principales objetivos, la SEE se propuso dotar de mayor solidez técnica y legitimidad pública a los datos producidos por las evaluaciones, así como garantizar la continuidad y sostenibilidad de las pruebas más allá de los cambios de gobierno. La creación de la Secretaría no solo implicó un cambio en la administración educativa, sino que también fortaleció el papel de la evaluación en la definición y control de las políticas públicas. Además, facilitó el acceso a financiamiento internacional para el sector.

Este reposicionamiento institucional también se reflejó en una participación más activa del país en las pruebas internacionales como PISA y regionales como las del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE). En efecto, un aspecto destacado de esta

etapa fue la incorporación, por primera vez, de la desagregación de los resultados por jurisdicción en PISA. Este hecho permitió analizar las disparidades entre provincias y reforzó el perfil técnico de los diagnósticos educativos, presentándolos como insumos para orientar políticas locales (Ministerio de Educación de la Nación, 2019). Al mismo tiempo, consolidó una lógica de rendición de cuentas territorializada, al visibilizar brechas de desempeño y alentar procesos de *benchmarking* entre sistemas educativos subnacionales.

En este escenario, las acciones orientadas a difundir los resultados de las evaluaciones adquirieron un rol estratégico y comenzaron a contar con el apoyo de organizaciones del sector privado.⁵ Presentadas como respuesta a la supuesta “opacidad” de gestiones anteriores y justificadas por la necesidad de reforzar la transparencia y el uso efectivo de los datos, se concretaron a través del llamado Sistema Abierto de Consulta. Este dispositivo buscó garantizar el acceso público a las bases de datos, facilitar la elaboración de informes diferenciados para diversos actores del sistema y promover materiales que acompañaran la interpretación de los resultados. Paralelamente, resurgió la discusión sobre la publicación de resultados por escuela, a pesar de la prohibición vigente en la legislación educativa nacional.⁶ Este debate fue impulsado por sectores que defienden la construcción de *rankings*, la promoción de la elección escolar por parte de las familias y la competencia entre instituciones como estrategias para incentivar la mejora educativa.

5 Por ejemplo, el Centro de Implementación de Políticas Públicas para la Equidad y el Crecimiento (CIPPEC), una organización no gubernamental fundada en 2000 y financiada principalmente por empresas y organismos internacionales, tuvo a su cargo el desarrollo de informes de evaluaciones como fue el caso de PISA 2012.

6 Nos referimos a la Ley de Educación Nacional (26.206) de 2006, cuyo artículo 97 establece que “El Ministerio de Educación, Ciencia y Tecnología y las jurisdicciones educativas harán públicos los datos e indicadores que contribuyan a facilitar la transparencia, la buena gestión de la educación y la investigación educativa. La política de difusión de la información sobre los resultados de las evaluaciones resguardará la identidad de los/as alumnos/as, docentes e instituciones educativas, a fin de evitar cualquier forma de estigmatización, en el marco de la legislación vigente en la materia”.

En conjunto, estas acciones evidencian la consolidación, en la política nacional, de un modelo de gobernanza orientado a la gestión por resultados, en el que la medición de aprendizajes y la difusión pública de los datos se erigen como ejes centrales de la intervención estatal. En este marco, las formas de responsabilidad dentro del sistema educativo se reconfiguran, trasladando el éxito o el fracaso directamente a estudiantes, docentes e instituciones según los resultados obtenidos. A la vez, el sentido de la educación se restringe a los aprendizajes medibles, en línea con el enfoque impulsado por el Banco Mundial. Esta concepción reduce la comprensión del hecho educativo, invisibilizando sus dimensiones sociales, políticas y culturales, y refuerza una lógica instrumental centrada en la eficiencia y la rendición de cuentas.

No obstante, más allá de la retórica oficial que presenta las pruebas como herramientas destinadas a orientar las políticas, el impacto de sus datos en la formulación de medidas concretas ha sido limitado o prácticamente nulo. En la práctica, estas evaluaciones cumplen sobre todo funciones diagnósticas y legitimadoras. Estudios recientes sobre el uso de la información generada por la evaluación Aprender durante el período 2016–2019 muestran que, si bien los resultados fueron empleados para diseñar o ajustar algunos programas y políticas, su utilización más significativa tuvo un carácter simbólico: legitimar decisiones, comunicar logros o problemas y construir narrativas públicas sobre la calidad educativa, incluso cuando las políticas no se encontraban plenamente sustentadas en evidencia técnica (López Leavy, 2023). Este uso político de la información se vincula con la necesidad de validar la propia evaluación, orientando las decisiones más por motivaciones simbólicas que por fundamentos técnicos.

En este sentido, la información producida por estas pruebas ha sido fundamental para instalar y consolidar en el debate público el discurso de “crisis”, que atraviesa tanto la agenda mediática como los marcos interpretativos desde los cuales se analiza el sistema escolar. De este modo, se evidencia la compleja relación entre evaluación, política y construcción de agendas públicas en el sistema educativo argentino.

Desafíos y tensiones en la evaluación de la educación durante la pandemia

En 2020, la pandemia de COVID-19 interrumpió abruptamente la continuidad pedagógica en América Latina, lo que generó una profundización de las desigualdades educativas (CEPAL, 2021). Estas condiciones no solo afectaron las trayectorias escolares, sino que también provocaron un aumento de las tasas de abandono escolar en varios países de la región (UNICEF, 2021). En este contexto de emergencia, resurgió la noción de “crisis de aprendizaje” y se lanzó la advertencia sobre una potencial “catástrofe generacional”. Este riesgo se vinculó a la supuesta pérdida de conocimientos y habilidades consideradas necesarias para la formación de capital humano (Banco Mundial, 2021). Este diagnóstico se fundamentó, una vez más, en los resultados de pruebas estandarizadas de desempeño.

En Argentina, la crisis sanitaria puso en primer plano la necesidad de revisar y transformar las estrategias de evaluación vigentes. En una primera etapa, se impulsaron propuestas que buscaban superar el enfoque reduccionista de las pruebas de aprendizaje estandarizadas. Entre estas iniciativas se destacó la Evaluación Nacional del Proceso de Continuidad Pedagógica, implementada por el Ministerio de Educación mediante encuestas dirigidas a equipos directivos, docentes y familias. El objetivo era obtener un diagnóstico amplio de la situación educativa durante la enseñanza remota en los niveles de la educación obligatoria (Diker, Briscioli y Stoppani, 2023). Estas acciones se inscribieron en una política educativa más vasta que, especialmente durante los primeros años del gobierno del Frente de Todos (2019–2023), intentó diferenciarse de la gestión previa, la cual se había centrado casi exclusivamente en la medición de aprendizajes individuales. En coherencia con esta línea, los planes nacionales de evaluación definidos entre 2021 y 2022 incorporaron la recolección de datos a través de múltiples instrumentos. Las propuestas se enfocaron en diversas dimensiones del proceso educativo — tales como la enseñanza, el aprendizaje, las trayectorias escolares, la organización institucional y el financiamiento —, utilizando una combinación de metodologías y estrategias de análisis para obtener una perspectiva más amplia y compleja del desempeño estudiantil.

No obstante, el alcance de este enfoque integral resultó limitado. Tras superarse la etapa más crítica de la pandemia, la política educativa tendió a focalizarse en la recuperación de los “aprendizajes perdidos”. Esta orientación coincidió con las recomendaciones de organismos multilaterales y organizaciones no gubernamentales, que en los últimos años han ejercido un rol central en la promoción y respaldo de estas medidas. En consecuencia, se reforzó la necesidad de implementar evaluaciones estandarizadas a gran escala, diseñadas por organismos con competencia nacional o jurisdiccional, y aplicadas a muestras o poblaciones completas en los niveles primario y secundario. Adicionalmente, se recomendó que estas pruebas incorporaran nuevas dimensiones de medición, como las habilidades socioemocionales (Banco Mundial *et al.*, 2022; Torre, Perusia y Xanthopoulos, 2023). De este modo, las evaluaciones estandarizadas recuperaron protagonismo y pasaron a ocupar un lugar dominante dentro de los planes nacionales de evaluación.

Consolidación del enfoque tecnocrático e instrumental en la evaluación educativa actual

El rumbo neoliberal y neoconservador en materia educativa se intensificó notablemente con la llegada al poder en Argentina de una fuerza política identificada con las nuevas derechas radicalizadas, caracterizada por la combinación de un discurso libertario y autoritario.⁷ Este sector impulsa una ofensiva directa contra los derechos sociales y la esfera pública, utilizando una narrativa que desdibuja la noción de derechos, presentándolos falazmente como meros privilegios (Barros, 2024). Desde esta perspectiva, se deslegitima cualquier política fundada en la igualdad y la justicia social, debilitando los principios democráticos que sustentan la noción de derechos e invalidando la igualdad como base para las políticas públicas (Ibidem).

7 Nos referimos al gobierno de Javier Milei, quien asumió la presidencia en diciembre de 2023 como líder de La Libertad Avanza.

En este contexto de erosión de derechos, la educación es reconfigurada de un bien público a un bien de consumo, susceptible de ser gestionado y ofertado en el mercado. Este cambio desplaza el horizonte igualitario de la política educativa y lo alinea con parámetros mercantiles, donde la ciudadanía se reduce a la figura de un consumidor y el mérito opera como el criterio moral fundamental para legitimar el acceso y la permanencia. Como consecuencia de esta despolitización, las demandas colectivas se diluyen, las obligaciones del Estado se transfieren a decisiones individuales y se socavan los fundamentos democráticos que sostienen a la educación como derecho (Barros, 2024). Este escenario configura una agenda educativa atravesada por una visión antiestatal, el avance de los procesos de privatización y una lógica de responsabilización individual (Ruiz, 2025). Las políticas de *accountability* se intensifican, subordinando aún más el sistema educativo a criterios de eficiencia, rendimiento y competencia, en detrimento de su dimensión social, pedagógica y democrática.

Una expresión concreta de este cambio de rumbo fue la reforma ministerial que degradó al Ministerio de Educación a una Secretaría, subordinándola al Ministerio de Capital Humano.⁸ Por un lado, esta medida supuso una pérdida significativa de jerarquía institucional y debilitó la capacidad del Estado para diseñar e implementar políticas educativas de alcance nacional. Por otra parte, esta subordinación revela la hegemonía de una perspectiva funcionalista que prioriza las demandas del mercado por sobre el bienestar social. Con ello se consolida una visión instrumental de la educación, orientada exclusivamente a la formación de “capital humano” y al desarrollo de competencias consideradas útiles para la economía, una tendencia que ya se había insinuado en gestiones previas (Feldfeber, 2024).

Entre 2024 y 2025, el desfinanciamiento del sistema alcanzó una magnitud inédita. Los recortes presupuestarios profundizan un proceso que ya estaba en curso, dejando a la actual Secretaría de Educación operando bajo un fuerte ajuste fiscal que limita su capacidad de acción (Curcio *et al.*,

8 Impulsada mediante el Decreto de Necesidad y Urgencia 8/2023.

2024).⁹ Paralelamente, se intensificaron los procesos de privatización mediante la transferencia de fondos públicos a actores privados y la creciente influencia de estos en el diseño de políticas, la provisión de servicios y los programas de evaluación (Puiggrós *et al.*, 2025). En línea con esta lógica, una de las primeras medidas fue la creación en 2024 del programa de asistencia Vouchers Educativos, diseñado para financiar parcialmente la matrícula de estudiantes en instituciones privadas con subvención estatal, trasladando recursos del ámbito público al privado.¹⁰

Asimismo, ese mismo año se presentó el Plan Nacional de Alfabetización, diseñado para garantizar que los estudiantes de nivel primario alcancen niveles adecuados de comprensión lectora. Este plan, impulsado desde 2023 por actores privados como Argentinos por la Educación, incorporó una unidad nacional de seguimiento y una prueba estandarizada para tercer grado. En línea con estas medidas, los planes nacionales de

9 Al observar la evolución del gasto nacional en educación como porcentaje del PBI, se advierte una trayectoria descendente, con altibajos, que refleja este proceso. El punto más alto de inversión se registró en 2015, cuando alcanzó el 1,59% del PBI. Le siguieron dos años de leve retracción (1,52% en 2016 y 2017), y desde 2018 la caída se intensificó: 1,28% en 2018 y apenas 1,08% en 2019, el valor más bajo del período. A partir de ese momento, se observó una recuperación paulatina que llevó la inversión al 1,15% en 2020, al 1,22% en 2021, al 1,26% en 2022 y al 1,48% en 2023, alcanzando así el mayor nivel desde 2015. Sin embargo, este proceso de recomposición se ve interrumpido en 2024. Según las estimaciones oficiales basadas en el crédito vigente y el PBI nominal proyectado, el gasto en educación caería al 0,91% del PBI. Para 2025, se prevé una nueva disminución, reduciendo la inversión al 0,88% y consolidando una tendencia de ajuste severo en el área. Esta caída no solo representa un retroceso en términos históricos, sino que también incumple el artículo 130 de la Ley de Presupuesto 2023 —aún vigente por prórroga— que establece un piso del 1,33% del PBI para la inversión educativa de la Administración Pública Nacional (Curcio *et al.* 2024).

10 El programa fue aprobado por Resolución del Ministerio de Capital Humano (61/2024) y consiste en una prestación económica mensual, no contributiva y de carácter temporario, destinada a familias con ingresos inferiores a siete salarios mínimos, cuyos hijos asisten a instituciones educativas de gestión privada con aporte estatal superior al 75%. Cubre los niveles inicial, primario y secundario, y equivale al 50% del valor de la cuota de jornada simple, según lo informado por las jurisdicciones o estimado por la Secretaría de Educación. La medida financia parcialmente la matrícula de estos estudiantes y se proyectó que alcanzaría al 20% del alumnado de los niveles obligatorios, especialmente de sectores medios y bajos.

evaluación para el período 2023–2025 reforzaron la centralidad de las pruebas estandarizadas. Aunque se contemplaron algunas estrategias cualitativas, el análisis cuantitativo de datos continuó siendo el eje principal para la formulación de diagnósticos y políticas.¹¹ Además, se amplió la participación de Argentina en evaluaciones como PISA y LLECE, y se implementó de manera piloto la prueba EGRA (*Early Grade Reading Assessment*), orientada a medir habilidades básicas de alfabetización en los primeros grados de la enseñanza obligatoria.¹²

El énfasis en la medición se consolidó con la propuesta de instaurar un examen censal obligatorio al finalizar el nivel secundario, cuyos resultados, certificados individualmente, serían también publicados por escuela. Aunque se aclaró que dicha evaluación no condicionaría la titulación, se planteó eliminar las restricciones legales que limitan la difusión de resultados institucionales, bajo el argumento de “informar a las familias”. Paralelamente, el gobierno promovió la evaluación docente obligatoria, periódica y pública como eje de la carrera profesional, vinculando ascensos y suplementos salariales al desempeño en lugar de la antigüedad.

Aunque varias de estas medidas aún se debaten, confirman el fortalecimiento de las políticas de *accountability*. En este contexto, se afianza un proyecto político educativo que traslada la responsabilidad a los individuos y naturaliza las desigualdades al presentarlas como producto exclusivo del mérito y el esfuerzo personal (Gluz; Karolinski; Kesler, 2024). Este proceso se articula con un discurso que exalta la supuesta superioridad del mercado como proveedor eficiente de servicios educativos, mientras que, al mismo tiempo, “demoniza” a la escuela pública y a sus trabajadores,

11 A partir del plan nacional de evaluación 2023–2024, se incorporaron modalidades de evaluación formativa y en línea. Destacan las plataformas Acompañar Educación Secundaria (2023) y Acompañar Primer Ciclo de la Escuela Primaria (2024). También se implementaron encuestas a estudiantes sobre percepciones, autonomía, uso de tecnologías y clima escolar.

12 Esta herramienta evalúa aspectos como el reconocimiento de sonidos del alfabeto, la lectura de palabras simples, la fluidez lectora y la comprensión de oraciones y párrafos.

describiéndolos como ineficientes, corruptos o derrochadores de recursos públicos (Feldfeber, 2024).

El protagonismo del sector privado y de la evaluación bajo principios de rendición de cuentas: el caso de Argentinos por la Educación

La principal política educativa impulsada por el gobierno de La Libertad Avanza, el Plan Nacional de Alfabetización, se inscribe en un contexto marcado por el creciente protagonismo de organizaciones no gubernamentales en la definición de la agenda educativa, un proceso que se ha intensificado a lo largo de la última década (Feldfeber *et al.*, 2020; Feldfeber; Gluz; Rodrigo, 2025). Este fenómeno es resultado de la profundización de convenios y alianzas entre los sectores público y privado, lo cual ha fortalecido la influencia empresarial en las políticas públicas y ha consolidado lógicas de gestión escolar propias del ámbito privado (Puiggrós *et al.*, 2025). En este marco, la rendición de cuentas y la responsabilidad por los resultados se han convertido en ejes del actual proyecto político, lo que se evidencia en el fortalecimiento de dispositivos de evaluación orientados a la medición de aprendizajes.

Esta dinámica no es reciente, sino que encuentra sus raíces en transformaciones institucionales iniciadas en 2015, momento en que las estructuras estatales vinculadas al sector educativo experimentaron cambios significativos. Entre estas, destaca la eliminación de áreas históricamente dedicadas a la producción y análisis de datos estadísticos, las cuales fueron reemplazadas por dependencias con una aparente mayor autonomía respecto del poder político, pero con vínculos más estrechos con el sector privado, como lo fue la creación de la entonces Secretaría de Evaluación Educativa.

Actualmente, el reordenamiento institucional que atraviesa el sector se inscribe en un escenario de profundización del desfinanciamiento y desmantelamiento de la administración pública. Esto se acompaña de políticas de desregulación económica y de la ocupación de áreas estratégicas de gestión por parte de funcionarios con vínculos directos con empresas privadas. En este contexto, se configura un terreno propicio para el avance de intereses corporativos sobre ámbitos deliberadamente

debilitados del aparato estatal, lo que incrementa la dependencia de actores externos y reduce la capacidad autónoma del Estado para definir la política educativa (Castellani, 2024).

La creación de la organización Argentinos por la Educación puede interpretarse como un caso paradigmático del desplazamiento de capacidades estatales hacia el sector privado, con un impacto directo en áreas estratégicas como la evaluación. Constituida en 2017 a partir de un convenio entre el Ministerio de Educación de la Nación y un grupo de empresarios y especialistas, la organización se presenta como un espacio dedicado al análisis y la difusión de indicadores sobre el sistema educativo. Bajo el lema “la Argentina del mañana depende de la educación de hoy”, instala una narrativa que asocia la mejora de la enseñanza con el crecimiento económico y el desarrollo social.¹³

Desde sus inicios, esta organización buscó posicionar la educación como una prioridad en la agenda pública, impulsando el uso de indicadores como herramienta estratégica para orientar políticas. A través de consignas como “lo que no se mide no se conoce”, ha fomentado una cultura de la evaluación orientada a la toma de decisiones, la asignación de recursos y la implementación de mecanismos de incentivos o sanciones, todo bajo los principios de eficiencia, calidad y transparencia.¹⁴ Además, ha mantenido desde sus comienzos vínculos estrechos con los poderes económicos y políticos del gobierno. Gran parte de su equipo directivo cuenta con experiencia en cargos clave de evaluación educativa durante las gestiones de Mauricio Macri, tanto a nivel nacional como en la Ciudad de Buenos Aires. Actualmente, destaca la presencia del Secretario de Educación en esta organización, junto a exministros, académicos e investigadores mayoritariamente alineados con un enfoque tecnocrático que ha

13 Véase al respecto el sitio institucional de la organización: <https://argentinosporlaeducacion.org/>

14 Cabe destacar que Argentinos por la Educación, a través de su Observatorio, no produce datos propios, sino que sistematiza, organiza y difunde información generada por fuentes oficiales, organismos nacionales e internacionales.

respaldado las políticas educativas neoliberales implementadas desde la década de 1990 hasta la actualidad (Puigrós *et al.*, 2025).

Las políticas basadas en evidencia como fundamento para la gestión educativa

Una de las primeras acciones de Argentinos por la Educación fue la firma de convenios con gobiernos locales para promover la apertura, el monitoreo y la “transparencia” de los datos educativos. Como resultado, en 2018, la organización lanzó una plataforma digital de visualización que integra información clave sobre el sistema: acceso, permanencia, eficiencia interna (tasas de repitencia y promoción), recursos humanos, infraestructura, financiamiento y resultados de evaluaciones estandarizadas.¹⁵ Bajo el lema “medir para mejorar”, esta herramienta fue presentada como un insumo técnico fundamental para el diagnóstico del sistema educativo y como un medio para fomentar un debate público sustentado en datos.

Desde sus inicios, los datos difundidos por este actor han incidido en la formulación de políticas públicas. La actual política nacional de alfabetización, por ejemplo, retoma una campaña impulsada por esta organización que advierte sobre una supuesta “crisis” en los aprendizajes de lectura y escritura. Este diagnóstico se apoya, entre otros insumos, en los resultados del Estudio Regional Comparativo y Explicativo (ERCE) 2019 del LLECE, que muestran que el 46% de los estudiantes de tercer grado no comprende lo que lee — cifra que asciende al 61,5% en los sectores más vulnerables —, mientras que solo el 14% alcanza el nivel más alto de desempeño, cifra por debajo del promedio regional del 21% (LLECE–OREALC/UNESCO, 2021). Asimismo, los datos de PISA 2022 indican que 7 de cada 10 estudiantes de 15 años pertenecientes al nivel socioeconómico más bajo no alcanzan el umbral mínimo en lectura, frente a 3 de cada 10 en los sectores más favorecidos.

Argentinos por la Educación interpreta estos resultados como la evidencia de una crisis profunda que exige una reforma estructural centrada

15 Desde sus inicios en 2017, estableció acuerdos y convenios con varias jurisdicciones, como la Ciudad de Buenos Aires, Mendoza y Córdoba.

en la mejora de la comprensión lectora. Para ello, sostiene la necesidad del “compromiso de todos”, situando a la ciudadanía — concebida en su dimensión individual, no organizada — como actor clave en la identificación de problemas y en la implementación de soluciones. En este marco, impulsa la construcción de redes, la formación de liderazgos y la promoción de campañas orientadas a movilizar a sectores sociales dispuestos a “ser parte del cambio”. Estas iniciativas, desarrolladas por fuera del Estado, se proponen consolidar un “movimiento social por la educación” capaz de incidir en las transformaciones del sistema educativo.

En esta línea, las acciones de Argentinos por la Educación se materializan a través de campañas de comunicación y participación ciudadana que, bajo el lema “potenciar la voz de la ciudadanía”, buscan incidir en la agenda pública. Para ello, la organización recopila y difunde datos, además de elaborar informes analíticos sobre diversas temáticas educativas: la educación durante la pandemia, las desigualdades regionales en el acceso, el financiamiento de los centros escolares, la relación entre pobreza y rendimiento académico, y el desempeño estudiantil por materia y región, entre otras. Estos reportes son rápidamente incorporados por los medios de comunicación, lo que amplifica su influencia en la configuración de la agenda educativa y en la formación de la opinión pública, posicionando a la organización como un actor central en el debate actual.

Los informes y reportes elaborados por la organización evidencian su consolidación en la promoción de una agenda centrada en la mejora educativa, en la que las acciones destinadas a la evaluación — entendida como la medición de aprendizajes — ocupan un lugar clave como herramientas para generar información objetiva que permita “iluminar” las decisiones políticas. Asimismo, se destaca la importancia de involucrar a la sociedad para avanzar en el cambio, subrayando la participación ciudadana, especialmente en momentos críticos como fue la crisis sociosanitaria provocada por la pandemia.

De hecho, durante este período, Argentinos por la Educación se distinguió por impulsar la reapertura de las escuelas y la aplicación de evaluaciones para cuantificar el impacto del COVID-19 en los aprendizajes (Observatorio Argentinos por la Educación, 2020). Paralelamente, fortaleció la participación social mediante la creación de la Red de Familias por

la Educación, destinada a informar y movilizar a la ciudadanía en torno a la defensa y mejora del sistema educativo, iniciativa que también surgió en el contexto de la pandemia.

De la campaña a la política pública: alfabetización, evaluación y participación del sector privado

Como hemos señalado, el caso de Argentinos por la Educación no solo ejemplifica cómo el creciente protagonismo del sector privado desplaza capacidades estatales estratégicas, sino que además consolida una concepción de la educación centrada en la rendición de cuentas y la responsabilidad por resultados, lo que impulsa la adopción de medidas orientadas a la medición estandarizada del aprendizaje.

En esta línea, en 2023, la organización, junto a otros actores del sector privado, lanzó y coordinó la Campaña Nacional por la Alfabetización, basada en un diagnóstico construido a partir de los resultados de evaluaciones estandarizadas. Esta iniciativa buscó generar presión política y social para priorizar la alfabetización inicial en la agenda pública. En un contexto electoral, la campaña alcanzó visibilidad y logró la adhesión de los seis principales precandidatos presidenciales — incluido el actual presidente Javier Milei —, quienes firmaron el compromiso correspondiente. Asimismo, varios gobernadores provinciales se sumaron a la iniciativa. El compromiso se estructuró en torno a tres ejes: en primer lugar, impulsar una política de alfabetización inicial con acciones coordinadas y sostenidas a nivel nacional; en segundo término, brindar apoyo técnico y financiero a las provincias, con énfasis en aquellas con mayores rezagos y en la implementación de políticas efectivas; y, finalmente, realizar periódicamente, con carácter censal, las pruebas Aprender en tercer grado enfocadas en comprensión lectora, y publicar sus resultados.

En 2024, Argentinos por la Educación relanzó la campaña con una estrategia orientada a movilizar a la ciudadanía mediante la difusión en redes sociales de fotografías de personas leyendo.¹⁶ Además, incorporó el apoyo

16 En esta oportunidad la consigna defendida se resumió en: #SiempreHayAlgoParaLeer.

de artistas y amplió su presencia en medios masivos, espacios públicos y plataformas digitales. En este marco, promovió la instauración de momentos diarios de lectura en los hogares y lanzó la iniciativa “Ponerse la camiseta por la alfabetización”, bajo la consigna: “Porque si no entendemos lo que leemos, no entendemos nada. Y nosotros queremos cambiarlo”.

A través de estas acciones, la organización no solo captó la atención pública sobre la problemática de la alfabetización, sino que también logró consolidarla como una prioridad nacional. Así, en 2024, luego de alcanzar consenso en el Consejo Federal de Educación, se aprobó el documento oficial que dio origen al Plan Nacional de Alfabetización y a los 24 planes jurisdiccionales. Este acuerdo brindó un respaldo político e institucional al plan promovido por el sector privado y estableció un marco estratégico para la implementación coordinada de la política a nivel nacional y provincial, incorporando la participación de organizaciones no estatales.

En el Plan Nacional de Alfabetización, la evaluación de aprendizajes en tercer grado de la educación primaria constituye la principal herramienta para medir los avances, estableciéndose la publicación anual de sus resultados como un mecanismo fundamental de rendición de cuentas que busca ofrecer diagnósticos actualizados. En este escenario, Argentinos por la Educación asume un rol protagónico en el seguimiento y acompañamiento de la política, elaborando informes periódicos que recopilan información sobre las metodologías de enseñanza, la cobertura territorial de las intervenciones, la formación docente, las instancias de evaluación implementadas y las articulaciones con actores locales.¹⁷

17 Se publicaron cuatro informes de monitoreo. El primero, resume de manera descriptiva los 24 planes jurisdiccionales y el plan nacional aprobados por el Consejo Federal de Educación (Lenzi; Pisani; Volman, 2024). El segundo abordó aspectos relacionados con la normativa vigente en cada provincia, las capacitaciones docentes llevadas a cabo, la entrega de materiales, las evaluaciones de seguimiento diseñadas y las alianzas establecidas con organizaciones, empresas o especialistas (Alzú; Sáenz Guillén; Volman, 2025). En el tercero se dieron a conocer los métodos y enfoques de alfabetización utilizados en las jurisdicciones en el marco del compromiso de mejorar la alfabetización y asegurar una educación de calidad (Abusamra; Alzú; Volman, 2025). El cuarto informe centró su atención en las evaluaciones nacionales y jurisdiccionales implementadas durante 2024 (Alzú; Sáenz Guillén; Volman, 2025).

Las funciones de la organización quedan estructuradas en tres ejes: el monitoreo sistemático tanto del Plan Nacional como de los planes jurisdiccionales; la articulación entre el Estado nacional, las provincias y la sociedad civil, fomentando la cooperación intersectorial; y la producción y sistematización de evidencia orientada a guiar las políticas públicas desde un enfoque basado en datos. Estas responsabilidades, en particular las vinculadas a la recolección, seguimiento y análisis de la información, evidencian un claro desplazamiento de la tradicional centralidad estatal en estos procesos, reflejando el creciente protagonismo del sector privado en la gobernanza educativa a través de la firma de convenios y asesorías, entre otras acciones.

En este contexto, la autodefinición de Argentinos por la Educación como una organización sin fines de lucro, políticamente independiente, plural y autónoma, cumple una función política y comunicativa estratégica: confiere legitimidad técnica y proyecta una imagen de neutralidad en su discurso. Al presentarse como un actor no gubernamental sin financiamiento estatal, fortalece la percepción de objetividad de sus diagnósticos y recomendaciones, facilitando que sus datos sean recibidos como “evidencia” confiable tanto en espacios públicos como mediáticos.

No obstante, la composición real de su financiación y sus redes de vínculo — que incluyen apoyos de conglomerados financieros, energéticos y digitales, así como participación en plataformas empresariales como Educar 2050 o REDUCA — revelan que su posición está integrada en un entramado de intereses privados. Esta estructura condiciona tanto los problemas que pone en agenda como las soluciones que propone: la producción de indicadores, la priorización de pruebas estandarizadas y la promoción de modelos de gestión inspirados en los principios de la Nueva Gestión Pública, tales como eficiencia, rendición de cuentas y métricas comparativas.

El efecto conjunto es doble. Por un lado, la presentación pública de la organización como una entidad “técnica” y “autónoma” le permite operar como intérprete legítimo de los datos y ejercer influencia en decisiones públicas a través de asesorías, convenios, plataformas de datos y campañas. Por otro lado, su dependencia del financiamiento privado y las alianzas

corporativas imprimen una dirección política a esa evidencia: lo que se mide, cómo se mide y con qué finalidad no son neutrales, sino que reflejan prioridades que suelen favorecer soluciones alineadas con intereses de mercado — por ejemplo, evaluar para gestionar recursos, promover la elección y competencia escolar, o legitimar intervenciones focalizadas. En suma, la imagen de independencia funciona como un capital simbólico que amplifica su influencia en un contexto político donde las posturas anties-tatales y las críticas al funcionamiento del sector público se han intensificado. Al mismo tiempo, su inserción en redes empresariales orienta sus agendas y propuestas hacia modelos tecnocráticos e instrumentales de gestión educativa, compatibles con lógicas de mercado y eficiencia.

Reflexiones finales

El análisis desarrollado evidencia que, durante la última década en Argentina, las propuestas de evaluación centradas en la medición estandarizada del aprendizaje han funcionado como instrumentos estratégicos en los procesos de privatización del sistema escolar. Las interrelaciones entre evaluación y privatización revelan que estos procesos no actúan aisladamente, sino que generan nuevas dinámicas que buscan transformar el funcionamiento del sistema educativo, orientándolo hacia la eficiencia, la competencia y la rendición de cuentas. A su vez, reflejan la ampliación del espacio para la intervención privada en la política educativa, que da lugar a mecanismos de gobernanza híbridos en los cuales lo público y lo privado convergen, a veces tensionando sus límites.

Las implicancias de esta convergencia trascienden la mera gestión y provisión de servicios, impactando en la definición y valoración de la calidad educativa y en la construcción de los sujetos involucrados — estudiantes y docentes. En este sentido, la evaluación se consolida como un dispositivo regulador que reproduce lógicas de competencia fundamentadas en el mérito y el esfuerzo individual, limitando la inclusión de perspectivas más amplias y contextualizadas sobre la educación, que queda reducida a la cuantificación de parámetros preestablecidos.

El caso de Argentinos por la Educación ejemplifica la tendencia creciente del sector privado a ocupar un rol central en la definición de la agenda

educativa en el país. Esta organización ha logrado consolidarse como un actor legítimo a partir de la producción y difusión de indicadores cuantitativos, principalmente derivados de pruebas estandarizadas de aprendizaje, que son utilizados para diagnosticar y orientar políticas públicas.

No obstante lo anterior, la presentación de Argentinos por la Educación como un ente técnico y autónomo debe analizarse críticamente, dado que su estructura institucional, las fuentes de financiamiento y su inserción en redes empresariales regionales revelan una orientación política. Esta se alinea con una visión de la educación subordinada a las demandas del mercado, cuyo foco principal es la formación de capital humano para responder a intereses económicos. En consecuencia, sus recomendaciones e intervenciones no pueden considerarse neutrales o desinteresadas, sino que reflejan y promueven ciertos marcos ideológicos y políticos específicos.

Esta posición se articula con la difusión persistente de una narrativa de “crisis”, la cual se fundamenta en los resultados de las pruebas y se interpreta a través del paradigma del capital humano. Bajo esta lógica, la educación es redefinida como un bien competitivo que debe ser evaluado, controlado y comparado en función del desempeño medible de sus actores — estudiantes, docentes y escuelas. Este enfoque implica que la responsabilidad por los resultados recae en los individuos y las instituciones, mientras que se invisibilizan o minimizan las condiciones estructurales de desigualdad social, económica y cultural que afectan los procesos cotidianos de enseñanza y aprendizaje en las escuelas argentinas. Así, organizaciones del sector privado como Argentinos por la Educación no solo reproducen un diagnóstico reducido, sino que también contribuyen a la difusión de un marco interpretativo que legitima políticas de *accountability*, alineadas con intereses de mercado.

Estas dinámicas, sumadas a la continua reducción del gasto público, el desmantelamiento de la administración estatal y la progresiva transferencia de funciones públicas al sector privado, evidencian un debilitamiento del liderazgo estatal en la definición y conducción de las políticas educativas. En este contexto de privatización de áreas sociales clave como la educación y de su evaluación en particular, se intensifica el riesgo de captura corporativa, donde organizaciones vinculadas al sector empresarial

adaptan sus indicadores a intereses particulares. Este fenómeno se materializa, por ejemplo, en la promoción de rankings escolares que incentivan la competencia interinstitucional y conducen a una profundización de la segmentación y exclusión dentro del sistema educativo.

Frente a este panorama, resulta fundamental fortalecer las resistencias que emergen desde ámbitos académicos, gremiales y sociales, las cuales advierten sobre los riesgos de fragmentación, pérdida de soberanía educativa, reduccionismo y vaciamiento del carácter público de la escuela que implican los procesos analizados. Estas voces críticas subrayan la necesidad de cuestionar las lógicas de mercado y competencia que subyacen a la medición estandarizada, proponiendo en cambio un enfoque que reconozca la complejidad y especificidad contextual de la educación. En este sentido, fortalecer dichas resistencias no solo constituye una estrategia para preservar la educación como derecho social, sino también para construir alternativas que desafíen su mercantilización y la reducción de la enseñanza a meros indicadores cuantitativos.

Referências

ABOITES, H. **La medida de una nación**. Los primeros años de la evaluación en México. Historia de poder, resistencia y alternativas (1982–2012). Buenos Aires: CLACSO, 2012.

ABUSAMRA, V.; ALZÚ, M. S.; VOLMAN, V. 3er Informe de monitoreo: ¿Qué métodos o enfoques de alfabetización utilizan las provincias? **Observatorio Argentinos por la Educación**, [S. l.], 2025. Disponible en: https://argentinosporlaeducacion.org/wp-content/uploads/2025/02/3er-informe-de-monitoreo_-%C2%BFQue-metodos-o-enfoques-de-alfabetizacion-utilizan-las-provincias_.pdf. Acceso: jun. 2025.

ALZÚ, M. S.; SÁENZ GUILLÉN, L.; VOLMAN, V. 4to Informe de monitoreo: Campaña Nacional por la Alfabetización. **Observatorio Argentinos por la Educación**, [S. l.], 2025. Disponible en: https://argentinosporlaeducacion.org/wp-content/uploads/2025/03/4to-informe-de-monitoreo_-Campaña-Nacional-por-la-Alfabetizacion.pdf. Acceso: jun. 2025.

BALL, S. J.; JUNEMANN, C. **Networks, new governance and education**. Bristol: Policy Press, 2012.

BALL, S.; YOUDELL, D. Hidden privatisation in public education. **Journal of Education Policy**, [S. l.], v. 22, n. 5, p. 419–438, 2007. Disponible en: https://www.researchgate.net/publication/228394301_Hidden_privatisation_in_public_education. Acceso: jun. 2025.

BANCO MUNDIAL. **Profesores excelentes**: cómo mejorar el aprendizaje en América Latina y el Caribe. Washington, DC: Banco Mundial, 2014. Disponible en: https://openknowledge.worldbank.org/bitstream/handle/10986/20488/Great%20Teachers_Spanish.pdf?sequence=8&isAllowed=y. Acceso: may. 2025.

BANCO MUNDIAL. **Informe sobre el desarrollo mundial**: Aprender para hacer realidad la promesa de la educación. Washington, DC: Banco Mundial, 2018. Disponible en: <https://www.worldbank.org/en/publication/wdr2018>. Acceso: may. 2025.

BANCO MUNDIAL. **Actuemos ya para proteger el capital humano de nuestros niños**. Los costos y la respuesta ante el impacto de la pandemia de covid-19 en el sector educativo de América Latina y el Caribe, Washington, DC: Banco Mundial, 2021. Disponible en: <https://documents1.worldbank.org/curated/en/803991615924634007/pdf/Acting-Now-to-Protect-the-Human-Capital-of-Our-Children-The-Costs-of-and-Response-to-COVID-19-Pandemics-Impact-on-the-Education-Sector-in-Latin-America-and-the-Caribbean.pdf>. Acceso: abr. 2025.

BANCO MUNDIAL; BILL & MELINDA GATES FOUNDATION; FCDO; UNESCO; UNICEF; USAID. **Guía para la recuperación y aceleración de los aprendizajes**: Usando el marco RÁPIDo para abordar las pérdidas de aprendizajes por COVID-19 y construir mejor en el futuro. Washington, DC: Banco Mundial, 2022.

BARROS, M. ¿Derechos o privilegios? El discurso libertario en su laberinto. **Revista Bordes**, José C. Paz, 2024. Disponible en: <https://revistabordes.unpaz.edu.ar/derechos-o-privilegios/>. Acceso: jul. 2025.

CASTELLANI, A. ¿Qué hay detrás de las fundaciones y ONGs educativas? Las redes de influencia público-privadas en torno a la educación argentina (2015–2018). Informe de investigación, n. 6. **Observatorio de las Elites**, CITRA, Argentina, 2019. Disponible en: https://citra.org.ar/wp-content/uploads/2020/01/2019_informe_6_Observatorio_de_las_Elites.pdf. Acceso: may. 2025.

CASTELLANI, A. Funcionarios públicos, intereses privados. Análisis de la participación en sociedades comerciales y asociaciones empresariales del gabinete nacional. Informe de Investigación, n. 12. **Observatorio de las Elites**, CITRA, Argentina, 2024. Disponible en: <https://citra.org.ar/informe-de-investigacion-n12-funcionarios-publicos-intereses-privados-analisis-de-la-participacion-en-sociedades-comerciales-y-asociaciones-empresariales-del-gabinete-nacional-argentina-202/>. Acceso: may. 2025.

CEPAL. **Informe especial COVID-19 N° 4: Educación y desigualdad en América Latina y el Caribe**. Santiago, CL: Comisión Económica para América Latina y el Caribe, 2021. Disponible en: <https://www.cepal.org/es/publicaciones/45337-america-latina-caribe-la-pandemia-covid-19-efectos-economicos-sociales>. Acceso: may. 2024.

CENTRO DE IMPLEMENTACIÓN DE POLÍTICAS PÚBLICAS PARA LA EQUIDAD Y EL CRECIMIENTO (CIPPEC). **Argentina en las pruebas PISA 2012: Análisis y recomendaciones para políticas educativas**. Buenos Aires: Secretaría de Evaluación Educativa; Ministerio de Educación de la Nación, 2017. Disponible en: https://www.argentina.gob.ar/sites/default/files/informe_pisa_2012.pdf. Acceso: may. 2025.

CURCIO, J.; ALZÚ, M. S.; SÁENZ GUILLÉN, L. **Presupuesto educativo nacional 2025**. Buenos Aires: Observatorio Argentinos por la Educación, 2024. Disponible en: <https://argentinosporlaeducacion.org/wp-content/uploads/2024/10/Presupuesto-Educativo-Nacional-2025.pdf>. Acceso: may. 2025.

DIKER, G.; BRISCIOLI, B.; STOPPANI, N. La institucionalización de la evaluación educativa en Argentina: disputas por sus sentidos y usos. **Jornal de Políticas Educacionais**, v. 17, 2023. Disponible en: <http://educa.fcc.org.br/pdf/jpe/v17/1981-1969-jpe-17-e92216.pdf>. Acceso: jun. 2025.

FELDFEBER, M. La nueva gestión pública en la “revolución educativa” del macrismo. Las políticas de formación docente en Argentina (2015–2019). In: MONARCA, H. (coord.). **Profesionalización docente: discursos, políticas y prácticas**. Madrid: Editorial Dykinson, 2024, p. 240–263.

FELDFEBER, M.; GLUZ, N.; RODRIGO, L. Gobernar a través de las emociones: intereses privados y redes políticas en Argentina: Gobernar a través de las emociones: intereses privados y redes de política en Argentina. **Revista Cocar**, [S. l.], n. 43, 2025. Disponible en: <https://periodicos.uepa.br/index.php/cocar/article/view/10934>. Acceso: nov. 2025.

FELDFEBER, M.; CARIDE, L.; DUHALDE, M. **Privatización y mercantilización educativa en Argentina**. Formación de subjetividades y construcción de sentido común durante el Gobierno de Cambiemos (2015–2019). Buenos Aires: Ediciones CTERA, 2020. Disponible en: <https://mediateca.ctera.org.ar/privatizacion-y-mercantilizacion-educativa-en-argentina/>. Acceso: abr. 2025.

FELDFEBER, M.; PUIGGRÓS, A.; ROBERTSON, S.; DUHALDE, M. **La privatización educativa en Argentina**. Buenos Aires: Ediciones CTERA, 2018. Disponible en: <https://www.ei-ie.org/es/item/25693:la-privatizacion-educativa-en-argentina>. Acceso: abr. 2025.

GLUZ, N.; KAROLINSKI, M.; KESLER, A. El anarcocapitalismo y la mercantilización de la vida: extremas derechas en la educación argentina a partir del gobierno de Javier Milei. In: GRUPO DE TRABAJO POLÍTICAS EDUCATIVAS Y DERECHO A LA EDUCACIÓN (CLACSO). **Estado y Derecho a la educación en América Latina y el Caribe**, n. 3, p. 8–24, set. 2024. Disponible en: https://www.clacso.org/wp-content/uploads/2024/10/V2_Estado-y-Derecho-a-la-educacion_N3-1.pdf. Acceso: jun. 2025.

LENZI, A.; PISANI, M.; VOLMAN, V. Primer Informe de monitoreo: Campaña Nacional por la Alfabetización. **Observatorio Argentinos por la Educación**, 2024. Disponible en: <https://argentinosporlaeducacion.org/informe/1er-informe-de-monitoreo-campana-nacional-por-la-alfabetizacion/>. Acceso: jul. 2025.

LLECE; OREALC/UNESCO. **Los aprendizajes fundamentales en América Latina y el Caribe**: Evaluación de logros de los estudiantes (ERCE 2019). Informe de resultados de logros de aprendizaje y factores asociados. Santiago de Chile, 2021. Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000380257>. Acceso: may. 2025.

LÓPEZ-LEAVY, M. Más allá de los números: los usos políticos de los datos de la evaluación nacional Aprender en la formulación de políticas educativas. Un análisis documental de políticas nacionales en Argentina (2016–2019). **Jornal de Políticas Educacionais**. V. 17, e92211, Dezembro de 2023. Disponible en: <http://educa.fcc.org.br/pdf/jpe/v17/1981-1969-jpe-17-e92211.pdf>. Acceso: may. 2025.

OBSERVATORIO ARGENTINOS POR LA EDUCACIÓN. **COVID-19**: ¿Qué hicieron los países para continuar con la educación a distancia? Buenos Aires: Observatorio Argentinos por la Educación, 2020. Disponible en: <https://argentinosporlaeducacion.org/informe/covid-19-que-hicieron-los-paises/>. Acceso: may. 2025.

OZGA, J. Knowledge, inspection and the work of governing. **Journal of Education**, [S. l.], v. 2, n. 1, p. 6–38, 2014. Disponible en: <https://www.redalyc.org/pdf/5757/575763889002.pdf>. Acceso: jun. 2025.

PECK, J.; THEODORE, N. Mobilizing policy: models, methods, and mutations. **Geoforum**, [S. l.], v. 41, n. 2, p. 169–174, 2010. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0016718510000047?via%3Dihub>. Acceso: abr. 2025.

POPKEWITZ, T. PISA: Números, estandarización de la conducta y la alquimia de las materias escolares. Profesorado. **Revista de Currículum y Formación del Profesorado**, [S. l.], v. 17, n. 2, p. 47–64, 2013. Disponible en: <https://www.redalyc.org/pdf/567/56729526004.pdf>. Acceso: jun. 2025.

PUIGGRÓS, A.; DUHALDE, M.; PASCUAL, L.; ALBERGUCCI, L.; ABAL MEDINA, M. D.; NUÑEZ, A.; MARTÍNEZ, G. **Informe de política educativa: participación y actores privados**. Buenos Aires: Ediciones CTERA, 2025. Disponible en: <https://ctera.org.ar/wp-content/uploads/2025/03/Privatizacion-y-mercantilizacion-educativa-en-la-Argentina.pdf>. Acceso: jul. 2025.

RAVELA, P. **Para comprender las evaluaciones educativas**. Fichas didácticas. Santiago de Chile: Ed. San Marino, 2006.

RAMOS-ZINCKE, C. Dispositivo de evaluación y gubernamentalidad del sistema educacional: entretejimiento de ciencia social y poder. **Cinta de Moebio**, [S. l.], n. 61, p. 41–55, 2018. Disponible en: <https://cintademoebio.uchile.cl/index.php/CDM/article/view/48545/53609>. Acceso: mayo 2025.

RODRÍGUEZ MOYANO, I.; RODRIGO, L. La evaluación de la calidad de la educación en Argentina. Tendencias y trayectoria de la política educativa nacional (1990–2022). **Revista Latinoamericana de Políticas y Administración de la Educación**, [S. l.], v. 9, n. 17, p. 123–145, 2022. Disponible en: <https://revistas.untref.edu.ar/index.php/relapae/article/view/1479>. Acceso: may. 2025.

RODRIGO, L.; RODRÍGUEZ MOYANO, I. La evaluación docente en Argentina (2015–2019): el programa “Enseñar” y la jerarquización de la cultura del rendimiento. **Revista ETD – Educação Temática Digital**, [S. l.], v. 25, p. 1–19, 2023. Disponible en: <https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/8671374/32591>. Acceso: may. 2025.

RODRÍGUEZ, L.; VIOR, S.; MÁZ ROCHA, S. M. Las políticas de evaluación de la calidad educativa en Argentina (2016–2018). **Educación Real**, [S. l.], v. 4, n. 43, p. 1405–1428, 2018. Disponible en: <https://www.scielo.br/j/edreal/a/QjK6jsgYVv-W6DNDfmHr87xQ/?lang=es>. Acceso: may. 2025.

RUIZ, G. R. La política educativa del gobierno nacional: el Plan de Alfabetización y la ESI como parte de la recomposición subsidiaria–paleolibertaria. **Praxis Educativa**, [S. l.], v. 29, n. 2, 2025. Disponible en: <https://cerac.unlpam.edu.ar/index.php/praxis/article/view/8878>. Acceso: jun. 2025.

SAURA, G. Redes políticas y redes de datos de gubernamentalidad neoliberal en educación. **Foro de Educación**, v. 19, n. 1, p. 1-10, 2021. Disponible en: <https://dialnet.unirioja.es/descarga/articulo/8023715.pdf>. Acceso: may. 2025.

TORRE, E.; PERUSIA, J. C.; XANTHOPOULOS, J. **Hacia una política nacional de evaluación educativa a gran escala al servicio de la mejora de los aprendizajes** (Documento de Política Pública n° 244). CIPPEC, 2023. Disponible en: <https://www.cippec.org/wp-content/uploads/2023/08/DPP-244-EDU-Hacia-una-politica-nacional-de-evaluacion-educativa-08.23.pdf>. Acceso: may. 2025.

UNESCO. **Informe de resultados regionales LLECE 2022**: La situación educativa en América Latina. París: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022. Disponible en: <https://www.unesco.org/es/fieldoffice/santiago/expertise/llece>. Acceso: may. 2025.

UNICEF (2021). COVID-19 and School Closures. One year of education disruption. Disponible en: <https://data.unicef.org/resources/one-year-of-covid-19-and-school-closures/> Acceso: may. 2025.

VERGER, A.; MOSCHETTI, M.; FONTDEVILA, C. **La privatización educativa en América Latina**. Una cartografía de políticas, tendencias y trayectorias. IEAL, 2017. Disponible en: <https://download.ei-ie.org/Docs/WebDepot/Privatizacion%201-Abril.pdf>. Acceso: may. 2025.

VERGER, A.; NORMAND, R. Nueva Gestión Pública y Educación: elementos teóricos y conceptuales para el estudio de un modelo de reforma educativo global. **Revista Educação & Sociedade**, [S. l.], v. 36, n. 132, p. 599-622, 2015. Disponible en: <https://www.scielo.br/j/es/a/fQMzP4tSxyBjFKrwMHZsVYB/?lang=es>. Acceso: abr. 2025.

ZIEGLER, S.; CATRI, G.; ORLICKI, E. ¿Qué opinan las familias sobre la educación argentina y la de sus hijos? **Observatorio Argentinos por la Educación**, 2022. Disponible en: <https://argentinosporlaeducacion.org/informe/ene-2022/>. Acceso: may. 2025.

CAPÍTULO 6

POLÍTICAS DE BONIFICAÇÃO E GESTÃO EDUCACIONAL: CENÁRIOS

Cristiane Machado¹

Paulo Henrique Arcas²

DOI: 10.52695/978-65-5456-163-1.6

Introdução

A centralidade das avaliações externas e em larga escala na formulação e gestão das políticas educacionais na Educação Básica tem se ampliado desde a implantação da primeira versão do Saeb na década de 1990. Observa-se, especialmente nos últimos 20 anos, após a criação da Prova Brasil (2005) e do Índice do Desenvolvimento da Educação Básica – Ideb (2007), a consolidação de tais avaliações como instrumentos de gestão educacional de sistemas, redes de ensino e escolas públicas em todos os níveis da federação.

Bonamino e Sousa (2012) analisaram diferentes gerações das avaliações externas e em larga escala no Brasil de acordo com os objetivos e desenhos de tais iniciativas; mais especificamente, as autoras destacam

1 Universidade Estadual de Campinas (UNICAMP). E-mail: crimacha@unicamp.br.

2 Universidade Federal de Lavras (UFLA).

três momentos distintos no percurso histórico de implementação das políticas de avaliação.

A primeira geração caracteriza-se pelo propósito de fornecer um diagnóstico e acompanhamento da qualidade da educação no país, sem implicações materiais diretas para as escolas. “De um modo geral, essas avaliações divulgam seus resultados na *Internet*, para consulta pública, ou utilizam-se da mídia ou de outras formas de disseminação, sem que os resultados da avaliação sejam devolvidos para as escolas” (Bonamino; Sousa, 2012, p. 375).

Já as avaliações de segunda geração, de acordo com as autoras, acrescentam à divulgação pública “a devolução dos resultados para as escolas, sem estabelecer consequências materiais” (Bonamino; Sousa, 2012, p. 375) como forma de estímulo a partir da divulgação pública dos resultados alcançados. Depreende-se que, nessas duas primeiras gerações das iniciativas de avaliação externa e em larga escala, as consequências dos resultados obtidos para as redes e escolas são somente simbólicas, sem implicações materiais diretas.

A terceira e última geração dessas avaliações, de acordo com as autoras, é aquela que geram consequências concretas para as escolas a partir dos resultados das avaliações, também conhecidas como políticas de responsabilização. Essas políticas de responsabilização se pautam pela instituição de prêmios e/ou bônus com base no alcance de metas pré-estabelecidas, comumente conhecidas como políticas de bonificação, nesse caso também chamadas de *high-stakes* ou de responsabilização forte; quando as consequências dessas políticas são somente simbólicas, elas são chamadas de *low-stakes* ou de responsabilização branda (Brooke, 2006).

É no contexto dessa terceira geração que este artigo se insere. O objetivo é, com base em levantamento de artigos científicos, identificar políticas de bonificação a partir dos resultados das avaliações externas e analisar como essas políticas incidem sobre a gestão da educação, expressando cenários no contexto educacional.

Procedimentos metodológicos

O levantamento dos artigos foi elaborado nos meses de maio e junho do corrente ano, nos repositórios do Scielo, Capes e Educ@, com o descritor “bonificação”, sem o uso de filtros. Foram encontrados 109 títulos; retirando as repetições, restaram 61 artigos. Foi feita a leitura dos resumos desse total e foram desprezados 39 títulos que não tratavam diretamente do tema em relação à gestão. Restaram 22 artigos que compuseram a amostra do estudo e foram lidos na íntegra para serem analisados.

Dentre os artigos selecionados, nove tiveram como objeto as políticas de bonificação da rede estadual paulista, dois tiveram como foco as políticas do estado do Ceará, um artigo pesquisou a política do estado de Pernambuco, um, a política do estado do Acre e outro, a política do município de Rio Branco (AC), além de um artigo que estudou a política de bonificação do município de Esteio (RS).

Os resultados do Saeb/Prova Brasil foram analisados em dois artigos, enquanto os municípios fluminenses do Rio de Janeiro e Duque de Caxias tiveram os dados das respectivas avaliações municipais e Saeb/Prova Brasil tomados como objeto de uma pesquisa. Uma pesquisa analisou a avaliação do desempenho docente (ADD) nas redes públicas de Educação Básica no Brasil, assim como dois artigos analisaram a influência dos organismos internacionais, Banco Mundial (BM) e Organização para a Cooperação e Desenvolvimento Económico (OCDE), nas políticas de bonificação. Por fim, um artigo discorreu sobre os embates teóricos e os diferentes posicionamentos sobre as políticas de incentivo e bonificação.

Dessa forma, pode-se dizer que a amostra selecionada contempla sistemas de avaliação estaduais, municipal e estadual, que analisam situações mais específicas, enquanto outros, a partir de uma análise mais documental, realizaram análises do Sistema de Avaliação da Educação Básica (Saeb), assim como análises da influência dos organismos internacionais. Apenas um artigo fez uma análise teórica dos usos da avaliação externa no Brasil e as políticas de bonificação e premiação dos profissionais da educação pelo desempenho nas avaliações.

Com relação aos anos de publicação dos artigos selecionados, a tabela a seguir apresenta o quantitativo por ano.

Tabela 1 – Ano de publicação dos artigos selecionados

2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
1	1	3	1	1	1	5	1	1	2	5

Fonte: Elaborado pelos autores.

A análise elaborada foi inspirada nos pressupostos de Bardin (2016), especialmente no que se refere à interpretação inferencial, e, a partir daí, foram construídos agrupamentos dos artigos por aproximação temática.

Descrição dos artigos da amostra

Bauer (2012) analisa a política educacional paulista tendo como foco os resultados do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saesp), procurando relacioná-los com o direcionamento de ações de formação de professores. A partir da análise dos documentos oficiais, a autora questiona a pertinência de propor programas e ações de formação e bonificação de professores a partir dos resultados do Saesp. Para Bauer (2012, p. 68):

[...] nos documentos analisados, não eram explicitados os motivos ou pressupostos teóricos que reforçavam a ideia de que haveria uma inter-relação entre os saberes dos alunos, expressados nos resultados do Saesp, e a formação dos professores, deixando a impressão de que tal relação, inquestionável, estava dada, ou seja, que havia uma relação entre a formação docente e os resultados obtidos nos testes.

A autora concluiu que não há uma relação direta entre as ações e programas de formação e bonificação de professores e as avaliações externas. Além disso, explicita a necessidade de outros estudos sobre a

temática que tragam mais evidências sobre os usos que as políticas educacionais centradas na gestão racionalizada do sistema educacional vêm fazendo das avaliações externas.

O artigo de Brooke (2013) analisa diferentes posicionamentos sobre as políticas de responsabilização e as próprias políticas, também tratadas pelo autor como políticas de incentivos monetários, ressaltando que o debate sobre o tema é eivado de argumentos ideológicos, impelindo alguns pesquisadores à formulação de críticas severas contra as bonificações, enquanto outros salientam aspectos positivos da implementação dos incentivos monetários como mecanismo de gestão e melhoria da qualidade. Em relação às críticas negativas na implementação da bonificação, tem relevo o argumento de que o estímulo financeiro pode levar os melhores professores a procurarem melhores escolas, o que geraria um incremento na já profunda desigualdade no sistema educativo nacional. Defende o autor que as consequências das políticas de incentivos, positivas ou negativas, dependem do seu desenho e da metodologia utilizada na formulação da política de incentivo e que, se esta for apropriada, é possível combater as desigualdades sociais do sistema.

A pesquisa de Sousa, Maia e Haas (2014) se insere no debate da relação entre o uso da avaliação externa como mecanismo de gestão da melhoria da qualidade do ensino ofertado. As autoras investigaram as variações do Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp), produzido com base nas aferições do Saresp, entre 2008 e 2012, das turmas do 3º ano do Ensino Médio de 41 unidades da rede pública de um município da Grande São Paulo, cotejando-os com as metas estabelecidas para as escolas a cada ano e o recebimento ou não da bonificação. O estudo de caráter exploratório teve o objetivo de responder às questões: "O cumprimento de metas pode ser interpretado como melhoria de qualidade das escolas? A premiação vem incidindo em escolas que evidenciam aprimoramento de qualidade?" (Sousa; Maia; Haas, 2014, p. 192).

Concluem as autoras, com base nos dados analisados no período, que há "inconsistência na relação entre Idesp e premiação anual de escolas, bem como põem em questão seus efeitos no sentido de induzir avanços cumulativos por parte das escolas" (Sousa; Maia; Haas, 2014, p. 204) e que,

portanto, as respostas para as duas questões que pautaram o desenvolvimento do estudo são negativas. Não é possível afirmar que há relação entre cumprimento de metas e evidências de melhoria de qualidade das escolas, como também não se verificou incidência positiva no aumento do Idesp das escolas com base na premiação anual a partir dos dados do Saresp.

O artigo de Novaes (2014) se insere no escopo de pesquisas que investigaram a relação do trabalho docente com a implementação das políticas de bonificação. O autor analisa, por meio de entrevistas semiestruturadas, as percepções de 21 professores que atuam nos Anos Iniciais do Ensino Fundamental em três escolas da rede estadual paulista com diferentes índices de desempenho apontado pelo Idesp, produzidos com base nas aferições do Saresp, sobre os impactos dessa avaliação no trabalho pedagógico. O exame das entrevistas permitiu ao autor afirmar que a avaliação externa, na percepção das professoras entrevistadas, tem servido como instrumento de controle do trabalho docente e do currículo ao adotar a hierarquização e a classificação de escolas por meio da produção de *rankings* dos resultados obtidos no Idesp, além de contribuir com a implementação de mecanismos de premiação/castigo de professores por meio dos bônus anuais.

Para o autor, a associação da avaliação externa com a política de bonificação tem resultado em “adoção de práticas questionáveis”, como a adequação do currículo escolar aos descritores do Saresp, retirando da escola a autonomia para gerir o currículo, além do “treinamento para a prova, mediante a manutenção de cursinhos informais e ‘provões’, o que desrespeita o professor e a escola na escolha dos seus modos de avaliação”. Ressalta o autor que esse cenário estimula a criação de “novas subjetividades e maneiras nem sempre claras de se relacionar com a profissão”, o que tem consequência na autoestima dessas profissionais e no aumento da angústia e da frustração com a carreira (Novaes, 2014, p. 304).

O trabalho de Rosistolato e Prado (2014) discute as contradições que se apresentam para as escolas que devem dar respostas às iniciativas governamentais ao implementarem sistemas de avaliação externa e lidarem com processos de aceitação e resistência. Portanto, cabe “indagar sobre as visões construídas durante sua recepção no cotidiano escolar” (Rosistolato; Prado, 2014, p. 312). Para tanto, o estudo teve como foco avaliação

externa nos municípios do Rio de Janeiro, com bonificação, e de Duque de Caxias, sem bonificação. No contexto da pesquisa, discutiu-se em ambos os municípios a avaliação nacional e as avaliações próprias das duas redes. Metodologicamente, a pesquisa se desenvolveu por meio da realização de grupos focais com gestores e professores. Segundo Rosistolato e Prado (2014, p. 316):

As falas revelam algumas especificidades das avaliações externas e, principalmente, dos processos de responsabilização. Professores e gestores resistiram menos às avaliações externas do que aos processos de responsabilização com oferecimento de bonificação. Houve falas contrárias à bonificação por parte daqueles que não a receberam, mas os que receberam também indicaram problemas, principalmente por não compreenderem quais seriam exatamente os critérios definidores da concessão da bonificação.

Concluíram que as avaliações externas vêm incidindo sobre a gestão escolar, principalmente em função da preocupação com os índices, mas também entre os docentes, explicitando influências positivas, como identificar lacunas nos processos de formação dos estudantes, e influências negativas no cotidiano escolar, com foco na preparação para as provas, visando alcançar os índices estabelecidos. Além disso, os autores explicitam que há espaços para resistências e lacunas na gestão reveladas por meio de “[...] de certo desalinhamento entre as instâncias de gestão, de forma que decisões estabelecidas na gestão central de Secretarias Municipais não são, necessariamente e, cumpridas nas instâncias intermediárias e escolares” (Rosistolato; Prado, 2014, p. 315).

Oshiro, Scorzafave e Dorigan (2015) desenvolveram um estudo estatístico com o uso da metodologia de *Propensity Score Matching* e de Diferenças em Diferenças, que consiste em encontrar um grupo de controle mais parecido com o grupo tratamento, com base em características observáveis para comparar os resultados ao longo do tempo e eliminar diferenças não observáveis fixas no tempo entre as escolas estaduais paulistas afetadas

pelo tratamento e as não afetadas. O estudo foi desenvolvido com base nos dados da Prova Brasil de 2005 a 2011, embora sejam os resultados do Saresp a base da bonificação na rede estadual paulista. Explicam os autores que essa escolha se baseou no fato de a Prova Brasil ser nacional e possibilitar a utilização de grupos de controle de outros estados, o que não seria possível com o Saresp. Os autores afirmam que encontraram:

[...] indícios de impacto positivo do programa de incentivo de professores apenas nas notas do quinto ano do ensino fundamental em 2009, de 6,3 pontos na escala SAEB em Matemática e de 2,6 pontos em Língua Portuguesa. No entanto, há forte decaimento desses efeitos entre 2009 e 2011. Para o nono ano, encontramos resultados negativos; mas nas análises de robustez para 2009, em alguns casos observou-se ausência de efeito da política. No entanto, para 2011, praticamente todas as estimativas confirmaram um impacto negativo (Oshiro; Scorzafave; Dorigan, 2015, p. 230).

Os autores tentaram também avaliar se a adoção do bônus alterou alguns indicadores escolares, como clima escolar, participação dos pais na escola, cumprimento do currículo, uso do tempo em sala de aula, experiência dos professores/diretores, etc., mas a análise foi inconclusiva.

Piovezan e Dal Ri (2016) efetivaram pesquisa com o objetivo de compreender as implicações das reformas educacionais no estado de São Paulo, nos últimos 20 anos, para a carreira da categoria docente. Concluíram as autoras, pelos documentos e legislações analisadas, a existência de uma intensificação e flexibilização das formas de contratação dos professores no período, bem como o arrocho salarial provocado por baixos salários e o incentivo à expansão de políticas que estimulam a remuneração dos docentes por meio de bonificação, promovendo um acelerado processo de precarização do trabalho docente. Especificamente sobre a bonificação, as autoras destacaram que a implementação de diversas avaliações para gerar o bônus tem a finalidade de controle e de vigilância do trabalho dos professores e, ao mesmo tempo, de camuflar o arrocho

salarial. “Os valores monetários distribuídos pelos bônus levaram os professores a ‘assimilarem’ com maior facilidade a defasagem dos salários percebidos” (Piovezan; Dal Ri, 2016, p. 191).

Koslinski, Ribeiro e Oliveira (2017) utilizaram as metodologias *Propensity Score Matching* e regressões múltiplas para análise exploratória do impacto do uso do Indicador de Desempenho Escolar – Alfabetização (IDE–Alfa) no aumento da proficiência dos estudantes pela política de incentivo financeiro às escolas chamada de “Prêmio Escola Nota Dez” do estado do Ceará, usado como parâmetro para a premiação para escolas com alto desempenho e apoio financeiro para escolas públicas com baixo desempenho no estado.

Para as autoras, com base na análise elaborada, os resultados indicam aumento no desempenho das escolas de todos os grupos: as que são premiadas por terem alto desempenho e as que são apoiadas por terem baixo desempenho. “Os resultados encontrados mostram que o recebimento do prêmio foi acompanhado por um efeito positivo sobre a proficiência média das escolas”, entretanto “tal tendência (efeito positivo sobre a proficiência) não foi verificada nas análises que focalizaram as escolas apoiadas” (Koslinski; Ribeiro; Oliveira, 2017, p. 842). Concluem que somente o apoio técnico–financeiro não tem se mostrado suficiente para promover melhoria na proficiência dos alunos das escolas de baixo rendimento, ainda de acordo com outras pesquisas desenvolvidas nos mesmos moldes.

Carvalho (2018) investigou a relação das professoras do Ensino Fundamental 1 com o pagamento de bonificação pelo cumprimento de metas por meio de estudo que combinou observação e entrevistas semiestruturadas. O debate se insere em pesquisa mais ampla sobre mudanças nas configurações de gênero no trabalho docente no Ensino Fundamental 1, a partir da adoção de formas de gestão empresariais na escola pública, e parte do princípio de que esse tipo de gestão vem exigindo dos educadores e educadoras posturas voltadas à eficiência e à prestação de contas, fortalecendo subjetividades individualizadas, competitivas, assertivas, planejadas, típicas do mercado e orientadas para a consecução de metas e para a ascensão na carreira por meio de recompensa monetária.

A autora conclui que há a coexistência e a combinação de dois modelos de trabalho docente: o da feminilidade, dimensão associada à dedicação, culpa e amor ao ensino; e o da lógica da gestão empresarial, que enfatiza a recompensa financeira, o individualismo e a competição.

Fonseca e Lima (2018) analisaram vários aspectos da política educacional implementada no estado de São Paulo a partir de 2008, com ênfase no debate sobre currículo. Em relação à bonificação, as autoras afirmam que os professores têm pautado o trabalho pedagógico com base nos conteúdos do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo, e não do currículo, como forma de incrementar a possibilidade de bonificação, além de adotarem a avaliação externa como referência do seu trabalho.

Furtado e Soares (2018) examinaram o impacto do Bônus de Desempenho Educacional, incentivo financeiro pago mediante o alcance de metas educacionais preestabelecidas aos funcionários da rede estadual de ensino de Pernambuco. A metodologia utilizada na pesquisa consistiu no uso de escores de propensão para produzir pareamentos, comparando-se uma escola da rede estadual sujeita à política de bonificação a uma escola da rede municipal similar à da rede estadual, porém sem política de bonificação.

Os resultados analisados, tanto em relação ao Sistema de Avaliação Educacional de Pernambuco (Saepe) quanto à Prova Brasil, segundo os autores, foram semelhantes e positivos, revelando que “[...] existem boas chances de a política de bonificação ter sido responsável pela variação positiva, ou ainda pela maior parte dos ganhos de desempenho da rede estadual no período analisado” e “por meio do pareamento, pode-se dizer que, desde o início do programa, em 2008, a bonificação demonstrou ganhos favoráveis ao grupo de intervenção” (Furtado; Soares, 2018, p. 71).

Melo e Braidí (2018), por meio de pesquisa documental, analisaram a concepção e a normatização da política de bonificação no Acre e concluíram que o estado tem adotado políticas baseadas na Nova Gestão Pública e que essas têm “gerado uma série de conflitos entre governo, sindicatos e profissionais da educação”, mesmo carecendo “de maiores estudos empíricos que possam avaliar os impactos desta política no processo de construção de novas subjetividades, se de fato se constituem

em instrumentos de controle e de culpabilização dos profissionais do magistério” (Melo; Braidi, 2018, p. 433).

O artigo de Araújo, Leite e Passon (2018), que tomou como objeto de exame a política educacional do estado do Ceará, com foco no Prêmio Escola Nota Dez, buscou identificar os efeitos do prêmio na dinâmica das escolas cearenses. O estudo de abordagem qualitativa fez uso da pesquisa bibliográfica e da pesquisa de campo, utilizando como procedimentos para a coleta de dados entrevistas semiestruturadas com representantes do núcleo de gestão das escolas, professores e alunos do 2º ano do Ensino Fundamental, ano em que é aplicada a prova de alfabetização do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaeece-Alfa). Segundo Araújo, Leite e Passon (2018, p. 101):

O Prêmio Escola Nota Dez caracteriza-se por ser uma política accountability educacional de alocação de recursos, com base em resultados das avaliações do Spaeece, com consequências fortes ou responsabilização high-stakes, visto que a distribuição de recursos, condicionada ao melhor desempenho escolar, provoca a competição entre as escolas, a reorientação das práticas pedagógicas e a preleção de conteúdos curriculares de acordo com as matrizes das avaliações externas.

Para a realização da pesquisa de campo, foram selecionados dois municípios cearenses: Sobral, que possui o maior número de escolas premiadas no Prêmio Escola Nota Dez, e o município de Caucaia, com escolas que recebem apoio por não alcançarem as metas. Em Sobral, foram selecionadas duas escolas premiadas e, em Caucaia, duas escolas apoiadas. O estudo indica que a melhoria dos resultados da alfabetização das crianças no estado do Ceará vem sendo associada, “sobretudo, ao modelo de gestão por resultados do Prêmio Escola Nota Dez, que faz frente para impulsionar e alavancar as melhorias no padrão de desempenho da alfabetização na idade certa” (Araújo; Leite; Passon, 2018, p. 106).

Os autores evidenciaram que o Prêmio Escola Nota Dez influencia na dinâmica das escolas, pois a pesquisa indicou que:

[...] nas escolas premiadas, as educadoras empreendem práticas docentes diversificadas, voltadas prioritariamente para desenvolver habilidades de leitura e de escrita, inclusive com ênfase nos gêneros textuais. Nas escolas apoiadas, por seu turno, as professoras fazem uso, predominantemente, da escrita no quadro para os alunos copiarem e do acompanhamento das atividades nos livros didáticos. Nas escolas nota dez de Sobral, as docentes fazem diariamente a audição da leitura com os educandos; nas escolas apoiadas de Caucaia, por sua vez, isso só ocorre em uma das instituições e por iniciativa da professora titular da sala de 2º ano do ensino fundamental (Araújo; Leite; Passon, 2018, p. 108).

Como conclusão, os autores sinalizam que o Prêmio Escola Nota Dez responsabiliza fortemente os atores educacionais pelos resultados escolares, utilizando-se, para tanto, das avaliações externas. A política de bonificação escolar estimula e provoca mudanças positivas no trabalho pedagógico das escolas, reorientando o planejamento, promovendo novas estratégias de alfabetização, letramento e avaliação; mas, por outra parte, pode pressionar os professores a alcançarem as metas, gerando tensões e desgastes nas relações entre os diferentes atores escolares.

Um achado interessante da pesquisa evidencia que “a avaliação escolar das escolas apoiadas de Caucaia não se encontra subordinada por completo aos padrões e aos objetivos das avaliações externas, como acontece com as escolas premiadas” (Araújo; Leite; Passon, 2018, p. 109), explicitando que, quanto mais a escola segue o modelo de gestão, currículo e avaliação implementado pelas políticas de *accountability*, maior padronização e limitação da autonomia se impõem às escolas.

Zatti e Minhoto (2019) examinaram dispositivos legais que instituíram a Avaliação do Desempenho Docente (ADD) nas redes públicas estaduais de Educação Básica no Brasil, relacionando-a com iniciativas de remuneração dos professores. Fizeram levantamento da legislação em 2015, tanto

nos sítios oficiais das secretarias de educação como por meio do envio de correspondência a esses órgãos, e constataram a existência dessa iniciativa em 24 estados brasileiros; porém, apenas em 14, a política estava efetivamente em vigor no período. Com base nas informações obtidas, os efeitos da ADD na remuneração docente foram classificados em progressão na carreira e bonificação.

Concluíram que, nas ADDs, predominam critérios que avaliam a conduta dos profissionais da educação, fomentando uma “adaptação dos indivíduos a definições prévias do que é ser um bom professor”, ao mesmo tempo que “retira-se dos docentes o exercício de autonomia, em virtude da padronização e da mecanização das condutas, prescritas como necessárias para o alcance de objetivos e metas exógenos e previamente estabelecidos”, aprofundando o controle externo, o que provoca nos docentes a perda do sentido sobre o próprio trabalho (Zatti; Minhoto, 2019, p. 14).

O artigo de Ramos e Moreira (2020) analisou, com base em pesquisa documental, as recomendações da Organização para Cooperação e Desenvolvimento Econômico (OCDE) para as políticas de remuneração de professores. Concluem os autores que as recomendações da OCDE estão pautadas na lógica da Teoria do Capital Humano e na Pedagogia das Competências, segundo as quais o professor ideal é aquele que tem um excelente desempenho e que deve receber reconhecimento financeiro por isso. O estímulo à bonificação docente compõe a concepção, expressa nos documentos, de educação a partir da lógica capitalista, devendo:

[...] trazer maiores benefícios econômicos e competitivos, em meio à mundialização do capital. Entre as palavras mais utilizadas nas recomendações da OCDE estão os termos gerencialismo, recursos, eficiência, todas constantemente presentes nos documentos oficiais (Ramos; Moreira, 2020, p. 1524).

O artigo de Guedes e Beiruth (2021) teve como foco a relação do desempenho de alunos de escolas de tempo integral no Espírito Santo com a concessão de incentivos financeiros aos professores. Para verificar os

efeitos dos incentivos financeiros, foram analisadas as notas das provas do Sistema de Avaliação da Educação Básica (SAEB) em 2017, aplicando, inicialmente, o método de pareamento PSM (*Propensity Score Matching*) e, em seguida, uma regressão Tobit. Evidenciou-se uma associação positiva e relacionada entre os incentivos financeiros (Bonificação por Desempenho) e as notas de língua portuguesa (leitura) e matemática (resolução de problemas) de estudantes do 3º ano do Ensino Médio de escolas de tempo integral. A conclusão atribui, como efeitos da concessão de incentivos financeiros, o sucesso do programa à redução do absenteísmo e à diminuição da rotatividade de professores.

Visando compreender os sentidos e os significados das políticas curriculares da rede estadual de ensino de São Paulo entre 2007 e 2018, Jacomini *et al.* (2021) realizaram pesquisa documental, analisando o Programa São Paulo Faz Escola (2007) e o Programa Escola de Tempo Integral (2012), que caracterizam a adoção de uma política que padroniza o currículo da rede estadual. Tais programas buscaram implementar um currículo mais moderno e remodelar as escolas na perspectiva de pedagogias de habilidades e competências e explicitam a centralidade das políticas públicas baseada em *accountability*, controle e padronização, reduzindo a qualidade da educação ao escopo da “aprendizagem”.

Com relação às avaliações externas e à bonificação, os autores sugerem que as políticas curriculares analisadas levam à “centralização dos processos educacionais, sua gestão na forma de resultados de desempenho e a desvalorização das dinâmicas da sala de aula que levem em consideração as singularidades tão próprias dos processos de aprendizagem” (Jacomini *et al.*, 2021, p. 10).

Também analisando as políticas educacionais do estado de São Paulo no período entre 2007 e 2018, Perrella e Alencar (2022) desenvolveram pesquisa documental sobre os programas Ler e Escrever, Qualidade da Escola e São Paulo Faz Escola, que preconizavam a melhoria da qualidade do ensino centrada nos resultados das avaliações externas. Os programas analisados se baseavam no controle do trabalho pedagógico a partir do uso de materiais padronizados no ensino, associação entre metas e índices,

desempenho dos alunos nas avaliações externas e bonificação e papel técnico do coordenador pedagógico como executor do currículo.

Os autores ressaltam que, no contexto do Programa Qualidade da Escola, “[...] o supervisor de ensino era o responsável pelos erros e falhas que se refletissem nos resultados do Saresp, caso não tivessem sido orientados adequadamente a serem superados” (Perrella; Alencar, 2022, p. 8). A supervisão de ensino também era mobilizada a orientar as equipes de gestão das escolas estaduais a se instrumentalizarem para ajustes em relação a diversas dimensões do aparato escolar, tendo em vista que as escolas atingissem as metas de Idesp (Perrella; Alencar, 2022).

A pesquisa concluiu que, no período analisado e por meio dos programas implementados, a Secretaria de Estado da Educação de São Paulo reforçou a gestão de resultados, centralizando o controle e limitando a participação da comunidade, o que revela uma contraposição à gestão democrática, como evidenciam os autores ao afirmarem que “a adoção da política de gestão do currículo definida de cima para baixo, afinada com a gestão para resultados que não garantem direitos, em grande medida invade a lacuna da gestão democrática gerada intencionalmente” (Perrella; Alencar, 2022, p. 13).

Estormovski e Silveira (2022), analisaram, por meio de estudo documental, as alterações realizadas no plano de carreira do magistério de uma rede municipal de ensino, no qual foi inserida a bonificação docente como resultado de uma avaliação discente. Concluem que a escola se desvincula de objetivos singulares e sociais para se adaptar a demandas macro definidas e se torna instrumento de reprodução de interesses economicistas. Segundo os autores, o anseio do poder público em solucionar a problemática da qualidade da educação resultou em um processo avaliativo que responsabiliza os docentes, direcionando a gestão escolar para o estímulo à obtenção de recompensas profissionais individuais, contrariando preceitos democráticos.

O artigo de Rodrigues e Barbosa (2022) investigou a trajetória do Bônus Mérito e Bonificação por Resultado, política de remuneração variável adotada, desde 2000, pela Secretaria da Educação do Estado de São Paulo. Com o propósito de se fazer um balanço crítico dessa política, foi realizada

uma pesquisa bibliográfico-documental, sendo analisados relatórios dos órgãos da Secretaria de Gestão Pública, responsável pela política de bonificação, assim como a análise da literatura nacional e internacional sobre o tema. Foram realizadas e analisadas entrevistas com professores e diretores de escolas. Tanto a análise da literatura quanto dos relatórios evidenciou a pouca efetividade da política de bonificação, pouco incidindo na melhoria do desempenho dos estudantes, não provocando, portanto, mudanças significativas que justifiquem a adoção de tal política.

A análise das entrevistas explicitou o descrédito de gestores e professores com o Bônus Mérito e Bonificação por Resultado e também repercussões negativas no trabalho docente. Quanto a essas conclusões, os autores comentam que:

[...] as entrevistas mostram que os professores não desenvolvem suas atividades pensando ou se orientando pelo bônus, mas sim por motivações associadas à prática pedagógica e ao desenvolvimento dos estudantes, reduzindo, portanto, a centralidade na crença oferecida pelos reformadores empresariais e pela própria Seduc de que o bônus é um forte indutor e motivador de mudanças (Rodrigues; Barbosa, 2022, p. 145).

Com relação às repercussões negativas no trabalho docente, tais como diferenciação entre os docentes de uma mesma escola em função do pagamento ou não do Bônus Mérito e Bonificação por Resultados, a política de bonificação não promove a valorização docente e, pelo contrário, pode levar ao abandono da docência, como evidenciado no grande número de exonerações de professores efetivos na rede pública estadual paulista.

O artigo de Cardoso e Melo (2022) apresenta os resultados de pesquisa de mestrado cujo foco foi a política de bonificação materializada por meio do Prêmio pela Elevação da Qualidade da Aprendizagem (PEQ), pago aos profissionais da rede municipal de ensino de Rio Branco/Acre. O estudo desenvolveu-se a partir da análise documental e da pesquisa empírica realizada com um grupo de professores. Concluiu-se que o Prêmio pela

Elevação da Qualidade da Aprendizagem (PEQ) é compreendido, por uma parcela dos professores, como uma motivação ao seu trabalho. Por outro lado, há o reconhecimento de que a política de bonificação limita a autonomia docente, sobrecarregando e exercendo mais controle e regulação sobre o trabalho dos professores.

Para os docentes, toda a responsabilidade pela elevação da qualidade da aprendizagem recai na figura do professor, como evidenciado pelos autores quando apontam que “é insustentável atribuir apenas ao docente a responsabilidade plena pelo aprendizado do aluno e podemos identificar que os professores participantes da pesquisa sentem tal encargo no exercício do trabalho” (Cardoso; Melo, 2022, p. 12).

Os autores ainda concluem que:

[...] as evidências empíricas revelam que pouco se pode dizer que o referido Prêmio tenha sido o responsável pelo avanço da qualidade educacional em contexto local. O que é possível afirmar é que a maioria dos sujeitos o veem como uma recompensa que em contrapartida controla a ação pedagógica, principalmente dos professores, retirando-lhes sua autonomia, elemento fundamental na atividade docente e que lhe é tão cara em suas lutas (Cardoso; Melo, 2022, p. 19).

Pereira e Morais (2022) discutem como a política adotada no governo de Aécio Neves (2003–2010) assimilou as propostas do Banco Mundial (BM) para a docência. Para tanto, os autores realizaram pesquisa documental e bibliográfica, analisando teses e dissertações que abordam as repercussões da política de bonificação e do “Prêmio por Produtividade”, que “configurou-se como uma bonificação concedida a partir dos resultados alcançados, mas também um impulsionador para que metas delineadas fossem atingidas. Tal programa foi revogado em 2016” (Pereira; Morais, 2022, p. 690).

Os autores apontam que a política educacional do período analisado se alinhou às propostas do BM, pois, para o Banco Mundial, “era preciso

atrair melhores estudantes para a carreira docente porque haveria melhora do quadro de professores e, conseqüentemente, melhoria na qualidade do ensino e da educação e que o oferecimento de bonificação atrairia bons profissionais” (Pereira; Morais, 2022, p. 695).

Nesse sentido, o governo de Aécio Neves utiliza as avaliações externas da aprendizagem dos estudantes para avaliar o desempenho docente, implementando programas e ações que promovem um Estado gerencial baseado no controle exercido por meio da centralidade da avaliação educacional, sendo o trabalho docente controlado por meio dos resultados das avaliações externas, na avaliação de desempenho e para o recebimento do “Prêmio por Produtividade”.

Os autores concluem que, em Minas Gerais, as propostas do Banco Mundial de modernização da gestão pública e implementação da bonificação foram assimiladas. Contudo, a premiação com base no mérito não levou à melhoria da qualidade da docência, mas trouxe a responsabilização dos agentes educacionais pelo sucesso ou não das metas estabelecidas para as escolas.

Cenários construídos a partir da análise dos artigos

A análise dos trabalhos selecionados possibilitou a construção de três categorias a partir leitura dos 22 artigos, as quais chamaremos de cenários. Um cenário revela os aspectos que, segundo os autores, evidenciam que as políticas de bonificação incidem positivamente na gestão e na melhoria da qualidade da educação. Nessa categoria, incluímos os artigos de Furtado e Soares (2018), Araújo, Leite e Passon (2018) e Guedes e Beiruth (2021).

Um segundo cenário que emergiu da leitura dos artigos apresenta resultados de pesquisas que indicam que as políticas de bonificação analisadas não vêm conseguindo resultados favoráveis; pelo contrário, vêm obtendo resultados negativos. Nessa categoria, incluem-se os artigos de Bauer (2012), Sousa, Maia e Haas (2014), Oshiro, Scorzafave e Dorigan (2015) e Melo e Braidí (2018).

O terceiro cenário observado congrega os 15 artigos e explicita resultados das pesquisas que demonstram mudanças na gestão, controle

do trabalho docente e desvalorização da profissão. Os artigos analisados nessa categoria caracterizam o que Bonamino e Sousa (2012) denominam de avaliações de 3ª geração, com consequências fortes para escolas e profissionais, implementadas por meio das políticas de bonificação associadas ao desempenho discente nas avaliações externas.

Ramos e Moreira (2020) analisaram os documentos da OCDE que vêm influenciando a política educacional no Brasil, com base na Teoria do Capital Humano e na Pedagogia das Competências e na lógica de que o professor ideal é aquele que tem um excelente desempenho e que deve receber reconhecimento financeiro por isso. Os efeitos esperados da gestão educacional gerencialista e da política de responsabilização, segundo a lógica neoliberal, deveriam incidir na melhoria da qualidade da educação. Porém, os artigos analisados, na maioria dos casos, sinalizaram diversos problemas gerados e dificuldade para alcançar os índices esperados.

Nesse sentido, dos 22 artigos selecionados, apenas quatro explicitaram resultados positivos das políticas de bonificação. Como Furtado e Soares (2018), que analisaram a política de incentivos financeiros para profissionais da educação a partir do alcance das metas e dos índices preestabelecidos para as escolas no estado de Pernambuco e explicitaram que a política de bonificação foi responsável por uma variação positiva ou ainda pela maior parte dos ganhos de desempenho da rede estadual.

Araújo, Leite e Passon (2018), ao analisarem o Prêmio Escola Nota Dez no estado do Ceará, concluíram que a política de bonificação escolar estimula e provoca mudanças positivas no trabalho pedagógico das escolas, reorientando o planejamento, promovendo novas estratégias de alfabetização, letramento e avaliação. Contudo, os autores não hesitam em afirmar que tais políticas podem gerar pressões sobre os professores para alcançarem as metas e promover o tensionamento das relações interpessoais, explicitando efeitos deletérios que as iniciativas de bonificação carregam.

Guedes e Beiruth (2021) também explicitam no seu artigo, ao analisarem a relação do desempenho de alunos de escolas de tempo integral no Espírito Santo a partir das notas das provas do Sistema de Avaliação da Educação Básica (SAEB), a associação positiva e relacionada com os

incentivos financeiros (Bonificação por Desempenho) e evidenciaram os efeitos da concessão de incentivos financeiros na redução do absenteísmo e a diminuição da rotatividade de professores.

No segundo cenário, temos um grupo de artigos cujas pesquisas revelaram que as políticas de bonificação não estão conseguindo alcançar seus objetivos, ou seja, a partir da responsabilização dos docentes e gestores de alcançarem as metas e índices definidos no contexto da gestão por resultados. Nesse sentido, Melo e Braidí (2018) explicitam que a pesquisa desenvolvida no estado do Acre não trouxe evidências que, de fato, permitem dizer que os incentivos financeiros a partir dos resultados da avaliação externa se constituem em instrumentos de controle e de culpabilização dos profissionais do magistério.

Sousa, Maia e Haas (2014), cujo artigo se insere no debate da relação do uso da avaliação externa como mecanismo de gestão da melhoria da qualidade do ensino ofertado, apresentam duas perguntas que a pesquisa desenvolvida pretendeu responder, ou seja, se o alcance das metas pode ser compreendido como melhoria da qualidade das escolas e se a premiação incide em escolas em que se observa o aprimoramento da qualidade. Contudo, as autoras concluem que, em ambos os casos, a resposta é negativa. Do mesmo modo, o artigo de Bauer (2012) não encontrou relação direta entre as ações e programas de formação e bonificação de professores e as avaliações externas analisadas pela autora.

Oshiro, Scorzafave, Dorigan (2015), ao analisarem os resultados da Prova Brasil de 2005 a 2011, buscaram avaliar se a adoção do bônus alterou indicadores escolares, tais como: clima escolar, participação dos pais na escola, cumprimento do currículo, uso do tempo em sala de aula, experiência dos professores/diretores, etc. Os dados analisados indicaram resultados inconclusivos e observou-se o impacto positivo do programa de incentivo de professores apenas nas notas do quinto ano do Ensino Fundamental em 2009 e resultados negativos para o nono ano.

Os dois cenários indicam que as políticas de bonificação não estão atingindo os objetivos de melhoria da gestão e dos resultados do desempenho dos estudantes nas avaliações externas. Os artigos no terceiro cenário, que

podemos chamar de mudanças na gestão, controle do trabalho docente e desvalorização da profissão, indicam resultados que explicitam as reais consequências das políticas de bonificação.

Nesse sentido, o artigo de Koslinski, Ribeiro e Oliveira (2017), que analisou o Indicador de Desempenho Escolar – Alfabetização (IDE-Alfa) no aumento da proficiência dos estudantes pela política de incentivo financeiro às escolas, chamada de “Prêmio Escola Nota Dez” do estado do Ceará, demonstrou que a política de incentivo não vem contribuindo para que as escolas apoiadas obtenham um efeito positivo sobre a proficiência dos estudantes, enquanto as escolas que são premiadas mantêm o alto desempenho. Desse modo, os autores alertam que o apoio técnico-financeiro não tem se mostrado suficiente para promover melhoria na proficiência dos alunos das escolas de baixo rendimento.

O artigo de Pereira e Moraes (2022), que analisou a implementação das políticas do Banco Mundial em Minas Gerais no período de 2003–2010, indica que a premiação com base no mérito não levou à melhoria da qualidade da docência, mas trouxe a responsabilização dos agentes educacionais pelo sucesso ou não das metas estabelecidas para as escolas.

Estormovski e Silveira (2022), ao analisarem a política de bonificação em Esteio (RS), concluem que os incentivos financeiros levam à responsabilização dos docentes, direcionando a gestão escolar para o estímulo à obtenção de recompensas profissionais individuais, contrariando preceitos democráticos. Novaes (2014) indica que a associação da avaliação externa com a política de bonificação tem resultado em adequação do currículo escolar aos descritores da avaliação externa, retirando da escola a autonomia para gerir o currículo, treinamento para a prova, mudanças no papel dos professores, podendo incidir na autoestima, no aumento da angústia e na frustração dos profissionais da educação com a carreira.

Nesse sentido, Zatti e Minhoto (2019), ao analisarem Avaliação do Desempenho Docente (ADD) nas redes públicas estaduais de Educação Básica no Brasil, sinalizam que o controle externo, por meio das avaliações externas e das políticas de incentivo financeiro, provoca, nos docentes, a perda do sentido sobre o próprio trabalho. Fonseca e Lima (2018), ao

analisarem o Saesp, indicam que os professores têm pautado o trabalho pedagógico com base nos conteúdos das avaliações e também indicam preocupação com a bonificação.

O artigo de Carvalho (2018) corrobora o que apontam os artigos anteriormente mencionados, destacando que a gestão empresarial na escola pública exige dos educadores e educadoras posturas voltadas à eficiência e à prestação de contas, fortalecendo subjetividades individualizadas, competitivas, assertivas, planejadas, típicas do mercado e orientadas para a consecução de metas e para a ascensão na carreira por meio de recompensa monetária.

Com relação às avaliações externas e à bonificação, Jacomini *et al.* (2021) sugerem que as políticas curriculares analisadas levam à gestão de resultados no desempenho das avaliações externas, na desvalorização das dinâmicas da sala de aula que levem em consideração as singularidades tão próprias dos processos de aprendizagem.

A gestão de resultados e a centralização do controle do trabalho docente, articuladas com a gestão do currículo definido de cima para baixo, de acordo com Perrella e Alencar (2022), geram limitações importantes para a gestão democrática e até mesmo se contrapõem à participação da comunidade e à gestão democrática. Com relação à bonificação, Piovezan e Dal Ri (2016) destacaram que a implementação de diversas avaliações para gerar o bônus tem a finalidade de controle e a vigilância do trabalho dos professores, o que é corroborado por outras pesquisas.

O artigo de Rosistolato e Prado (2014), que aborda a avaliação externa nos municípios do Rio de Janeiro, com bonificação, e Duque de Caxias, sem bonificação, explicita influências positivas, como identificar lacunas nos processos de formação dos estudantes, mas também influências negativas no cotidiano escolar, com foco na preparação para as provas, visando alcançar os índices estabelecidos e articulados com a bonificação.

Rodrigues e Barbosa (2022) apontam, no seu artigo, a pouca efetividade da política de bonificação, pouco incidindo na melhoria do desempenho dos estudantes, não provocando, portanto, mudanças significativas

que justifiquem a adoção de tal política. Denotam também o descrédito de gestores e professores com o Bônus Mérito e Bonificação por Resultado e repercussões negativas no trabalho docente, além de não promover a valorização docente.

No artigo de Cardoso e Melo (2022), é analisada a política de bonificação adotada pela rede municipal de ensino de Rio Branco (Acre). Segundo os autores, uma parcela dos professores vê a bonificação como uma motivação ao seu trabalho. Porém, professores e gestores indicam que a política de bonificação limita a autonomia docente, leva a um controle maior e regulação sobre o trabalho docente, responsabilizando os professores pela elevação da qualidade da aprendizagem.

Considerações finais

A partir dos artigos selecionados e analisados, encontramos dados de pesquisas cujo foco esteve na política de incentivos financeiros aos profissionais da educação em diferentes redes de ensino no país. Os sistemas de avaliação externa estão articulados com a política educacional, e pensar a educação atualmente sem levar em conta os resultados produzidos por tais sistemas de avaliação é praticamente impossível.

A ênfase na gestão por resultados articulada com as políticas de bonificação proporcionaria, do ponto de vista da proposição e da justificativa para a implementação da política, melhoria da qualidade do ensino oferecido, entretanto não é o que fica evidenciado com a análise dos artigos que se detiveram no exame aprofundado de tais políticas.

Os artigos selecionados, na sua maioria, demonstram a fragilidade do uso dos resultados das avaliações externas e em larga escala na implementação de políticas de incentivos financeiros e bonificação para os profissionais da educação a partir dos resultados dos estudantes, uma vez que não é possível afirmar a incidência na melhoria da qualidade do ensino.

Ao que parece, as políticas de bonificação têm mais incrementado o controle e a responsabilização das escolas e dos profissionais pelos resultados escolares do que proporcionado avanços na qualidade, produzindo

um ambiente escolar angustiante (Novaes, 2014), competitivo (Carvalho, 2018) e sem sentido para o trabalho docente e da gestão (Zatti; Minhoto, 2019), provocando a adoção de mecanismos de individualização de recompensas e prêmios, o que contraria os preceitos democráticos constitucionais (Estormovski; Silveira 2022). Nesse sentido, torna-se imprescindível o debate amplo e coletivo sobre as políticas de bonificação e suas consequências efetivas na escola e na gestão.

Referências

ARAÚJO, K. H.; LEITE, R. H.; PASSONE, E. F. K. Política de *Accountability* Educacional no Estado do Ceará: Repercussões nas Dinâmicas Pedagógicas. **Revista Teias**, [S. l.], v. 19, n. 54, jul./set. 2018. Disponível em: <https://www.e-publicacoes.uerj.br/revistateias/article/view/35985>. Acesso em: 18 jun. 2025.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2016.

BAUER, A. É Possível Relacionar Avaliação Discente e Formação de Professores? A Experiência de São Paulo. **Educação em Revista**, Belo Horizonte, v. 28, n. 2 p. 61–82, jun. 2012. Disponível em: <https://www.scielo.br/j/edur/a/TPPxVX-CHHdMW6MG96twZJyG/abstract/?lang=pt>. Acesso em: 07 jun. 2025.

BONAMINO, A.; SOUSA, S. Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. **Educação e Pesquisa**, São Paulo, v. 38, n. 2, p. 373–388, abr./jun. 2012. Disponível em: <https://www.scielo.br/j/ep/a/rtQkYDSjky4mXG9TCrgRSqJ/?format=pdf&lang=pt>. Acesso em: 10 mai. 2025.

BROOKE, N. O futuro das políticas de responsabilização no Brasil. **Cadernos de Pesquisa**, São Paulo, v. 36, n. 128, p. 377– 401, 2006. Disponível em: <https://www.scielo.br/j/cp/a/xpH7dBYJxQ9wdrN5NLNyD8n/abstract/?lang=pt>. Acesso em: 18 jun. 2025.

BROOKE, N. Sobre a equidade e outros impactos dos incentivos monetários para professores. **Estudos em Avaliação Educacional**, [S. l.], v. 24, n. 55, p. 34–62, 2013. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2719>. Acesso em: 07 jun. 2025.

CARDOSO, P. F.; MELO, L. F. A política de bonificação e responsabilização educacional no município de Rio Branco/Acre: prêmio pela elevação da qualidade da aprendizagem. **Jornal de Políticas Educacionais**, [S. l.], v. 16, p. 1–21, jan. 2022. Disponível em: <https://revistas.ufpr.br/jpe/article/view/83559>. Acesso em: 10 mai. 2025.

CARVALHO, M. P. de. As professoras e o pagamento de bônus por resultado: o caso da rede estadual de São Paulo. **Educar em Revista**, Curitiba, v. 34, n. 72, p. 187-207, nov./dez. 2018. Disponível em: <https://www.scielo.br/j/er/a/ydB7LK6SYHWZHVqD7Fv8Rxm/?format=pdf&lang=pt>. Acesso em: 11 mai. 2025.

ESTORMOVSKI, R. C.; SILVEIRA, C. L. A. da. A redefinição da gestão escolar a partir da concessão de bonificações docentes. **Educação: Teoria e Prática**, Rio Claro, v. 32, n. 65, p. 1-18, 2022. Disponível em: <https://www.periodicos.rc.biblioteca.unesp.br/index.php/educacao/article/view/14928>. Acesso em: 18 jun. 2025.

FONSECA, V. A.; LIMA, P. G. Sobre a política curricular do Estado de São Paulo (2007-2009): algumas reflexões e encaminhamentos. **Revista e-Curriculum**, São Paulo, v. 16, n. 3, p. 544-565, jul./set. 2018. Disponível em: <http://educa.fcc.org.br/pdf/curriculum/v16n3/1809-3876-curriculum-16-03-544.pdf>. Acesso em: 07 jun. 2025.

FURTADO, C. S. V.; SOARES, T. M. Impacto da Bonificação Educacional em Pernambuco. **Estudos em Avaliação Educacional**, São Paulo, v. 29, n. 70, p. 48-76, jan./abr. 2018. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/3940>. Acesso em: 18 jun. 2025.

GUEDES, H. G.; BEIRUTH, A. X. O impacto dos bônus dos professores no desempenho dos alunos em escolas de tempo integral: o caso do Espírito Santo. **Contextus – Revista Contemporânea de Economia e Gestão**, [S. l.], v. 119, p. 374-385, 2021. Disponível em: https://repositorio.ufc.br/bitstream/riufc/63986/1/2021_art_hguedes.pdf. Acesso em: 11 mai. 2025.

JACOMINI, M. A. *et al.* Padronização, controle e *accountability* na política curricular paulista (2007-2018). **Revista Brasileira de Educação**, [S. l.], v. 26, p. 1-24, 2021. Disponível em: <https://www.scielo.br/j/rbedu/a/rBL5bpVD4JzrqXT8cfGFj-Gz/?format=html&lang=pt>. Acesso em: 31 mai. 2025.

KOSLINSKI, M. C.; RIBEIRO, E.; OLIVEIRA, L. X. de. Indicadores educacionais e responsabilização escolar: um estudo do “Prêmio Escola Nota Dez”. **Estudos em Avaliação Educacional**, [S. l.], v. 28, n. 69, p. 804-846, 2017. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/4087>. Acesso em: 18 jun. 2025.

MELO, L. de F.; BRAIDI, D. M. P. A política de bônus para os profissionais da educação no estado do Acre: concepção e normatização. **Revista Teias**, [S. l.], v. 19, n. 54, p. 422-437, 2018. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/revistateias/article/view/37804>. Acesso em: 24 jun. 2025.

NOVAES, L. C. Os impactos da avaliação externa sobre o trabalho de professores na rede estadual paulista. **Roteiro**, [S. l.], v. 39, n. 2, p. 283–309, 2014. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_abstract&pid=S2177-60592014000200002&lng=es&nrm=iso. Acesso em: 31 mai. 2025.

OSHIRO, C. H.; SCORZAFAVE, L. G.; DORIGAN, T. A. Impacto sobre o desempenho escolar do pagamento de bônus aos docentes do ensino fundamental do Estado de São Paulo. **Revista Brasileira de Economia**, Rio de Janeiro, v. 69, n. 2, p. 213–249, abr./jun. 2015. Disponível em: <https://www.scielo.br/j/rbe/a/Db-7fCKW7RXkbvQBFgyKXGKM/?lang=pt>. Acesso em: 11 mai. 2025.

PEREIRA, M. G.; MORAIS, S. P. A bonificação em Minas Gerais e a confluência com as recomendações do Banco Mundial: repercussões no trabalho docente. **Revista Educação e Política em Debate**, [S. l.], v. 11, n. 2, p. 688–704, mai./ago. 2022. Disponível em: <https://seer.ufu.br/index.php/revistaeducaopoliticas/article/view/64522>. Acesso em: 24 jun. 2025.

PERRELLA, C. S. S.; ALENCAR, F. Gestão para resultados e ações de controle na política educacional paulista. **Educação em Revista**, Belo Horizonte, v. 38, 2022. Disponível em: <https://www.scielo.br/j/edur/a/6XzVzvqy5SKmjVZf-n76Ppgg/?format=html&lang=pt>. Acesso em: 18 jun. 2025.

PIOVEZAN, P. R.; DAL RI, N. M. A precarização do trabalho docente no estado de São Paulo: vinte anos de reformas. **ETD – Educação Temática Digital**, Campinas, v. 18, n. 1, p. 178–197, 2016. Disponível em: <https://www.e-publicacoes.uerj.br/revistateias/article/view/35985>. Acesso em: 10 mai. 2025.

RAMOS, C. V.; MOREIRA, J. A. S. Política de remuneração de professores e as recomendações da Organização para Cooperação e Desenvolvimento Econômico (OCDE). **EDUCA – Revista Multidisciplinar em Educação**, [S. l.], v. 7, n. 17, p. 1510–1526, 2020. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/8635567>. Acesso em: 11 mai. 2025.

RODRIGUES, J. D. Z.; BARBOSA, A. Política de Remuneração Variável na Rede Estadual de Ensino Paulista: um balanço dos 20 anos do “Bônus” do Magistério. **Revista Iberoamericana de Educación**, [S. l.], v. 90, n. 1, p. 133–149, 2022. Disponível em: <https://siteal.iiep.unesco.org/pt/node/4467>. Acesso em: 18 jun. 2025.

ROSISTOLATO, R.; PRADO, A. P. Os profissionais da educação e as avaliações externas de aprendizagem: uma comparação entre Rio de Janeiro e Duque de Caxias. **Roteiro**, Joaçaba, v. 39, n. 2, p. 311–330, 2014. Disponível em: <http://educa.fcc.org.br/pdf/roteiro/v39n02/v39n02a03.pdf>. Acesso em: 31 mai. 2025.

SOUSA, S. Z.; MAIA, M. M. V. ; HAAS, C. M. Avaliação, índices e bonificação: controvérsias suscitadas por dados da rede estadual paulista. **Estudos em Avaliação Educacional**, v. 25, n. 58, p. 188-209, 2014. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2772/2779>. Acesso em: 10 mai. 2025.

ZATTI, A. M.; MINHOTO, M. A. P. Minhoto. Política de avaliação do desempenho docente nas redes estaduais de ensino do Brasil. **Educação e Pesquisa**, São Paulo, v. 45, p. 1-21, 2019. Disponível em: <https://www.scielo.br/j/ep/a/jGQCm-DyNY39SQHMhfcQhhNb/abstract/?lang=pt>. Acesso em: 24 jun. 2025.

CAPÍTULO 7

ANÁLISE DO IDEB COMO INDICADOR DE MONITORAMENTO DA QUALIDADE DA EDUCAÇÃO BÁSICA

Girlene Ribeiro de Jesus¹
Francisco Augusto da Costa Garcia²
Gizelle Alves Xavier dos Santos³
DOI: 10.52695/978-65-5456-163-1.7

Introdução

O Plano Nacional de Educação (Brasil, 2014) pretende, na meta 7, fomentar a qualidade da Educação Básica no Brasil. Para tanto, elenca 36 estratégias a serem perseguidas ao longo do decênio, que vão desde o incremento dos níveis de aprendizagem e melhoria do fluxo escolar até a garantia de transporte escolar, a universalização do acesso à *internet*, a melhoria da infraestrutura das escolas, a criação de novos indicadores contextuais, o combate à violência nas escolas, a implementação de políticas de inclusão, a consolidação da educação indígena e do campo e o fortalecimento da colaboração entre os entes federativos para a promoção de uma educação de qualidade,

1 Universidade de Brasília. E-mail: girlene@unb.br.

2 Universidade de Brasília.

3 Universidade de Brasília.

entre outras. Apesar de as estratégias serem bastante abrangentes, a meta 7 utiliza, como termômetro para o monitoramento da qualidade da educação, a melhoria do fluxo e da aprendizagem por meio do atingimento de metas nacionais para o Índice de Desenvolvimento da Educação Básica, o Ideb.

A ideia de um Plano Nacional de Educação teve suas primeiras raízes no Manifesto dos Pioneiros da Educação Nova de 1932, aparecendo, *a posteriori*, como uma meta na constituição de 1934 e na LDB de 1961, e tomando diferentes configurações, como na reforma do ensino, Lei nº 5.692/71 (Brasil, 1971) e no Plano Decenal de Educação para todos de 1993. Em 2001, foi aprovado o primeiro Plano Nacional de Educação, Lei nº 10.172 (Brasil, 2001), previsto na constituição e na LDB em vigência, Lei nº 9.394 de 1996 (Brasil, 1996). Tal plano era constituído por 295 metas distribuídas em 11 áreas, com aproximadamente 27 metas cada (Saviani, 2007).

O primeiro plano decenal teve duração prevista até 2011 e foi reestruturado em 2014, com restrição de 20 metas, sendo um total de 254 estratégias distribuídas entre elas. Entre as metas, observaram-se a preocupação com a universalização da Educação Básica, a erradicação do analfabetismo, o oferecimento de educação em diversas modalidades e localidades, a ampliação do acesso ao Ensino Superior, a melhoria da qualidade da educação, a formação e valorização dos professores, o financiamento educacional e a gestão democrática (Brasil, 2014).

O PNE está referido no artigo 214 da Constituição Federal de 1988 (Brasil, 1988), que determina a sua elaboração de acordo os princípios fundamentais da educação brasileira: I) erradicação do analfabetismo; (II) universalização do atendimento escolar; (III) melhoria da qualidade do ensino; (IV) formação para o trabalho; e (V) promoção humana, científica e tecnológica do país. Já a sua regulamentação foi determinada através da Lei de Diretrizes e Bases da Educação Nacional (Brasil, 1996), que deixou a cargo da União, em colaboração com estados e municípios, a incumbência de organizar o PNE, que, posteriormente, foi aprovado pela Lei nº 10.172, de 09/01/2001, com vigência decenal.

O PNE 2014–2024 foi elaborado com os compromissos largamente debatidos e apontados como estratégicos pela sociedade na CONAE 2010,

os quais foram aprimorados na interação com o Congresso Nacional. Há 20 metas estabelecidas no atual PNE 2014–2024, que contemplam todos os níveis e modalidades da educação. O tripé que sustenta a educação pode ser visto ao longo das 20 metas traçadas, a saber: acesso, permanência e qualidade. Tendo em vista esse tripé, no presente estudo, será dado enfoque ao eixo da qualidade por se tratar de um tema que acaba congregando os demais, pois não se tem qualidade em educação sem acesso e permanência. No PNE 2014–2024, a meta 7 especifica o fomento à qualidade da Educação Básica em todas as etapas e modalidades, de modo a atingir as seguintes médias nacionais para o Ideb:

Quadro 1 – Médias projetadas para o Ideb no PNE 2014–2024

IDEB	2015	2017	2019	2021
Anos Iniciais do Ensino Fundamental	5,2	5,5	5,7	6,0
Anos Finais do Ensino Fundamental	4,7	5,0	5,2	5,5
Ensino Médio	4,3	4,7	5,0	5,2

Fonte: PNE 2014–2024 (Brasil, 2014).

O monitoramento da Meta 7 é realizado por meio de três indicadores nacionais:

- Indicador 7A: Ideb dos Anos Iniciais do Ensino Fundamental.
- Indicador 7B: Ideb dos Anos Finais do Ensino Fundamental.
- Indicador 7C: Ideb do Ensino Médio.

Embora a meta 7 do PNE 2014–2024 tenha como foco a qualidade da Educação Básica, esse conceito não é trazido na meta 7 do plano. Dessa forma, surge o questionamento sobre o que se está considerando como qualidade da educação. Uma vez que o Ideb é utilizado como um indicador de monitoramento da qualidade da educação, acaba sendo, muitas vezes, considerado como a operacionalização desse conceito.

Para compreender melhor o Ideb, trata-se de um indicador educacional que relaciona em um único valor informações sobre: (a) o rendimento

escolar ou fluxo escolar (considerando os dados sobre aprovação, reprovação e abandono), obtido por meio do censo escolar; e (b) o desempenho (proficiência em língua portuguesa e matemática) obtido por meio do Sistema de Avaliação da Educação Básica – Saeb (Fernandes, 2007; Fernandes; Gremaud, 2009). O Ideb é atualmente o principal indicador de qualidade da Educação Básica no Brasil e sintetiza, em um único índice, dois conceitos relevantes que têm sido utilizados para aferir a qualidade do ensino no país, fluxo escolar e aprendizado.

Fernandes e Gremaud (2009) justificam o uso tanto do rendimento escolar, aferido pelas escolas, quanto do resultado do Saeb para compor o Ideb, indicando que, caso fosse considerado apenas o fluxo, por exemplo, haveria a possibilidade de melhoria do fluxo escolar sem que, de fato, o desempenho dos alunos melhorasse. Em outras palavras, poderia haver um efeito negativo de aprovações indiscriminadas. Por outro lado, se apenas os resultados dos exames padronizados, como o Saeb, fossem considerados, o incentivo seria o oposto, pois as escolas poderiam elevar o padrão de aprovação, fazendo com que boa parte dos estudantes abandonasse a escola antes de completar a Educação Básica, o que não é desejável. Dessa forma, o Ideb ficou composto pela multiplicação simples de dois componentes: (N) proficiência média dos alunos da escola e (P) taxa média de aprovação:

$$\text{IDEB} = \text{N} * \text{P}$$

A escala do Ideb foi construída variando de 0 a 10, teoricamente; mas, empiricamente, esses valores são impossíveis de acontecer, ficando as médias, em geral, entre 4,0 e 8,0 (Soares; Xavier, 2013). Para o Brasil, tendo como ponto inicial de referência o ano de 2005, foi estabelecida uma meta de Ideb igual a 6,0, para todas as regiões do país, até o ano de 2021. Todavia, com o passar do tempo, essa meta foi sendo alterada, especialmente para os Anos Finais do Ensino Fundamental e para o Ensino Médio.

Fernandes e Gremaud (2009) indicam que o valor do Ideb para a educação brasileira igual a 6,0 foi estabelecido em 2007, procurando ter correspondência com a média dos países membros da Organização de Cooperação e Desenvolvimento Econômico (OCDE). Foi utilizada a ordenação

dos alunos brasileiros no Programa Internacional de Avaliação de Alunos (Pisa) e no Saeb e a ordenação dos alunos da OCDE no Pisa. Tendo por suposto que os exames ordenam os alunos de forma similar, foi realizada a conversão das escalas.

A respeito da pertinência do Ideb, Horta Neto (2010) enfatiza que esse indicador é um instrumento favorável para se obter o conhecimento da radiografia da Educação Básica. O indicador, em cada série histórica, além de indicar falhas, também aponta reflexões para as escolas sobre como as fragilidades podem ser corrigidas ao replanejar o fazer pedagógico. Ao ser administrado como parâmetro, acaba sendo um instrumento favorável para ações em prol da melhoria do espaço escolar.

Por sua vez, Saviani (2007) ressalta que o Ideb tem, indiscutivelmente, um caráter de política educacional, visando à busca de resoluções para problemas encontrados que influenciam na qualidade do ensino ofertado por escolas públicas do país. Não há dúvidas de que o Ideb tem provocado discussões quanto aos seus propósitos e relevância para a educação brasileira.

Chirinéa e Brandão (2015) destacam que o Ideb passou a ser propo- sitor de políticas públicas e indutor de ações das escolas e sistemas de ensino, visando alcançar a desejada qualidade da educação. Para esses mesmos autores, o Ideb pode ser considerado indicador de resultado e não de qualidade, pois é um indicador incapaz de retratar a realidade das instituições escolares, uma vez que há muitas outras variáveis que exer- cem influência na qualidade da educação, como gestão escolar, formação e condições de trabalho dos professores, ambiente educativo, prática pe- dagógica, acesso à e permanência na escola.

Por sua vez, Soares e Xavier (2013) apontam outra inconsistência do Ideb, pois os autores afirmam que as proficiências padronizadas em ma- temática das escolas são muito maiores do que as proficiências padroni- zadas em leitura, o que acaba dando maior proeminência no resultado do Ideb à proficiência em matemática. E também leva alguns a considerarem, erroneamente, que os resultados de matemática são superiores aos re- sultados de leitura.

Por fim, Figueiredo *et al.* (2018) realizou um estudo com dados do Ideb de 2015 e concluiu que os resultados do Ideb são inflados pelos resultados crescentemente positivos do fluxo, sem que, necessariamente, os resultados do Saeb tenham evoluído.

Segundo Schneider e Nardi (2014), é preciso ter atenção na hora de compreender o Ideb e introduzi-lo na rede pública de ensino, tendo em vista que ele coloca em causa apenas o desempenho dos alunos e das escolas, excluindo outros fatores que intervêm nas condições de qualidade (Charlot, 2021; Dourado; Oliveira; Santos, 2007; D'Escoffier, 2021).

Frente ao exposto, considerando o relevante papel do Ideb no monitoramento da meta 7 do PNE 2014–2024, decidiu-se realizar o presente estudo, que tem como objetivo analisar os resultados do Ideb entre os anos de 2005 e 2021, evidenciando seus limites e possibilidades enquanto indicador de monitoramento da qualidade da Educação Básica no Brasil.

Método

Para o alcance do objetivo proposto, foram utilizadas as planilhas do Ideb dos Anos Iniciais e Finais do Ensino Fundamental e do Ensino Médio, com dados disponíveis de 2005 a 2021, organizadas por escola, município, região e unidade da federação, disponibilizadas pelo INEP.⁴

Foram utilizados os dados dos Anos Iniciais e Finais do Ensino Fundamental, exclusivamente aqueles referentes à rede pública de ensino, de 2005 a 2021, e os dados do Ensino Médio, exclusivamente aqueles referentes à rede estadual de ensino, de 2005 a 2021.

Para a confecção dos gráficos, bem como para o cálculo dos coeficientes angulares e de correlação, foi utilizado o *Microsoft Excel*.

4 As informações podem ser acessadas no link: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/ideb/resultados>.

Resultados e discussão

Os gráficos de N (proficiência no Saeb) foram feitos com os dados coletados bianualmente, de 2005 a 2021, por região, da rede pública de ensino para os Anos Iniciais e Finais do Ensino Fundamental e para o Ensino Médio. A nossa expectativa era de uma curva de (N) crescente e com o maior coeficiente angular possível, indicando a velocidade de aproximação das metas.

Depois, tendo em vista que o Ideb é função do fluxo (P) obtido a partir do censo escolar realizado anualmente e tentando avaliar a força de P no resultado final do Ideb, foram feitos os gráficos de P, calculados bianualmente de 2005 a 2021, por estado e região, com a média regional, para a rede pública de ensino. Esperava-se, a despeito de algumas flutuações em torno de uma média estável, uma curva de (P) com um coeficiente angular baixo, mostrando essa mesma estabilidade no número de aprovações, tomadas a cada dois anos.

Segundo Fernandes e Gremaud (2009), com a composição do Ideb pelo fluxo escolar e pelo resultado do Saeb, há uma moderação entre o efeito negativo de aprovações indiscriminadas para aumentar o fluxo (P), por exemplo, e a retenção de alunos, com alta reprovação ou evasão, para forçar resultados com valor mais alto nos exames padronizados (N). No entanto, se por um lado P limita Ideb pela nota Saeb, quando $P=1$, pode-se crescer P até 1, influenciando fortemente Ideb sem contrapartida de N e alterando as impressões de Ideb e da sua comparação com a sua projeção de Ideb.

A expectativa era de ver o crescimento linear do Ideb em direção à meta, principalmente fruto do crescimento de N. Ou seja, P constante, já que, objetivamente, o fluxo, se há variação, é em torno de uma média e não pretendia ser crescente, e N crescente, já que representa objetivamente o resultado medido em uma avaliação de desempenho.

Para uma melhor compreensão do impacto de P e N no Ideb, e como isso se relaciona entre as etapas dos Anos Iniciais e Finais do Ensino Fundamental e Ensino Médio, calculamos uma matriz de correlações, que resolvemos apresentar na forma dos seguintes gráficos: 1) a correlação

do Ideb dos Anos Iniciais com os Anos Finais do Ensino Fundamental por estado e região; 2) a correlação do Ideb dos Anos Finais do Ensino Fundamental com o Ensino Médio por estado e região; 3) a correlação do Ideb dos Anos Iniciais do Ensino Fundamental com o Ensino Médio por estado e região; 4) para os Anos Iniciais do Ensino Fundamental, a correlação de P com Ideb, correlação de N com Ideb e correlação de P com N; 5) para os Anos Finais do Ensino Fundamental, a correlação de P com Ideb, correlação de N com Ideb e correlação de P com N; e 6) para os Anos Iniciais do Ensino Fundamental, a correlação de P com Ideb, correlação de N com Ideb e a correlação de P com N.

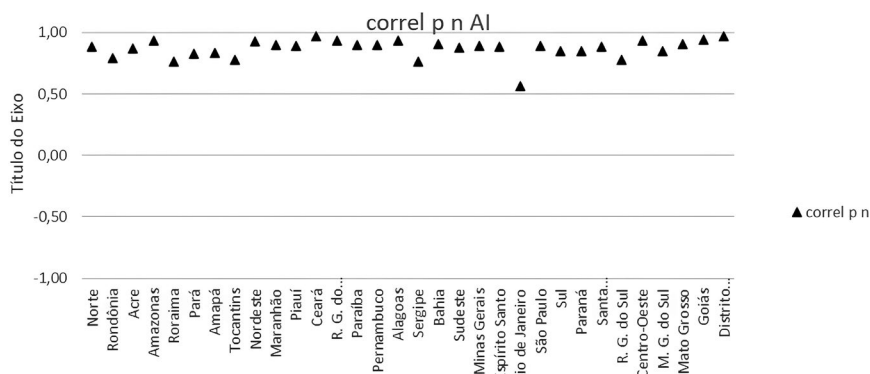
A correlação pode mostrar, por estado e por região, como o Ideb dos Anos Iniciais se relaciona com o Ideb dos Anos Finais e do Ensino Médio e o quanto o Ideb cresce em função do crescimento de P e de N. Se mais perto de 1, o Ideb cresce fortemente em função de P e N; se próximo de 0, o Ideb cresce pouco em função de P e N.

As análises, a seguir, estão divididas da seguinte forma: (a) correlação entre a nota padronizada do Saeb (N) e o fluxo (P) para os três níveis de ensino, Anos Iniciais do Ensino Fundamental, Anos Finais do Ensino Fundamental e Ensino Médio; (b) correlação entre o Ideb e seus componentes, N e P, nos três níveis de ensino; (c) correlação entre o Ideb dos diferentes níveis de ensino. Por fim, são apresentadas as metas projetadas pelo PNE 2014–2024 e os resultados aferidos até 2021.

Correlação entre a nota padronizada do Saeb (N) e o fluxo (P)

Espera-se que, quanto maior o desempenho nos testes padronizados (N), maior também seja o aprendizado mensurado internamente pela escola por meio das taxas de aprovação ou fluxo (P). A seguir, são apresentados os gráficos que apresentam, por nível de ensino, nos estados, a correlação entre N e P.

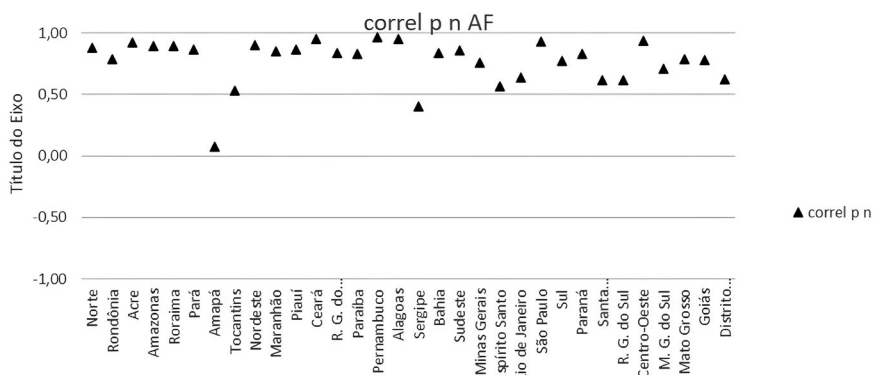
Gráfico 1 – Correlação entre P e N – Anos Iniciais do Ensino Fundamental (AI) por estado



Fonte: Elaboração própria.

Ao analisar a correlação entre N e P nos Anos Iniciais do Ensino Fundamental, observa-se que, em geral, há uma correlação alta e positiva entre N e P nos estados. Todavia, alguns estados chamam a atenção, especialmente o Rio de Janeiro, onde a correlação é $r = 0,57$; Sergipe e Roraima ($r = 0,76$) e Tocantins ($r = 0,78$).

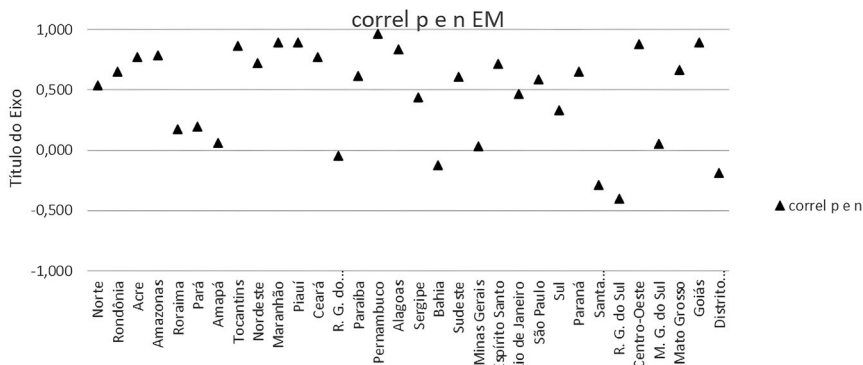
Gráfico 2 – Correlação entre P e N – Anos Finais do Ensino Fundamental (AF) por estado



Fonte: Elaboração própria.

Nos Anos Finais, observa-se um resultado diferente quando comparado aos Anos Iniciais do Ensino Fundamental, uma vez que as correlações são menores e irregulares. Por exemplo, Amapá ($r = 0,08$), Sergipe ($r = 0,40$), Tocantins ($r = 0,53$), Espírito Santo ($r = 0,57$), Santa Catarina, Distrito Federal e Rio Grande do Sul ($r = 0,62$).

Gráfico 3 – Correlação entre P e N – Ensino Médio por estado



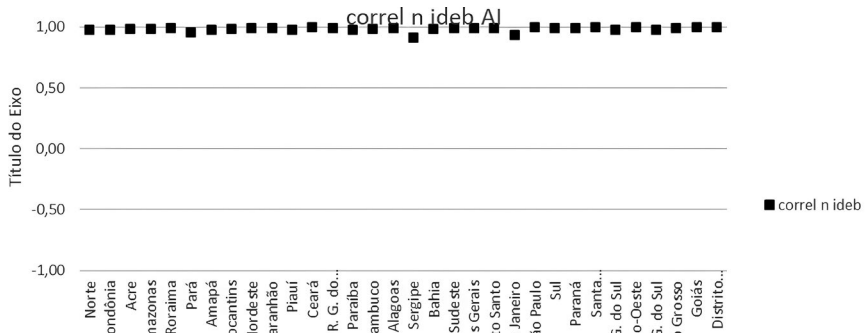
Fonte: Elaboração própria.

No Ensino Médio, observa-se um resultado estranho, pois há, inclusive, correlações negativas entre P e N em alguns estados: Santa Catarina, Rio Grande do Sul, Distrito Federal, Bahia e Rio Grande do Norte, fato que não faz sentido.

Correlação entre o Ideb e seus componentes: a nota padronizada do Saeb (N) e o fluxo (P)

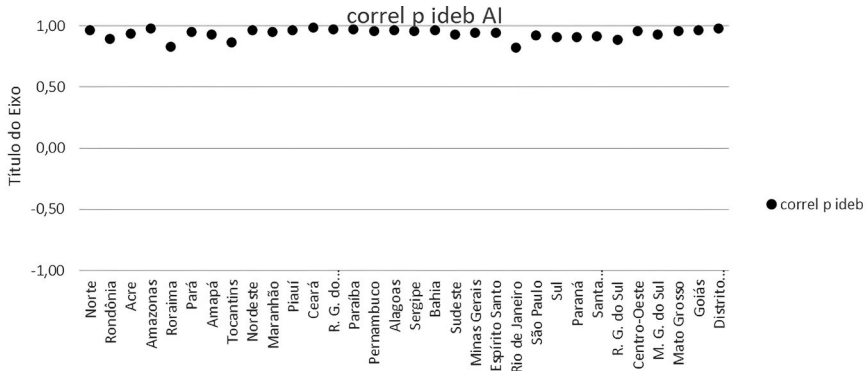
Com o intuito de verificar a relação entre o Ideb e os seus componentes, também foram calculadas correlações. Esperava-se encontrar uma alta relação entre o Ideb e N, assim como entre o Ideb e P, em todos os estados para os três níveis de ensino.

Gráfico 4 – Correlação entre Ideb e N – Anos Iniciais do Ensino Fundamental (AI) por estado



Fonte: Elaboração própria.

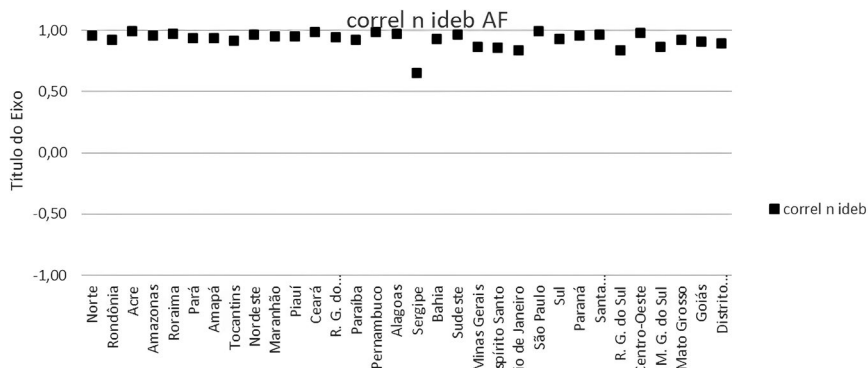
Gráfico 5 – Correlação entre Ideb e P – Anos Iniciais do Ensino Fundamental (AI), por estado



Fonte: Elaboração própria.

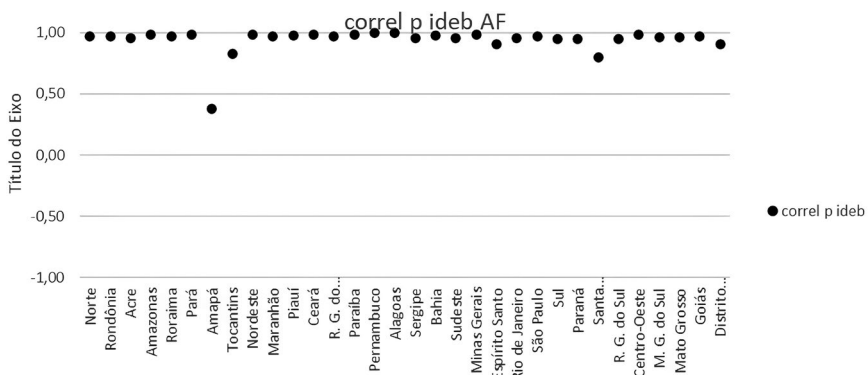
Os gráficos 4 e 5 mostram correlações altas e consistentes nos Anos Iniciais ($r > 0,82$), em todos os estados, entre o Ideb e seus componentes, N e P. Ambos os componentes se relacionam fortemente com os resultados da série histórica do Ideb. Os próximos gráficos apresentam os resultados para os Anos Finais do Ensino Fundamental.

Gráfico 6 – Correlação entre Ideb e N – Anos Finais do Ensino Fundamental (AF) por estado



Fonte: Elaboração própria.

Gráfico 7 – Correlação entre Ideb e P – Anos Finais do Ensino Fundamental (AF) por estado

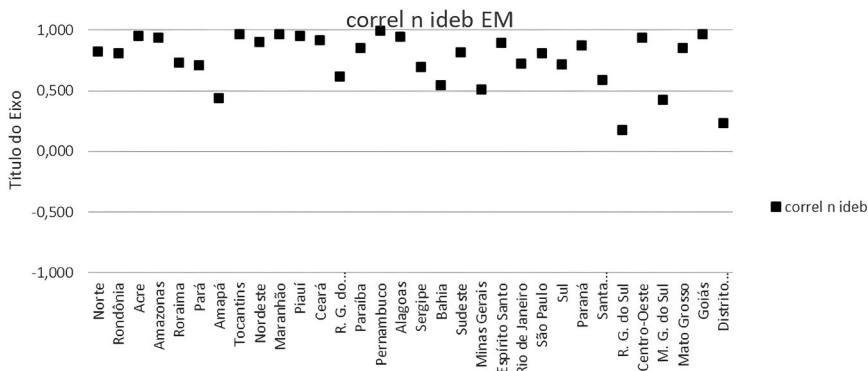


Fonte: Elaboração própria.

Observa-se, nos gráficos 6 e 7, uma correlação alta e consistente entre o Ideb e seus componentes em quase todos os estados. Chama a atenção o estado de Sergipe, que possui a correlação mais baixa entre o Ideb e N ($r = 0,65$) e Amapá, com a menor correlação entre o Ideb e P ($r = 0,38$). Os próximos gráficos apresentam os resultados para o Ensino Médio, que é o único

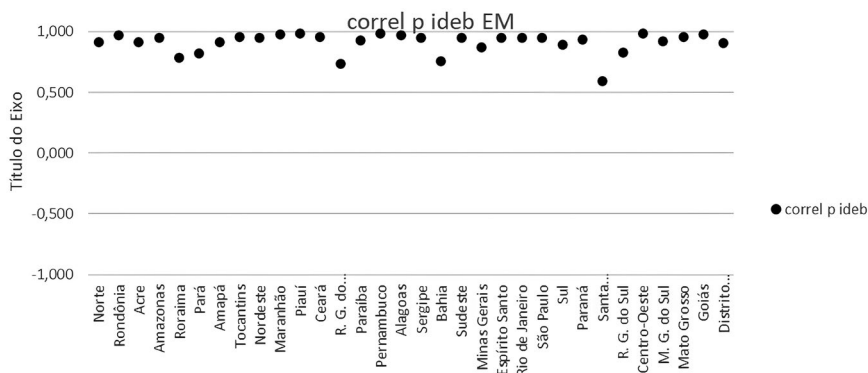
nível de ensino em que não há consistência no padrão de correlações entre os estados, especialmente entre Ideb e N, como pode ser visto a seguir.

Gráfico 8 – Correlação entre Ideb e N – Ensino médio por estado



Fonte: Elaboração própria.

Gráfico 9 – Correlação entre Ideb e P – Ensino Médio por estado



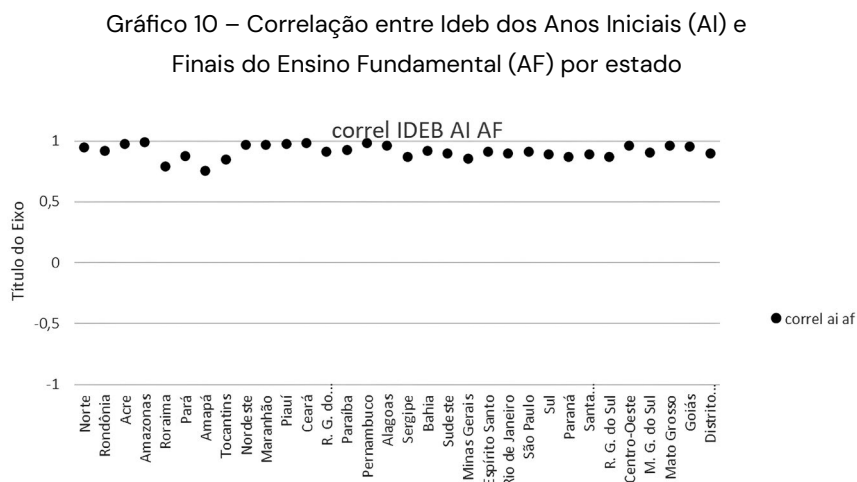
Fonte: Elaboração própria.

No Ensino Médio, observa-se que as correlações entre o Ideb e P são todas superiores a $r \geq 0,59$. Todavia, no que se refere às correlações entre o Ideb e N, verifica-se que há valores muito baixos, por exemplo, no Rio Grande do Sul ($r = 0,17$) e no Distrito Federal ($r = 0,23$). Esse resultado demonstra que, em unidades da federação como essas, não se

pode relacionar o resultado do Ideb aos resultados obtidos nos testes padronizados aplicados pelo Saeb, mas se relacionam fortemente às taxas de aprovação.

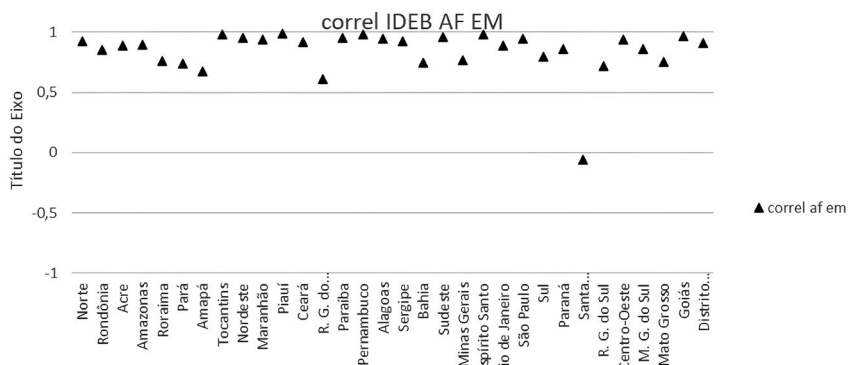
Correlação entre Ideb e Ideb

Por fim, foram realizadas correlações entre os resultados do Ideb de diferentes níveis, visando observar se há consistência entre os estados nos resultados obtidos no Ideb dos diferentes níveis de ensino, ou seja, se um bom resultado no Ideb nos Anos Iniciais do Ensino Fundamental se relaciona a bons resultados também nos Anos Finais do Ensino Fundamental e, por conseguinte, no Ensino Médio. Os próximos gráficos apresentam os resultados.



Fonte: Elaboração própria.

Gráfico 11 – Correlação entre o Ideb dos Anos Finais do Ensino Fundamental (AF) e Ensino Médio (EM) por estado



Fonte: Elaboração própria.

Quanto à relação entre o Ideb dos Anos Iniciais e Finais do Ensino Fundamental, observa-se que há uma boa relação entre os resultados desses dois níveis de ensino nos estados. Não obstante, a relação entre o Ideb dos Anos Finais do Ensino Fundamental e do Ensino Médio não segue uma regularidade entre os estados. Chama a atenção Santa Catarina, com correlação negativa entre o resultado obtido no Ideb dos Anos Finais do Ensino Fundamental e o obtido no Ensino Médio ($r = -0,06$).

Projeções e resultados das metas do PNE 2014-2024

Por fim, foi feita a comparação entre as metas projetadas pelo PNE 2014-2024 e os resultados obtidos até o ano de 2021 para o Ideb. O quadro a seguir apresenta os resultados.

Quadro 2 – Médias projetadas para o Ideb no PNE
2014-2024 e os resultados obtidos

IDEB	Meta	Saeb	Meta	Saeb	Meta	Saeb	Meta	Saeb
	2015	2015	2017	2017	2019	2019	2021	2021
Anos Iniciais do Ensino Fundamental	5,2	5,5	5,5	5,8	5,7	5,9	6,0	5,8
Anos Finais do Ensino Fundamental	4,7	4,5	5,0	4,7	5,2	4,9	5,5	5,1
Ensino Médio	4,3	3,7	4,7	3,8	5,0	4,2	5,2	4,2

Fonte: PNE 2014-2024 (Brasil, 2014) e INEP.⁵

Como pode ser observado, no fechamento do ciclo de acompanhamento do Ideb, ano de 2021, nenhuma das metas projetadas foi alcançada. Embora tenha sido um ano ainda afetado pela pandemia de Covid-19, o que certamente deve ter influenciado os resultados de N de forma negativa, mas também deve ter afetado o resultado de P com a política de não re-provação implementada durante a pandemia, entre os anos de 2020 e 2021.

Conclusão

O presente estudo teve como finalidade analisar o Ideb entre os anos de 2005 e 2021, evidenciando seus limites e possibilidades enquanto indicador de monitoramento da qualidade da Educação Básica no Brasil.

Muitas são as críticas que vêm sendo feitas em relação ao uso do Ideb como indicador de qualidade da Educação Básica, especialmente por desconsiderar qualquer informação contextual no seu cálculo, sendo restrito apenas à avaliação da aprendizagem (Almeida; Dalben; Freitas, 2013; Pestana, 2016; Travitzki, 2020; Ramos *et al.*, 2023; Soares; Soares; Santos, 2023).

5 Os dados podem ser acessados no link: https://download.inep.gov.br/ideb/resultados/apresentacao_ideb_2021.pdf

Neste estudo, os resultados encontrados indicam que há algumas inconsistências nos resultados apresentados pelo Ideb nos estados. Ao analisar as correlações de toda a série histórica proposta para o acompanhamento do Ideb, observaram-se alguns estados com correlações baixas nos Anos Finais do Ensino Fundamental entre os dois componentes do indicador, desempenho no Saeb (N) e taxas de aprovação (P).

Não obstante, no Ensino Médio, os resultados foram ainda mais inconsistentes com o que se esperava, uma vez que cinco estados apresentaram correlações negativas entre P e N, ou seja, o aumento nas taxas de aprovação correspondendo à diminuição do aprendizado avaliado pelo Saeb. Quando Fernandes e Gremaud (2009) propuseram a composição do Ideb pelo fluxo escolar e pelo resultado do Saeb, esperavam que um fator moderasse o efeito de outro, uma vez que ambos avaliam o mesmo construto, a aprendizagem, mas sob olhares diferenciados, interno e externo.

Ao analisar a relação entre o Ideb e seus componentes, N e P, observou-se neste estudo que, no Ensino Médio o fluxo (P) se relaciona de forma mais consistente com o Ideb do que o desempenho no Saeb (N). Esse resultado pode indicar que o fluxo infla artificialmente o Ideb no ensino médio, resultado semelhante ao encontrado por Figueiredo *et al.* (2018). Tal fato pode resultar de políticas voltadas para o incremento da aprovação nesse nível de ensino, ou o descompromisso dos estudantes desse nível de ensino com as avaliações externas.

Esperava-se encontrar uma relação positiva entre os resultados do Ideb nos diferentes níveis de ensino. No Ensino Fundamental, foi possível observar uma consistência maior entre os estados nessa relação, entre o Ideb dos Anos e Finais do Ensino Fundamental. Todavia, é interessante observar, novamente no Ensino Médio, correlacionado com o resultado do Ideb dos Anos Finais do Ensino Fundamental, que há diversas correlações moderadas e há, inclusive, uma pequena correlação negativa (estado de Santa Catarina) entre o Ideb dos Anos Finais do Ensino Fundamental e o do Ensino Médio. Esse resultado é, no mínimo, curioso.

Nesse contexto, considerando os resultados encontrados no presente estudo em relação ao comportamento do Ideb e de seus componentes

nos diferentes estados, questiona-se até que ponto esse indicador é uma ferramenta adequada para o monitoramento da meta 7 do PNE 2014-2024, que pretende fomentar a qualidade da Educação Básica em todos os níveis e modalidades, olhando apenas o fluxo indicado pelas escolas e a aprendizagem avaliada por meio do Saeb. Os resultados encontrados no presente estudo evidenciam muitas incoerências entre os resultados de fluxo e aprendizagem, além de mostrar muitas relações inconsistentes entre o Ideb do Ensino Fundamental e o do Ensino Médio.

Entre 2015 e 2021, as metas traçadas no PNE 2014-2024 nunca foram alcançadas para os Anos Finais do Ensino Fundamental e para o Ensino Médio. Embora alguns estados se destaquem com bons resultados ao longo da série histórica (Gramani, 2017), apenas nos Anos Iniciais do Ensino Fundamental, foi possível alcançar as metas propostas entre 2015 e 2019. Esse fato pode levar alguns a pensarem que, ao longo da vigência do atual PNE, não foi possível fomentar uma educação de qualidade para o Ensino Fundamental e Médio no Brasil. Ou esse pensamento está correto, ou o PNE 2014-2024 tem utilizado medidas não adequadas para monitorar a qualidade da Educação Básica brasileira.

A qualidade da educação é um conceito multifatorial (Charlot, 2021; Dourado; Oliveira; Santos, 2007). Dessa forma, é de se esperar que qualquer tentativa de o mensurar envolva muitas fontes de informação e, conseqüentemente, diversificadas ferramentas de coleta de informação, tanto quantitativas quanto qualitativas (Tolentino-Neto; Amestoy, 2023).

Ao utilizar um indicador para monitorar a qualidade, é imprescindível que se considere a multifatorialidade desse construto. Nesse sentido, ao invés de se utilizar um único indicador sintético, como o Ideb, para monitorar a qualidade da educação, o ideal seria utilizar um painel de indicadores, contemplando fatores intra e extraescolares, inclusive fatores relacionados ao acesso à e permanência na escola.

Referências

- ALMEIDA, L. C.; DALBEN, A.; FREITAS, L. C. O Ideb: Limites e ilusões de uma política educacional. **Educação e Sociedade**, [S. l.], v. 34, n. 125, p. 1153–1174, 2013. Disponível em: <https://www.scielo.br/j/es/a/FGHLWhm47PZpFShWNxs5GLM/abstract/?lang=pt>. Acesso em: 26 nov. 2025.
- BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996**. Estabelece as diretrizes e bases da educação nacional. Brasília, DF: Diário Oficial da União, 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 26 nov. 2025.
- BRASIL. **Lei nº 13.005, de 25 de junho de 2014**. Aprova o Plano Nacional de Educação – PNE e dá outras providências. Brasília, DF: Diário Oficial da União, 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm. Acesso em: 26 nov. 2025.
- BRASIL. **Lei nº 5.692, de 11 de agosto de 1971**. Fixa Diretrizes e Bases para o ensino de 1º e 2º graus. Brasília, DF: Diário Oficial da União, 1971. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l5692.htm. Acesso em: 26 nov. 2025.
- BRASIL. **Lei nº 10.172, de 9 de janeiro de 2001**. Aprova o Plano Nacional de Educação. Brasília, DF: Diário Oficial da União, 2001. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10172.htm. Acesso em: 26 nov. 2025.
- BRASIL. **Constituição da República Federativa do Brasil**. Brasília, DF: Senado, 5 out. 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 26 nov. 2025.
- CHARLOT, B. “Qualidade da educação”: o nascimento de um conceito ambíguo. **Educar em Revista**, Curitiba, v. 37, p. 1–15, 2021. <https://www.scielo.br/j/er/a/wBJC4X4xMX9t4S3bcLxNJtk/?format=html&lang=pt>. Acesso em: 26 nov. 2025.
- D’ESCOFFIER, L. N. *et al.* Avaliação educacional na produção científica vista por meio de análise de redes sociais. **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 30, n. 115, p. 285–311, 2021. Disponível em: <https://www.scielo.br/j/ensaio/a/BCKB3KnPWdzYyQxvLYy6zcv/?lang=pt>. Acesso em: 26 nov. 2025.
- DOURADO, L. F.; OLIVEIRA, J. F.; SANTOS, C.A. A qualidade da educação: conceitos e definições. **Série Documental: Textos para Discussão**, Brasília, DF, v. 24, n. 22, p. 5–34. 2007. Disponível em: <https://www2.unifap.br/gpcem/files/2011/09/A-Qualidade-na-educacao-DISCUSS%C3%83O-N%C2%BA-24.pdf>. Acesso em: 26 nov. 2025.

FERNANDES, R. **Índice de Desenvolvimento da Educação Básica (Ideb)**. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2007.

FERNANDES, R.; GREMAUD, A. P. Qualidade da educação: avaliação, indicadores e metas. In: VELOSO, F.; PESSOA, S.; HENRIQUES, R.; GIAMBIAGI, F. (orgs.). **Educação básica no Brasil** – construindo o país do futuro. Rio de Janeiro: Elsevier, 2009. p. 213–238.

FIGUEIREDO, D. *et al.* Os cavalos também caem: Tratado das inconsistências do IDEB. **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 26, n. 100, p. 552–572, 2018. DOI: <https://doi.org/10.1590/s0104-403620180026001178>. Disponível em: <https://www.scielo.br/j/ensaio/a/7x9tyCtnSPdwJ9fCZwgr9ZF/abstract/?lang=pt>. Acesso em: 26 nov. 2025.

GRAMANI, M. C. Análise dos determinantes de eficiência educacional do estado do Ceará. **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 25, n. 95, p. 507–526, 2017. Disponível em: <https://www.scielo.br/j/ensaio/a/ymFRRCY7KQF3kX8bBtJxzvB/abstract/?lang=pt>. Acesso em: 26 nov. 2025.

HORTA NETO, J. L. Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 91, n. 227, p. 84–104, jan./abr. 2010. Disponível em: <https://rbep.inep.gov.br/ojs3/index.php/rbep/article/view/2836>. Acesso em: 26 nov. 2025.

JANNUZZI, P. de M. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil. **Revista do Serviço Público**, [S. l.], v. 56, n. 2, p. 137–160, 2014. DOI:10.21874/rsp.v56i2.222. Disponível em: <https://revista.enap.gov.br/index.php/RSP/article/view/222>. Acesso em: 26 nov. 2025.

PESTANA, M. I. Trajetória do SAEB: criação, amadurecimento e desafios. **Em Aberto**, Brasília, v. 29, n. 96, p. 71–84, 2016. <https://doi.org/10.24109/2176-6673.emaberto.29i96%25p>. <https://rbep.inep.gov.br/ojs3/index.php/emaberto/article/view/3152>. Acesso em: 26/11/2025

RAMOS, M. N.; SILVA, L.; SILVA, J. B. P.; COSTA FILHO, J. Uma análise longitudinal do Ideb em municípios de médio porte do estado de São Paulo. **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 31, n. 118, p. 1–22, 2023. DOI: <https://doi.org/10.1590/S0104-40362023003103818>. Disponível em: <https://www.scielo.br/j/ensaio/a/QFQKGrvsZhFdsnM7WHztbFK/?format=html&lang=pt>. Acesso em: 26 nov. 2025.

RIBEIRO, R. M.; SOUSA, S. Z. A controvérsia sobre avaliações em larga escala no Brasil: continuum argumentativo. **Educação e Pesquisa**, [S. l.], v. 49, p. 1-19, 2023. DOI: <https://doi.org/10.1590/S1678-4634202349250287>. Disponível em: <https://www.scielo.br/j/ep/a/QDntdnHTXyfm3vpSCk4s46t/?format=html&lang=pt>. Acesso em: 26 nov. 2025.

SCHNEIDER, M. P.; NARDI, E. L. O Ideb e a construção de um modelo de accountability na educação básica brasileira. **Revista Portuguesa de Educação**, Braga, v. 27, n. 1, p. 7-28, jun. 2014. Disponível em: <https://revistas.rcaap.pt/rpe/article/view/4295>. Acesso em: 26 nov. 2025.

SOARES, J. F.; XAVIER, F. P. Pressupostos educacionais e estatísticos do IDEB. **Educação & Sociedade**, [S. l.], v. 34, n. 124, p. 903-23. 2013. DOI: <https://doi.org/10.1590/S0101-7330201300030001>. Disponível em: <https://www.scielo.br/j/es/a/JLzr4qdx89rjrNXnydNcvy/abstract/?lang=pt>. Acesso em: 26 nov. 2025.

SOARES, M. J. D.; SOARES, T. E. A.; SANTOS, W. O algoritmo do Ideb e as metas projetadas para a educação brasileira: uma análise estatístico-matemática. **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 31, n. 118, p. 1-25, 2023. Disponível em: <https://www.scielo.br/j/ensaio/a/PjCsK4tm53TfqMtLqmYxXFM/?format=pdf&lang=pt>. Acesso em: 26 nov. 2025.

TOLENTINO-NETO, L. C. B.; AMESTOY, M. B. **Avaliações externas na educação básica**: contextos, políticas e desafio. São Paulo: Cortez Editora, 2023.

SAVIANI, D. O Plano de Desenvolvimento da Educação: análise do projeto do MEC. **Educ. Soc.**, Campinas, v. 28, n. 100, p. 1231-1255, 2007.

TRAVITZKI, R. Qual é o grau de incerteza do Ideb e por que isso importa? **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 28, n. 107, p. 500-520, 2020. Disponível em: <https://www.scielo.br/j/ensaio/a/pQL73NQVkgPcWSxDtz-Fb5Tj/?lang=pt>. Acesso em: 26 nov. 2025.

SOLIGO, V. Indicadores: conceito e complexidade do mensurar em estudos de fenômenos sociais. **Estudos em Avaliação Educacional**, São Paulo, v. 23, n. 52, p. 12-25, 2012. DOI: [10.18222/eaee235220121926](https://doi.org/10.18222/eaee235220121926). Disponível em: <https://publicacoes.fcc.org.br/eaee/article/view/1926>. Acesso em: 26 nov. 2025.

CAPÍTULO 8

AVALIAÇÕES EXTERNAS E IDEB: PERCEPÇÕES SOBRE USOS E IMPLICAÇÕES NO TRABALHO ESCOLAR

Daiane Ansaloni Mendes¹
Daniel Abud Seabra Matos²
Breyner Ricardo de Oliveira³
DOI: 10.52695/978-65-5456-163-1.8

Introdução

Neste capítulo, abordamos um tema que vem sendo muito discutido na literatura educacional: as avaliações externas e o Ideb. Nosso objetivo foi analisar as percepções acerca dos usos e das implicações das avaliações externas e do Ideb no trabalho escolar.

Em relação ao percurso metodológico, na primeira etapa, realizamos uma revisão da literatura, com o intuito de mapear os estudos sobre as percepções dos usos e implicações das avaliações externas e do Ideb no cotidiano de trabalho dos profissionais da educação. A busca pelos trabalhos sobre os

1 Universidade Federal de Ouro Preto. E-mail: daianemendes20@yahoo.com.br.

2 Universidade Federal de Ouro Preto.

3 Universidade Federal de Ouro Preto.

usos e as implicações foi realizada nos seguintes bancos de dados: *Scientific Electronic Library Online* (SciELO), Associação Nacional de Pós-graduação e Pesquisa em Educação (ANPEd), Biblioteca Brasileira de Teses e Dissertações (BBTD), Associação Brasileira de Avaliação Educacional (ABAVE) e Portal de Periódicos da CAPES. No total, foram encontrados 23 estudos sobre o Ideb e 28 sobre avaliações externas, especificamente voltados para os usos e as implicações no cotidiano de trabalho escolar.

Este capítulo foi organizado em quatro seções: introdução; revisão da literatura referente ao Ideb e às avaliações externas; análise das percepções e implicações dessas avaliações e do Ideb no cotidiano do trabalho escolar; e, por fim, as considerações finais.

Avaliações externas e Ideb: revisão da literatura

Índice de Desenvolvimento da Educação Básica (IdEB)

O Índice de Desenvolvimento da Educação Básica (Ideb), criado em 2007, é um indicador de qualidade educacional a ser utilizado para orientar a criação de políticas públicas educacionais, o financiamento da educação e para informar a população em geral sobre a qualidade educacional (Almeida; Dalben; Freitas, 2013). A sua criação associa-se ao Plano de Metas Compromisso Todos pela Educação. O Plano de Metas Compromisso Todos pela Educação foi instituído por meio do Decreto nº 6.094, de 24 de abril de 2007 e está inserido no Plano de Desenvolvimento da Educação (PDE). O plano buscou unir ações entre os municípios, o Distrito Federal e os estados na tentativa de alcançar melhorias para a Educação Básica.

O indicador assume o papel de indutor de políticas para a melhoria da Educação Básica através do termo de adesão ao Plano de Metas Compromisso Todos pela Educação pelos municípios e estados. As metas almejam que o Brasil alcance um Ideb igual a 6,0 na média nacional (Alves; Soares, 2013). A adesão ao compromisso se configura de forma voluntária.

É importante destacar que o princípio básico do Ideb é o de que a qualidade da educação se caracteriza quando o aluno aprende e passa

de uma série para a outra, ou seja, é aprovado (Franco; Alves; Bonamino, 2007). Nessa mesma direção, Fernandes (2007, p. 7) afirma:

Um sistema educacional que reprova sistematicamente seus estudantes, fazendo que grande parte deles abandone a escola antes de completar a Educação Básica, não é desejável, mesmo que aqueles que concluem essa etapa atinjam elevadas pontuações nos exames padronizados. Por seu lado, um sistema em que os alunos concluem o ensino médio no período correto não é de interesse caso eles aprendam muito pouco. Em suma, um sistema ideal seria aquele no qual todas as crianças e adolescentes tivessem acesso à escola, não desperdiçassem tempo com repetências, não abandonassem os estudos precocemente e, ao final de tudo, aprendessem.

A forma que o Ministério da Educação (MEC) encontrou para monitorar e fazer as intervenções necessárias, de maneira que todos os envolvidos se comprometessem com o desafio de melhorar o ensino, foi por meio do acordo Compromisso Todos pela Educação ao estabelecer que cada governo, na sua localidade, desenvolvesse estratégias para a melhoria do Ideb. As escolas que alcançassem as metas iriam receber um aumento de recursos no Programa Dinheiro Direto na Escola (PDDE) (Macêdo, 2011).

Nesse contexto, na tentativa de alcançar a qualidade educacional, a meta estabelecida para as escolas seria alcançar a média 6,0 no Ideb até o ano de 2021 (Fernandes, 2007). Essa média é considerada compatível com as dos países desenvolvidos.

O Ideb é calculado a partir de dois componentes: a média de desempenho e a taxa de rendimento. O desempenho da escola, item para o cálculo do índice, é obtido por meio da média de proficiência em leitura e matemática dos alunos no Saeb. Já o rendimento é calculado a partir do valor médio das taxas de aprovação da escola. Como demonstram Soares e Xavier (2013,

p. 911), “o Ideb sintetiza os indicadores de rendimento e desempenho em um único número, através do produto dos dois indicadores”.

Ainda em relação aos componentes do Ideb, Soares e Xavier (2013) ampliam nosso entendimento:

a) O indicador de desempenho: o desempenho de uma determinada escola é adquirido por meio das médias das proficiências em leitura e matemática, obtidas pelos seus alunos no Saeb. As escalas das medidas são diferentes; assim, é realizada uma padronização dessas medidas antes de calcular a sua média.

b) O indicador de rendimento: é a síntese da aprovação dos alunos de uma escola. O indicador de rendimento pode ser entendido como a razão entre o número de anos da etapa e o número de anos letivos que um aluno gasta para fazer uma determinada série. Ainda segundo Soares e Xavier (2013, p. 910): “A soma das taxas de aprovação, reprovação e abandono é 100%”.

Em uma escala de 0 a 10, o Ideb informa a qualidade da educação através de um único número e torna possível comparar as escolas da rede de ensino estaduais e municipais, bem como acompanhar a evolução dos resultados ao longo dos anos. A aparente “simplicidade” do resultado, por meio de um único número, gerou uma rápida aceitação pelos gestores da educação (Alves; Soares, 2013).

O Ideb representa um avanço na educação por enfatizar informações sobre o aprendizado dos alunos. De acordo com Soares e Xavier (2013), conhecer o rendimento e o aprendizado dos alunos é uma forma de o Estado verificar se o direito à educação está sendo alcançado. Os autores ainda afirmam que, “na ausência de um sistema como esse, o direito público subjetivo estabelecido no texto constitucional não pode ser monitorado e, eventualmente, exigido” (Soares; Xavier, 2013, p. 905).

Outra significativa contribuição é que o índice levantou uma importante questão: as escolas não devem ser avaliadas somente pelo processo de ensino e gestão, mas também pelo aprendizado e pela trajetória escolar dos alunos (Soares; Xavier, 2013). Assim, “o Ideb, sem questionar a

necessidade de novos recursos e expansões, coloca o aprendizado e a regularidade na trajetória escolar dos alunos como elementos essenciais de um sistema educacional” (Soares; Xavier, 2013, p. 904).

Freitas e Ovando (2015, p. 969) ressaltam que o Ideb tem sido uma das ferramentas utilizadas pela União na coordenação nacional da Educação Básica, já que torna possível:

- a. Articular ferramentas de medida, avaliação e formação de que dispunha desde os anos 1990, canalizando-as para as funções de indução e regulação da política de melhoria da qualidade do ensino no País;
- b. Mapear a situação das redes e unidades escolares, identificando os seus patamares no ano de 2005, com base nos quais pode projetar avanços mínimos que traduziu em metas bienais a serem alcançadas até 2021;
- c. Fixar parâmetro médio nacional da qualidade de ensino, nos indicadores priorizados, expresso na média 6,0 que, referida no Programa de Avaliação Internacional Comparada (Pisa), estreita conexões da política educacional interna com “guiões” internacionais;
- d. Precisar compromissos de melhoria a serem atendidos pelos entes subnacionais, estabelecendo em que melhorar, quanto (ao menos), em que ritmo (mínimo), em que prazo (máximo).

Diante do que foi exposto, não podemos negar a importância e as contribuições do Ideb, mas também é necessário levar em conta seus limites e críticas. Assim, Soares e Xavier (2013) discutem algumas das limitações deste indicador:

a) Só os alunos presentes ao teste são considerados: uma das limitações dos dados para o cálculo do indicador de desempenho é o fato de considerar apenas os alunos presentes no dia da Prova Brasil. Nesse

sentido, nem todos os alunos fazem os testes que foram enviados para a escola. De acordo com o MEC, o indicador é calculado desde que tenha 50% dos alunos presentes. Assim, há o risco de as escolas escolherem os seus melhores alunos;

b) Proficiência em matemática tem mais peso do que proficiência em leitura: os algoritmos utilizados para padronizar as proficiências têm como consequência o fato de que o domínio da competência matemática pelos alunos é maior do que o domínio da compreensão leitora. Essa estatística pode ser utilizada como estratégia para melhorar o Ideb, bastando enfatizar o ensino de matemática para determinados alunos;

c) Assume substituições questionáveis entre os diferentes componentes: sendo o numerador do Ideb uma média, o indicador considera que o bom desempenho de um aluno compensa o mau desempenho de outro e que um melhor desempenho compensa uma taxa de reprovação mais alta. Esses dois fatores podem gerar a exclusão educacional, pois, na tentativa de se conseguir alcançar melhores valores, podem-se escolher os alunos para realizar as provas;

d) Não se pode usar a metáfora da nota escolar para analisar o Ideb: uma vez que o Ideb tem uma apresentação simples, variando de 0 a 10, é frequente interpretá-lo como conceitos escolares. Mas a escala é de difícil interpretação. Por exemplo: um Ideb 7 é excelente. Além disso, os valores extremos da escala do Ideb, acima de 6 e abaixo de 3, são raros e valores acima de 8 e abaixo de 2 são quase impossíveis;

e) Naturaliza baixos desempenhos de muitos alunos: o resultado do Ideb igual a 6,0 e o seu uso isolado fazem com que as dificuldades dos alunos não tenham nenhuma relevância, porque esse resultado está em um nível tido como adequado;

f) O Ideb é muito correlacionado com o nível socioeconômico da escola: a literatura tem demonstrado que o desempenho dos alunos está associado ao seu nível socioeconômico. Assim, é esperado que a associação entre o Ideb e o NSE das escolas seja alta.

Existem muitas críticas também com relação a esse índice educacional. Para Camargo, Pinto e Guimarães (2008, p. 832), o Ideb apresenta alguns limites estruturais. E os autores ainda assinalam que:

Falta ao Ideb uma dimensão que leve em conta os insumos educacionais (qualificação dos professores, recursos colocados à disposição dos alunos etc.). Da maneira como o índice está montado, ao invés de investir em medidas de longo prazo que sabidamente melhoram a qualidade do ensino, os municípios, para não verem minguar o repasse de recursos do governo federal, tendem a adotar medidas cosméticas, ou mesmo a valer-se de fraude. A imprensa noticiou casos de escolas que diziam a determinados alunos, aqueles que tinham as maiores dificuldades, que não precisariam vir à escola no dia da Prova Brasil, por exemplo.

Para Chirinéa e Brandão (2015), o índice torna-se insuficiente para a medida de qualidade da educação por levar em conta apenas duas variáveis e por desconsiderar outros fatores, como a cultura organizacional da escola, a prática docente, o nível socioeconômico e cultural das famílias e o estilo de gestão e liderança. Nesse sentido, não demonstra a realidade das escolas. Sobre essa insuficiência do Ideb, os mesmos autores apontam para outro caminho, que mostra a importância de articular o desempenho escolar aos contextos intraescolares para a efetiva intervenção pedagógica na escola:

Essa integração dará condições para a gestão escolar traçar suas próprias metas e estratégias, e definir intervenções alicerçadas na democratização e na autonomia, sem que isso signifique realizar manobras pedagógicas para atender as metas pré-estabelecidas pelo índice [...] (Chirinéa; Brandão, 2015, p. 479).

Outra questão é que, apesar de os documentos oficiais não mencionarem, o índice pode estar atrelado aos mecanismos de controle interno

e social, no que se refere aos recursos da educação, e pode ser apontado como a efetivação da política de *accountability*⁴ educacional (Vidal; Vieira, 2011). Assim, segundo as autoras, “a forma velada de como tais mecanismos vêm sendo implementados pode estar associada ao fato desse modelo de responsabilização ter sido desenvolvido e adotado por países cuja ideologia é de cunho assumidamente neoliberal [...]” (Vidal; Vieira, 2011, p. 426). Nesse cenário, também se configuram as formas de o estado monitorar as escolas.

Para Freitas (2007), por exemplo, o Ideb não é um padrão de referência para verificar a qualidade educacional, pois não considera as desigualdades sociais e os diferentes contextos (perfil de alunos, clima da escola, nível socioeconômico, etc.). Além disso, o “equilíbrio ótimo entre desempenho e aprovação para elevar a média do Ideb pode ter como efeito perverso o aumento da desigualdade entre escolas e, principalmente, dentro das escolas, por meio da segregação de alunos com desempenho mais fraco” (Alves; Soares, 2013, p. 182).

Portanto, as escolas têm pontos de partida desiguais: escolas favorecidas culturalmente, economicamente e com um bom clima escolar, por exemplo, podem apresentar melhor desempenho no Ideb e também nas avaliações externas. Assim, “espera-se, portanto, que uma escola que atende a alunos com maior poder aquisitivo e com mais recursos culturais atinja valores mais altos do Ideb [...]” (Alves; Soares, 2013, p. 183).

Tendo em vista as questões apontadas, não se nega a importância dos indicadores para o desenvolvimento da educação. No entanto, é necessário ter cautela para que o resultado do Ideb obtido por uma escola

4 Sobre o termo *accountability* educacional, Vidal e Vieira (2011, p. 425) afirmam que “é proveniente da língua inglesa e vem sendo traduzido como prestação de contas e responsabilização. Neste caso, a prestação de contas se refere ao processo pelo qual os sujeitos informam e/ou exigem informações acerca do uso dos recursos financeiros, humanos, materiais etc. para a obtenção de um determinado objetivo. A responsabilização está intimamente associada à prestação de contas, uma vez que ser responsável por um processo ou um resultado implica – sobretudo no setor público em uma sociedade democrática – em ter que prestar contas dos resultados e do uso dos recursos aplicados”.

não seja interpretado como um reflexo absoluto de sua realidade em relação à aprendizagem. Ferreira e Vieira (2010, p. 248) também alertam para essa limitação:

Nesse processo, não podemos esquecer um dos aspectos fundamentais relativos à construção de indicadores que possam ser confiáveis. Para tanto, não podemos esquecer que os mesmos precisam ser considerados em função de seus limites e possibilidades como identificadores de recortes da realidade que se pretende avaliar. Só relativizando os modelos de avaliação adotados é que poderemos pensar em possíveis cruzamentos e articulações que favoreçam a construção de olhares sob os diversos aspectos e resultados encontrados decorrentes dos recortes feitos sobre a realidade concreta.

Assim, precisamos analisar as escolas para além dos resultados, já que alguns fatores internos e externos podem contribuir ou interferir na qualidade do aprendizado. Os resultados do Ideb devem ser verificados a partir da contextualização com a real situação escolar. Nesse sentido, “um sistema educacional só pode ser dito de qualidade se suas desigualdades são também consideradas na análise de seu desempenho [...]” (Alves; Soares, 2013, p. 190).

Avaliações externas

A avaliação externa, que também é abordada na literatura como avaliação em larga escala ou como avaliação de sistemas educacionais, é definida da seguinte forma:

É um instrumento de acompanhamento global de redes de ensino com o objetivo de traçar séries históricas do desempenho dos sistemas, que permitam verificar tendências ao longo do tempo, com a finalidade de reorientar políticas públicas

(Freitas, 2009, p. 47 *apud* Alavarse; Bravo; Machado, 2012, p. 2).

A década de 1990 é considerada um marco para as avaliações externas, pois foi quando, de fato, ocorreu a sua implantação nas escolas públicas. Nesse mesmo período, o Brasil passou por uma reforma de Estado que resultou em uma reforma educacional. A necessidade de regulação e de fiscalização por parte do Estado se deu por meio das avaliações externas nas escolas (Polato, 2014). Portanto, as mudanças e as reformas no cenário educacional brasileiro foram acompanhadas pelo crescimento das avaliações externas.

Uma importante mudança, ocasionada pela reforma educacional, foi a promulgação da Lei de Diretrizes e Bases (LDB), de 1996, que colocou a educação diante de um novo formato e de novas finalidades, já que, até então, o ensino brasileiro estava baseado na reforma desenvolvida pelo regime militar entre 1964 e 1970 (Durham, 2010). A LDB fortaleceu a responsabilidade dos estados e municípios: a universalização da Educação Infantil e do Ensino Fundamental tornou-se responsabilidade dos municípios, e o Ensino Médio ficou a cargo dos Estados (Durham, 2010). A LDB (Brasil, 1996, art. 9, inc. V-VI) reforça as avaliações externas na escola e incube a União de:

V – Coletar, analisar e disseminar informações sobre a educação;

VI – Assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino.

Uma das avaliações que contribuiu para solidificar a cultura das avaliações externas no Brasil foi o SAEB (1990), que será mais bem detalhado nos trechos a seguir. Ele foi criado para coletar informações sobre o desempenho da Educação Básica de todo país, abrangendo a rede estadual e a rede municipal (Castro, 2000). A Portaria nº 447, de 24 de maio de 2017, explicita um dos principais objetivos do SAEB:

Parágrafo único. O SAEB, por meio da coleta de dados junto aos sistemas de ensino e às escolas brasileiras, tem como um dos seus principais objetivos avaliar a qualidade da educação nacional e, assim, oferecer subsídios para a formulação, reformulação e monitoramento das políticas públicas educacionais (Brasil, 2017, art. 1).

O SAEB passou por algumas reformulações ao longo dos anos, como, por exemplo, as mudanças das séries e das disciplinas avaliadas. As alterações podem ser observadas no quadro 1.

Quadro 1 – Evolução do Saeb ao longo dos anos (continua)

Edição	Série avaliada	Disciplina
1990	1ª, 3ª, 5ª e 7ª séries do EF	Língua Portuguesa, Matemática, Ciências Naturais, Redação
1993	1ª, 3ª, 5ª e 7ª séries do EF	Língua Portuguesa, Matemática, Ciências Naturais, Redação
1995	1ª, 3ª, 5ª e 7ª séries do EF	Língua Portuguesa, Matemática, Ciências Naturais, Redação
1997	4ª, 8ª séries do EF; 3ª série do EM	Língua Portuguesa, Matemática, Ciências Naturais, Física, Química e Biologia
1999	4ª, 8ª séries do EF; 3ª série do EM	Língua Portuguesa, Matemática, Ciências Naturais, Física, Química e Biologia História e Geografia
2001	4ª, 8ª séries do EF; 3ª série do EM	Língua Portuguesa, Matemática
2013	5º ano, 9º ano do EF; 3ª série do EM	Língua Portuguesa, Matemática
2015	5º ano, 9º ano do EF; 3ª série do EM	Língua Portuguesa, Matemática
2016	3º ano do EF	Leitura, Escrita e Matemática

Quadro 1 – Evolução do Saeb ao longo dos anos (conclusão)

2017	5º e 9º ano do EF (Escolas públicas (censitária) Escolas privadas (amostral) e 3ª e 4ª série do EM Escolas públicas (censitária) Escolas privadas (amostral + adesão)	Língua Portuguesa e Matemática
2019	Início do alinhamento à BNCC - Creche e pré-escola da educação infantil; 2º ano do EF; 5º e 9º ano do EF; 9º ano do EF; 3ª e 4ª série do EM Educação Infantil	Língua Portuguesa e Matemática
2021		Em 2021, a aplicação do Saeb manteve as mesmas características da edição anterior, com a adoção de medidas para garantir as condições sanitárias necessárias. Nesse ciclo, foi implementada, pela primeira vez, a avaliação da Educação Infantil.

Fonte: Adaptado de INEP (2017–2023).

No primeiro ciclo, ainda em 1990, 23 Estados participaram da sua aplicação. Apenas no ano de 1995, o SAEB se tornou um sistema nacional ao contemplar o Ensino Fundamental e Médio da rede municipal, estadual e particular (Castro, 2000).

Os resultados dessa avaliação externa, desde o ano de 1990, apontaram para um baixo desempenho cognitivo médio em língua portuguesa dos alunos do Ensino Fundamental. Diante desse quadro, o governo desenvolveu algumas propostas para mudar a situação e, considerando que o Saeb não avaliava as habilidades do processo de alfabetização, criou-se a Provinha Brasil de Português, em 2008 (Brasil, 2015). Além dos resultados em língua portuguesa, os alunos também não apresentaram desempenho cognitivo médio satisfatório nas habilidades básicas de matemática (Pirola; Moraes; Sander, 2011). Portanto, para oferecer aos professores

informações sobre as habilidades de alfabetização/letramento em língua portuguesa e em matemática e para possibilitar as intervenções necessárias nas dificuldades encontradas no aprendizado dos alunos, foi criada a Provinha Brasil, que atendeu a essas duas áreas (a Provinha Brasil de Matemática foi criada em 2011).

Destaca-se que o SAEB passou por outras importantes reformulações. Entre elas, podemos destacar que, no ano de 2005, passou a ser composto por duas avaliações: a Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc), conhecida como Prova Brasil. Em 2013, a Avaliação Nacional da Alfabetização (ANA) também passou a constituir-lo (INEP, 2017).⁵

Em 2019, importantes modificações foram implementadas no sistema de avaliação da Educação Básica. Entre elas, destaca-se a unificação das siglas: as avaliações anteriormente conhecidas como Aneb (Avaliação Nacional da Educação Básica) e Anresc (Avaliação Nacional do Rendimento Escolar) foram incorporadas sob uma única nomenclatura. A partir de então, todas as avaliações passaram a ser identificadas como Saeb (Sistema de Avaliação da Educação Básica), seguidas das etapas de ensino, áreas de conhecimento e tipos de instrumentos utilizados. Essa mudança contribuiu para a padronização e mais clareza na comunicação e análise dos resultados educacionais em âmbito nacional.

Em 2023, o Sistema de Avaliação da Educação Básica (Saeb) passou por outras importantes atualizações, formalizadas pela Portaria INEP nº 267, de 10 de julho de 2023,⁶ que estabeleceu as diretrizes para a aplicação da avaliação naquele ano. O documento detalha as modificações implementadas, como a ampliação das áreas avaliadas e a inclusão de novos públicos-alvo. Dentre as principais mudanças, destacam-se: (1) a incorporação das áreas de Ciências da Natureza e Ciências Humanas para o 5º ano do Ensino Fundamental, ampliando o escopo das disciplinas

5 Disponível em: <http://portal.inep.gov.br/educacao-basica/saeb>. Acesso em: 5 jul. 2017.

6 Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/saeb/portaria-estabelece-diretrizes-para-o-saeb-2023>.

contempladas; (2) a inclusão, no público respondente, dos professores de turmas com aplicação amostral do 2º ano do Ensino Fundamental, bem como dos docentes das áreas mencionadas no 9º ano; e (3) o alinhamento das avaliações à Base Nacional Comum Curricular (BNCC), com o objetivo de refletir as competências e habilidades previstas nesse documento orientador da educação brasileira.

É importante dizer que, após a criação do SAEB, muitas outras avaliações externas foram implementadas no Brasil, incluindo avaliações municipais, estaduais, nacionais e internacionais. São alguns exemplos: SIMAVE (Sistema Mineiro de Avaliação e Equidade da Educação Pública), Provinha Brasil, ENEM (Exame Nacional do Ensino Médio), ENCCEJA (Exame Nacional para Certificação de Competências de Jovens e Adultos); PISA (Programa Internacional de Avaliação de Estudantes), Estudos Regionais Comparativos (ERCE/LLCE).

Em síntese, a avaliação externa foi se consolidando durante um período histórico recente. O Plano de Desenvolvimento da Educação (PDE, 2007), por exemplo, reafirma a importância das políticas de avaliação. Esse documento fortalece a presença da União no que se refere à regulação da Educação Básica, tornando possível a fiscalização até mesmo do financiamento de programas e projetos (Silva, 2010).

Portanto, ao mesmo tempo que as avaliações externas oferecem um diagnóstico da qualidade da educação, também indicam alguns avanços na área da avaliação educacional. Tendo em vista o contexto histórico, podemos perceber que muitos progressos vêm sendo constatados ao longo dos anos, e alguns deles podem ser identificados, como, por exemplo, a ampliação dos instrumentos de avaliação para a coleta e a divulgação de informações sobre o desempenho dos alunos.

No entanto, as avaliações externas também carregam alguns limites em seu percurso histórico. Trazendo essas avaliações para o ambiente escolar, a literatura evidencia que, apesar de elas trazerem contribuições significativas, também têm efeitos que impactam negativamente os profissionais da educação. Nesse sentido, existem diferentes percepções sobre os usos

das avaliações externas e as implicações no trabalho escolar, que serão apresentados a seguir.

Percepções sobre avaliações externas e o Ideb: revisão da literatura sobre seus usos e implicações no trabalho escolar

Avaliações externas: usos, limites e implicações na literatura educacional

Silva *et al.* (2012) identificaram vários usos e apropriações possíveis. Os autores apresentam e discutem os usos das avaliações externas em dois eixos: 1) Geração de ações por parte das escolas a partir do desenho de avaliação adotado; e 2) Geração de ações nas unidades escolares a partir dos resultados da avaliação externa:

1) Geração de ações por parte das escolas a partir do desenho de avaliação adotado (Silva *et al.*, 2012)

a) Realização de provas simuladas: é compreendida como qualquer avaliação ou prova elaborada pela escola que tenha como base as avaliações externas. A elaboração de algumas provas são adaptações das avaliações em larga escala desenvolvidas pelo governo com o intuito de imitá-las ou até mesmo sendo idênticas às questões delas. No estudo, foram verificados dois usos diferentes: i) treinar os alunos para a realização das avaliações externas, como, por exemplo, preenchimento de gabaritos; ii) compreender os conhecimentos que precisam ser trabalhados em sala de aula;

b) Alterações na concepção/forma de avaliação da aprendizagem dos professores: as avaliações externas desencadeiam alterações nas concepções e na forma das avaliações utilizadas em sala de aula, sendo possíveis mudanças na elaboração das questões nas provas. Assim, as avaliações externas tornam-se mais próximas das avaliações da sala de aula, já que a primeira é utilizada pelos docentes como um instrumento que serve como inspiração para elaborações de questões;

c) Planejamento de atividades de acordo com as matrizes de referência da avaliação externa: elaboração do planejamento que tem como

instrumento norteador os descritores e as matrizes de referências das avaliações externas. Muitos gestores orientam os professores para a utilização deles. Em algumas escolas, esse uso parte do entendimento de preparar os alunos para as avaliações externas, o que proporciona o conhecimento dos conteúdos.

2) Geração de ações nas unidades escolares a partir dos resultados da avaliação externa (Silva et al., 2012)

a) Análise dos resultados da avaliação: refere-se a cinco possibilidades de usos, sendo: i) identificar os alunos com baixo desempenho, sendo verificadas as demandas específicas para redirecionar o trabalho pedagógico; ii) identificar os conteúdos a serem trabalhados em sala de aula com os alunos; iii) acompanhar o desempenho dos alunos por meio de fichas individuais, o que pode contribuir com o trabalho dos professores que assumem uma turma pela primeira vez; iv) identificar os erros mais frequentes nas provas para facilitar a tomada de decisão no que diz respeito aos erros; v) autorreflexão: refere-se a uma avaliação por parte do professor da sua própria atuação profissional e do processo de ensino;

b) Uso dos resultados como critérios para formação de grupos de alunos: refere-se ao encaminhamento de alunos para turmas de reforço e para a organização de turmas na escola. A primeira trabalha as dificuldades dos alunos a fim de contribuir para o desempenho significativo e a segunda, para a separação de alunos com distintos níveis de proficiência;

c) Diversificação e/ou intensificação de atividades pedagógicas: mudanças na organização pedagógica, entre elas a lição de casa, em que os alunos intensificam os estudos dos conteúdos que precisam ser apropriados. Outras práticas são elaboração de exercícios relacionados às habilidades a serem desenvolvidas nas avaliações externas e a discussão com os alunos dos itens que mais erraram;

d) Exposição dos resultados aos alunos e pais: uma das ações é apresentar aos pais os resultados alcançados pelos alunos nas avaliações, sendo encaminhado o boletim individual.

e) Elaboração de Plano de Ação Escolar: nesse caso, o intuito é realizar a interpretação pedagógica dos resultados para a elaboração de metas a alcançar, muitas vezes sendo de responsabilidade dos gestores;

f) Cobrança/reconhecimento interno à escola em relação aos resultados da avaliação: refere-se à cobrança a partir dos resultados das avaliações externas por parte da escola ou da Secretaria de Educação. Em alguns casos, essa cobrança é realizada a partir da exposição dos resultados nas salas dos professores para que se sintam responsáveis pelos resultados dos alunos.

Nessa mesma direção, Brooke e Cunha (2011) analisam os impactos das avaliações externas, especialmente do Ideb, na dinâmica do trabalho escolar. Os autores realizaram um estudo dividido em duas partes: 1ª) levantamento de todas as informações sobre a avaliação educacional na *internet* e nos *sites* do governo (estadual e municipal); 2ª) realização de 20 entrevistas semiestruturadas com gestores, técnicos e especialistas em avaliação de sete estados diferentes.

O resultado do estudo indicou diversos usos das avaliações externas para direcionar a gestão educacional: 1) avaliar e orientar a política educacional; 2) informar as escolas sobre a aprendizagem dos alunos e definir as estratégias de formação continuada; 3) informar ao público; 4) alocar recursos; 5) para formular políticas de incentivos salariais; 6) como componente da política de avaliação docente; 7) certificar de alunos e escolas.

Sousa *et al.* (2015) realizaram um estudo com o objetivo de verificar os usos dos resultados das avaliações externas pela gestão escolar em cinco redes municipais de educação do estado de São Paulo. Utilizaram uma investigação qualitativa, de caráter exploratório, por meio de questionários aplicados a 131 coordenadores pedagógicos.

A pesquisa evidenciou que os resultados das avaliações externas têm sido utilizados pelos coordenadores para a organização do trabalho pedagógico, perpassando pelos seguintes usos: 1) replanejar o trabalho pedagógico; 2) comparar resultados de sua escola com os resultados obtidos por outras escolas da rede municipal; 3) estabelecer metas de melhoria do desempenho dos alunos nas próximas avaliações; 4) planejar ações peda-

gógicas visando à melhoria de aprendizagem dos alunos; 5) rever, com os professores, suas práticas pedagógicas; 6) realizar estudos e pesquisas visando ao aprimoramento profissional; 7) orientar a formação continuada de professores na escola; 8) orientar os professores que apresentem dificuldades de ensino; 9) valorizar e incentivar o uso de registros, na prática pedagógica, como uma forma de acompanhamento da aprendizagem dos alunos; 10) rever documentos pedagógicos, como a proposta pedagógica da escola, o regimento escolar, entre outros, e divulgar os resultados aos pais, responsáveis e aos demais membros da comunidade escolar.

Um estudo realizado por Borges (2016) buscou analisar a percepção de diretores de escolas estaduais de Minas Gerais sobre as implicações do Programa de Avaliação da Rede Pública de Educação Básica (PROEB) no contexto escolar. A pesquisa utilizou a abordagem qualitativa e quantitativa e teve, como instrumento para coleta de dados, o questionário. O questionário foi enviado para os diretores de todas as escolas públicas estaduais de Minas Gerais e foi respondido por 978 diretores (26,7%). O resultado indicou que, na percepção dos diretores, as avaliações externas são relevantes para as escolas. Em relação aos seus efeitos, destacam, de maneira geral, que: 1) as avaliações atuam para melhorar o ensino; 2) os professores ensinam de acordo com a proposta curricular da SEE-MG e não ocorre diminuição do tempo destinado às outras dimensões da formação do educando; 3) as avaliações não interferem de forma negativa nas práticas colaborativas entre os profissionais da escola.

Monno (2006), citado por Freitas (2014), ressalta que a Prova Brasil e o Ideb fazem parte da implementação da política de *accountability* (responsabilização) no país. Essa expressão é entendida na literatura como uma forma de responsabilizar as próprias escolas pelos resultados e pelas ações. Afonso (2012) apresenta uma abordagem crítica sobre o conceito de *accountability*. Segundo ele, a responsabilização tem um efeito negativo por “forçar” a determinadas práticas e resultados e desconsiderar alguns objetivos, a complexidade das escolas e a maneira de pensar dos profissionais da educação. Nesse sentido, “a responsabilização é mais facilmente reduzida à ameaça ou imputação negativa de culpa sobre determinadas ações e seus supostos resultados [...]” (Afonso, 2012, p. 480).

Outro exemplo de análise dos usos e implicações das avaliações externas é a pesquisa de Cerdeira, Almeida e Costa (2014), que realizaram um estudo com o objetivo de compreender a percepção sobre as influências das avaliações externas no currículo dos profissionais da educação que atuam no Ensino Fundamental. Para tornar possível a pesquisa, analisaram oito grupos focais, que tinham, como participantes, 27 gestores e 36 professores de 27 escolas, sendo 12 da cidade do Rio de Janeiro-RJ e 15 de Duque de Caxias-RJ. O grupo focal é uma técnica de caráter exploratório que visa gerar debates em torno de temas específicos. Entre os resultados sobre a percepção dos gestores e professores, destacamos: 1) os professores e os gestores são a favor das avaliações externas; 2) afirmaram ser contra as políticas de responsabilização; 3) são contra a divulgação das avaliações externas que promovam o ranqueamento; 4) os professores e os gestores acreditam que as avaliações externas contribuem para a organização pedagógica da escola.

Em relação às influências das avaliações externas no currículo e no cotidiano escolar, alguns dos resultados assinalaram que: 1) os gestores e os professores estavam de acordo que a Secretaria de Educação fornecesse às escolas os materiais didáticos baseados nas diretrizes curriculares do SAEB/Prova Brasil; 2) os profissionais concordavam que as avaliações interferem no currículo, como, por exemplo, “enxugando” os conteúdos oferecidos aos alunos; 3) alguns professores acreditam que o “enxugamento” dos conteúdos poderia interferir na qualidade da formação dos alunos; 4) existia uma concordância de alguns profissionais de que as avaliações externas auxiliam o trabalho do professor.

Assim, tendo em vista os estudos apresentados, as informações geradas pelas avaliações externas podem ser utilizadas como instrumento de reflexão pelos diretores, vice-diretores, coordenadores pedagógicos e professores na tomada de decisões necessárias para avançar em direção à qualidade da educação e à tomada de consciência da sua própria realidade. Isso também mostra que “um dos argumentos favoráveis às avaliações externas considera a relevância do conhecimento das redes e do seu nível” (Amaro, 2013, p. 37).

Em síntese, se, de um lado, as avaliações externas podem contribuir com o trabalho pedagógico escolar, por outro, podem gerar alguns impactos negativos. Isso pode exigir da escola cada vez mais o desenvolvimento de um trabalho integrado entre os envolvidos no processo de ensino e aprendizagem a fim de superar os desafios colocados pelas avaliações externas. Também identificamos, na literatura, muitos usos dos resultados das avaliações externas pelos profissionais da educação na escola, como, por exemplo, o planejamento das atividades pedagógicas (Silva *et al.*, 2012; Souza *et al.*, 2015) ou até mesmo a dificuldade de apropriação dos resultados (Martins, 2010).

Vale ainda destacar que os resultados das avaliações externas precisam ser articulados com os resultados das avaliações internas da escola para o surgimento de possibilidades de usos que estabeleçam as ações pedagógicas necessárias. Essa articulação também mostra que existem diversas formas para verificar a qualidade de uma escola e não apenas uma. Nevo (1998, p. 94) ressalta:

Ao combinarmos as avaliações externa e interna para avaliar a escola como um todo, com o objetivo de responsabilização ou de aprimoramento escolar, supomos que não exista uma única forma objetiva de avaliar a qualidade da escola. A chamada verdade sobre a qualidade da escola não é prerrogativa de um avaliador externo, nem da própria escola, que atesta pela sua qualidade. Mesmo que exista uma verdade única sobre a escola, estará em algum nível intermediário, e, portanto, somente poderá ser revelada através de um diálogo entre as duas partes, as duas recolhendo dados e interpretando-os.

Por fim, as avaliações externas podem direcionar o trabalho escolar e colocar o desafio de estabelecer ações a partir dos resultados gerados. Assim, é preciso compreender essas avaliações não como um fim em si mesmas, mas como uma possibilidade de realizar transformações na educação e de articulação com as avaliações criadas pela própria escola.

Ideb: percepções sobre usos e implicações

Um estudo realizado por Chirinéa (2016) buscou compreender como o sistema municipal de Bauru-SP utiliza os resultados do Ideb para a realização de ações no interior das escolas. A pesquisa utilizou uma abordagem qualitativa. O percurso metodológico utilizado foi a análise documental e a entrevista semiestruturada. A entrevista foi realizada com 15 gestores de escolas (nível fundamental) e na Secretaria de Educação.

Os resultados indicaram os seguintes usos do Ideb: 1) 11 diretores entrevistados utilizam o Ideb para o planejamento escolar; 2) 14 utilizam o índice para orientar o próprio trabalho enquanto gestor; 3) 15 para orientar o trabalho do professor em sala de aula; 4) 12 para melhorar o desempenho da escola mediante aumento da nota do índice; 5) oito gestores articulam ações a partir do Ideb visando ao reconhecimento da escola em termos de qualidade; 6) 12 deles utilizam os dados para aprimorar políticas e ações em prol da qualidade dos processos formativos; 7) 10 gestores informaram que utilizam os dados para melhorar a aprendizagem dos alunos (Chirinéa, 2016). Assim, o Ideb traz como consequência algumas implicações para o cotidiano da escola, o que afeta os sujeitos.

Outros estudos também demonstram isso, como, por exemplo, o de Saçço (2015), que investiga a influência do Ideb sobre a satisfação profissional. O principal objetivo do estudo foi verificar se os resultados do Ideb influenciam a satisfação profissional dos docentes do 1º ao 5º ano em seis escolas estaduais de Juiz de Fora-MG. A pesquisa adotou uma abordagem qualitativa, e o instrumento usado para a geração de dados foi o questionário. Sobre o resultado, o Ideb afeta a satisfação profissional dos sujeitos investigados por meio de fatores como: cobranças, falta de autonomia no trabalho, pressão para alcançar índices, mudança de currículos e forma de trabalhar o conteúdo. Assim, o Ideb seria percebido pelos sujeitos dessa pesquisa como um causador de insatisfação docente, uma vez que os professores eram levados a desenvolver o seu trabalho para o alcance de metas.

Alguns estudos demonstram implicações positivas do Ideb na escola e no trabalho dos profissionais que nela atuam. O estudo realizado por

Mesquita (2012) buscou investigar o impacto da divulgação do resultado do Ideb sobre uma escola do Rio de Janeiro-RJ. A metodologia adotada foi uma abordagem qualitativa de inspiração etnográfica. O pesquisador ficou seis meses em campo para coletar informações por meio da observação e de entrevistas com os gestores e professores, além de acompanhar atividades realizadas nas aulas e analisar documentos da escola. Os resultados evidenciaram que a escola buscou estratégias para melhorar os baixos resultados atingidos e obter mais envolvimento da comunidade escolar. Nesse sentido, o impacto foi positivo ao gerar mais engajamento para estabelecer ações eficazes para o ensino e aprendizado na escola.

Um estudo realizado por Oliveira (2015) buscou investigar a concepção de qualidade da educação no Ideb e seu impacto no trabalho de seis escolas da rede municipal de Francisco Beltrão-PR. Participaram do estudo 10 professoras, oito gestoras e seis coordenadoras. Para a coleta de dados, foram utilizadas entrevistas semiestruturadas. Entre os resultados, destacamos que as participantes do estudo demonstraram incômodo e resistência ao Ideb na escola por considerarem uma avaliação que homogeneiza todo o sistema, que desconsidera as subjetividades dos alunos, as deficiências e as disparidades socioeconômicas.

Boas (2016) realizou um estudo com o intuito de analisar as interpretações dos professores sobre os resultados recebidos das avaliações externas, focando a sua análise no Ideb. Assim, procurou compreender os limites e as potencialidades desse índice no trabalho desenvolvido na escola. Participaram do estudo os professores dos Anos Iniciais e Finais do Ensino Fundamental de três escolas da rede municipal da cidade de Pouso Alegre-MG. O percurso metodológico utilizado foi a técnica de grupo focal, e a análise de dados coletados foi norteadada pela análise de conteúdo.

Os resultados apontaram para algumas implicações do Ideb na escola: 1) nas três escolas participantes do estudo, o planejamento pedagógico era ajustado a partir dos resultados do Ideb; 2) existia um esforço da equipe para superar as metas estipuladas pelo governo; 3) nos professores de duas escolas, prevalecia um sentimento de alegria por seus resultados, mas também um cansaço por terem que sempre superar metas; 4) para

a outra escola, muita frustração e sentimento de exclusão perante a rede por não conseguirem alcançar um bom resultado.

Uma pesquisa desenvolvida por Silva (2012) buscou identificar, por meio da percepção de gestores e docentes, as mudanças na organização, na gestão escolar e no trabalho docente a partir da divulgação do Ideb, da adesão de municípios do estado da Paraíba ao Plano de Metas Compromisso Todos pela Educação e da elaboração do plano de ações articuladas (PAR). O percurso metodológico foi realizado por meio de revisão da literatura, pesquisa documental e de questionários. Participaram do estudo os professores e os gestores de 40 escolas municipais de oito municípios da Paraíba.

Foram aplicados os questionários em duas escolas de cada município que ofertavam os Anos Iniciais do Ensino Fundamental e que alcançaram o menor e o maior Ideb em 2007. Os dados demonstraram o impacto do Ideb: a divulgação dos seus resultados tem interferido na organização escolar e no trabalho docente, como, por exemplo, ao contribuir para a melhoria da infraestrutura material da escola e para a realização de atividades diferenciadas para o aprendizado dos alunos. Diversas ações desenvolvidas nas instituições participantes do estudo têm voltado seu trabalho para a elevação desse índice.

Um estudo realizado por Assis (2014) buscou compreender as possíveis interferências produzidas pelo Ideb na realidade das escolas da rede municipal de ensino de Jataí-GO. Participaram do estudo 27 escolas de Ensino Fundamental dessa rede. Os instrumentos para a coleta de dados foram o questionário e a entrevista semiestruturada. A entrevista foi realizada com três diretores das escolas selecionadas e com a assessora pedagógica das séries iniciais da Secretaria Municipal de Educação de Jataí-GO.

Responderam ao questionário seis coordenadores e 44 professores de três escolas, que atuavam no Ensino Fundamental, além de diretores das escolas municipais urbanas e rurais. Os resultados indicaram que o Ideb implicou em algumas mudanças e ações na rede municipal de ensino de Jataí-GO, com o intuito de alcançar as metas previstas no Ideb. A organização e a gestão também sofreram influências desse índice, pois foram

estabelecidas outras formas de avaliação no contexto escolar, projetos, ações e atividades pedagógicas com a finalidade de melhorar o desempenho dos estudantes.

Em síntese, a maioria desses estudos evidenciou que, nas escolas analisadas, os profissionais da educação têm a percepção de que os resultados do Ideb vêm sendo utilizados de diferentes maneiras. As ações mais recorrentes, identificadas em mais de um estudo, incluem: 1) melhoria da aprendizagem dos alunos (Farias, 2015; Chirinéa, 2016); 2) melhoria do desempenho dos alunos (Audino, 2014; Chirinéa, 2016). Por outro lado, determinados estudos também assinalaram que alguns gestores, por exemplo, não utilizam o resultado do Ideb (Laurentino, 2013; Pagan, 2016).

Esse índice também tem trazido algumas implicações negativas para o trabalho escolar ao criar alguns sentimentos na equipe profissional, como: 1) pressão para alcançar índices e insatisfação docente (Saçço, 2015); 2) sentimento de frustração e de exclusão perante a rede (Boas, 2016); 3) sentimento de culpa, de frustração e responsabilização pelos resultados alcançados (Costa, 2015). É importante considerar que as pesquisas também evidenciaram diversas implicações positivas, como: 1) geração de mais engajamento a favor do estabelecimento de ações eficazes para o ensino e aprendizado da escola (Mesquita, 2012); 2) organização do trabalho pedagógico (Farias, 2015); 3) utilização do índice para melhorar a aprendizagem dos alunos (Chirinéa, 2016); 4) repensar a prática pedagógica (Audino, 2014).

Conclusão

Na revisão da literatura, foram identificadas percepções predominantemente favoráveis em relação às avaliações externas e ao Índice de Desenvolvimento da Educação Básica (Ideb), destacando-se: a necessidade de repensar a prática pedagógica; a adoção de metodologias diferenciadas para o ensino; o planejamento de atividades voltadas à melhoria do desempenho dos estudantes; a análise e a discussão dos resultados obtidos; a utilização desses resultados como referência para a condução do trabalho pedagógico; a formulação de estratégias para superar baixos

índices de aprendizagem; a busca por maior envolvimento da comunidade escolar; o planejamento de atividades com base nas matrizes de referência das avaliações externas; a elaboração de Planos de Ação Escolar; e o replanejamento das práticas pedagógicas.

Por outro lado, nas pesquisas analisadas, também emergiram percepções críticas quanto aos impactos das avaliações externas e do Ideb, entre as quais se destacam: a responsabilização exclusiva das escolas pelos resultados obtidos; a exposição das unidades escolares a comparações por meio de *rankings*; e a geração de sentimentos negativos entre os profissionais da educação, como descontentamento, sobrecarga de trabalho e angústia.

Portanto, as instituições educacionais vêm superando uma postura passiva diante dos resultados das avaliações e avançando na adoção de medidas concretas com base nas informações produzidas. Esse movimento evidencia um compromisso coletivo com a melhoria da qualidade do ensino, colocando a escola diante do desafio de mobilizar esforços em conjunto. Nesse processo, diretores, vice-diretores e coordenadores pedagógicos assumem um papel central na viabilização das ações, articulando-se com a comunidade escolar para promover intervenções fundamentadas a partir dos dados das avaliações externas e do Ideb.

Tomados em conjunto, os resultados da nossa pesquisa evidenciam um forte efeito de indução do Ideb e das avaliações externas sobre o cotidiano das instituições escolares, influenciando diretamente suas práticas organizacionais e pedagógicas. Nesse sentido, o Ideb e as avaliações externas têm modificado o cotidiano de trabalho das escolas e exigido novas ações dos profissionais. Sobre isso, Oliveira e Sousa (2010) discutem o efeito de indução das políticas utilizadas pelo governo federal. Esse efeito está diretamente ligado às modificações geradas pelas ações externas no trabalho dos profissionais de educação nas escolas. Portanto, a nossa revisão da literatura confirma esse poder de indução que o Ideb e as avaliações externas possuem no cotidiano de trabalho das instituições escolares, uma vez que diversas ações são implementadas a partir dos seus resultados.

Vale ainda destacar que a qualidade educacional de uma escola deve ser compreendida a partir de uma análise contextualizada, que considere múltiplos indicadores e resultados, dentre os quais se destaca o nível socioeconômico dos estudantes. Nesse sentido, é fundamental evitar a redução da qualidade de uma instituição escolar a um único indicador, como o Ideb. Tal perspectiva pode favorecer uma leitura mais sensível à singularidade de cada escola, permitindo a identificação mais precisa de seus avanços ou eventuais retrocessos no que diz respeito à aprendizagem dos alunos ao longo de um determinado período.

Por fim, a revisão e a análise das pesquisas contribuíram para compreender como os profissionais da educação têm percebido as avaliações externas e o Ideb no cotidiano de trabalho, mais especificamente sobre os seus usos e implicações. Isso indica a relevância do presente estudo, uma vez que as avaliações externas, assim como o Ideb, são instrumentos fundamentais no monitoramento da qualidade educacional no Brasil.

Referências

AFONSO, Almerindo Janela. Para uma conceitualização alternativa de *accountability* em educação. **Educação e Sociedade**, Campinas, v. 33, n. 119, p. 471-484, jun. 2012. Disponível em: https://www.redalyc.org/pdf/873/87323122007.pdf?utm_source=chatgpt.com. Acesso em: 26 nov. 2025.

ALAVARSE, Ocimar M.; BRAVO, Maria Helena.; MACHADO, Cristiane. Avaliação como diretriz das políticas educacionais dos governos federal, estaduais e municipais: o caso brasileiro. *In*: CONGRESSO IBERO AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 3., 2012, Zaragoza. **Anais [...]**. Zaragoza: ANPAE, 2012.

ALMEIDA, Luana Costa; DALBEN, Adilson; FREITAS, Luiz Carlos de. O Ideb: limites e ilusões de uma política educacional. **Educ. Soc.**, Campinas, v. 34, n. 125, p. 1153-1174, dez. 2013. Disponível em: <https://www.scielo.br/pdf/es/v34n125/O8.pdf>. Acesso em: 26 nov. 2025.

ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. **Educação e Pesquisa**, [S. l.], v. 39, n. 1, p. 177-194, mar. 2013. Disponível em: <https://www.scielo.br/j/ep/a/PkVXrTbnCJDktQxLZNK7dD-j/?format=html&lang=pt>. Acesso em: 26 nov. 2025.

AMARO, Ivan. Avaliação externa da escola: repercussões, tensões e possibilidades. **Estudos em Avaliação Educacional**, São Paulo, v. 24, n. 54, p. 32–55, jan./abr. 2013. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/1901>. Acesso em: 26 nov. 2025.

ASSIS, Renata Machado de. **O Ideb nas escolas da rede municipal de ensino de Jataí-GO**: qualidade, avaliação e interferências. 2014. 302 f. Tese (Doutorado em Educação) – Universidade Federal de Goiás, Goiânia, 2014.

AUDINO, Janaina Franciscatto. **Movimentos de apropriação do Ideb na gestão escolar em duas escolas da rede pública estadual em Porto Alegre-RS**. 2014. 113 f. Dissertação (Mestrado em Gestão Educacional) – Universidade do Vale do Rio dos Sinos, São Leopoldo, 2014. Disponível em: <https://www.repositorio.jesuita.org.br/handle/UNISINOS/3743>. Acesso em: 26 nov. 2025.

BOAS, Ivanilda Vilela Vilas. **A percepção dos professores sobre o Ideb de sua escola**: ele reflete o trabalho desenvolvido? 2016. 96 f. Dissertação (Mestrado em Educação) – Universidade do Vale do Sapucaí, Pouso Alegre, 2016.

BORGES, Edna Martins. **Avaliações externas em larga escala no contexto escolar**: percepção dos diretores escolares da rede estadual de ensino de Minas Gerais. 2016. 327 f. Tese (Doutorado em Educação) – Programa de Pós-graduação em Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996**. Estabelece as diretrizes e bases da educação nacional. Diário Oficial da União: Brasília, DF, 23 dez. 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/19394.htm. Acesso em: 26 nov. 2025.

BRASIL. **Decreto nº 6.094, de 24 de abril de 2007**. Estabelece as diretrizes e metas para a melhoria da educação básica. Diário Oficial da União: seção 1, Brasília, DF, 25 abr. 2007. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6094.htm. Acesso em: 26 nov. 2025.

BRASIL. **Provinha Brasil**: guia de aplicação – leitura. Brasília, DF: MEC; INEP, 2011. Disponível em: https://download.inep.gov.br/educacao_basica/provinha_brasil/kit/2011/1_semestre/guia_aplicacao_leitura_1_2011.pdf. Acesso em: 26 nov. 2025.

BRASIL. **Portaria nº 447, de 24 de maio de 2017**. Dispõe sobre o Sistema de Avaliação da Educação Básica – SAEB. Brasília, DF: Diário Oficial da União, 2017. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/KujrwOTZC2Mb/content/id/19142540. Acesso em: 26 nov. 2025.

BROOKE, Nigel; CUNHA, Maria Amália de A. A avaliação externa como instrumento da gestão educacional nos Estados. **Estudos e Pesquisas Educacionais**, São Paulo, v. 2, p. 17–79, 2011. Disponível em: https://fvc.org.br/wp-content/uploads/2018/04/estudos_e_pesquisas_educacionais_vol_2.pdf?utm_source=chatgpt.com. Acesso em: 26 nov. 2025.

CAMARGO, Rubens Barbosa de; PINTO, José Marcelino de Rezende; GUIMARAES, José Luiz. Sobre o financiamento no Plano de Desenvolvimento da Educação. **Cadernos de Pesquisa**, São Paulo, v. 38, n. 135, p. 817–839, dez. 2008. Disponível em: <https://publicacoesfcc.emnuvens.com.br/cp/article/view/335>. Acesso em: 26 nov. 2025.

CASTRO, Maria Helena Guimarães de. Sistemas nacionais de avaliação e de informações educacionais. **São Paulo Perspectiva**, São Paulo, v. 14, n. 1, p. 121–128, mar. 2000. Disponível em: <https://www.scielo.br/j/spp/a/Lfc37RDdj4czGv-94gpbz4Xg/?format=pdf&lang=pt>. Acesso em: 26 nov. 2025.

CERDEIRA, Diana Gomes da Silva; ALMEIDA, Andrea Baptista de; COSTA, Marcio da. Indicadores e avaliação educacional: percepções e reações a políticas de responsabilização. **Estudos em Avaliação Educacional**, São Paulo, v. 25, n. 57, p. 198–225, jan./abr. 2014. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2845>. Acesso em: 26 nov. 2025.

CHIRINÉA, Andréia Melanda; BRANDÃO, Carlos da Fonseca. O Ideb como política de regulação do Estado e legitimação da qualidade: em busca de significados. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 23, n. 87, p. 461–484, jun. 2015. Disponível em: <https://www.scielo.br/j/ensaio/a/L3FwTBZ79fLPRRwHfVgmkP/?format=pdf&lang=pt>. Acesso em: 26 nov. 2025.

CHIRINÉA, Andréia Melanda. **Gestão da escola pública municipal e utilização do Ideb**: as “traduções” no contexto da prática. 2016. 181 f. Tese (Doutorado em Educação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista (UNESP), Marília, 2016.

COSTA, Loany Larissa Ferreira da. **Avaliação da educação básica**: o que pensam professores e gestores de uma escola pública amazônica sobre o Ideb. 2015. 121 f. Dissertação (Mestrado em Psicologia) – Departamento de Psicologia, Universidade Federal de Rondônia, Porto Velho, 2015. Disponível em: <https://ri.unir.br/jspui/handle/123456789/1366/>. Acesso em: 26 nov. 2025.

DURHAM, Eunice Ribeiro. A política educacional do governo Fernando Henrique Cardoso: uma visão comparada. **Novos Estudos** – CEBRAP, São Paulo, n. 88, p. 153–179, dez. 2010. Disponível em: <https://novosestudos.com.br/produto/edicao-88/>. Acesso em: 26 nov. 2025.

FARIAS, Maria Adalgiza de. **Avaliação externa e gestão da escola**: apropriações e usos dos dados do Ideb na gestão de escolas públicas municipais de fortaleza. 2015. 227 f. Dissertação (Mestrado em Educação) – Universidade Estadual do Ceará, Fortaleza, 2015.

FERNANDES, Reynaldo. **Índice de Desenvolvimento da Educação Básica (Ideb)**. Brasília: INEP, 2007.

FERREIRA, Rosilda Arruda; VIEIRA, Marcos. Contribuições do Ideb para a avaliação da educação no município de Teodoro Sampaio. *In*: TENÓRIO, Robinson; MACHADO, Cristiane Brito; LOPES, Uaçai Magalhães (orgs.). **Indicadores da Educação Básica**: avaliação para uma gestão sustentável. Salvador: EDUFBA, 2010, p. 243–262.

FRANCO, Creso; ALVES, Fátima; BONAMINO, Alícia. Qualidade do ensino fundamental: políticas, suas possibilidades, seus limites. **Educação & Sociedade**, Campinas, v. 28, n. 100, p. 989–1014, out. 2007. Disponível em: <https://www.scielo.br/j/es/a/ZqbnNkCG3YtpSs9VGhYzKxn>. Acesso em: 26 nov. 2025.

FREITAS, Dirce Nei Teixeira de; OVANDO, Nataly Gomes. A avaliação educacional em contextos municipais. **Educação e Sociedade**, Campinas, v. 36, n. 133, p. 963–984, out./dez. 2015. Disponível em: <https://www.scielo.br/j/es/a/dLBgv-vJncbFtTHVCVgSfqkh/?lang=pt&format=pdf>. Acesso em: 26 nov. 2025.

FREITAS, Luiz Carlos de. Eliminação adiada: o ocaso das classes populares no interior da escola e a ocultação da (má) qualidade do ensino. **Educação & Sociedade**, Campinas, v. 28, n. 100, p. 965–987, out. 2007. Disponível em: <http://www.redalyc.org/articulo.oa?id=87313704016>. Acesso em: 26 nov. 2025.

FREITAS, Pâmela Félix. **Usos das avaliações externas**: concepções de equipes gestoras de escolas da rede municipal de ensino de São Paulo. 2014. 193 f. Dissertação (Mestrado em Educação) – Universidade de São Paulo, São Paulo, 2014.

LAURENTINO, Joselma Nunes. **Utilização dos resultados do Ideb pelos gestores escolares da rede municipal de ensino de Cajazeiras-PB**. 2013. 128 f. Dissertação (Mestrado em Educação) – Universidade Federal da Paraíba, João Pessoa, 2013.

MACÊDO, Seandra Doroteu de. **Gestão pedagógica em tempos de Ideb**. 2010. 189 f. Dissertação (Mestrado em Educação) – Universidade Estadual do Ceará, Fortaleza, 2010.

MARTINS, S. M. **O papel do coordenador pedagógico na melhoria da educação básica**: uma reflexão sobre o uso da avaliação externa. 2010. 93 f. Dissertação (Mestrado em Ciências Humanas) – Unoeste, Presidente Prudente, 2010.

MESQUITA, Silvana. Os resultados do Ideb no cotidiano escolar. **Ensaio**: Avaliação e Políticas Públicas em Educação, Rio de Janeiro, v. 20, n. 76, p. 587–606, set. 2012. Disponível em: <https://www.redalyc.org/articulo.oa?id=399538142009>. Acesso em: 26 nov. 2025.

MINAS GERAIS. **Programa de avaliação da alfabetização**: PROALFA, Boletim pedagógico. Belo Horizonte: Secretaria de Educação de Minas Gerais, 2009.

MOTTA, Alfredo Carnevalli. **Uso de resultados do SIMAVE/PROEB pela Superintendência Regional de Ensino de Caxambu**. 2014. 130 f. Dissertação (Mestrado Profissional em Gestão e Avaliação da Educação Pública) – Universidade Federal de Juiz de Fora, Juiz de Fora, 2014.

NEVO, D. Avaliação por diálogos: uma contribuição possível para o aprimoramento escolar. *In*: SEMINÁRIO INTERNACIONAL DE AVALIAÇÃO EDUCACIONAL, 1997, Rio de Janeiro. **Anais [...]**. Brasília, DF: INEP, 1998.

OLIVEIRA, Quelli Cristina da Silva. **O Ideb e a qualidade da educação**: a política do Ideb nas escolas da rede municipal de Francisco Beltrão-PR, no período de 2007–2013. 2015. 195 f. Dissertação (Mestrado em Educação) – Universidade Estadual do Oeste do Paraná, Francisco Beltrão, 2015.

OLIVEIRA, Romualdo Portela; SOUSA, Sandra Sákia. Introdução. *In*: SANTANA, Wagner (org.). **Educação e federalismo no Brasil**: combater as desigualdades, garantir a diversidade. Brasília: Unesco, 2010, p. 7–297.

PAGAN, Ariane Rodrigues. **Como os gestores escolares utilizam os resultados do Ideb**: um estudo em escolas municipais de Campo Grande – MS. 2016. Dissertação (Mestrado em Educação) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2016.

PIROLA, Nelson Antônio; MORAES, Mara Sueli Simões; SANDER, Giovana Pereira. Provinha Brasil de Matemática: Expectativas dos professores e utilização dos seus resultados. *In*: REUNIÃO DA ABAVE, 6., 2011, Fortaleza. **Anais [...]**. Fortaleza: Associação Brasileira de Avaliação Educacional, 2011.

POLATO, Amanda. **A relação entre as avaliações em larga escala e a organização do trabalho escolar em duas escolas públicas estaduais do interior de São Paulo**. 2014. 172 f. Dissertação (Mestrado em Educação) – Instituto de Biotecnologia, Universidade Estadual Paulista (UNESP), Rio Claro, 2014.

PRADO, Maria Lúcia Viana do. **Análise dos cursos de capacitação para divulgação e apropriação de resultados do SIMAVE/PROALFA e PROEB da Superintendência Regional de Teófilo Otoni-MG**. 2014. 135 f. Dissertação (Mestrado Profissional em Gestão e Avaliação da Educação Pública) – Universidade Federal de Juiz de Fora, Juiz de Fora, 2012.

SAÇÇO, Thays Alessandra Silva. **Influências do Ideb sobre a satisfação profissional na visão de professores do 1º ao 5º anos do Ensino Fundamental, em seis escolas estaduais em Juiz de Fora-MG**. 2015. 220 f. Tese (Doutorado em Educação) – Universidade Federal de Juiz de Fora, Juiz de Fora, 2015.

SILVA, Isabelle Fiorelli. O sistema nacional de avaliação: características, dispositivos legais e resultados. **Estudos em Avaliação Educacional**, São Paulo, v. 21, n. 47, p. 427-448, set./dez. 2010. Disponível em: <https://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1602/1602.pdf>. Acesso em: 26 nov. 2025.

SILVA, Vandrê Gomes da *et al.* **Uso da avaliação externa por equipes gestoras e profissionais docentes: um estudo em quatro redes de ensino público**. São Paulo: Fundação Carlos Chagas, 2012.

SOARES, José Francisco; XAVIER, Flávia Pereira. Pressupostos educacionais e estatísticos do Ideb. **Educação e Sociedade**, [S. l.], v. 34, n. 124, p. 903-924, 2013. Disponível em: <https://www.redalyc.org/pdf/873/87328534013.pdf>. Acesso em: 26 nov. 2025.

SOUSA, Sandra Zákia *et al.* Uso de dados de avaliações externas por redes municipais de educação paulistas. **Revista Diálogo Educacional**, [S. l.], v. 15, n. 44, p. 37-60, jan. 2015.

VIDAL, Eloisa Maia; VIEIRA, Sofia Lerche. Gestão educacional e resultados no Ideb: um estudo de caso em dez municípios cearenses. **Estudos em Avaliação Educacional**, São Paulo, v. 22, n. 50, p. 419-434, set./dez. 2011. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/1963>. Acesso em: 26 nov. 2025.

CAPÍTULO 9

LACUNAS NA APRENDIZAGEM E DESIGUALDADES EDUCACIONAIS NA PANDEMIA: EVIDÊNCIAS DA REDE ESTADUAL DE MINAS GERAIS

Joyce Soares Rodrigues Petrus¹
Flavia Pereira Xavier²
DOI: 10.52695/978-65-5456-163-1.9

Introdução

A garantia do direito à educação não se limita ao acesso à escola, mas envolve também a permanência e a aprendizagem em condições de igualdade, conforme estabelecido pelos marcos legais que estruturam o sistema educacional brasileiro. A Constituição Federal de 1988, em seus artigos 205, 206 e 210, afirma que a educação deve ser oferecida com qualidade e igualdade de condições, assegurando o acesso, a permanência e o desenvolvimento de uma formação básica comum, com respeito aos valores culturais e artísticos nacionais e regionais (Brasil, 1988).

1 Universidade Federal de Minas Gerais. E-mail: joyceufmg@gmail.com.

2 Universidade Federal de Minas Gerais.

Esses princípios são reafirmados na Lei de Diretrizes e Bases da Educação Nacional (Brasil, 1996). Os Planos Nacionais de Educação de 2001 e 2014 (Brasil, 2001, 2014) detalham essas diretrizes ao proporem metas voltadas à universalização do acesso, à redução da distorção idade-série e à elevação dos níveis de aprendizagem. Mais recentemente, o novo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb) incorporou, de maneira explícita, o princípio da equidade ao atrelar o financiamento à aprendizagem dos estudantes, considerando desigualdades socioeconômicas, raciais e de condição de deficiência (Brasil, 2020a).

O monitoramento desse direito mobiliza diversas fontes de dados. O acesso à escola é acompanhado por pesquisas domiciliares realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), como a Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) e o Censo Demográfico. A permanência e as trajetórias escolares são analisadas com base no Censo Escolar, conduzido anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) (Soares; Alves; Fonseca, 2021; Fonseca *et al.*, 2024). Já a aprendizagem é aferida por meio das avaliações em larga escala do Sistema de Avaliação da Educação Básica (Saeb), também organizadas pelo INEP, além das avaliações externas desenvolvidas por diversas redes estaduais e municipais (Soares; Xavier, 2013; Brooke *et al.*, 2011).

Essas avaliações externas têm sido objeto de críticas por parte da literatura especializada. Entre os principais limites apontados, estão o foco restrito às disciplinas de língua portuguesa e matemática (Bonamino; Sousa, 2012; Bauer; Alavarse; Oliveira, 2015), o predomínio de itens de múltipla escolha, que reduzem a complexidade dos processos de aprendizagem avaliados (Alavarse, 2013; Bonamino, 2016), e o uso dos resultados em políticas de responsabilização que desconsideram as condições concretas de ensino (Freitas *et al.*, 2014; Gatti, 2014; Libâneo, 2018).

Ainda assim, em um país de dimensões continentais e marcantes desigualdades, essas avaliações seguem sendo instrumentos valiosos para revelar padrões estruturais de exclusão e orientar políticas públicas (Alavarse; Bravo; Machado, 2013; Machado; Alavarse, 2014; Moriconi; Bauer, 2017). Como

destacam Alves, Soares e Xavier (2016), mesmo com a elevação das médias de desempenho ao longo da série histórica do Saeb, as desigualdades entre grupos sociais não diminuíram, evidenciando o papel central dessas ferramentas no enfrentamento das iniquidades educacionais.

Nesse contexto, a pandemia de Covid-19 representou um ponto de inflexão, aprofundando os desafios já existentes para a efetivação do direito à educação. Declarada em março de 2020 pela Organização Mundial da Saúde (OMS), a crise sanitária resultou no fechamento prolongado das escolas no Brasil, por 178 dias apenas no primeiro ano, número expressivamente superior à média de 48 dias observada nos países da OCDE (OECD, 2021). A transição para o ensino remoto ocorreu de forma abrupta e profundamente desigual entre redes, escolas e estudantes brasileiros, revelando deficiências no acesso a equipamentos, conectividade e apoio pedagógico familiar (Koslinski; Bartholo, 2021; Barros *et al.*, 2021³). Estudantes dos anos iniciais da Educação Básica, em especial, enfrentaram barreiras maiores à aprendizagem, dada sua menor autonomia e maior dependência de mediação presencial, sobretudo aqueles em contextos de maior vulnerabilidade social.

A combinação entre formatos emergenciais de ensino, limitações da infraestrutura domiciliar e desigualdades preexistentes resultou em realidades educacionais heterogêneas, tanto em termos de oferta quanto de aproveitamento escolar. Estudos apontam que grande parte dos estudantes não aprendeu os conteúdos esperados para sua etapa de escolarização (São Paulo, 2021), com efeitos mais acentuados entre os mais novos e em componentes curriculares como matemática (Kuhfeld *et al.*, 2020; Koslinski; Bartholo, 2021; Lichand *et al.*, 2022). A crise educacional gerada pela pandemia, portanto, não apenas comprometeu a aprendizagem, mas também aprofundou as desigualdades já existentes no sistema de ensino.

3 Estudo realizado por pesquisadores do Instituto Unibanco (Insper), cujo relatório sintético denominado “Perda de aprendizagem na pandemia” foi publicado em 2021 e encontra-se disponível em: <https://repositorio.insper.edu.br/entities/publication/d0c6156b-749e-4241-9701-70423fe13d7b>. Acesso em: 26 nov. 2025.

Este capítulo toma como caso empírico a rede estadual de Minas Gerais para estimar os efeitos da pandemia sobre a aprendizagem dos estudantes, com ênfase nos padrões de desigualdade e nas lacunas de desempenho associadas ao período. A análise baseia-se nos resultados do Programa de Avaliação da Rede Pública de Educação Básica (Proeb), parte integrante do Sistema Mineiro de Avaliação e Equidade da Educação Pública (Simave), aplicado anualmente desde 2000.

Apesar de o retorno às aulas presenciais ter ocorrido há quatro anos, os efeitos da pandemia ainda são perceptíveis nas escolas e seguem evidenciados pelas avaliações externas. Compreender a magnitude desses impactos, especialmente nas redes públicas, é essencial para subsidiar políticas de recomposição da aprendizagem, garantir a efetividade do direito à educação e aprofundar a reflexão sobre como situações de crise podem acirrar desigualdades estruturais.

Ao longo do capítulo, discutimos a distinção entre perda e atraso na aprendizagem no contexto da suspensão das aulas presenciais, apresentamos um panorama da oferta de ensino na rede estadual de Minas Gerais durante o período, detalhamos a metodologia adotada na análise e discutimos os principais achados do estudo. Por fim, refletimos sobre as implicações desses resultados, seus limites e possibilidades que oferecem para futuras pesquisas e políticas educacionais.

Entre o imprevisto e a desigualdade: a oferta de ensino durante a pandemia

Em dezembro de 2017, a Base Nacional Comum Curricular (BNCC) foi homologada, estabelecendo as aprendizagens essenciais a serem garantidas ao longo da Educação Básica em todas as redes de ensino do país (Brasil, 2018). Sua implementação vinha sendo conduzida por estados e municípios em ritmos variados, com prazos de até dois anos para a adequação dos currículos locais. No entanto, esse processo foi abruptamente interrompido com a suspensão das atividades presenciais nas escolas em março de 2020, no início do ano letivo, como resposta à emergência sanitária causada pela pandemia de Covid-19.

Diante da continuidade da obrigação constitucional de oferta educacional, redes estaduais e municipais adotaram o ensino remoto como estratégia emergencial para manter o vínculo com os estudantes e garantir, ainda que parcialmente, o direito à educação. Essa oferta assumiu formas distintas, combinando recursos digitais (como *internet* e aplicativos), canais televisivos, rádio e, em muitos contextos, materiais impressos, especialmente quando o acesso a tecnologias era limitado ou inexistente.

Como apontam Segati e Jordão (2022), tratava-se de uma transposição improvisada das práticas pedagógicas presenciais para outros meios, realizada em caráter emergencial e sem o planejamento necessário para uma mediação sistemática da aprendizagem. Essa reorganização alterou significativamente a rotina das comunidades escolares. Diretores passaram a acompanhar o trabalho dos professores e a gerenciar afastamentos por motivos de saúde; professores tiveram que dominar rapidamente o uso de tecnologias; estudantes passaram a frequentar a escola de forma remota, muitas vezes sem as condições adequadas; e famílias se viram diante da tarefa de compartilhar o espaço e o tempo de trabalho com a mediação das atividades escolares. De forma geral, o período implicou em uma sobrecarga de trabalho e na necessidade de adaptação por parte de todos os envolvidos.

Mesmo como estratégia viável diante das circunstâncias, o ensino remoto apresentou limitações significativas. Nem todos os estudantes foram alcançados, e, entre os que foram, as formas de acesso e os recursos disponíveis variaram de modo a comprometer a equidade na aprendizagem. Enquanto alguns contavam com *internet*, equipamentos e apoio docente, outros dependiam de materiais impressos e tinham menos interação com os professores. A falta de coordenação de políticas nacionais estruturadas gerou muitas disparidades na implementação das estratégias, como mostram Alves *et al.* (2022) e Senkevics e Bof (2022).

Ademais, a modalidade exigia um grau de autonomia nos estudos nem sempre compatível com as possibilidades das crianças, sobretudo nos anos iniciais da escolarização. A situação foi ainda mais crítica para estudantes com deficiência, especialmente para aqueles que necessitavam

de profissionais de apoio. Em muitos casos, o acesso aos materiais adaptados dependia da atuação de outros profissionais e da obtenção prévia dos conteúdos que seriam trabalhados, o que resultava em atrasos adicionais e, por vezes, na ausência de recursos adequados.

Além dos desafios pedagógicos e operacionais, as redes de ensino enfrentaram a complexa tarefa da regulamentação. Era necessário definir regras sobre o calendário letivo, as formas de contabilização da carga horária, os critérios de avaliação e promoção dos estudantes, entre outros aspectos. Sem uma coordenação nacional imediata, algumas redes iniciaram suas atividades de forma autônoma. A Lei nº 14.040/2020 (Brasil, 2020b), aprovada apenas no segundo semestre daquele ano, forneceu o respaldo legal necessário à oferta emergencial de atividades não presenciais na educação básica.

O Conselho Nacional de Educação (CNE) também desempenhou um papel relevante ao emitir pareceres que flexibilizaram exigências legais, como o número mínimo de dias letivos, desde que cumprida a carga horária anual, e orientaram a priorização de aprendizagens essenciais em vez da cobertura integral dos currículos. A perspectiva do “*continuum* curricular 2020–2021” foi central nesse processo: as aprendizagens não consolidadas em 2020 deveriam ser incorporadas ao planejamento do ano seguinte (Brasil, 2020d, 2020e, 2020f, 2020g).

Em Minas Gerais, a suspensão das aulas presenciais na rede estadual ocorreu em 22 de março de 2020, conforme Deliberação do Comitê Extraordinário COVID-19 (Minas Gerais, 2020). Em abril do mesmo ano, foi instituído o Regime Especial de Atividades Não Presenciais (REANP), regulamentado pela Resolução SEE nº 4.310/2020. A proposta combinava diferentes recursos: aulas veiculadas pela televisão, vídeos no *YouTube*, aplicativo Conexão Escola e Planos de Estudos Tutorados (PETs), que reuniam atividades por bimestre e por componente curricular, distribuídos no formato digital ou impresso.

Os PETs, que, no primeiro ano do REANP, foram distribuídos mensalmente, tiveram entregas bimestrais em 2021 — segundo ano de implementação. A gestão escolar ficou responsável por imprimir e distribuir os

PETs para os estudantes sem acesso digital, bem como recolher os materiais para correção. A elaboração dos PETs foi orientada por Planos de Curso construídos pela Secretaria de Estado de Educação, com base nas habilidades priorizadas.

Esses planos também serviram de referência para as avaliações aplicadas ao longo do ano letivo de 2021, incluindo uma avaliação diagnóstica inicial e duas avaliações trimestrais, de caráter formativo, além de uma avaliação final. As avaliações foram aplicadas a todos os estudantes regularmente matriculados do 2º ano do Ensino Fundamental ao 3º ano do Ensino Médio, incluindo estudantes da Educação de Jovens e Adultos (EJA).

Os resultados dessas avaliações subsidiaram estratégias de reforço escolar, acompanhamento do aprendizado e intervenções pedagógicas nas unidades de ensino. Também foram utilizados para definir escolas prioritárias, que passaram a receber acompanhamento mais próximo dos técnicos da Secretaria de Estado de Educação, e para iniciativas de reconhecimento da gestão escolar, como o Prêmio Escola Transformação (Minas Gerais/SEE, 2021a). O retorno às aulas presenciais em Minas Gerais teve início em junho de 2021, de forma gradativa, considerando as condições de infraestrutura das escolas, que deveriam permitir o distanciamento físico mínimo de um metro e meio entre os estudantes, bem como a situação dos municípios, já que muitos ainda estavam sob decretos que impediam a reabertura de suas unidades de ensino (Minas Gerais/SEE, 2021b).

Apesar das respostas emergenciais adotadas pelas redes de ensino, o impacto da pandemia sobre as aprendizagens, rapidamente, evidenciou a insuficiência dessas medidas improvisadas. Em 2022, o Ministério da Educação (MEC) apresentou diretrizes voltadas à recomposição das aprendizagens, como o fortalecimento das avaliações diagnósticas, a reorganização curricular e a ampliação do tempo de aprendizagem. Com a persistência dos impactos educacionais, ficou evidente a necessidade de superar o caráter pontual das ações e construir uma política mais estruturada de apoio técnico e financeiro para estados e municípios.

Nesse contexto, o Pacto Nacional pela Recomposição das Aprendizagens foi lançado em junho de 2024, pelo MEC, sob a coordenação da Secretaria

da Educação Básica (SEB), com o objetivo de assegurar padrões adequados de aprendizagem, reduzir desigualdades e promover a equidade. Apesar do lançamento em 2024, o Decreto nº 12.391/2025, que institui formalmente o Pacto, foi publicado apenas no início de 2025. Trata-se de uma iniciativa de cooperação federativa que articula ações como avaliação diagnóstica contínua, reorganização curricular com base na BNCC, formação de professores e suporte pedagógico intensivo, contando com apoio técnico e financeiro da União.

Minas Gerais aderiu ao Pacto ainda em 2024, comprometendo-se com ações de recomposição das aprendizagens, formação docente e ampliação do apoio técnico-pedagógico às escolas. Em julho de 2025, o MEC disponibilizou materiais de apoio com orientações práticas aos gestores e professores, como os *Guias de Reorganização Curricular, Avaliação Diagnóstica e Mediação Pedagógica*. A efetividade do Pacto, contudo, depende da adesão voluntária dos entes federados, do investimento consistente e da capacidade das redes locais de contextualizar essas diretrizes diante das desigualdades educacionais historicamente acumuladas.

Diante do cenário de reorganização emergencial do ensino e das desigualdades no acesso às condições de aprendizagem, é fundamental esclarecer alguns conceitos que permitem interpretar adequadamente os efeitos da pandemia, ou de outras situações que envolvem a suspensão das aulas presenciais, sobre o aprendizado dos estudantes. Trata-se de compreender: houve, de fato, uma perda ou apenas um atraso na aprendizagem?

Perda ou atraso na aprendizagem? Algumas considerações conceituais

A hipótese de que houve perda de aprendizagem pelo fato de os estudantes não estarem frequentando a escola é ancorada não só pela concepção de que a escola é essencial para o aprendizado de muitas habilidades, mas também por pesquisas internacionais que se dedicaram a estudar as férias de verão, o fechamento de escolas devido ao clima intenso ou desastres naturais e o absenteísmo dos estudantes. Em todas essas situações, observou-se perda de aprendizado (Kuhfeld *et al.*, 2020;

Koslinsk; Bartholo, 2021; Brooke; Soares, 2008; Lichand *et al.*, 2022). Esses estudos também apontam que o tamanho dessa perda depende da série cursada, da disciplina avaliada e de questões demográficas. Adicionalmente, encontraram mais perdas em matemática do que em leitura.

As pesquisas que utilizaram dados nacionais anteriores ao término da pandemia estimaram diferentes perdas de aprendizagem através de simulações, como, por exemplo, o estudo de Barros *et al.* (2021) e Petrus *et al.* (2021). Além dessas simulações, há o trabalho encomendado pelo governo do estado de São Paulo ao Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF), no início de 2021, que realizou um estudo amostral na rede estadual de ensino, avaliando cerca de 21 mil estudantes matriculados no 5º ano e no 9º ano do Ensino Fundamental e no 3º ano do Ensino Médio, nas competências de leitura e matemática (São Paulo, 2021).

Os resultados da amostra representativa de todas as regiões do estado e por estratos sociais indicaram perda entre 11 e 29 pontos em leitura e entre 14 e 47 pontos em matemática para as séries avaliadas, sendo mais pronunciadas nos anos iniciais, seguido dos anos finais e médio, respectivamente. Mais recentemente, foi publicado um estudo que olhou para a Educação Infantil, com resultados de testes com crianças de 5 e 6 anos, e reforçou uma lacuna de aprendizagem maior em matemática quando comparada com língua Portuguesa (Koslinski; Bartholo, 2021), durante a pandemia.

Estudos como os de Barros *et al.* (2021), Bartholo *et al.* (2023) e Kuhfeld *et al.* (2020) utilizam o termo “perda de aprendizagem” de forma mais genérica para se referirem tanto ao que se “desaprendeu” quanto ao que não se aprendeu em determinado período. Contudo, há uma diferença semântica entre o que se “perde” e o que “não se aprende”. A “perda” pressupõe que se tinha algo no passado que não se tem mais no presente.

Seguindo essa distinção, estudos que avaliam as férias de verão tratam de perda de aprendizagem (Quinn; Polikoff, 2017; Atteberry; Mceachin, 2020; Von Hippel *et al.*, 2018). Ou seja, estudantes deixam de praticar alguma habilidade e perdem conhecimentos no período, ainda que possam adquirir novas habilidades pelas experiências vividas no verão. Já no caso

da suspensão das aulas presenciais, não é correto tratar somente como uma “perda”, mas, sim, como um atraso na aprendizagem. Os estudantes que tiveram acesso à educação ofertada de maneira não presencial deixaram de aprender e desenvolver habilidades que eram esperadas para determinada etapa escolar. Não é o caso de terem “desaprendido”, mas, sim, de não terem sido expostos aos conteúdos de maneira satisfatória. Esse argumento é reforçado, inclusive, pela diretriz de flexibilização curricular orientada pelo CNE (Brasil, 2020c), que possibilita a priorização de habilidades a serem desenvolvidas.

Feita a distinção entre perda e atraso escolar, seguimos para uma segunda distinção importante de ser feita: aquela entre atraso na aprendizagem e atraso escolar. Embora estejam associados de alguma forma, o termo “atraso na aprendizagem” não pode ser visto como sinônimo de “atraso escolar”. São coisas distintas. O atraso na aprendizagem implica que determinados estudantes não adquiriram, de maneira exitosa, os conhecimentos exigidos para a série que estavam matriculados. Já o atraso escolar se refere a estar cursando uma série inferior à recomendada para sua idade, o que ocorre, geralmente, por motivos de abandono e/ou reprovação escolar.

É importante ressaltar que atrasos na aprendizagem já existiam antes mesmo da pandemia. Estudantes terminavam a Educação Básica com proficiências inferiores à recomendada. Prova disso é que 63% e 65,8% dos estudantes da 3ª série do Ensino Médio brasileiro, em 2019, apresentaram desempenho básico em língua portuguesa (até 300 pontos) e em matemática (até 300 pontos), respectivamente (INEP, 2021). Já em Minas Gerais, aferido pelo Simave, 37% e 59% dos estudantes apresentavam aprendizado considerado baixo, na escala local, em língua portuguesa (até 250 pontos) e matemática (até 275 pontos), respectivamente.

A diferença é que, com a suspensão das aulas presenciais, a garantia do direito a uma educação de qualidade para todos, que já era desafiadora, tornou-se algo ainda mais distante.

Como garantir grupos comparáveis para avaliar o impacto da pandemia? Uma abordagem com o pareamento exato de estudantes

Para avaliar os efeitos da pandemia na aprendizagem, adotamos a estratégia de comparar dois grupos de estudantes: um que passou pela suspensão das aulas presenciais e outro que não vivenciou essa experiência. Os grupos em comparação são definidos como “grupo de tratamento” e “grupo de controle”. O grupo tratamento é aquele que passou pela intervenção – no caso, a suspensão das aulas presenciais – enquanto o grupo controle é composto por indivíduos com características semelhantes, mas que não passaram por tal intervenção.

Em experimentos aleatorizados, os resultados dos grupos de tratamento podem ser comparados diretamente, pois suas unidades são usualmente similares. No entanto, em experimentos não aleatorizados, como é o caso deste estudo, essa comparação direta pode introduzir vieses, uma vez que os indivíduos do grupo de tratamento tendem a diferir significativamente dos demais (Rosenbaum; Rubin, 1983).

Segundo dados disponíveis no Portal SIMAVE, a participação dos estudantes no Proeb de 2019 foi de 94%, 89% e 85% no 5º ano do Ensino Fundamental, 9º ano do Ensino Fundamental e 3º ano do Ensino Médio, respectivamente. Já em 2021, os percentuais foram de 82%, 74% e 59%, para as mesmas etapas de ensino. Uma das justificativas para a queda na participação é que nem todas as escolas estavam autorizadas a retornar com o ensino presencial devido aos decretos municipais vigentes.

Como estamos diante de um estudo do tipo caso–controle, não é suficiente comparar as proficiências dos estudantes em 2019 (antes da pandemia) e em 2021 (após quase dois anos de suspensão). É necessário controlar variáveis que influenciam essa comparação a fim de reduzir os vieses e evitar conclusões equivocadas. A própria queda na taxa de participação em 2021 indica a possibilidade de viés de seleção nos dados – por exemplo, a exclusão, da amostra final, de estudantes de menor nível socioeconômico. Esse problema seria minimizado se, em ambos os anos,

tivéssemos os dados de avaliação dos mesmos estudantes, o que não é viável, considerando as séries e os anos avaliados no Proeb.

Vale destacar que trabalhar com dados de mais de um ponto no tempo não equivale, necessariamente, a realizar um estudo longitudinal. Ainda que a escala de proficiência seja comparável, comparações diretas dos resultados, especialmente quando se busca estimar um efeito causal, exigem cautela. Estudos com dados longitudinais, que acompanham os mesmos indivíduos ao longo do tempo, partem da suposição de que as variáveis não observadas permanecem fixas durante o período analisado. Essa suposição permite que parte da variabilidade não observada seja controlada na análise.

Construindo grupos comparáveis com base em variáveis observáveis

Para realizar a comparação entre as proficiências de 2019 e 2021 e estimar o atraso na aprendizagem, mesmo com estudantes distintos avaliados em cada ano, utilizamos a metodologia de pareamento exato. Esse método permite identificar, em 2019, estudantes com “exatamente” as mesmas características observáveis dos estudantes de 2021. É como se estivéssemos comparando estudantes idênticos, com a diferença de que apenas os de 2021 foram expostos à suspensão das aulas presenciais. Assim, as diferenças nas proficiências podem ser atribuídas a esse fator.

O pareamento de grupos é utilizado para estimar o efeito de tratamentos com dados não experimentais. Um dos métodos mais conhecidos é o *Propensity Score Matching* (PSM) (Ho *et al.*, 2011; Fundação Itaú Social, 2017), no qual se calcula a probabilidade de o indivíduo pertencer ao grupo de tratamento a partir das variáveis observáveis. Cada indivíduo recebe um escore de propensão, que é utilizado para encontrar pares com base em uma distância máxima. No entanto, há críticas ao PSM, como a de King e Nielsen (2019), que apontam limitações do método ao não considerar a covariância entre as variáveis de pareamento.

Para contornar essas limitações e reduzir os riscos de vieses no pareamento, optamos pelo uso do pareamento exato. Nesse método, os estu-

dantes são considerados pares apenas se todas as variáveis apresentarem exatamente os mesmos valores. Por exemplo, um estudante branco, do sexo masculino, com 12 anos, nível socioeconômico no 2º quartil, sem reprovações ou abandonos escolares e que iniciou a Educação Básica no 1º ano do Fundamental será pareado apenas com outro estudante que tenha exatamente esse mesmo perfil. Quando mais de um estudante apresenta o mesmo padrão de respostas, é atribuído um peso ao perfil, correspondente à quantidade de estudantes pareados em ambos os grupos.

Para a construção dos perfis de pareamento, foram utilizadas sete características observáveis dos estudantes, coletadas a partir dos questionários contextuais aplicados junto aos testes: i) sexo (feminino, masculino); ii) raça (branco, pardo, preto, outros); iii) idade; iv) nível socioeconômico (em quintis); v) número de reprovações (nenhuma, uma, duas ou mais); vi) número de abandonos escolares (nenhum, um, dois ou mais); vii) série em que iniciou a Educação Básica (creche, pré-escola, 1º ano do Ensino Fundamental, depois do 1º ano). Importante destacar que as variáveis utilizadas para o pareamento não incluem a proficiência, justamente porque o objetivo é comparar o desempenho entre grupos semelhantes nessas características.

O indicador de nível socioeconômico foi calculado por meio da Teoria de Resposta ao Item (TRI) (Andrade; Tavares; Valle, 2000) a partir de 14 variáveis: escolaridade da mãe e do pai ou responsáveis, recebimento de Bolsa Família, acesso à *internet*, quarto próprio, número de quartos e banheiros e posse de bens, como máquina de lavar, automóvel, TV, *smartphone*, computador, *videogame* e quantidade de livros. Com as variáveis de controle definidas, a etapa seguinte consiste na organização da base de dados para o processo de pareamento dos estudantes.

Essa organização implica a exclusão de parte da amostra por diferentes motivos. Inicialmente, são selecionados apenas os estudantes que foram avaliados e possuem um valor de proficiência, uma vez que essa é a variável de interesse na análise após o pareamento. Em seguida, são removidos os estudantes com dados ausentes nas variáveis de controle, conforme os critérios exigidos pelo método de pareamento adotado. Por fim, realiza-se o pareamento, e os estudantes cujo perfil não foi encontrado nos dois anos

analisados também são excluídos. A tabela 1 apresenta a redução gradativa da base ao longo dessas etapas. No total, cerca de 45 mil estudantes não foram pareados em ambos os anos, o que corresponde a aproximadamente 11% da amostra disponível em cada ano.

Tabela 1 – Quantitativo de estudantes por grupo de seleção

Agrupamento	Ano	Série			Total
		5º ano	9º ano	3ª série	
Total de estudantes na base de dados	2019	79.282	154.529	175.474	409.285
	2021	71.653	170.563	192.184	434.400
Grupo 1: Estudantes avaliados em geral	2019	74.404	137.187	149.284	360.875
	2021	59.009	125.964	113.935	298.908
Grupo 2: Estudantes avaliados e sem <i>missing</i> nas variáveis de controle	2019	72.017	133.093	143.656	348.766
	2021	55.816	120.868	107.444	284.128
Grupo 3: Estudantes excluídos da primeira etapa (por terem deixado de responder questões do questionário contextual) = Grupo 1 - Grupo 2	2019	2.387	4.094	5.628	12.109
	2021	3.193	5.096	6.491	14.780
Grupo 4: Estudantes Pareados	2019	63.541	120.952	132.369	316.862
	2021	48.513	108.621	94.990	252.124
Grupo 5: Estudantes excluídos da segunda etapa (por não terem um par, com as mesmas características, encontrado) = Grupo 2 - Grupo 4	2019	8.476	12.141	11.287	31.904
	2021	7.303	12.247	12.454	32.004
Grupo 6: Total de estudantes excluídos = Grupo 3 + Grupo 5	2019	10.863	16.235	16.915	44.013
	2021	10.496	17.343	18.945	46.784

Fonte: Elaborado pelas autoras, 2025, com base nos microdados disponibilizados.

Para operacionalizar o pareamento dos alunos, utilizamos o pacote *MatchIt* (Ho *et al.*, 2011; Sekhon, 2011), desenvolvido para o *software R-project*, com especificação do método exato na fórmula de ajuste. Diferentemente do método PSM, o pareamento exato não exige uma correspondência um para um. Cada perfil de estudante recebe um peso e um número

índice que identifica quais estudantes de 2019 estão pareados com os de 2021 (grupo 4).

Com o pareamento concluído por série e disciplina, obtivemos grupos de estudantes de 2021 bastante semelhantes aos de 2019, com a diferença de que estiveram sujeitos à suspensão das aulas presenciais.

Nem todos os estudantes foram pareados, o que ocorre quando não há um correspondente com o mesmo perfil. Nos dois anos analisados, aproximadamente 90% dos estudantes com dados completos nas variáveis de controle foram pareados, totalizando 252.124 estudantes de 2021 e 316.862 estudantes de 2019. Esses números representam 58% dos previstos na avaliação de 2021 e 77,4% dos previstos na de 2019. A análise por disciplina revela diferenças mínimas, indicando que poucos estudantes fizeram apenas uma das provas (matemática ou língua portuguesa).

Como etapa complementar à técnica de pareamento, foi realizada uma análise comparativa entre o perfil dos estudantes incluídos (grupo 2) e excluídos (grupo 6) da amostra final. O grupo 2, composto por 632.894 estudantes com dados completos de proficiência e questionário contextual para ambos os anos, permitiu a formação de grupos comparáveis com base em sexo, raça/cor, idade, trajetória escolar e nível socioeconômico (NSE).

Em 2021, esse grupo apresentou um perfil levemente distinto daquele de 2019, com maior proporção de estudantes mais velhos, brancos, de NSE mais alto e histórico escolar mais irregular, o que reforça a importância do pareamento exato para controlar variáveis observáveis. Já o grupo 6, com 90.797 estudantes excluídos por ausência de dados ou falta de correspondência entre os anos, revelou um padrão de exclusão desigual: em 2019, foram majoritariamente estudantes negros e de NSE mais baixo; em 2021, destacou-se a exclusão de estudantes com trajetórias escolares atípicas, especialmente no Ensino Médio.

A análise do NSE mostrou que essa variável foi a que mais influenciou a composição dos grupos, tanto antes quanto depois do pareamento. Testes estatísticos para médias revelaram que os estudantes de 2021 tinham, em média, NSE superior aos de 2019 e que os estudantes não pareados

(grupo 6) apresentavam NSE inferior aos pareados (grupo 4). Esses resultados sugerem que o processo de pareamento excluiu proporcionalmente mais estudantes em situação de maior vulnerabilidade, o que pode ter gerado um pequeno viés de seleção na amostra final. Embora esse viés não invalide os achados sobre o impacto da pandemia, é possível que os resultados subestimem a magnitude real da defasagem de aprendizagem, sobretudo entre os grupos mais desfavorecidos.

Desigualdades e atraso na aprendizagem: resultados empíricos

Esta seção apresenta os principais achados empíricos do estudo, com base nos dados do PROEB 2019 e 2021. Antes da análise dos resultados, é importante caracterizar a base utilizada. A base de dados reúne 843.685 estudantes distintos previstos nas avaliações de 2019 e 2021. Em 2019, dos 409.285 estudantes previstos, 360.875 foram efetivamente avaliados (88,2%). Em 2021, a taxa de participação caiu para 68,8%, com 298.908 estudantes avaliados dos 434.400 previstos, em razão de decretos municipais que impediram a realização das provas presenciais. Em 172 escolas (5,05%), nenhum estudante foi avaliado, o que representa 15.106 alunos fora da amostra efetiva de 2021.

Análises do atraso na aprendizagem durante a pandemia

Para mensurar o impacto da pandemia nas aprendizagens, utilizamos os dados dos estudantes pareados (grupo 4), que permitem uma comparação mais equilibrada entre os grupos de 2019 e 2021. A proficiência dos estudantes no Proeb é calculada com base na TRI e transformada para uma escala que varia entre 0 e 500 pontos, compatível com a do Saeb. Assim, é possível comparar diretamente os resultados entre os anos e as etapas da Educação Básica.

A hipótese norteadora desta análise é de que os estudantes expostos à suspensão das aulas presenciais ao longo de 2020 e parte de 2021 apresentaram desempenho inferior àqueles que cursaram o mesmo ano escolar em condições regulares, em 2019. Essa hipótese é sustentada por

pesquisas anteriores, nacionais e internacionais, que já apontaram “perdas” de aprendizagem no período pandêmico, especialmente em matemática (Kuhfeld *et al.*, 2020; Koslinski; Bartholo, 2021).

A tabela 2 apresenta os resultados do teste de comparação de médias de proficiência entre os estudantes do grupo controle (2019) e do grupo tratamento (2021) por etapa de ensino e componente curricular. Em todas as séries e disciplinas, observa-se queda no desempenho médio em 2021, com variações que vão de -0,7 a -7,5 pontos em língua portuguesa e de -5,7 a -11,1 pontos em matemática. As reduções mais acentuadas foram observadas no 5º ano do Ensino Fundamental e na 3ª série do Ensino Médio, sobretudo em matemática, evidenciando um padrão já descrito em outras pesquisas: os estudantes mais novos foram mais vulneráveis aos efeitos da descontinuidade pedagógica.

Esses resultados confirmam a hipótese de atraso escolar generalizado, mas revelam variações relevantes conforme a etapa de ensino. Nos Anos Iniciais do Ensino Fundamental (5º ano), o atraso estimado equivale a 22,7% de um ano letivo em língua portuguesa e 55,5% em matemática, tomando como referência a estimativa de que um ano escolar corresponde a 20 pontos de proficiência (Alves; Soares; Xavier, 2016). Já nos Anos Finais do Ensino Fundamental (9º ano), as lacunas parecem ter sido menores, de apenas 0,69 pontos em língua portuguesa e 5,68 em matemática. No Ensino Médio, os dados voltam a indicar quedas expressivas: 5,34 pontos em língua portuguesa e 11,04 em matemática, valores próximos aos verificados no 5º ano.

Tabela 2 – Teste de médias das proficiências entre 2019 e 2021

Etapa	Disciplina	Média do Grupo Controle (2019)	Média do Grupo Tratamento (2021)	Diferença (2021-2019)	Atraso em % do ano letivo	Intervalo de Confiança de (95%)	P-valor
5º ano EF	Português	226,75	219,20	-7,54	22,7%	[-8,17; -6,92]	< 0,000
	Matemática	234,43	223,34	-11,09	55,5%	[-11,67; -10,51]	< 0,000
9º ano EF	Português	250,98	250,29	-0,69	3,5%	[-1,14; -0,24]	< 0,000
	Matemática	258,12	252,44	-5,68	28,4%	[-6,11; -5,26]	< 0,000
3º ano EM	Português	268,45	263,12	-5,34	26,7%	[-5,77; -4,90]	< 0,000
	Matemática	271,88	260,84	-11,04	55,2%	[-11,47; -10,61]	< 0,000

Nota: EF=Ensino Fundamental; EM=Ensino Médio.

Fonte: Elaborada pelas autoras, 2025, com base nos microdados disponibilizados.

Sob outra perspectiva, observa-se que as reduções de proficiência entre 2019 e 2021 foram, na maioria dos casos, muito superiores às variações registradas entre 2018 e 2019 (tabela 3). No período pré-pandêmico, com exceção da língua portuguesa nos Anos Iniciais, todas as demais etapas já apresentavam queda de proficiência, com variações entre -0,50 e -6,9 pontos nos componentes avaliados.

No período pandêmico, essas reduções se intensificaram de forma expressiva. Em matemática, mesmo após o ajuste temporal da variação (dividindo-se a queda por dois), a redução superou em mais de dez vezes a observada entre 2018 e 2019 nos Anos Iniciais do Ensino Fundamental e no Ensino Médio. Nos Anos Finais, embora a queda também tenha superado a do período anterior, sua magnitude foi sensivelmente menor (1,2 vezes).

Em língua portuguesa, por sua vez, observa-se um padrão distinto daquele verificado em matemática entre as etapas de ensino. Nos Anos Iniciais, a redução entre 2019 e 2021 voltou a superar em mais de dez vezes

a variação do período pré-pandêmico. Já nos Anos Finais do Ensino Fundamental e no Ensino Médio, a queda no contexto pandêmico foi menos intensa que a registrada entre 2018 e 2019. Infelizmente, não foi possível estender a comparação para outros anos pré-pandêmicos em razão da indisponibilidade de dados públicos.

Tabela 3 – Comparação da queda de proficiência entre 2019 e 2021 com variação pré-pandêmica

Disciplina	Etapa de ensino	2018	2019	Variação na pré-pandemia (2019-2018) (A)	Variação na pandemia (2021-2019) (3 edições) (B)	Variação na pandemia estimada (entre 2 edições) (C=B/2)	Variação na pandemia em relação à variação na pré-pandemia (C/A)
Língua Portuguesa	5º ano EF	222,11	222,39	0,28	-7,54	-3,77	13,5
	9º ano EF	251,87	246,46	-5,41	-0,69	-0,35	0,1
	3º ano EM	272,08	265,19	-6,89	-5,34	-2,67	0,4
Matemática	5º ano EF	230,85	230,35	-0,50	-11,09	-5,55	11,1
	9º ano EF	256,07	253,77	-2,30	-5,68	-2,84	1,2
	3º ano EM	268,93	268,61	-0,32	-11,04	-5,52	17,3

Nota: EF=Ensino Fundamental; EM=Ensino Médio.

Fonte: Elaborada pelas autoras, 2025, com base nos dados públicos disponibilizados em: <https://avaliacoes.educacao.mg.gov.br/dados-abertos/participa%C3%A7%C3%A3o-e-desempenho>. Acesso em: 10 dez. 2025.

Independentemente da etapa de ensino, em ambas as perspectivas, as reduções nas proficiências em matemática foram maiores que nas de língua portuguesa. Adicionalmente, em ambos os componentes, os Anos Iniciais do Ensino Fundamental demonstraram mais atrasos na aprendizagem, seguidos do Ensino Médio e, por último, dos Anos Finais do Ensino Fundamental.

A estimativa pontual das proficiências médias carrega uma interpretação pedagógica, definida para cada série e componente curricular. De acordo com a interpretação da escala de proficiências do Saeb (INEP, 2020) e considerando a componente de língua portuguesa, a proficiência 226,75 (aferida em 2019), no 5º ano do Ensino Fundamental, localiza-se no nível 5 da escala do Saeb. De maneira geral, isso significa que os estudantes dominam habilidades que refletem uma compreensão dos elementos linguísticos e contextuais presentes nos diversos gêneros textuais.

Já a proficiência 219,20 (aferida em 2021) retrocede o desenvolvimento de algumas competências e se posiciona no nível 4 da escala do Saeb. Nesse nível, os alunos dominam habilidades que refletem uma compreensão multifacetada dos diversos aspectos da linguagem e da narrativa presentes nos textos analisados. São capazes de identificar elementos, mas habilidades de reconhecimento e de inferências ainda estão menos desenvolvidas.

No 9º ano, a média de proficiência em língua portuguesa foi de 250,98, em 2019, e de 250,29, em 2021. Ambas as pontuações situam os estudantes no nível 3 da escala do Saeb, sugerindo estabilidade na posição relativa do grupo. No entanto, mesmo com uma diferença numérica pequena (-0,69), é possível que esse leve decréscimo oculte efeitos cumulativos no processo de aprendizagem. Estudantes nesse nível conseguem identificar informações explícitas, interpretar relações semânticas simples e reconhecer a função de elementos como conjunções e pontuação em textos narrativos e expositivos. Contudo, permanecem com dificuldades de realizar inferências mais complexas e lidar com múltiplas vozes discursivas.

Na 3ª série do Ensino Médio, a média de proficiência em língua portuguesa caiu de 268,45, em 2019, para 263,12, em 2021, mantendo-se, ainda assim, no nível 2 da escala. Os estudantes desse nível demonstram capacidade de localizar informações principais em textos diversos, reconhecer pronomes e conectivos e interpretar o sentido de palavras ou expressões a partir do contexto. No entanto, ainda revelam limitações em análises críticas, síntese de argumentos e leitura de textos multimodais — habilidades fundamentais para o prosseguimento nos estudos e para o exercício da cidadania. A permanência no mesmo nível, portanto, não deve ser

interpretada como estabilidade no desempenho, mas como manutenção de um patamar ainda insuficiente para muitos dos objetivos da etapa.

Em matemática, os resultados apontam perdas mais expressivas do que em língua portuguesa, com deslocamentos relevantes nos níveis de desempenho. No 5º ano do Ensino Fundamental, a média de proficiência em 2019 foi de 234,43 pontos, situando os estudantes no nível 5 da escala do Saeb. Em 2021, a média caiu para 223,34, o que corresponde ao nível 4, indicando uma perda de complexidade nas habilidades consolidadas.

No 9º ano, as médias em matemática foram de 258,12 pontos, em 2019, e 252,44, em 2021. Ambas as pontuações situam os estudantes no nível 3 da escala do Saeb, embora a pontuação de 2021 esteja próxima do limite inferior desse nível. O nível 3 representa um patamar intermediário de aprendizado na etapa, e a queda de quase 6 pontos, ainda que mantendo os estudantes no mesmo nível, pode indicar um empobrecimento do repertório conceitual e procedimental necessário para avançar para os níveis mais altos, especialmente em tópicos como proporcionalidade, frações e interpretação de representações gráficas mais complexas.

Na 3ª série do Ensino Médio, a média de proficiência em matemática caiu de 271,88, em 2019, para 260,84, em 2021 (uma diferença de mais de 11 pontos), embora ambas permaneçam no nível 2 da escala. No entanto, essa manutenção no mesmo nível não deve ser interpretada como ausência de impacto. A queda reforça a hipótese de que, mesmo nos Anos Finais da Educação Básica, a suspensão das atividades presenciais comprometeu aprendizagens importantes para a progressão acadêmica e a preparação para exames e transições educacionais.

Efeitos diferenciais da pandemia na aprendizagem por nível socioeconômico

Nesta seção, analisamos se a pandemia afetou de forma diferente os estudantes conforme seu nível socioeconômico. O objetivo é compreender se os atrasos na aprendizagem foram mais intensos entre aqueles com NSE mais baixo, conforme aponta a literatura (Koslinski; Bartholo, 2021; Senkevics; Bof, 2022; Senkevics; Alcantara, 2024). Para tanto, calculamos a proficiência

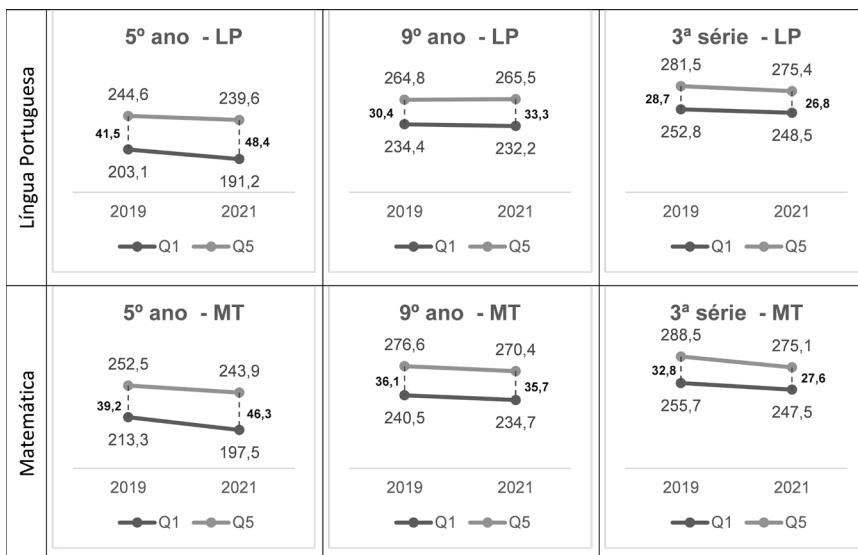
média dos estudantes da rede estadual de Minas Gerais em língua portuguesa (LP) e matemática (MT), para os anos de 2019 e 2021, segmentada por série escolar (5º ano, 9º ano e 3ª série do Ensino Médio) e quintis de NSE, sendo Q1 o mais baixo e Q5 o mais alto (quadro 1).

Os resultados das proficiências médias em língua portuguesa e matemática, por etapa de ensino e quintil de nível socioeconômico (Q1 e Q5), revelam padrões distintos de impacto da pandemia sobre a aprendizagem. Nos Anos Iniciais (5º ano), observa-se uma queda generalizada, mais acentuada entre os estudantes de menor NSE (Q1), o que ampliou a desigualdade entre os quintis. Em língua portuguesa, o desempenho do Q1 caiu de 203,1 para 191,2 pontos, enquanto o Q5 teve uma redução menor, de 244,6 para 239,6 pontos. Em matemática, a diferença entre os grupos se alargou ainda mais: o Q1 perdeu 15,8 pontos e o Q5, 8,6. Esse cenário indica que, no 5º ano, os efeitos da pandemia foram mais severos para os estudantes em situação de maior vulnerabilidade, reforçando um padrão de aprofundamento das desigualdades educacionais. No entanto, esse padrão não se repete nas demais etapas.

No 9º ano, por exemplo, a proficiência média em língua portuguesa se manteve praticamente estável para ambos os grupos, com ligeiro aumento no Q5 (de 264,8 para 265,5 pontos) e pequena redução no Q1 (de 234,4 para 232,2 pontos), resultando em um pequeno acréscimo das desigualdades. Já em matemática, observou-se queda tanto no Q1 quanto no Q5, mas mais acentuada entre os estudantes de maior NSE, sugerindo um leve estreitamento das desigualdades.

Na 3ª série do Ensino Médio, o comportamento se inverte: as quedas de proficiência foram mais expressivas entre os estudantes do Q5 do que entre os do Q1. Em língua portuguesa, o Q5 perdeu 6,1 pontos, enquanto o Q1 teve uma redução menor, de 4,3 pontos. Em matemática, a diferença é ainda mais notável, já que o Q5 caiu de 288,5 para 275,1 (-13,4 pontos) e o Q1, de 255,7 para 247,5 (-8,2 pontos). Apesar das quedas de proficiências em ambos os grupos de NSE, o movimento superior no Q5 ocasionou uma redução nas desigualdades. Diferentemente dos Anos Iniciais, os estudantes mais favorecidos do Ensino Médio podem ter sido mais afetados.

Quadro 1 – Proficiências médias dos estudantes por quintis de NSE, etapa de ensino, disciplina e ano



Fonte: Elaborado pelas autoras, 2025, com base nos microdados disponibilizados.

Esses achados indicam que os efeitos da pandemia sobre a aprendizagem não foram homogêneos entre os grupos sociais nem entre as etapas de ensino. Enquanto, nos Anos Iniciais, houve um claro aprofundamento das desigualdades, nos Anos Finais e especialmente no Ensino Médio, os dados sugerem uma relativa estabilidade ou até uma inversão no padrão esperado, com os estudantes de maior NSE apresentando quedas mais pronunciadas. Esses resultados reforçam a necessidade de análises desagregadas, sensíveis às especificidades de cada etapa e grupo, para orientar políticas de recomposição da aprendizagem de forma mais equitativa e eficaz.

Em conjunto, os resultados confirmam que a pandemia não afetou de forma homogênea o desempenho escolar. Assim como em estudos anteriores (Koslinski; Bartholo, 2021; Senkevics; Bof, 2022; Senkevics; Alcantara, 2024), também encontramos que os impactos foram particularmente mais intensos em matemática e que os estudantes de NSE mais baixo

experimentaram perdas mais acentuadas, especialmente quando matriculados no 5º ano, se comparadas às demais etapas, sinalizando um aprofundamento das desigualdades já existentes.

Tal padrão é confirmado por estudos que evidenciam as dificuldades enfrentadas pelas escolas para garantir o vínculo pedagógico e a continuidade da aprendizagem, sobretudo em territórios com menos infraestrutura digital, baixa escolaridade familiar e ausência de políticas nacionais e articuladas de apoio aos municípios (Alves *et al.*, 2022). Nesse sentido, a gestão do sistema de ensino não foi capaz de assegurar o pleno direito à educação, sobretudo na dimensão do aprendizado. As desigualdades educacionais foram acentuadas na contramão dos marcos legais que definem a educação como direito (Brasil, 1988, 1996).

Considerações finais

Durante o período de suspensão das aulas presenciais, o governo do estado de Minas Gerais implementou diferentes estratégias com o objetivo de garantir a continuidade do ensino, considerando a diversidade de condições tecnológicas e culturais das famílias. Entre as ações adotadas, destacam-se: a criação do aplicativo “Conexão Escola” e de plataformas *online* (como páginas no *site* da Secretaria e o *hotsite* “Estude em Casa”); a disponibilização de *internet* patrocinada para acesso aos conteúdos; a transmissão de aulas por canais de televisão e *YouTube*; além da distribuição de materiais impressos por meio dos PETs. No entanto, este estudo não tem como objetivo avaliar o impacto dessas políticas na aprendizagem, razão pela qual não são feitas inferências causais nesse sentido.

Em relação às lacunas de aprendizagem observadas, os resultados apontam diferenças importantes entre as etapas de ensino. Os estudantes dos Anos Iniciais apresentaram os maiores atrasos (7,54 pontos em língua portuguesa e 11,09 em matemática), resultado coerente com a menor autonomia dessa faixa etária no ensino remoto. O Ensino Médio, embora composto por estudantes teoricamente mais adaptados ao sistema escolar, também apresentou atrasos expressivos, especialmente em matemática (5,34 pontos em língua portuguesa e 11,04 em matemática).

O comportamento mais atípico foi observado nos Anos Finais do Ensino Fundamental, cuja defasagem foi significativamente menor, sobretudo em língua portuguesa (0,69 ponto em língua portuguesa e 5,68 em matemática). Apesar das perdas observadas nas médias de proficiência, os níveis de desempenho permaneceram estáveis entre 2019 e 2021, tanto para os Anos Finais quanto para o Ensino Médio, o que levanta hipóteses sobre a efetividade do ensino remoto ou, alternativamente, sobre a baixa qualidade do ensino anterior à pandemia – ou seja, a manutenção de um patamar de desempenho historicamente baixo.

Esse último argumento é reforçado pelos dados do Proeb de 2019, que já indicavam baixos níveis de aprendizagem, mesmo antes da crise sanitária. Em língua portuguesa, os percentuais de estudantes nos níveis “baixo” e “intermediário”, somados, foram de 33% no 5º ano, 67% no 9º ano e 73% na 3ª série do Ensino Médio. Em matemática, esses percentuais foram ainda mais elevados: 45%, 82% e 93%, respectivamente. Com exceção dos Anos Iniciais, praticamente sete em cada dez estudantes da rede estadual não aprendiam o conteúdo adequado à etapa cursada, o que reforça a hipótese de que o desempenho observado em 2021 pode refletir a persistência de um problema estrutural.

Ao se examinarem os resultados por NSE, observa-se que os efeitos da pandemia também se manifestaram de maneira desigual. Nos Anos Iniciais, os estudantes do primeiro quintil de NSE (Q1) apresentaram perdas mais acentuadas do que os do quinto quintil (Q5), ampliando a distância entre os grupos. No entanto, nos Anos Finais e, sobretudo, no Ensino Médio, esse padrão se inverte: em várias situações, os estudantes de menor NSE apresentaram quedas menores do que os mais favorecidos, especialmente em matemática, sugerindo que o impacto do período pode ter afetado também o desempenho de estudantes de Q5. Em outras palavras, os efeitos da pandemia não apenas variaram por etapa, mas também por faixa socioeconômica, indicando que as desigualdades se reconfiguraram ao invés de se ampliarem de maneira uniforme.

A queda de proficiência em matemática, que foi mais acentuada entre os estudantes de maior nível socioeconômico dos Anos Finais do Ensino

Fundamental e do Ensino Médio, sugere que a instituição escolar desempenha um papel relevante no desenvolvimento dessas habilidades, especialmente para os grupos que conseguem se beneficiar mais plenamente de suas práticas pedagógicas. Por outro lado, a redução menos expressiva entre os estudantes de menor NSE indica que, embora a escola seja central para as aprendizagens, suas práticas já apresentavam, mesmo antes da pandemia, alcance limitado sobre os alunos mais vulneráveis.

Outro fator a ser considerado é o possível viés de seleção decorrente da menor participação nas avaliações em 2021, especialmente entre estudantes de menor NSE. A análise do pareamento evidenciou que muitos estudantes foram excluídos por não possuírem correspondentes entre os dois anos analisados. Testes de comparação entre os grupos pareados (grupo 4) e não pareados (grupo 6) revelaram que os não pareados apresentavam, em média, NSE inferior, com diferenças de até 0,27 desvios-padrão em 2019. Isso sugere que os estudantes efetivamente analisados representam uma parcela levemente mais favorecida da rede, o que pode ter suavizado os resultados e subestimado o atraso real de aprendizagem. Considerando que parte dos estudantes de maior vulnerabilidade nem sequer foi avaliada em 2021, é plausível supor que as perdas tenham sido ainda maiores do que aquelas estimadas neste estudo.

Com essas ressalvas, os achados levantam pelo menos três questões relevantes para investigações futuras: (i) por que os Anos Finais do Ensino Fundamental apresentaram menor atraso em comparação às outras etapas? Teria havido um viés de seleção específico? (ii) Esse comportamento é observado em outras unidades da federação ou é uma característica particular de Minas Gerais? (iii) Será que os testes de proficiência estão realmente aferindo as habilidades que os estudantes deveriam desenvolver em cada etapa? Ou haveria um descompasso entre a matriz de avaliação e os conteúdos efetivamente trabalhados?

Por fim, é importante reconhecer que não é possível isolar o impacto da suspensão das aulas presenciais dos efeitos das estratégias de ensino remoto implementadas no período. Assim, os resultados apresentados refletem a combinação desses dois fatores: a incerteza quanto à efetividade

do ensino remoto e o grau de êxito na execução das ações emergenciais adotadas. Ainda assim, as estimativas apontam para uma lacuna real de aprendizagem na rede estadual de ensino de Minas Gerais, que deve orientar a formulação de políticas de recomposição das aprendizagens e reforçar o compromisso com a garantia do direito à educação de qualidade e com equidade.

Referências

- ALAVARSE, O. M. Desafios da avaliação educacional: ensino e aprendizagem como objetos de avaliação para a igualdade de resultados. **Cadernos Cenpec**, São Paulo, v. 3, n. 1, p. 135–153, jun. 2013. Disponível em: <https://pdfs.semanticscholar.org/3513/bc0bc2ff5aeccab109524845d122d0ccbae7.pdf>. Acesso em: 26 nov. 2025.
- ALAVARSE, O. M.; BRAVO, M. H.; MACHADO, C. Avaliações externas e qualidade na educação básica: articulações e tendências. **Estudos em Avaliação Educacional**, São Paulo, v. 24, n. 54, p. 12–31, jan./abr. 2013. Disponível em: <https://publicacoes.fcc.org.br/ea/article/view/1900>. Acesso em: 26 nov. 2025.
- ALVES, M. T. G.; SOARES, J. F.; XAVIER, F. P. Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais. **Revista Brasileira de Sociologia**, São Cristóvão, v. 4, n. 7, p. 49–81, 2016. Disponível em: <https://rbs.sbsociologia.com.br/index.php/rbs/article/view/181>. Acesso em: 26 nov. 2025.
- ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria de Resposta ao Item**: conceitos e aplicações. São Paulo: Associação Brasileira de Estatística, 2000.
- ATTEBERRY, A.; MCEACHIN, A. **School's out**: The role of summers in understanding achievement disparities. Providence: Ed. Working Paper, 2020.
- BAUER, A.; ALAVARSE, O. M.; OLIVEIRA, R. P. Avaliações em larga escala: uma sistematização do debate. **Educação e Pesquisa**, São Paulo, v. 41, n. esp, p. 1367–1382, dez. 2015. Disponível em: <https://www.scielo.br/j/ep/a/PgMHxD3BYhzBr-6B7CpB5BjS/?format=pdf&lang=pt>. Acesso em: 26 nov. 2025.
- BARTHOLO, T. L.; KOSLINSK, M. C.; TYMMS, P.; CASTRO, D. L. Learning loss and learning inequality during the Covid-19 pandemic. **Ensaio: Avaliação de Políticas Públicas em Educação**, Rio de Janeiro, v.31, n.119, p. 1–24, 2023. Disponível em: <https://www.scielo.br/j/ensaio/a/8sNJkg9syT5dXMP9wrBtbDc/?format=html&lang=en>. Acesso em: 26 nov. 2025.

BONAMINO, A. A evolução do SAEB: desafios para o futuro. **Em Aberto**, Brasília, v. 29, n. 96, p. 113–126, mai./ago. 2016. Disponível em: <https://rbep.inep.gov.br/ojs3/index.php/emaberto/article/view/3155>. Acesso em: 26 nov. 2025.

BONAMINO, A; SOUSA, S. Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. **Educação e Pesquisa**, São Paulo, v. 38, n. 2, p. 373–388, abr./jun. 2012. Disponível em: <https://www.scielo.br/j/ep/a/rtQkYDSjky4mXG9TCrgRSqJ/?format=pdf&lang=pt>. Acesso em: 26 nov. 2025.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Senado Federal, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 26 nov. 2025.

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996**. Estabelece as diretrizes e bases da educação nacional. Brasília, DF: Diário Oficial da União, 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 26 nov. 2025.

BRASIL. **Lei nº 10.172, de 9 de janeiro de 2001**. Aprova o Plano Nacional de Educação e dá outras providências. Brasília, DF: Diário Oficial da União, 2001. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10172.htm. Acesso em: 26 nov. 2025.

BRASIL. **Lei nº 13.005, de 25 de junho de 2014**. Aprova o Plano Nacional de Educação. Brasília, DF: Diário Oficial da União, 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm. Acesso em: 26 nov. 2025.

BRASIL. **Base Nacional Comum Curricular**. Brasília: MEC, 2018. Disponível em: <https://basenacionalcomum.mec.gov.br/>. Acesso em: 26 nov. 2025.

BRASIL. **Lei nº 14.113, de 25 de dezembro de 2020**. Regulamenta o Fundo de Manutenção e Desenvolvimento da Educação Básica. Brasília, DF: Diário Oficial da União, 2020a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14113.htm. Acesso em: 26 nov. 2025.

BRASIL. **Lei nº 14.040, de 18 de agosto de 2020**. Estabelece normas educacionais excepcionais a serem adotadas durante o estado de calamidade pública. Brasília, DF: Diário Oficial da União, 2020b. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14040.htm. Acesso em: 26 nov. 2025.

BRASIL. **Decreto Legislativo nº 6, de 2020**. Brasília, DF: Diário Oficial da União, 2020c. Disponível em: https://www.planalto.gov.br/ccivil_03/portaria/dlg6-2020.htm. Acesso em: 26 nov. 2025.

BRASIL. **Parecer CNE/CP nº 19/2020**. Aprovado em 8 de dezembro de 2020. Brasília, DF: CNE, 2020d. Disponível em: <https://www.gov.br/mec/pt-br/cne/parecer-cp-2020>. Acesso em: 26 nov. 2025.

BRASIL. **Parecer CNE/CP nº 5/2020**. Brasília, DF: CNE, 2020e. Disponível em: <https://www.gov.br/mec/pt-br/cne/parecer-cp-2020>. Acesso em: 26 nov. 2025.

BRASIL. **Parecer CNE/CP nº 9/2020**. Brasília, DF: CNE, 2020f. Disponível em: <https://www.gov.br/mec/pt-br/cne/parecer-cp-2020>. Acesso em: 26 nov. 2025.

BRASIL. **Parecer CNE/CP nº 11/2020**. Brasília, DF: CNE, 2020g. Disponível em: <https://www.gov.br/mec/pt-br/cne/parecer-cp-2020>. Acesso em: 26 nov. 2025.

BRASIL. **Decreto nº 12.391, de 28 de fevereiro de 2025**. Institui o Pacto Nacional pela Recomposição das Aprendizagens. Diário Oficial da União: Seção 1, Brasília, DF, 5 mar. 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/d12391.htm. Acesso em: 26 nov. 2025.

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

BROOKE, N. *et al.* **A avaliação externa como instrumento da gestão educacional nos estados**. São Paulo: Fundação Victor Civita, 2011. Disponível em: https://fvc.org.br/wp-content/uploads/2018/04/estudos_e_pesquisas_educacionais_vol_2.pdf. Acesso em: 26 nov. 2025.

FONSECA, I. C. da; RODRIGUES, C. G.; MACANA, E. C.; ALVES, M. T. G.; SOARES, J. F. Monitoramento das trajetórias educacionais nos municípios. **Estudos em Avaliação Educacional**, v. 35, p. 1-25, 2024. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0103-68312024000100020. Acesso em: 26 nov. 2025.

FREITAS, L. C. *et al.* **Avaliação educacional**: caminhando pela contramão. Petrópolis: Vozes, 2014.

FUNDAÇÃO ITAÚ SOCIAL. **Avaliação econômica de projetos sociais**. 3. ed. São Paulo: Fundação Itaú Social, 2017.

GATTI, B. Avaliação: contexto, história e perspectivas. **Olh@res**, Guarulhos, v. 2, n. 1, p. 08-26, maio 2014. Disponível em: <https://periodicos.unifesp.br/index.php/olhares/article/download/202/76>. Acesso em: 26 nov. 2025.

HO, D.; IMAI, K.; KING, G.; STUART, E. A. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. **Journal of Statistical Software**, [S. l.], v. 42, n. 8, p. 1–28, 2011. Disponível em: <https://www.jstatsoft.org/article/view/v042i08>. Acesso em: 26 nov. 2025.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Escala de proficiência do Saeb**. Brasília, DF: INEP, 2020. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_examenes_da_educacao_basica/escalas_de_proficiencia_do_saeb.pdf. Acesso em: 26 nov. 2025.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Relatório de resultados do Saeb 2019**. Volume 1: 5º e 9º anos do Ensino Fundamental e séries finais do Ensino Médio. Brasília, DF: INEP, 2021. Disponível em: https://download.inep.gov.br/educacao_basica/saeb/2019/resultados/relatorio_de_resultados_do_saeb_2019_volume_1.pdf. Acesso em: 26 nov. 2025.

KING, G.; NIELSEN, R. Why Propensity Scores Should Not Be Used for Matching. **Political Analysis**, Cambridge, v. 27, p. 435–454, 2019. Disponível em: <https://gking.harvard.edu/publications/why-propensity-scores-should-not-be-used-for-matching>. Acesso em: 26 nov. 2025.

KOSLINSKI, M. C.; BARTHOLO, T. L. A pandemia e as desigualdades de oportunidades de aprendizagem na educação infantil. **Estudos em Avaliação Educacional**, São Paulo, v. 32, p. 2–27, 2021. DOI: 10.18222/eae.v32.8314. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/8314>. Acesso em: 26 nov. 2025.

KUHFELD, M.; SOLAND, J.; TARASAWA, B.; JOHNSON, A.; RUZEK, E.; LIU, J. Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. **Educational Researcher**, [S. l.], v. 49, n. 8, p. 549–565, 2020. Disponível em: <https://journals.sagepub.com/doi/10.3102/0013189X20965918>. Acesso em: 26 nov. 2025.

LIBÂNIO, J. C. Políticas educacionais neoliberais e escola: uma qualidade de educação restrita e restritiva. In: LIBÂNIO, José Carlos; FREITAS, Raquel Aparecida Marra da Madeira (orgs.). **Políticas educacionais neoliberais e escola pública**. Goiânia: Espaço Acadêmico, 2018. p. 44–87.

LICHAND, G.; DORIA, C. A.; LEAL-NETO, O.; FERNANDES, J. P. C. The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behaviour*. **Nature Publishing Group**, [S. l.], v. 6, n. 8, p. 1079–1086, ago. 2022. Disponível em: <https://www.nature.com/articles/s41562-022-01350-6>. Acesso em: 26 nov. 2025.

MACHADO, C.; ALAVARSE, O. M. Qualidade das escolas: tensões e potencialidades das avaliações externas. **Educação & Realidade**, Porto Alegre, v. 39, n. 2, p. 413–436, abr./jun. 2014. Disponível em: <https://www.scielo.br/j/edreal/a/fyKb65xtFvXhMw3Hhvp9vNk/abstract/?lang=pt>. Acesso em: 26 nov. 2025.

MINAS GERAIS. **Deliberação do Comitê Extraordinário COVID-19 nº 15, de 20 de março de 2020**. Belo Horizonte: Secretaria de Estado de Fazenda, 2020. Disponível em: <https://www.fazenda.mg.gov.br/coronavirus/instrumentos-normativos/DELIBERACAO-DO-COMITE-EXTRAORDINARIO-COVID-19-N-15-DE-20-DE-MARCO-DE-2020.pdf>. Acesso em: 26 nov. 2025.

MINAS GERAIS. **Resolução SEE nº 4.524, de 11 de março de 2021**. Belo Horizonte: Secretaria de Estado de Educação, 2021a. Disponível em: <https://www.educacao.mg.gov.br/documentos-legislacao/resolucao-see-no-4-524-de-11-de-marco-de-2021/>. Acesso em: 26 nov. 2025.

MINAS GERAIS. **Resolução SEE nº 4.644, de 25 de outubro de 2021**. Belo Horizonte: Secretaria de Estado de Educação, 2021b. Disponível em: <https://www.educacao.mg.gov.br/documentos-legislacao/resolucao-see-no-4-644-de-25-de-outubro-de-2021/>. Acesso em: 26 nov. 2025.

MORICONI, G.; BAUER, A. Avaliações e pesquisas educacionais internacionais: entrevista com Andreas Schleicher. **Estudos em Avaliação Educacional**, São Paulo, v. 28, n. 68, p. 630–640, maio/ago. 2017. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/4848/3391>. Acesso em: 26 nov. 2025.

OECD. **The State of Global Education: 18 Months into the Pandemic**. Paris: OECD Publishing, 2021.

PETRUS, J. S. R.; CAMINHAS, D. A.; BARROS, J. V. S.; MENDONÇA, C. C.; SOUZA, A. V. Simulação de desigualdades educacionais acirradas pela pandemia da Covid-19. **Revista Brasileira de Avaliação**, [S. l.], v. 10, n. 1, p. 1–15, 2021. Disponível em: <https://rbaval.org.br/article/doi/10.4322/rbaval202110008>. Acesso em: 26 nov. 2025.

QUINN, D.; POLIKOFF, M. **Summer learning loss: What is it, and what can we do about it**. Washington, D.C.: Brookings Institution, 2017.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects, **Biometrika**, [S. l.], v. 70, n. 1, p. 41–55, abr. 1983. Disponível em: <https://www.jstor.org/stable/2335942>. Acesso em: 26 nov. 2025.

SÃO PAULO (ESTADO). **O impacto da pandemia na educação**: Avaliação amostral da aprendizagem dos estudantes. São Paulo: Governo do Estado de São Paulo, 2021.

SEKHON, J. S. Matching: multivariate and propensity score matching with balance optimization. **Journal of Statistical Software**, [S. l.], v. 42, n.7, p. 1–52, 2011. Disponível em: <https://www.jstatsoft.org/article/view/v042i07>. Acesso em: 26 nov. 2025.

SEGATI, A. F.; JORDÃO, R. D. S. Os Contextos do Ensino Remoto e Remoto/Presencial sob a Perspectiva dos Alunos do Ensino Médio durante a Pandemia da COVID-19. **EaD Em Foco**, [S. l.], v. 12, n. 2, p. 1–15, 2022. Disponível em: <https://eademfoco.cecierj.edu.br/index.php/Revista/article/view/1664>. Acesso em: 26 nov. 2025.

SENKEVICS, A. S.; BOF, A. M. Desigualdades educacionais na pandemia. **Cadernos de Estudos e Pesquisas em Políticas Educacionais**, [S. l.], v. 7, p. 173–209, 2022. Disponível em: <https://cadernosdeestudos.inep.gov.br/ojs3/index.php/cadernos/article/view/5574>. Acesso em: 26 nov. 2025.

SENKEVICS, A. S.; ALCANTARA, V. G. Nivelando por baixo: Pandemia e queda de aprendizado no ensino fundamental. **Estudos em Avaliação Educacional**, [S. l.], v. 35, p. 1–34, 2024. Disponível em: <https://publicacoes.fcc.org.br/eaee/article/view/10525>. Acesso em: 26 nov. 2025.

SOARES, J. F.; ALVES, M. T. G.; FONSECA, J. A. Trajetórias educacionais como evidência da qualidade da educação básica brasileira. **Revista Brasileira de Estudos de População**, [S. l.], v. 38, p. 1–21, 2021. Disponível em: <https://www.scielo.br/jj/rbepop/a/9ZRM8LBTqQMhMDQNJdWjQZQ/?format=html&lang=pt>. Acesso em: 26 nov. 2025.

SOARES, J. F.; XAVIER, F. P. Pressupostos educacionais e estatísticos do Ideb. **Educação & Sociedade**, [S. l.], v. 34, n. 124, p. 903–923, 2013. Disponível em: <https://www.scielo.br/j/es/a/JLzr4qdx89rjrNXnydNvcy/abstract/?lang=pt>. Acesso em: 26 nov. 2025.

VON HIPPEL, P. T.; WORKMAN, J.; DOWNEY, D. B. Inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of “Are schools the great equalizer?” **Sociology of Education**, v. 91, n. 4, p. 323–357, 2018. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/0038040718801760>. Acesso em: 26 nov. 2025.

CAPÍTULO 10

DIF EM TESTES EDUCACIONAIS: COMPARAÇÃO DE APLICAÇÕES EM PAPEL E COMPUTADOR

Alexandre Jaloto¹

Gustavo Henrique Martins²

DOI: 10.52695/978-65-5456-163-1.10

Introdução

Diferenças entre indivíduos e entre grupos são esperadas em avaliações educacionais. No entanto, quando essas diferenças estão associadas ao formato de aplicação (e.g., papel ou computador), surge uma questão relevante: as disparidades observadas refletem diferenças na proficiência ou estão associadas ao meio de aplicação do teste? Caso a resposta a essa pergunta se deva ao formato de aplicação do teste, podemos estar diante de uma injustiça na avaliação educacional, pois as diferenças na medida de proficiência dos sujeitos devem ser explicadas somente pelo nível de habilidade do sujeito. Portanto, se observarmos diferenças nas

1 Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. E-mail: alexandrejaloto@gmail.com.

2 Universidade São Francisco. CATvante – Laboratório de Testagem Computadorizada.

proficiências entre os grupos que realizaram o teste em papel e em computador, podemos comparar diretamente essas proficiências? A resposta é: não, sem antes verificar a invariância da medida.

A invariância de medida se refere à ideia de que um teste deve medir o mesmo construto de maneira equivalente em diferentes grupos ou contextos (Lubke *et al.*, 2003). Isso significa que, independentemente da forma como o teste é aplicado — por exemplo, no papel ou no computador — os itens devem funcionar da mesma forma e ter o mesmo significado para todos os participantes. Quando essa condição é atendida, podemos afirmar que os escores obtidos refletem de forma justa a habilidade que está sendo avaliada, sem interferência de fatores externos como o meio de aplicação, o gênero, a idade ou a origem cultural dos indivíduos.

Garantir a invariância de medida é fundamental para assegurar interpretações válidas e comparações justas entre grupos (*American Educational Research Association; American Psychological Association; National Council On Measurement In Education*, 2014). Quando essa condição não é atendida, corre-se o risco de tirar conclusões equivocadas, pois diferenças nos escores podem ser consequência de algum viés do instrumento, e não de verdadeiras diferenças na proficiência. Por isso, verificar a invariância é uma etapa essencial em medidas educacionais que envolvem comparações entre grupos. É também uma exigência ética, especialmente em contextos diversos como o brasileiro, em que desigualdades sociais, regionais e de acesso à tecnologia podem influenciar a forma como os testes são compreendidos e respondidos.

Uma das formas de se testar a invariância de medida é por meio da análise de Funcionamento Diferencial do Item (em inglês, *Differential Item Functioning* – DIF). Essa análise é comumente presente no contexto educacional, especialmente em avaliações de larga escala e frequentemente faz uso da Teoria de Resposta ao Item (TRI) como um arcabouço metodológico para essa investigação. Dada a importância de se verificar a invariância da medida ao se aplicar um teste em diferentes formatos, o objetivo deste capítulo é investigar, por meio de análise de DIF, se os itens de um teste educacional de larga escala mantêm suas propriedades psicométricas quando aplicados em papel ou em computador.

Todos os comandos e os dados utilizados nas análises estão disponíveis no repositório https://github.com/alexandrejaloto/DIF_avaliacao_educacional/. Para clonar o projeto no seu computador, use:

```
library(usethis)
create_from_github('alexandrejaloto/DIF_avaliacao_educacional',
'seu_diretorio_de_trabalho')
```

Funcionamento Diferencial do Item (DIF)

Diferenças entre grupos e entre indivíduos existem. Pessoas diferentes podem apresentar diferentes magnitudes de uma variável latente. Investigar essas diferenças é relevante, mas antes é preciso verificar se há invariância da medida. Isso significa que é preciso investigar se as diferenças entre grupos residem no nível da medida, ou seja, se elas se associam à produção da medida da variável latente (às respostas observadas aos itens do teste). Em outras palavras, é preciso se certificar de que o item funciona de maneira equivalente nos dois grupos.

Um item se comporta de forma diferente entre dois grupos quando ele apresenta propriedades psicométricas distintas entre esses grupos. Quando isso ocorre, tem-se o comportamento diferencial do item — DIF. Quando um item apresenta DIF, um sujeito pertencente a um grupo possui probabilidade de acerto diferente de um sujeito com mesmo nível de proficiência pertencente a um segundo grupo.

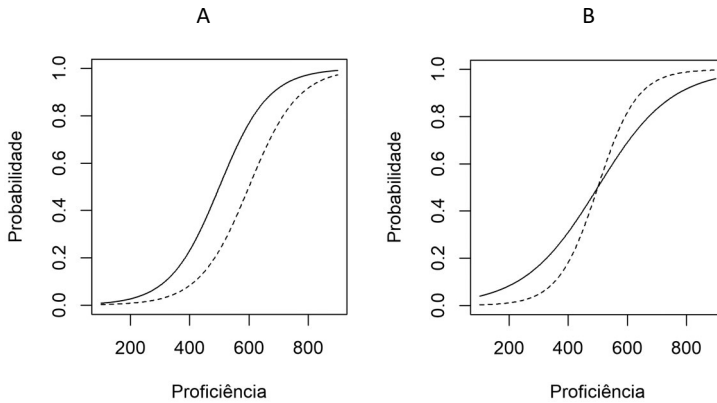
A título de ilustração, considere a comparação entre grupos que responderam a um mesmo teste em dois formatos: papel e computador. Tomemos dois sujeitos de mesma proficiência, mas que responderam ao teste em formatos diferentes. Se o que respondeu em papel tiver probabilidade maior de acertar determinado item do que aquele que respondeu no computador, temos um indicativo de DIF — e não de uma diferença real em proficiência. Nesse caso, o item se comporta diferente entre os dois grupos, e é mais fácil para o que respondeu em papel. É importante destacar que esse fenômeno é distinto das diferenças observadas entre médias

de grupos, pois, mesmo após se controlar a proficiência, a diferença de probabilidade de acerto entre os grupos persiste.

O DIF pode ser do tipo uniforme ou não uniforme. No caso de DIF uniforme, o item é sempre mais difícil para um grupo do que para o outro, independente da proficiência dos sujeitos. A figura 1 ilustra essa situação. Suponha que essa figura se refira a um item que se mostrou mais difícil para quem respondeu à prova em computador. Nesse caso, a curva tracejada representa esse grupo. No eixo x do gráfico, temos a proficiência (estimada por meio da Teoria de Resposta ao Item – TRI). Note que as curvas não se cruzam.

No caso de DIF não uniforme, o item apresenta valores diferentes de discriminação para os grupos. Como consequência, as curvas do item para cada grupo possuem inclinações diferentes. A figura 1 também apresenta uma ilustração de situação de DIF não uniforme. Nessa ilustração, o item é mais discriminativo para o grupo que respondeu no computador. É possível notar como a curva desse grupo é mais inclinada em relação à contínua, que representa o grupo que respondeu no papel. Repare que, para proficiências menores do que 500, o grupo do computador tem probabilidades menores de acertar o item. Essa diferença se inverte para proficiências maiores do que 500.

Figura 1 – Representação das curvas características de um item fictício com DIF uniforme e de outro com DIF não uniforme



Nota. (A) Representação de um item com DIF uniforme; (B) Representação de um item com DIF não uniforme. Linha contínua: grupo que respondeu no papel; linha tracejada: grupo que respondeu no computador.

Fonte: Elaboração própria.

Para lidar com a presença de DIF nos itens, duas soluções são comumente adotadas no arcabouço da TRI. A primeira é excluir o item com DIF. A exclusão do item elimina diretamente o possível viés, porém pode reduzir a cobertura do conteúdo do construto do teste e impactar na sua validade. A segunda solução é recalibrar os itens, permitindo que os parâmetros do item com DIF variem entre grupos. Essa solução preserva o item no teste, mas exige uma interpretação diferenciada dos escores para os grupos e pode impactar pontos de corte e normas baseadas em itens. Assim, ambas as estratégias têm vantagens e limitações que precisam ser cuidadosamente ponderadas na prática psicométrica.

Existem diversos métodos para análise de DIF, que podem ser estudados em literatura nacional e internacional (Damásio, 2013; Gamera; Gonçalves; Soares, 2015; Klein, 2025; Magis *et al.*, 2010; Rusch; Mair; Hatzinger, 2013). A título de ilustração, temos Análise Fatorial Confirmatória Multigrupo (AFCMG), *Multiple Indicators and Multiple Cause* (MI-MIC), Mantel-Hanszel, Raju e Regressão Logística. Este capítulo apresenta

uma análise de DIF com o método da Raiz do Desvio Quadrático Médio (RMSD; em inglês, *Root Mean Square Deviation*).

RMSD

O RMSD é uma estatística de ajuste que tem sido utilizada em avaliações internacionais de larga escala (como o *Programme for International Student Assessment, PISA*; e o *Trends in International Mathematics and Science Study, TIMSS*) para detectar DIF entre grupos (Fishbein; Yin; Bezirhan, 2024; OECD, 2024). Essencialmente, o RMSD mede a discrepância entre o que é *observado* e o que é *esperado* sob um modelo teórico. O método quantifica, em cada grupo, o grau de discrepância entre a curva característica do item (CCI) observada nos dados e a CCI obtida com os parâmetros do item. Quanto menor o valor do RMSD, melhor o ajuste, indicando que as respostas observadas estão alinhadas com a modelagem. Valores maiores de RMSD indicam maior desvio entre o comportamento do item no grupo e o que seria esperado de acordo com o modelo.

Formalmente, o RMSD é definido como:

$$\text{RMSD} = \sqrt{\int (P_o(\theta) - P_e(\theta))^2 f(\theta) d\theta}$$

onde $P_o(\theta)$ é a CCI observada, $P_e(\theta)$ é a CCI baseada no modelo e $f(\theta)d\theta$ é a função de densidade da distribuição de proficiência. Isso significa que o cálculo do valor de RMSD envolve a verificação da distância entre a curva observada e a esperada ponderada pela quantidade de observações nas regiões da escala de proficiência. A seguir, descrevemos como cada uma dessas três partes é obtida.

A CCI observada representa o desempenho dos examinandos em um item para diferentes níveis de proficiência. No entanto, diferentemente de uma simples proporção de acertos, a obtenção da CCI observada envolve um processo mais sofisticado que utiliza o algoritmo *Expectation-Maximization* (EM). A probabilidade observada de acerto é baseada nas pseudo-contagens do algoritmo EM (Bock; Aitkin, 1981).

Esse algoritmo itera duas etapas principais: Etapa E e Etapa M. Na Etapa E, o programa calcula a probabilidade de que cada padrão de resposta tenha sido gerado em cada ponto de proficiência (ponto de quadratura), dados os parâmetros provisórios dos itens. Com base nessas probabilidades, as frequências observadas dos padrões são redistribuídas pelos pontos de quadratura, produzindo o número esperado de sujeitos em cada ponto (pseudocontagens). Na Etapa M, o programa calcula, para cada item, o número esperado de acertos em cada ponto de quadratura e divide esses valores pelas pseudocontagens correspondentes, obtendo a probabilidade observada de acerto em cada ponto. No cálculo do RMSD, essa curva observada será comparada à CCI obtida pelo modelo.

A CCI baseada no modelo é obtida a partir da equação do modelo de TRI utilizado. Essa equação calcula a probabilidade de um sujeito acertar um item, dada a sua proficiência e os parâmetros desse item. Para itens dicotômicos, essa probabilidade é dada por:

$$P(x) = c + \frac{1 - c}{1 + e^{-a(\theta - b)}}$$

onde a é o parâmetro de discriminação, b é o parâmetro de dificuldade, c é o parâmetro de pseudochute, θ é a proficiência do sujeito e e é a constante de Euler (cerca de 2,72).

A função de densidade da distribuição de proficiência corresponde à distribuição prévia assumida para as habilidades dos examinandos. Normalmente utiliza-se uma distribuição normal padrão (média 0 e desvio padrão 1). Em linhas gerais, essa distribuição atua como uma ponderação inicial, representando a proporção plausível de sujeitos em cada região da escala de proficiência. Essa função de densidade é utilizada no cálculo do RMSD para ponderar a diferença entre a CCI observada e a CCI do modelo. Nas regiões da escala com mais sujeitos, o peso dessa diferença será maior e, nas regiões com menos sujeitos, ele será menor. Isso garante que o RMSD foque nas discrepâncias que são mais impactantes para o grupo em questão, pois diferenças nas regiões com mais sujeitos têm mais relevância para o cálculo.

O protocolo de análise de DIF com o método RMSD envolve, em linhas gerais, os seguintes passos: calibração dos itens em múltiplos grupos, sob a restrição de que os parâmetros dos itens sejam invariantes entre esses grupos. Em seguida, o valor de RMSD é calculado para cada item em cada grupo, quantificando o desajuste entre as curvas características observadas e as esperadas. Posteriormente, esses valores de RMSD são comparados com um ponto de corte predefinido, e os itens que apresentam valores acima desse limiar em um ou mais grupos são sinalizados com DIF. Por último, calibra-se um novo modelo, permitindo que os parâmetros desses itens sinalizados variem livremente entre os grupos, enquanto os demais itens mantêm suas restrições de invariância. O ciclo se repete até que nenhum item seja sinalizado com DIF.

O RMSD oferece algumas vantagens em relação a outros métodos de detecção de DIF. A primeira é que ele fornece uma medida quantitativa da magnitude do DIF, o que permite priorizar itens com viés mais significativo. A segunda é que ele é menos sensível ao tamanho da amostra do que outros testes de significância estatística, o que torna esse método útil em comparações entre grupos de tamanhos variados. Outra vantagem é a possibilidade de aplicá-lo a diferentes tipos de itens e modelos de TRI. Além disso, o RMSD está implementado em alguns pacotes em R, como *mirt*, *CDM* e *TAM*, o que amplia o acesso a esse método. Neste capítulo, utilizamos as funcionalidades disponíveis no pacote *mirt* para conduzir as análises.

Uma limitação do RMSD é a inexistência de um ponto de corte universal para sinalizar itens com DIF. O PISA adota 0,15 para itens cognitivos novos na edição e 0,20 para itens comuns entre edições na detecção de DIF entre países, nos testes de desempenho (OECD, 2024). Porém, existe recomendação de maior severidade e adoção de 0,10 como ponto de corte (Klein, 2025), valor adotado neste capítulo.

Este capítulo utiliza o método do RMSD para verificar a existência de DIF em itens do Enem ao se comparar a aplicação no papel e no computador. Antes, aplicamos o método a um banco simulado para fins de ilustração.

Análise de DIF em um banco simulado

A análise do banco simulado requer os pacotes `data.table` e `mirt`. Os comandos para carregar esses pacotes seguem:

```
# caso os pacotes não estejam instalados, é preciso instalá-los
# install.packages('data.table')
# install.packages('mirt')

# carregar os pacotes
library(data.table)
library(mirt)
```

O banco utilizado nesta ilustração é uma simulação das respostas de 2.000 pessoas (1.000 de cada grupo fictício) a 45 itens. O desenho da simulação prevê DIF para os itens 1 (uniforme) e 2 (não uniforme). O banco está disponível no repositório deste capítulo e pode ser baixado diretamente com o seguinte comando:

```
dados_sim <- fread('https://raw.githubusercontent.com/alexandrejaloto/DIF_avaliacao_educacional/refs/heads/master/dados/banco_simulado.csv')
```

O primeiro passo da análise é calibrar os itens, ou seja, estimar seus parâmetros no modelo de TRI. É nessa fase que o algoritmo EM entra em ação para estimar os parâmetros dos itens. A calibração é feita usando múltiplos grupos e sob restrições de invariância, ou seja, informa-se ao programa que os parâmetros dos itens devem ser invariantes (devem ser iguais) entre os grupos.

Inicialmente, gera-se uma tabela com as configurações iniciais de todos os parâmetros do modelo para cada item e grupo. Essa tabela não é requisito para a calibração dos itens nem para a execução do método RMSD. No entanto, neste capítulo, optamos por ajustar as distribuições prévias (*priors*) de alguns parâmetros do modelo para tentar aumentar a estabilidade do processo de estimação. A tabela com as configurações dos parâmetros permite a inspeção e a modificação manual das suas distribuições prévias.

```
tab.pars <- multipleGroup(dados_sim[,1:45], 1, group = dados_sim$grupo,  
  pars = 'values',  
  itemtype = '3PL', TOL = .01,  
  invariance = c('free_means', 'free_var',  
    colnames(dados_sim)[1:45]))
```

Os argumentos usados são esses:

- `dados_sim[,1:45]`: o banco de dados simulado, já selecionando somente as colunas de 1 a 45, que correspondem às respostas aos itens
- `1`: indica que o modelo é unidimensional
- `group = dados_sim$grupo`: especifica a variável do banco que define os dois grupos (G1 e G2)
- `pars = 'values'`: esse é o argumento responsável por gerar a tabela com a configuração dos parâmetros em vez de o programa ajustar o modelo
- `itemtype = '3PL'`: define o modelo da TRI (modelo logístico de três parâmetros, 3PL)
- `TOL = .01`: define a tolerância para a convergência do algoritmo EM
- `invariance = c('free_means', 'free_var', colnames(dados_sim)[1:45])`: define o que será invariante entre os grupos. 'free_means' fixa a média do Grupo 1 (G1) em zero e permite que ela varie livremente no Grupo 2 (G2). 'free_var' fixa a variância do G1 em 1,0 e permite que ela varie livremente no G2. `colnames(dados_sim)[1:45]` impõe uma invariância completa dos parâmetros dos itens entre os grupos. Ou seja, assume-se que os parâmetros de todos os itens são iguais para G1 e G2.

Para ajustar as distribuições prévias, é necessário alterar o objeto `tab.pars` (a tabela com a configuração dos parâmetros).

```

tab.pars[tab.pars$name == 'a1', 'prior.type'] <- 'lnorm'
tab.pars[tab.pars$name == 'a1', 'prior_1'] <- log(1.7)
tab.pars[tab.pars$name == 'a1', 'prior_2'] <- 0.5
tab.pars[tab.pars$name == 'a1', 'value'] <- 1.7
tab.pars[tab.pars$name == 'g', 'prior.type'] <- 'expbeta'
tab.pars[tab.pars$name == 'g', 'prior_1'] <- 5
tab.pars[tab.pars$name == 'g', 'prior_2'] <- 17
tab.pars[tab.pars$name == 'g', 'value'] <- .2

```

Para o parâmetro de discriminação (a), um prior lognormal é especificado ('lnorm') com média e desvio padrão definidos ($\log(1.7)$ e 0,5). Para o parâmetro de pseudochute (g, que é o c no modelo 3PL), um prior beta ('expbeta') é definido com seus parâmetros (5 e 17). Os valores iniciais dos parâmetros dos itens (*values*) também são ajustados. Com as configurações de invariância e os *priors* definidos no objeto `tab.pars`, o modelo de múltiplos grupos é finalmente ajustado.

```

fit1 <- multipleGroup(dados_sim[,1:45], 1, group = dados_sim$grupo,
  pars = tab.pars,
  itemtype = '3PL', TOL = .01,
  invariance = c('free_means', 'free_var',
    colnames(dados_sim)[1:45]))

```

Essa chamada de *multipleGroup* é idêntica à primeira, exceto que agora o argumento `pars = tab.pars` passa o objeto `tab.pars` com os *priors* configurados. O objeto `fit1` resultante contém o modelo calibrado sob a suposição de que nenhum item apresenta DIF, ou seja, todos os parâmetros dos itens são os mesmos entre G1 e G2. A título de curiosidade, os parâmetros dos itens podem ser obtidos com a função `coef`. A seguir, a chamada da função e as primeiras linhas do resultado para cada grupo. É importante notar que, como a invariância completa foi imposta, os parâmetros estimados para G2 são idênticos aos apresentados para o G1.

```
coef(fit1, IRTpars = TRUE, simplify = TRUE)
## $G1
## $items
##
##      a    b    g u
## Item_1 1.662 0.709 0.166 1
## Item_2 0.630 -1.815 0.193 1
## Item_3 1.580 0.450 0.173 1
## Item_4 1.352 0.558 0.179 1
## Item_5 2.083 -1.609 0.205 1
## Item_6 0.684 -0.663 0.193 1
##
## $G2
## $items
##
##      a    b    g u
## Item_1 1.662 0.709 0.166 1
## Item_2 0.630 -1.815 0.193 1
## Item_3 1.580 0.450 0.173 1
## Item_4 1.352 0.558 0.179 1
## Item_5 2.083 -1.609 0.205 1
## Item_6 0.684 -0.663 0.193 1
```

De posse do objeto `fit1` (modelo com os itens calibrados sob invariância completa entre os grupos), podemos aplicar a função `RMSD_DIF` para quantificar a discrepância de ajuste para cada item em cada grupo. A seguir, a chamada para a função e o resultado do RMSD para os seis primeiros itens em cada grupo.

```

dif.rmsd <- RMSD_DIF(fit1)
dif.rmsd
## $G1
##      P.0 P.1
## Item_1 0.188 0.188
## Item_2 0.138 0.138
## Item_3 0.016 0.016
## Item_4 0.021 0.021
## Item_5 0.009 0.009
## Item_6 0.037 0.037
##
## $G2
##      P.0 P.1
## Item_1 0.184 0.184
## Item_2 0.124 0.124
## Item_3 0.016 0.016
## Item_4 0.023 0.023
## Item_5 0.010 0.010
## Item_6 0.026 0.026

```

Observa-se que os valores de RMSD dos itens 1 e 2 são maiores do que o valor de corte (0,10), o que indica a existência de DIF para esses itens. Cabe ressaltar que, no caso de itens dicotômicos, as colunas P.0 e P.1 terão o mesmo valor, então é indiferente qual delas utilizar. Além disso, nesta análise, os valores de RMSD dos Itens 1 e 2 para ambos os grupos superaram o limite estabelecido. Porém, ainda que o valor de somente um grupo ultrapasse o critério, marca-se o item como indicativo de DIF.

Quando um item possui parâmetros psicométricos intrinsecamente diferentes entre os grupos (devido ao DIF uniforme ou não uniforme), mas o modelo é calibrado sob a restrição de que esses parâmetros são os mesmos para todos os grupos, espera-se que ele apresente um ajuste inadequado (maior discrepância entre as CCI's observada e esperada) para, pelo menos, um grupo. Dessa forma, os valores de RMSD elevados para os Itens 1 e 2 servem como um indicativo da presença do DIF. Cabe

destacar que, ainda que não se utilizasse um ponto de corte universal para o RMSD, os valores para os Itens 1 e 2 são notavelmente mais elevados quando comparados aos demais itens, tanto para o G1 quanto para o G2, o que reforça o indicativo de DIF.

Outra forma de obter os valores de RMSD é já indicar, na função, o ponto de corte para que o programa retorne somente os valores acima desse ponto.

```
dif.rmsd <- RMSD_DIF(fit1, flag = .1)
dif.rmsd
## $G1
## P.0 P.1
## Item_1 0.188 0.188
## Item_2 0.138 0.138
## Item_3 NA NA
## Item_4 NA NA
## Item_5 NA NA
## Item_6 NA NA
##
## $G2
## P.0 P.1
## Item_1 0.184 0.184
## Item_2 0.124 0.124
## Item_3 NA NA
## Item_4 NA NA
## Item_5 NA NA
## Item_6 NA NA
```

Após a detecção dos itens com DIF, o próximo passo da análise é permitir que os parâmetros desses itens variem livremente entre os grupos. Isso é feito ajustando um novo modelo de múltiplos grupos em que as restrições de invariância são relaxadas apenas para esses itens. A calibração dos demais itens (sem DIF) permanece sob a restrição de invariância, garantindo a escala comum.

No exemplo simulado, os Itens 1 e 2 apresentaram DIF. Portanto, calibraremos um novo modelo em que todos os itens, exceto 1 e 2, terão seus parâmetros forçados a serem iguais entre os grupos. Para os Itens 1 e 2, os parâmetros serão estimados independentemente para G1 e G2. O comando para essa calibração é o seguinte:

```
fit2 <- multipleGroup(dados_sim[,1:45], 1, group = dados_sim$grupo,
  pars = tab.pars,
  itemtype = '3PL', TOL = .01,
  invariance = c('free_means', 'free_var',
    colnames(dados_sim)[3:45]))
```

A diferença dessa chamada da função *multipleGroup* para a anterior, quando o objeto *fit1* foi criado, é que o argumento *invariance* foi alterado para *colnames(dados_sim)[3:45]*. Isso significa que a invariância de parâmetros será imposta do Item 3 ao Item 45, ou seja, todos os itens, exceto os que apresentaram DIF. Dessa forma, o modelo *fit2* irá estimar parâmetros de item separados para G1 e G2 apenas para o Item 1 e o Item 2, enquanto os parâmetros dos demais 43 itens permanecerão iguais entre os grupos.

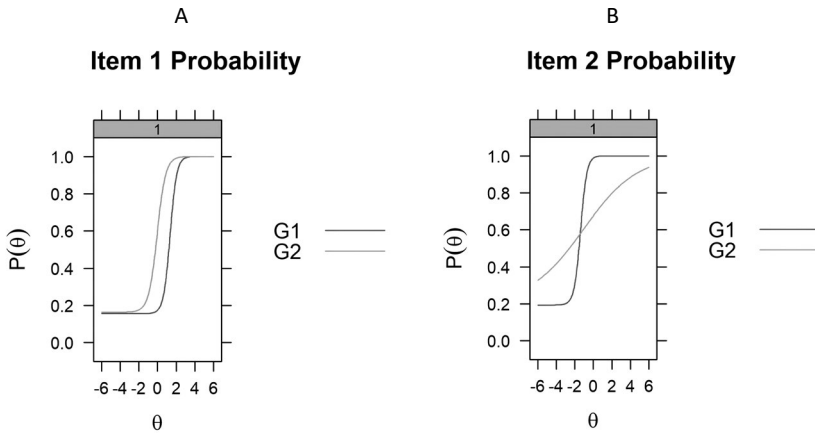
O objeto *fit2* agora contém o modelo calibrado com a suposição de que o DIF pode ocorrer nos Itens 1 e 2. Para examinar diretamente como os parâmetros desses itens diferem entre os grupos, podemos inspecionar os coeficientes do modelo. Os parâmetros dos Itens 1 e 2 devem ser diferentes entre os grupos, porém os dos demais itens devem ser iguais. A seguir, a chamada da função e as primeiras linhas do resultado para cada grupo:

```
coef(fit2, IRTpars = TRUE, simplify = TRUE)
## $G1
## $items
##      a    b    g u
## Item_1 2.842 1.346 0.156 1
## Item_2 2.943 -1.399 0.194 1
## Item_3 1.566 0.455 0.172 1
## Item_4 1.337 0.562 0.178 1
##
## $G2
## $items
##      a    b    g u
## Item_1 2.172 -0.051 0.164 1
## Item_2 0.348 -1.057 0.207 1
## Item_3 1.566 0.455 0.172 1
## Item_4 1.337 0.562 0.178 1
```

A função `itemplot` do pacote `mirt` gera as CCIs para os itens de interesse. Essa função permite visualizar como a probabilidade de acerto de um item varia em função da proficiência para cada grupo, revelando visualmente as discrepâncias identificadas numericamente. A figura 2 mostra as CCIs dos dois itens com DIF geradas a partir do comando a seguir.

```
itemplot(fit2, 1)
itemplot(fit2, 2)
```

Figura 2 – Curvas características dos Itens 1 e 2 após calibração do modelo com parâmetros livres entre os grupos



Nota. (A) Curva característica do Item 1 (DIF uniforme); (B) Curva característica do Item 2 (DIF não uniforme).

Fonte: Elaboração própria.

Ao observar o gráfico do Item 1, é possível notar que as curvas características para o G1 e para o G2 estão separadas verticalmente ao longo da escala de proficiência e não se cruzam. Essa configuração visual é uma indicação de DIF uniforme. A curva do G1 está consistentemente **à direita** da curva do G2 — ou seja, para qualquer nível de proficiência, o item é mais difícil para o G1 e os sujeitos desse grupo têm uma probabilidade de acerto menor do que os sujeitos do G2. Isso é reflexo das diferenças nos parâmetros a e b desse item para cada grupo. O parâmetro a é semelhante entre os grupos, porém o parâmetro b é maior para o G1 (a diferença é de cerca de 1,4).

No gráfico do Item 2, as curvas características para o G1 e para o G2 apresentam inclinações visivelmente diferentes e se cruzam em algum ponto da escala de proficiência. Isso é uma indicação de DIF não uniforme. A diferença nas inclinações reflete uma diferença nos parâmetros de discriminação (a), em que o item discrimina de maneira distinta entre os grupos. Como esse parâmetro é substancialmente menor para o G2

(0,348) em comparação com o G1 (2,943), a curva do Item 2 para o G2 é consideravelmente menos íngreme. A consequência, no caso desse item, é que, para proficiências muito baixas, o G2 tem maior probabilidade de acerto, o que se inverte a partir de proficiências próximas de $-2,0$.

A inspeção gráfica reforça a necessidade de se calibrarem esses dois itens livremente entre os grupos. Após a nova calibração, é necessário confirmar se os valores de RMSD estão abaixo do valor crítico.

```
dif.rmsd2 <- RMSD_DIF(fit2, flag = .1)
dif.rmsd2
## $G1
##      P.0 P.1
## Item_1 NA NA
## Item_2 NA NA
## Item_3 NA NA
## Item_4 NA NA
## Item_5 NA NA
## Item_6 NA NA
##
## $G2
##      P.0 P.1
## Item_1 NA NA
## Item_2 NA NA
## Item_3 NA NA
## Item_4 NA NA
## Item_5 NA NA
## Item_6 NA NA
```

Todos os valores de RMSD estão abaixo do valor crítico para ambos os grupos, dada a ausência de valores numéricos e a presença de NA para todos os itens quando $\text{flag} = .1$ é utilizado. Isso indica que todos os RMSD calculados foram menores ou iguais ao limiar estabelecido. A conclusão é que não há mais evidência de DIF para os itens. Na próxima seção, apresentamos a análise de DIF no Enem 2021 entre os sujeitos que responderam em papel e os que responderam no computador.

Análise de DIF para o formato de aplicação do Enem

O Exame Nacional do Ensino Médio (Enem) é uma avaliação de larga escala composta por uma redação e quatro provas de 45 itens de múltipla escolha: Ciências Humanas e suas Tecnologias (CH); Ciências da Natureza e suas Tecnologias (CN); Linguagens, Códigos e suas Tecnologias (LC); e Matemática e suas Tecnologias (MT). O exame é utilizado principalmente como processo de seleção para ingresso em cursos de Ensino Superior.

Tradicionalmente, o Enem é aplicado em formato impresso. No entanto, em 2020, 2021 e 2022, o Enem também foi aplicado em versão digital para um número limitado de participantes. A comparação entre o desempenho dos estudantes nesses dois formatos de aplicação é de interesse para a psicometria, pois pode revelar a presença de DIF associado ao formato do teste. Se confirmado, esse DIF revelaria que as diferenças de desempenho estão associadas ao método de aplicação — e não à proficiência dos estudantes. Adicionalmente, essa comparação é socialmente relevante, pois, caso a existência de DIF se confirme, o uso da medida assumindo sua invariância pode acarretar injustiças na seleção para o Ensino Superior. Este capítulo investiga o DIF em itens da prova de CN do Enem 2021, comparando as respostas de estudantes que realizaram o exame nos formatos papel e computador.

Os dados utilizados neste estudo são secundários e foram extraídos em 01/11/2022 dos microdados do Enem, disponibilizados no portal do INEP.³ Para compor a amostra, selecionamos os sujeitos presentes na prova de CN ($TP_PRESENCIA_CN == 1$) que declararam estar concluindo o Ensino Médio em 2021 ($TP_ST_CONCLUSAO == 2$) no ensino regular ($TP_ENSINO == 1$), possuir pelo menos 17 anos no final de 2021 ($TP_FAIXA_ETARIA >= 2$) e que responderam à prova regular sem adaptações (para o formato em papel, os códigos de 909 a 912; para o formato digital, os códigos de 1011 a 1014). Após aplicação do filtro pelos critérios descritos, restaram 799.587 sujeitos que

3 INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados**. [s.l.: s.n.]. Disponível em: <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/>. Acesso em: 14 mar. 2023.

responderam à prova em papel e 15.421 que responderam em computador. Para otimizar o tempo de processamento da análise, selecionamos aleatoriamente 15.421 sujeitos que responderam à prova em papel. A amostra utilizada na análise está disponível no repositório deste capítulo, bem como os comandos para gerá-la. A variável grupo no banco de dados diferencia esses grupos: 'DIG' para os que responderam no computador (digital) e 'REG' para os que responderam no papel (regular).

Para investigar a presença de DIF, empregou-se o método RMSD. O primeiro passo envolveu a calibração de um modelo de TRI (neste caso, 3PL) sob a restrição de invariância completa dos itens entre os grupos (formato papel e formato computador). Os comandos R utilizados para carregar os pacotes para as análises, importar os dados, configurar os *priors* e calibrar o modelo são apresentados a seguir:

```
library(data.table)
library(mirt)
dados_enem <- fread('https://raw.githubusercontent.com/alexandrejaloto/DIF_ava-
liacao_educacional/refs/heads/master/dados/banco_enem.csv')
tab.pars <- multipleGroup(dados_enem[,1:45], 1,
  group = dados_enem$grupo,
  pars = 'values',
  itemtype = '3PL',
  invariance = c('free_mean',
    'free_var',
    colnames(dados_enem)[1:45]))
tab.pars[tab.pars$name == 'a1', 'prior.type'] <- 'lnorm'
tab.pars[tab.pars$name == 'a1', 'prior_1'] <- log(1.7)
tab.pars[tab.pars$name == 'a1', 'prior_2'] <- 0.5
tab.pars[tab.pars$name == 'a1', 'value'] <- 1.7
tab.pars[tab.pars$name == 'g', 'prior.type'] <- 'expbeta'
tab.pars[tab.pars$name == 'g', 'prior_1'] <- 5
tab.pars[tab.pars$name == 'g', 'prior_2'] <- 17
tab.pars[tab.pars$name == 'g', 'value'] <- .2
fit1 <- multipleGroup(dados_enem[,1:45],
```

```

1, group = dados_enem$grupo,
pars = tab.pars,
itemtype = '3PL',
invariance = c('free_mean',
               'free_var',
               colnames(dados_enem)[1:45]),
TOL = .01)

```

Com o modelo fit1 calibrado sob a suposição de ausência de DIF, a função RMSD_DIF foi aplicada para calcular os valores de RMSD para cada item em cada grupo, utilizando um ponto de corte de 0.10 para sinalizar itens com potencial DIF. A seguir, a chamada da função e as primeiras linhas do resultado:

```

dif.rmsd <- RMSD_DIF(fit1, flag = .1)
dif.rmsd
## $DIG
##      P.0 P.1
## X117867 NA NA
## X63095  NA NA
## X68355  NA NA
## X85781  NA NA
## X117843 NA NA
##
## $REG
##      P.0 P.1
## X117867 NA NA
## X63095  NA NA
## X68355  NA NA
## X85781  NA NA
## X117843 NA NA

```

A inspeção dos resultados armazenados no objeto dif.rmsd revelou que nenhum item de CN apresentou valores de RMSD acima do ponto de corte de 0.10, tanto para o grupo que realizou a prova em papel quanto

para o grupo que a realizou em computador. Isso é inferido pela ausência de valores numéricos e presença de NA para todos os itens quando flag = .1 é utilizado, indicando que todos os RMSD calculados foram menores ou iguais ao limiar estabelecido.

Esse achado sugere que os itens se comportam de maneira psicometricamente equivalente entre os dois formatos de aplicação (papel e computador), considerando o limiar de RMSD adotado. Em outras palavras, não foi detectado DIF relevante relacionado ao formato de aplicação da prova, indicando que as diferenças observadas entre os grupos, se existirem, residem predominantemente no nível de proficiência dos estudantes, e não em um viés sistemático dos itens associado ao formato do teste.

Considerações finais

Este capítulo explorou o DIF em testes educacionais, destacando a relevância desse fenômeno para a garantia da validade e equidade das avaliações. A presença de DIF, especialmente entre modalidades de aplicação (papel *versus* computador), tem implicações significativas para a interpretação dos resultados dos testes. Como apresentado, um item que funciona diferentemente entre grupos devido ao formato de aplicação compromete a validade das inferências feitas a partir dos escores obtidos (*American Educational Research Association; American Psychological Association; National Council On Measurement In Education, 2014*).

Embora diferenças entre as modalidades papel e computador tenham se mostrado invariantes em alguns contextos (Schroeders; Wilhelm, 2011; Terluin *et al.*, 2018; Yu; Yu, 2007), aplicações em papel ou em computador podem não apenas refletir diferenças técnicas, mas também afetar o desempenho real dos estudantes. Por exemplo, realizar cálculos matemáticos utilizando lápis e papel é uma atividade qualitativamente distinta da realização desses mesmos cálculos em ambiente digital (Borba; Scucuglia; Gadanidis, 2014; Gu; Drake; Wolfe, 2006). Da mesma forma, testes que avaliam atenção ou inteligência podem gerar desempenhos distintos quando aplicados em papel ou em dispositivos como *tablets*, pois fatores como navegação na interface, tamanho da tela e interação com o dispositivo podem

influenciar os resultados obtidos, uma vez que o tempo de resposta é também avaliado nesses testes (Latendorf *et al.*, 2021; Tosi, 2022).

Além disso, a multiplicidade de dispositivos utilizados em aplicações digitais, como computadores, *notebooks*, *tablets* e *smartphones*, adiciona uma complexidade à análise do DIF. Diferenças nos sistemas operacionais (por exemplo, *Android*, *iOS*, *Windows*) e nos navegadores também podem interferir diretamente na interação dos estudantes com os itens e, conseqüentemente, em seu desempenho (Bridgeman; Lennon; Jackenthal, 2003). Portanto, pesquisadores(as) e aplicadores(as) precisam estar atentos não somente às diferenças entre papel e computador, mas também às nuances que surgem da diversidade tecnológica contemporânea.

Este capítulo contém algumas limitações que devem ser reconhecidas. Primeiramente, as análises conduzidas foram ilustrativas e limitadas a exemplos específicos (dados simulados e Enem 2021), o que pode não representar plenamente outros contextos educacionais ou culturais. Além disso, o método RMSD, apesar das suas vantagens de detectar e quantificar DIF, carece de um ponto de corte universalmente aceito, o que demanda cautela nas decisões tomadas com base nessa análise. Em resumo, é essencial que pesquisadores(as) e demais profissionais da área educacional continuem atentos(as) à detecção e manejo do DIF para garantir que os testes educacionais cumpram seu papel com justiça e validade em contextos cada vez mais diversos e tecnologicamente variados.

Declaração de conflito de interesses

As opiniões expressas nesta publicação são de exclusiva e integral responsabilidade dos autores, não necessariamente expressando o ponto de vista do INEP ou do Ministério da Educação do Brasil.

Referências

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; AMERICAN PSYCHOLOGICAL ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. **Standards for educational and psychological testing**. Washington: American Educational Research Association, 2014.

BOCK, R. D.; AITKIN, M. Marginal Maximum Likelihood estimation of item parameters: application of an EM algorithm. **Psychometrika**, Marginal Maximum Likelihood Estimation of Item Parameters, v. 46, n. 4, p. 443–459, dez. 1981.

BORBA, M. C.; SCUCUGLIA, R. R. S.; GADANIDIS, G. **Fases das tecnologias digitais em educação matemática**: sala de aula e internet em movimento. Belo Horizonte: Autêntica, 2014.

BRIDGEMAN, B.; LENNON, M. L.; JACKENTHAL, A. Effects of screen size, screen resolution, and display rate on computer-based test performance. **Applied Measurement in Education**, [S. l.], v. 16, n. 3, p. 191–205, jul. 2003.

DAMÁSIO, B. F. Contribuições da Análise Fatorial Confirmatória Multigrupo (AF-CMG) na avaliação de invariância de instrumentos psicométricos. **Psico-USF**, Bragança Paulista, v. 18, n. 2, p. 211–220, 2013. Disponível em: <https://www.scielo.br/j/psuf/a/8cvmNWqFB6XqWdV9ZGJzsLF/?lang=pt>. Acesso em: 25 nov. 2025.

FISHBEIN, B.; YIN, L.; BEZIRHAN, U. **TIMSS 2023 achievement scaling implementation**. [S. l.]: TIMSS & PIRLS International Study Center, Boston College, 2024. Disponível em: <https://timss2023.org/methods/chapter-12>. Acesso em: 21 jun. 2025.

GAMERMAN, D.; GONÇALVES, F. B.; SOARES, T. M. Differential item functioning. In: VAN DER LINDEN, W. J. (org.). **Handbook of item response theory**. Boca Raton: CRC Press, 2015. v. 3p. 67–86.

GU, L.; DRAKE, S.; WOLFE, E. W. Differential item functioning of GRE mathematics items across computerized and paper-and-pencil testing media. **The Journal of Technology, Learning and Assessment**, [S. l.], v. 5, n. 4, 2006. Disponível em: <https://ejournals.bc.edu/index.php/jtla/article/view/1643>. Acesso em: 27 jul. 2025.

KLEIN, R. Medidas de qualidade de ajuste e de DIF (differential item functioning). **Ensaio: Avaliação e Políticas Públicas em Educação**, [S. l.], v. 33, n. 126, p. 1–25, mar. 2025. Disponível em: <https://www.scielo.br/j/ensaio/a/yehY8ZFm3CXdg9b-587wjy6g/?lang=pt>. Acesso em: 15 jun. 2025.

LATENDORF, A. *et al.* Digitization of neuropsychological diagnostics: a pilot study to compare three paper-based and digitized cognitive assessments. **Agng Clinical and Experimental Research**, Digitization of neuropsychological diagnostics, [S. l.], v. 33, n. 6, p. 1585–1597, jun. 2021.

LUBKE, G. H. *et al.* On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. **Intelligence**, [S. l.], v. 31, n. 6, p. 543–566, dez. 2003.

MAGIS, D. *et al.* A general framework and an R package for the detection of dichotomous differential item functioning. **Behavior Research Methods**, [S. l.], v. 42, n. 3, p. 847–862, 2010. Disponível em: <https://link.springer.com/article/10.3758/BRM.42.3.847>. Acesso em: 25 nov. 2025.

OECD. **PISA 2022 Technical Report**. Paris: OECD, 2024. Disponível em: https://www.oecd.org/en/publications/pisa-2022-technical-report_01820d6d-en.html. Acesso em: 15 jun. 2025.

RUSCH, T.; MAIR, P.; HATZINGER, R. *et al.* **Psychometrics with R: a review of CRAN packages for item response theory**. Vienna: Center for Empirical Research Methods, 2013. Disponível em: <https://epubdev.wu.ac.at/4010/1/resre-pIRThandbook.pdf>. Acesso em: 29 fev. 2020.

SCHROEDERS, U.; WILHELM, O. Equivalence of reading and listening comprehension across test media. **Educational and Psychological Measurement**, [S. l.], v. 71, n. 5, p. 849–869, out. 2011.

TERLUIN, B. *et al.* Assessing the equivalence of web-based and paper-and-pencil questionnaires using differential item and test functioning (DIF and DTF) analysis: a case of the Four-Dimensional Symptom Questionnaire (4DSQ). **Quality of Life Research**, [S. l.], v. 27, n. 5, p. 1191–1200, mai. 2018. Disponível em: <https://link.springer.com/article/10.1007/s11136-018-1816-5>. Acesso em: 10 jun. 2025.

TOSI, C. M. G. **Avaliação neuropsicológica remota das funções executivas em adultos com e sem TDAH: evidências de validade do Stroop app**. 2022. 112 f. Dissertação (mestrado em Psicologia) — Universidade Federal de Santa Catarina, Florianópolis, 2022. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/240949/PPSI1003-D.pdf?sequence=-1&isAllowed=y>. Acesso em: 10 jun. 2025.

YU, S.-C.; YU, M.-N. Comparison of internet-based and paper-based questionnaires in Taiwan using multisample invariance approach. **CyberPsychology & Behavior**, [S. l.], v. 10, n. 4, p. 501–507, ago. 2007.

CAPÍTULO 11

AVALIAÇÃO EM LARGA ESCALA E DESEMPENHO ESTUDANTIL: UMA ANÁLISE COMPARATIVA ENTRE SRES CAPIXABAS NO PAEBES (2021-2024)

Lara Nunes Barbosa¹

Ronildo Stieg²

Wagner dos Santos³

DOI: 10.52695/978-65-5456-163-1.11

Introdução

O processo de avaliação constitui um componente fundamental para o aprimoramento do sistema educacional e para a promoção da aprendizagem dos estudantes. Sua função extrapola a simples verificação de resultados, contribuindo significativamente para o desenvolvimento de competências, bem como para o fortalecimento da equidade no ambiente escolar. Conforme apontam Libâneo e Freitas (2018), avaliar é uma atividade complexa que envolve a produção de dados tanto quantitativos como qualitativos, permitindo compreender de forma mais ampla o desempenho discente.

1 Universidade Federal do Espírito Santo.

2 Universidade Federal do Espírito Santo.

3 Universidade Federal do Espírito Santo. E-mail: wagner.santos@ufes.br.

Nesse contexto, destacam-se as avaliações externas, caracterizadas por serem elaboradas, coordenadas e aplicadas por agentes e instituições que não integram diretamente a rotina das escolas. Em virtude de sua amplitude e abrangência, essas avaliações são também denominadas avaliações em larga escala. Sua finalidade principal é oferecer um panorama do funcionamento das redes de ensino, permitindo a identificação de desafios e potencialidades, bem como subsidiar a formulação e o redirecionamento de políticas públicas educacionais fundamentadas em evidências concretas.

Avaliações em larga escala, portanto, são instrumentos avaliativos que têm como um dos principais objetivos medir o conhecimento e a capacidade dos alunos de acordo com o nível escolar que estão. Esses exames são considerados como instrumentos reguladores desenvolvidos por órgãos governamentais e não governamentais (Lopes; Stieg; Santos, 2024). Nos últimos anos, essas avaliações têm tido um protagonismo que vai além dos âmbitos acadêmicos, inserindo-se no âmbito político como forma de melhoria da qualidade educacional. Cronbach (1963) define a avaliação como compilação e uso de informação para a tomada de decisões. Nesse sentido, os exames standardizados são modelos de avaliações em que os resultados obtidos podem ser utilizados para o desenvolvimento ou análise de políticas educacionais no âmbito escolar e comunitário.

Nesse contexto, o que define esses exames são as decisões que são tomadas a partir dos resultados. Eles são divididos em duas classificações: *high stakes* (alto risco/efeito) e *low stake* (baixo risco/efeito) (Marques; Stieg; Santos, 2020; Lopes; Stieg; Santos, 2024). De acordo com esses autores, os exames de alto risco são aqueles que impactam diretamente na vida do estudante ou docentes, como, por exemplo, a atribuição de notas, conclusão do Ensino Médio ou o ingresso no Ensino Superior. Já os exames de baixo risco não afetam diretamente a vida estudantil, profissionais ou instituições. Eles podem ser realizados como avaliações diagnósticas, porém geram informações significativas para ações de políticas educacionais (Marques; Stieg; Santos, 2020).

Orfield e Wald (2000) explicam que esses dois efeitos estão relacionados à forma como os resultados dessas avaliações são aplicados, seja

para promoção ou certificação de alunos. Esse fator inclui as bonificações de professores, como ocorre em diferentes estados brasileiros cujas políticas de avaliação têm incentivado essas práticas (Santos *et al.*, 2021).

As tomadas de decisões a partir dessas avaliações possuem uma influência do Movimento de Educacional Global (da sigla em inglês GERM), que surgiu na conjuntura internacional, desde 2001, com o objetivo de reforçar reformas educacionais que promovam e fortaleçam o discurso da eficiência dos sistemas (Oliveira; Canuto; Lemos, 2023). Com isso, de acordo com Carrasco e Gunter (2019), as ideias de medidas para reivindicações e tomadas de decisões educacionais foram espalhadas em vários países, promulgando uma mudança na educação, incluindo a interpretação de contas baseadas em testes padronizados e a competição entre escolas devido a essas interpretações. Alguns dos exemplos das tomadas de decisões a partir dos resultados desses exames são a certificação de estudantes, bonificação aos docentes e classificações das escolas, gerando, como consequência, as competições escolares.

No âmbito da rede estadual de ensino do Espírito Santo, a participação em avaliações em larga escala ocorre em diferentes esferas: internacional, nacional e estadual. No que se refere à esfera estadual, a coordenação e a regulamentação desses processos são de responsabilidade da Secretaria de Estado da Educação (SEDU), conforme determinado por diversos instrumentos normativos. Entre os principais documentos que orientam essa política avaliativa, destacam-se: a) Portaria nº 064-R, de 24 de maio de 2017; b) Portaria nº 168-R, de 23 de dezembro de 2020; c) Portaria nº 138-R, de 23 de junho de 2022a; d) Portaria nº 139-R, de 23 de junho de 2022b; e) Portaria nº 181-R, de 09 de agosto de 2022c; f) Portaria nº 003-R, de 13 de janeiro de 2023a; g) Portaria nº 021-R, de 30 de janeiro de 2023b.

Esses dispositivos legais delinham os parâmetros da política de avaliação em larga escala no estado, garantindo sua institucionalização e orientando sua implementação nas escolas da rede pública capixaba.

Desse modo, uma hipótese a ser confirmada neste estudo é de que as notas obtidas por meio dos exames standardizados são diferentes entre as diferentes regiões do estado do Espírito Santo. Tendo em vista essa

questão, as análises presentes neste texto buscam estudar sobre essas variações entre as localidades. De modo específico, objetivamos identificar, por meio das notas do Programa de Avaliação da Educação Básica do Espírito Santo (Paebes), especialmente no que se refere ao componente curricular de matemática, quais as regiões que obtiveram déficit maior nos índices de acertos, quais alcançaram um resultado melhor e quais as influências que podem estar por trás desses resultados. Com base nesses questionamentos, estabelecemos uma análise das médias obtidas no transcurso dos últimos quatro anos (2021–2024).

O Paebes foi criado em 2000 e é destinado aos 5º e 9º anos do Ensino Fundamental e 3ª série do Ensino Médio. O principal objetivo é analisar se os estudantes desenvolveram ou não as habilidades vinculadas à matriz de referência da rede do estado do Espírito Santo (SEDU).

Inicialmente, o Paebes foi criado com a classificação de baixo efeito para a população alvo (*low stake*), considerando que seus resultados não apresentavam, de forma clara, os efeitos de aprimoramento sobre as práticas educativas dos profissionais da educação. Contudo, com a criação do Indicador de Desenvolvimento das Escolas Estaduais do Espírito Santo (IDE), que é utilizado como política de bonificação para professores e servidores administrativos associados à Secretaria de Estado da Educação (SEDU), essa avaliação teve uma mudança em seus efeitos, passando a ser considerada como um exame de alto risco (*high stakes*) (Soares; Soares; Santos, 2022).

No entanto, essa avaliação traz implicações diretas sobre o trabalho docente, uma vez que os professores são frequentemente considerados os principais responsáveis pelo desempenho dos estudantes. Segundo Lima e Fassarella (2022), as avaliações externas podem ser empregadas como instrumentos de controle do governo sobre o trabalho dos professores. Nesse contexto, observa-se um controle por parte da direção escolar para que os docentes sejam responsáveis por solucionar os problemas relacionados ao baixo desempenho dos alunos. No entanto, desconsidera-se a limitação do tempo disponível para desenvolver o trabalho dos conteúdos previsto na matriz referencial, o que leva muitos professores a recorrerem a práticas pedagógicas tradicionais (Lima; Fassarella, 2022).

O cálculo da nota dos estudantes nesse exame tem como base o modelo psicométrico: Teoria de Resposta ao Item (TRI). Esse modelo tem como objetivo mostrar a probabilidade do estudante de responder corretamente à questão, seus conhecimentos a respeito da área que está sendo avaliada e as características dos itens (Pasquali, 2017). O TRI não calcula a nota apenas pela quantidade de questões respondidas corretamente, mas, sim, pela particularidade de cada questão (Rocha Junior, 2018).

Portanto, essa pesquisa se justifica pela necessidade da continuação do estudo sobre o tema, tendo em vista que já foram realizados estudos anteriores derivados dos projetos guarda-chuvas *Avaliação estandarizada nos estados do Espírito Santo/Brasil e Baja California/México: dilemas e tensões do Paebes e do Exeims-BC*, Edital Universal FAPES processo nº 11367/2021 e *Dos exames estandarizados ao direito à aprendizagem: experiências para a melhoria da qualidade do Ensino Médio na Rede Estadual do Espírito Santo*, Edital Universal FAPES nº 28/2022 (Dutra; Negreiros; Santos, 2024; Lopes; Stieg; Santos, 2024; Loureiro et al., 2024; Santos et al., 2024; Lopes et al., 2025).

Reconhecendo a importância dos exames estandarizados para realizações de melhorias educacionais, o presente trabalho busca contribuir com esses estudos. Nesse sentido, o objetivo desta pesquisa consiste em analisar os índices de desempenhos dos estudantes em matemática nas provas do Paebes nas 11 Superintendências Regionais de Educação (SREs) do estado do Espírito Santo, comparando esse desempenho entre 2021-2024. Além disso, a proposta consiste em identificar quais SRE têm apresentado um déficit maior nos índices de acerto no Paebes e em qual/quais seria um ponto importante para realizar uma pesquisa-ação-intervenção (Chizzotti, 2014), oferecendo um mapa de SREs e municípios nos quais deveriam ser realizadas políticas de intervenção.

Metodologia

De natureza mista, combinando abordagens qualitativas e quantitativas (Silveira; Córdova, 2009), esta pesquisa fundamenta-se na análise crítico-documental (Bloch, 2001) e no método comparado em educação

(Ferrer, 2002). As principais fontes de dados foram disponibilizadas pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd), vinculado à Universidade Federal de Juiz de Fora (UFJF), que reúne os resultados das provas do Paebes aplicadas aos estudantes do 3º ano do Ensino Médio, com ênfase especial nas questões de matemática.

Para obtenção dos dados, foram consultadas informações disponíveis no Painel da Gerência de Avaliação (GEA), uma plataforma pública mantida pelo governo do Estado do Espírito Santo que reúne dados das diferentes avaliações de larga escala ofertadas no âmbito da Sedu, incluindo do Paebes referentes aos anos de 2021 a 2024. Esse sistema permite o acesso a médias de desempenho por região, conforme recorte de ano/série, disciplina e localização geográfica. A partir dessa base de dados, os resultados foram sistematizados em tabelas do *Excel* e, posteriormente, organizados no gráfico 1.

Com o intuito de aprofundar a análise dos dados obtidos, foram elaboradas três figuras que sintetizam e ilustram aspectos relevantes do desempenho dos estudantes no Paebes. A figura 1 apresenta os níveis de desempenho em matemática, permitindo uma visualização da distribuição dos estudantes nos diferentes patamares de proficiência. A figura 2 retrata a distribuição dos padrões de desempenho no Paebes, evidenciando como os resultados se organizam em relação às categorias definidas pela avaliação. Por fim, a figura 3 reúne os indicadores de desempenho da proficiência média da rede estadual, possibilitando uma análise comparativa ao longo dos anos e entre diferentes regiões. Essas representações gráficas foram construídas com base nos dados sistematizados anteriormente e contribuem significativamente para a compreensão dos resultados e suas implicações no contexto educacional capixaba.

Além disso, o gráfico 2, mesmo não tendo sido elaborado pelos autores deste texto, apresenta o percentual de docentes da Educação Básica na rede estadual segundo três dimensões: situação funcional, regime de contratação e tipo de vínculo, no Brasil, no ano de 2024. Essa representação permite observar a composição do corpo docente da rede pública estadual, evidenciando a proporção de professores efetivos, temporários e com outras formas de vínculo, bem como a predominância de

determinados regimes de contratação. Os dados contribuem para compreender a estrutura e o perfil docente e suas possíveis implicações na continuidade das políticas educacionais, na estabilidade profissional e na qualidade do ensino ofertado.

Em todas as ilustrações apresentadas e com base na aplicação do método comparado, buscou-se identificar elementos passíveis, ou não, de comparação, bem como evidenciar resultados que revelam aproximações ou divergências entre os dados (Ferrer, 2002). As informações coletadas foram organizadas e interpretadas por meio de tabelas e gráficos, com o apoio dos *softwares Microsoft Word e Excel*. Essa sistematização permitiu analisar as semelhanças e as diferenças entre as médias regionais, contribuindo diretamente para a condução e o aprofundamento das análises e para a discussão dos resultados obtidos.

A análise seguiu os princípios da abordagem crítico-documental (Bloch, 2001), envolvendo a organização, o registro e a interpretação sistemática das fontes. Um aspecto relevante para essa análise foi a estrutura das questões da avaliação, uma vez que o Paebes apresenta a matemática como um componente curricular específico, o que exige atenção à forma como os conteúdos e habilidades são distribuídos e cobrados nos instrumentos avaliativos.

Resultados e discussão

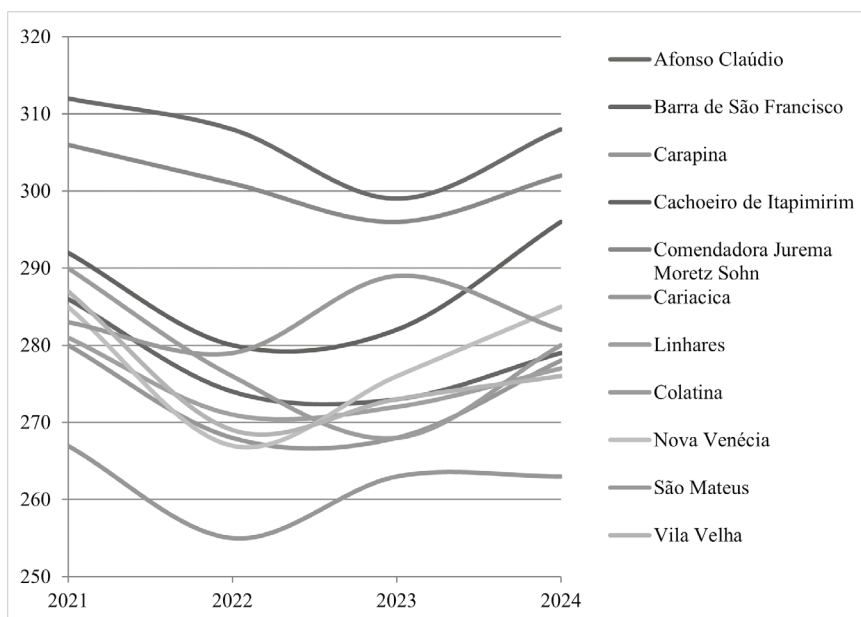
As avaliações padronizadas, como é o caso do Paebes, desempenham um papel relevante na análise do desenvolvimento dos estudantes, com destaque para o componente curricular de matemática, frequentemente priorizado nesses exames. Essa ênfase possibilita a realização de estudos mais aprofundados com base nos resultados obtidos nas provas, contribuindo para a compreensão do desempenho discente nessa área do conhecimento.

O Paebes é uma avaliação externa aplicada exclusivamente no estado do Espírito Santo, sem abrangência em nível nacional. Trata-se de um exame que não possui implicações diretas para os estudantes, ou seja, seu desempenho individual não é utilizado para compor a nota trimestral ou

gerar qualquer tipo de consequência acadêmica direta. Seu principal propósito é avaliar o desempenho das redes públicas estadual e municipal, gerando dados que possam orientar a formulação, o monitoramento e a reformulação de políticas educacionais. Dessa forma, busca-se minimizar a interferência de motivações individuais dos alunos sobre os resultados, preservando o caráter diagnóstico e institucional da avaliação.

No primeiro movimento de comparação, foi elaborado o gráfico 1, que apresenta as 11 SREs e as respectivas médias de desempenho por ano, no período de 2021 a 2024. O objetivo desse gráfico é possibilitar uma análise comparativa dos resultados da avaliação nos anos finais do Ensino Médio, considerando as diferentes regiões do estado do Espírito Santo. Essa visualização permite identificar tendências, aproximações e disparidades no desempenho entre as SREs ao longo do tempo, contribuindo para o aprofundamento da análise proposta neste estudo.

Gráfico 1 – Médias do Paebes por SRE do Espírito Santo (2021-2024)



Fonte: Elaboração própria.

A partir do gráfico 1, é possível observar o comportamento das médias de desempenho em matemática obtidas pelos estudantes do 3º ano do Ensino Médio nas 11 SREs do Espírito Santo, no período de 2021 a 2024. A análise evidencia dinâmicas importantes que ajudam a compreender tanto as desigualdades regionais como os possíveis efeitos de políticas educacionais no estado.

De maneira geral, percebe-se uma tendência comum entre as regiões: queda nos desempenhos entre os anos de 2021 e 2022, seguida por uma recuperação progressiva em 2023 e 2024. Esse padrão pode estar relacionado aos efeitos do período pós-pandêmico, com o retorno das aulas presenciais e o processo de recomposição das aprendizagens, especialmente em componentes curriculares considerados sensíveis, como a matemática.

Algumas regiões se destacam positivamente ao longo do período. É o caso da SRE de Afonso Cláudio, que apresentou os melhores desempenhos médios em todos os anos analisados, encerrando 2024 com média acima de 310 pontos. Também se sobressaem Comendadora Jurema Moretz Sohn (Vila Velha) e Cachoeiro de Itapemirim, que demonstraram crescimento consistente nos últimos dois anos da série, aproximando-se da média de 300 pontos.

Em contrapartida, regiões como Cariacica apresentaram desempenhos significativamente inferiores ao longo do período analisado, com médias que chegaram próximas a 255 pontos em 2022, e apenas uma leve recuperação nos anos seguintes. A persistência desses baixos indicadores sugere a presença de desafios estruturais e pedagógicos complexos, demandando atenção especial das políticas públicas educacionais. Curiosamente, segundo dados do IBGE de 2021, os municípios pertencentes a essa SRE não apresentavam situação socioeconômica desfavorável, indicando que fatores além do contexto econômico, como resistência dos alunos ou da comunidade escolar ao processo avaliativo, podem influenciar esses resultados. Essa hipótese ganha suporte na análise de Marques *et al.* (2024), que ressaltam que avaliações como o ENEM apresentam mais adesão e motivação dos estudantes por seu caráter decisivo no ingresso ao Ensino Superior, diferentemente do Paebes, cuja importância percebida pelos alunos pode ser menor, afetando a participação e o desempenho.

Outras SREs, como Linhares, Nova Venécia, Colatina e Carapina, oscilaram entre médias intermediárias, situadas em torno de 270 a 280 pontos. Nessas regiões, os resultados sugerem relativa estabilidade, embora com variações pontuais que também devem ser analisadas em conjunto com fatores locais.

Ao analisar mais detidamente o desempenho destacado da SRE de Afonso Cláudio, nota-se que seu resultado discrepante pode estar relacionado a fatores estruturais e contextuais. Localizada no interior do estado, essa SRE conta com menor presença de zonas periféricas e um número reduzido de escolas, apenas cinco estabelecimentos de Ensino Médio em 2023, conforme dados do IBGE. Além disso, abrange municípios, como Santa Maria de Jetibá, que apresentam uma boa situação econômica, com PIB *per capita* em 2021 de R\$ 39.228,24, ocupando o décimo lugar no estado. Esses aspectos econômicos e demográficos parecem favorecer o desempenho escolar local, indicando a influência do contexto socioeconômico no processo educacional.

Outro ponto relevante refere-se à SRE Comendadora Jurema Moretz Sohn, que obteve bons resultados em todos os anos, ficando apenas abaixo da SRE de Afonso Cláudio. Diferentemente da relação tradicional entre condições socioeconômicas e desempenho, esse caso pode estar associado ao significado atribuído à avaliação Paebes pela comunidade escolar. Conforme apontam Soares, Soares e Santos (2022), a criação do Índice de Desenvolvimento da Educação (IDE) e sua vinculação ao cálculo do bônus pago aos servidores da Secretaria de Educação do Espírito Santo (SEDU) resultou em uma mudança significativa na percepção da avaliação, que passou de um exame de baixo risco para uma ferramenta de alto risco para professores e gestores. Essa responsabilização dos resultados nas mãos dos docentes promoveu negociações e tensões no cotidiano escolar, reforçando a ideia de que o Paebes, embora não seja uma preocupação direta dos alunos, reflete a atuação e o engajamento dos professores.

Em contraste, a SRE de Cariacica, como já mencionado, obteve resultados inferiores ao longo de todo o período, o que não pode ser diretamente atribuído a fatores socioeconômicos, dado o contexto favorável

do município segundo o IBGE. A resistência dos alunos e a percepção reduzida da importância do Paebes podem ser fatores explicativos para esse desempenho, reforçando as conclusões de Marques, Stieg e Santos (2020) e Lopes, Stieg e Santos (2024), que apontam que avaliações como o Paebes funcionam mais como instrumentos de responsabilização docente, ao passo que outras, como o ENEM, exercem mais influência motivacional sobre os estudantes.

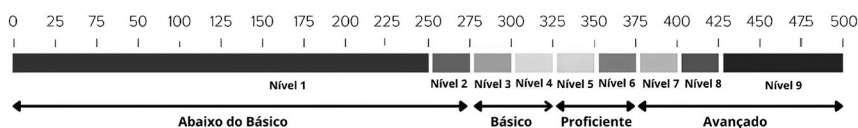
Outro aspecto destacado pela análise do gráfico 1 é o comportamento da SRE de São Mateus, que, ao contrário das demais, não seguiu o padrão de queda seguida por recuperação. Essa superintendência apresentou um aumento nas médias em 2022 e 2023, seguido por uma queda em 2024.

Esse declínio recente pode estar associado ao processo de adaptação à reforma do Ensino Médio, cujas dificuldades são apontadas por Zluhan, Santos e Corrêa (2024). Entre as fragilidades desse novo modelo, destacam-se a redução da carga horária da Base Nacional Comum Curricular (Brasil, 2018), disparidades na implementação dos itinerários formativos entre escolas públicas e privadas, insuficiência e superficialidade dos materiais didáticos, falta de recursos pedagógicos, formação docente inadequada, infraestrutura escolar precária, dificuldades dos alunos na transição entre itinerários, problemas decorrentes de transferências e alocação insuficiente de recursos financeiros (Zluhan; Santos; Corrêa, 2024). Esses desafios podem ter impactado negativamente o desempenho dos estudantes, evidenciando a complexidade das mudanças educacionais em curso.

Essa leitura comparativa dos dados permite levantar hipóteses sobre as desigualdades educacionais regionais e os efeitos das estratégias de gestão e políticas públicas voltadas à melhoria da aprendizagem. A retomada do crescimento das médias em 2023 e 2024 pode estar associada a investimentos em formação docente, recomposição curricular e ações de acompanhamento pedagógico implementadas em nível estadual. No entanto, as disparidades persistentes entre as SREs indicam a necessidade de intervenções mais localizadas e equitativas, voltadas especialmente para as regiões que enfrentam mais fragilidades no desempenho escolar.

No âmbito do Paebes, os resultados dos estudantes são apresentados em níveis de desempenho, que indicam o grau de apropriação dos conhecimentos e habilidades esperados para a etapa avaliada. A pontuação varia de 0 a 500 pontos, distribuída em nove níveis organizados em uma escala de proficiência. Cada nível descreve um conjunto específico de competências e habilidades que os estudantes foram capazes de demonstrar, e esses níveis são agrupados em quatro grandes faixas interpretativas: Abaixo do Básico, Básico, Proficiente e Avançado.

Figura 1 – Régua de níveis de desempenho em matemática no Paebes



Fonte: Elaboração própria.

A régua de níveis de desempenho em matemática (figura 1) é uma ferramenta gráfica e descritiva que auxilia na interpretação dos resultados obtidos por estudantes em avaliações padronizadas, como o próprio Paebes. Através dessa régua, torna-se possível identificar os diferentes níveis de aprendizagem alcançados, o que contribui para diagnósticos pedagógicos mais precisos, comparações entre grupos e o monitoramento da qualidade educacional ao longo do tempo.

A faixa Abaixo do Básico, que compreende os níveis 1 e 2 (até aproximadamente 275 pontos), representa estudantes com domínio incipiente dos conteúdos e habilidades previstos para sua etapa de escolarização. Esses alunos, geralmente, apresentam dificuldades para realizar operações matemáticas simples, interpretar problemas ou aplicar conceitos em contextos cotidianos. Trata-se de um grupo que exige atenção pedagógica urgente, pois pode estar acumulando lacunas significativas em sua trajetória escolar. A amplitude dessa faixa evidencia que uma parcela considerável dos estudantes pode se encontrar em situação de vulnerabilidade educacional, o que demanda políticas públicas focadas em recomposição das aprendizagens.

A faixa Básico, que abrange os níveis 3 e 4 (entre 275 e 325 pontos), inclui alunos que apresentam domínio parcial das habilidades esperadas. Embora tenham alguma compreensão dos conceitos matemáticos, ainda demonstram fragilidades ao aplicar conhecimentos de forma autônoma e em situações mais complexas. O desafio, nesse caso, consiste em consolidar as aprendizagens fundamentais e promover condições para que esses estudantes avancem em direção à proficiência.

Já a faixa Proficiente, formada pelos níveis 5 e 6 (325 a 375 pontos), representa a meta desejável para os sistemas de ensino. Os estudantes que se encontram nesse patamar demonstram domínio sólido dos conteúdos curriculares, são capazes de resolver problemas com raciocínio lógico e aplicar os conhecimentos em diferentes contextos. Estar na faixa da proficiência indica que o estudante possui condições para prosseguir com sucesso sua trajetória educacional. Portanto, ampliar a proporção de alunos nesse grupo é uma prioridade estratégica para as políticas de qualidade na Educação Básica.

A faixa Avançado, composta pelos níveis 7 a 9 (acima de 375 pontos), corresponde a estudantes com domínio aprofundado e ampliado dos conhecimentos matemáticos. São capazes de lidar com situações-problema mais complexas, operar com abstrações e empregar diferentes representações matemáticas com fluência. Apesar de numericamente menos representativo, esse grupo evidencia o potencial de excelência presente na rede pública, apontando para a necessidade de estratégias que promovam enriquecimento curricular e estimulem o desenvolvimento de altas habilidades.

A régua de níveis de desempenho é, portanto, um recurso importante para a leitura pedagógica dos resultados das avaliações em larga escala. Ela permite uma análise mais qualificada do desempenho médio de escolas, municípios ou regiões, indo além de médias aritméticas que podem mascarar realidades diversas. Por exemplo, uma SRE pode apresentar média de 290 pontos, o que a situaria no nível Básico, mas concentrar a maioria dos alunos no nível 2. Tal cenário exigiria intervenções pedagógicas específicas, mesmo quando a média geral sugira um desempenho mais favorável.

Adicionalmente, a régua permite identificar desigualdades educacionais e monitorar as trajetórias de aprendizagem em diferentes contextos socioterritoriais. Como ressaltam Werle (2011) e Falabella (2023), a simples comparação de médias entre redes ou regiões tende a ocultar a diversidade de perfis e os desafios específicos enfrentados por certos grupos de estudantes. A desagregação dos dados por nível de desempenho, portanto, favorece a formulação de políticas educacionais mais equânimes, que reconheçam os diferentes pontos de partida dos alunos e garantam o direito à aprendizagem para todos (Afonso, 2009).

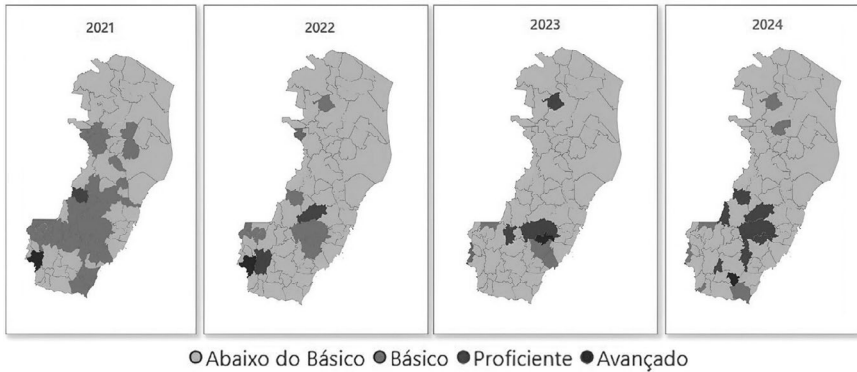
Nesse sentido, como mostra a figura 1, os níveis 1 e 2 pertencem à faixa Abaixo do Básico; os níveis 3 e 4 são classificados como Básico; os níveis 5 e 6 compõem a faixa Proficiente; e os níveis 7 a 9 correspondem à faixa Avançado. Cada nível contempla um conjunto específico de habilidades. Por exemplo, no nível 1, uma das 16 habilidades avaliadas é *Reconhecer a planificação usual do cubo a partir de seu nome*. Com base nessa estrutura, as equipes pedagógicas e os gestores escolares podem analisar com mais precisão o desenvolvimento dos estudantes, identificando habilidades que necessitam de mais ênfase, reconhecendo as principais dificuldades enfrentadas e planejando ações pedagógicas alinhadas às evidências diagnósticas. Os dados também podem ser utilizados para orientar a formação docente contínua e embasar o planejamento do trabalho pedagógico.

É importante destacar que o uso pedagógico da régua deve ser acompanhado por uma formação docente consistente, por ações de planejamento escolar baseadas em evidências e por um acompanhamento sistemático das aprendizagens. Mais do que um instrumento de ranqueamento, a régua de desempenho deve ser compreendida como uma ferramenta de diagnóstico e promoção da equidade. Em um contexto educacional pós-pandêmico, marcado por mais heterogeneidade nas turmas, instrumentos como esse ganham ainda mais relevância na tarefa de reconstruir trajetórias educacionais interrompidas ou fragilizadas.

Analisando a Distribuição de Padrões de Desempenhos no Paebes em todos os municípios do Espírito Santo, observamos algumas variações. Os municípios estão divididos em cores que determinam as classificações:

Abaixo do Básico, Básico, Proficiente e Avançado. Os mapas a seguir mostram de forma detalhada essas oscilações.

Figura 2 – Distribuição dos Padrões de Desempenho no Paebes



Fonte: Adaptado do Painel GEA.

A figura 2 retrata a evolução da distribuição geográfica dos padrões de desempenho em matemática dos estudantes do 3º ano do Ensino Médio, com base nos resultados do Paebes, no período de 2021 a 2024. As diferentes tonalidades no mapa do Espírito Santo representam as faixas de desempenho classificadas como Abaixo do Básico (cinza claro), Básico (azul claro), Proficiente (azul médio) e Avançado (azul escuro), de acordo com a régua de níveis adotada pela avaliação.

A análise comparativa entre os quatro anos revela dinâmicas significativas de mudança na proficiência média dos municípios, refletindo tanto o impacto das políticas educacionais quanto os efeitos do contexto pandêmico e pós-pandêmico sobre o desempenho escolar.

Em 2021, observa-se uma prevalência de tonalidades mais claras, especialmente em regiões centrais e do norte do estado, indicando ampla concentração de municípios com desempenho Abaixo do Básico e Básico. Esse padrão pode ser interpretado como um reflexo direto das restrições pedagógicas impostas pela pandemia da Covid-19, com forte comprometimento das aprendizagens, sobretudo em áreas com menor estrutura tecnológica e dificuldades de acesso ao ensino remoto. Além disso, nesse

ano, o mapa apresentou um município como proficiente e um como avançado. Esses dois municípios são, respectivamente, Laranja da Terra e Guaçuí, pertencentes às duas SREs que melhor se desenvolveram no Paebes, como mostrado no gráfico 1. Ademais, os poucos que estão categorizados como básico também pertencem a essas SREs.

No ano seguinte, 2022, percebe-se uma intensificação das regiões em azul escuro, indicando que algumas poucas SREs obtiveram avanço significativo, atingindo o padrão Avançado, enquanto outras permaneceram nos níveis mais baixos. Isso indicia (Ginzburg, 1989) uma ampliação das desigualdades educacionais regionais, pois as melhorias não foram distribuídas de forma equitativa. Municípios localizados ao sul e sudoeste, por exemplo, apresentam destaque positivo, enquanto o norte do estado permanece com desempenho predominantemente frágil. Santa Maria de Jetibá, pertencente à SRE de Afonso Cláudio, por exemplo, passou de Básico para Proficiente e Alegre, da SRE Comendadora Jurema Moretz Sohn, teve um bom avanço e passou de Abaixo do Básico para Proficiente.

Os mapas da figura 2 mostram que há um leve aumento na tonalidade azul médio (Proficiente), especialmente em áreas antes marcadas por baixos níveis, o que sinaliza uma tendência de recomposição das aprendizagens em algumas regiões. Isso pode ser atribuído à retomada das atividades presenciais, aos investimentos em formação continuada, à recomposição curricular e às ações direcionadas pelas superintendências regionais e escolas. Por outro lado, os mapas evidenciam também uma regressão ainda maior do desempenho estudantil em matemática. É o caso de Guaçuí: continuou classificado como Abaixo do Básico, sendo que, em 2021 e 2022, alcançou o nível Avançado. Poucos evoluíram como Conceição do Castelo, Domingos Martins, Marechal Floriano e Vila Pavão.

Em contrapartida, em 2024, a distribuição espacial evidencia avanços mais consistentes na proficiência, com maior presença de municípios nas faixas Proficiente e Avançado. Observa-se também um padrão de dispersão mais equilibrado, indicando que os efeitos das políticas públicas de reforço, formação docente e acompanhamento pedagógico começam a se consolidar em escala mais ampla. Entretanto, ainda persistem bolsões de baixo desempenho, sobretudo em regiões de mais vulnerabilidade social.

A análise longitudinal da figura 2 oferece pistas de que o sistema educacional capixaba respondeu progressivamente aos efeitos adversos da pandemia, apresentando melhorias contínuas, mas com fortes indícios de desigualdade regional. A permanência de regiões em faixas mais baixas sinaliza que a recuperação das aprendizagens não ocorreu de maneira uniforme, exigindo políticas de intervenção territorializada e equitativa.

Autores como Santos *et al.* (2021) argumentam que avaliações em larga escala, como o Paebes, quando bem utilizadas, oferecem subsídios importantes para o diagnóstico e a formulação de ações pedagógicas baseadas em evidências. No entanto, sua eficácia depende da articulação entre análise dos dados e ação pedagógica concreta, o que requer formação docente continuada, apoio técnico às escolas e políticas de valorização do magistério (Soares *et al.*, 2023).

Além disso, como alertam Lopes *et al.* (2025), os efeitos das avaliações também estão vinculados aos mecanismos de responsabilização institucional, como bonificações por desempenho, que podem induzir tanto à mobilização pedagógica quanto a práticas de ensino voltadas exclusivamente para os testes (*teaching to the test*). Assim, os avanços observados em 2024 devem ser analisados também à luz de como as escolas e SREs vêm internalizando os objetivos das avaliações, e não apenas como um indicativo direto da qualidade do ensino.

A figura 3 apresenta a evolução do desempenho médio dos estudantes do 3º ano do Ensino Médio no Paebes, entre os anos de 2023 e 2024, nas disciplinas de língua portuguesa e matemática. A variação positiva em ambas as áreas do conhecimento aponta para avanços significativos no processo de recomposição das aprendizagens e na atuação das políticas públicas de gestão educacional no Espírito Santo.

Figura 3 – Indicadores de desempenho da Proficiência Média da Sedu



Fonte: Adaptado do Painel GEA.

De acordo com a figura 3, na disciplina de língua portuguesa, a média passou de 273 pontos, em 2023, para 284 pontos, em 2024, representando um acréscimo de 11 pontos. Em matemática, a média evoluiu de 276 para 284 pontos no mesmo período, o que corresponde a um ganho de 8 pontos. Embora ainda situadas na faixa classificada como Básico, conforme os níveis de proficiência definidos pelo Paebes, essas médias se aproximam gradualmente do nível Proficiente, cuja pontuação se inicia em 325 pontos. Esse crescimento, ainda que modesto em termos absolutos, tem relevância significativa quando analisado no contexto de um sistema público de ensino, marcado por históricos desafios estruturais e desigualdades regionais.

A melhora nos resultados pode ser atribuída a um conjunto de fatores. Em primeiro lugar, destaca-se a implementação de políticas de recomposição das aprendizagens, intensificadas após o retorno às aulas presenciais no período pós-pandêmico. Autores como Ruiz Martins e Mendes Sant’Ana (2024) enfatizam a importância de medidas como tutoria pedagógica, adaptação curricular, reforço escolar e monitoramento frequente das trajetórias de aprendizagem. Tais ações, quando bem articuladas com o trabalho docente, contribuem para reduzir as lacunas acumuladas durante o ensino remoto e favorecer a retomada do processo educativo.

Além disso, é possível que esse avanço esteja associado ao fortalecimento da cultura de avaliação e ao maior engajamento das escolas com os resultados do Paebes. Em algumas regiões, como demonstrado em

estudos de Soares, Soares e Santos (2022), a utilização do Índice de Desempenho Escolar (IDE) como critério para bonificação por desempenho provocou uma mudança de percepção sobre o exame, que passou a ser encarado como instrumento de alto risco e de impacto direto sobre a rotina de professores e gestores. Esse cenário contribui para um maior alinhamento entre os objetivos avaliativos e as práticas pedagógicas, ainda que tal responsabilização demande reflexões críticas quanto aos seus efeitos sobre o trabalho docente.

Por outro lado, autores como Lopes, Stieg e Santos (2024) alertam para os limites das políticas de bonificação baseadas em resultados, destacando os riscos de intensificação do trabalho, padronização curricular e exclusão de sujeitos com trajetórias escolares mais vulneráveis. Assim, é necessário que os avanços observados sejam acompanhados de políticas educacionais mais equitativas e sustentadas por diagnósticos precisos da realidade de cada escola.

Outro aspecto a considerar é a possível estabilização do novo Ensino Médio, cujas mudanças estruturais vêm sendo gradualmente assimiladas pelas redes de ensino. Conforme apontado por Zluhan, Santos e Corrêa (2024), os primeiros anos da implementação do novo modelo revelaram inúmeras fragilidades, como a redução da carga horária da Base Nacional Comum, a desigualdade no acesso aos itinerários formativos, a falta de materiais pedagógicos adequados e a insuficiência de formação docente. No entanto, com o tempo, parte dessas dificuldades pode ter sido mitigada, permitindo uma reorganização do trabalho pedagógico.

Apesar dos resultados positivos, é importante ressaltar que as médias apresentadas ainda indicam que a maior parte dos estudantes se encontra em níveis abaixo da proficiência esperada para o final da Educação Básica. Ou seja, ainda há um longo caminho a ser percorrido para que os estudantes desenvolvam plenamente as competências e habilidades previstas nas diretrizes curriculares. Como defendem Afonso (2011) e Falabella (2023), a leitura agregada das médias pode ocultar desigualdades internas significativas entre escolas e grupos sociais. Por isso, a interpretação dos dados deve considerar a desagregação por níveis de desempenho, possibilitando uma análise mais precisa das necessidades pedagógicas.

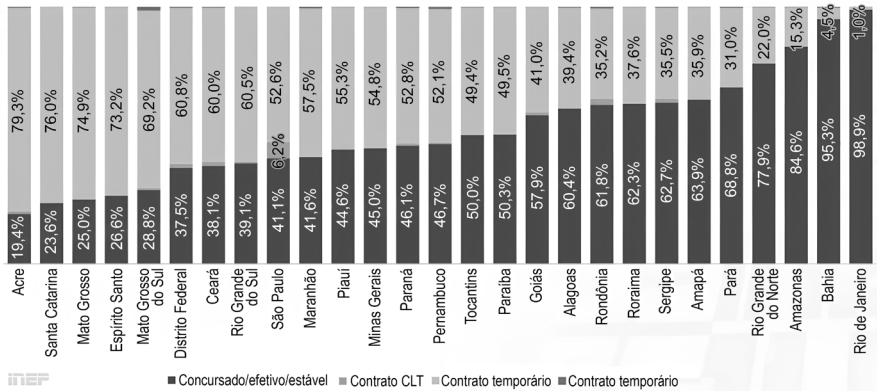
No entanto, o elevado número de municípios classificados na faixa Abaixo do Básico, bem como a tímida evolução observada em outros, pode refletir fatores que envolvem tanto os estudantes quanto os professores. Uma das possíveis causas desses resultados está relacionada à diminuição do interesse dos alunos pelos estudos (Mattei; Cunha, 2021). A escolarização, tradicionalmente entendida como um ponto de partida para a conquista de oportunidades profissionais e culturais, vem sendo ressignificada por muitos jovens.

Segundo Mattei e Cunha (2021), é cada vez mais comum que estudantes enfrentem a difícil tarefa de conciliar a rotina escolar com o trabalho, impulsionados por necessidades socioeconômicas. Além disso, muitos desejam conquistar autonomia financeira e ingressar precocemente na vida adulta como uma forma de reconhecimento e respeito social, optando, por exemplo, por atuar como jovens aprendizes. Nesse contexto, a escola passa a ser percebida, por alguns, como um entrave, o que pode comprometer o engajamento e, conseqüentemente, o desempenho acadêmico.

Ademais, outro fator que pode estar diretamente relacionado aos resultados insatisfatórios no desempenho escolar é o elevado número de contratações temporárias de docentes. A alta rotatividade dificulta a construção de projetos pedagógicos consistentes de médio e longo prazo, comprometendo a continuidade das ações educativas e o acompanhamento sistemático das aprendizagens. Nesse cenário, torna-se desafiador utilizar os resultados de avaliações externas, como o Enem, para promover melhorias efetivas no processo de ensino-aprendizagem, uma vez que o tempo reduzido de permanência dos professores nas escolas limita a elaboração e implementação de planejamentos pedagógicos mais robustos.

O Censo Escolar de 2024 evidencia essa realidade: no Espírito Santo, o número de professores contratados temporariamente supera o de docentes efetivos, revelando um quadro preocupante de instabilidade nas equipes pedagógicas. Para ilustrar essa situação, apresentamos o gráfico 2, que explicita a distribuição entre vínculos efetivos e temporários na rede estadual (INEP, 2024).

Gráfico 2 – Percentual de docentes da Educação Básica nas redes estaduais, por situação funcional, regime de contratação ou tipo de vínculo; Brasil



Fonte: INEP (2024).

O gráfico 2 evidencia uma expressiva disparidade entre os estados brasileiros no que se refere à forma de contratação dos professores da Educação Básica nas redes estaduais. Os vínculos contratuais foram agrupados em quatro categorias: concursado/efetivo/estável, contrato CLT, contrato temporário e contrato terceirizado.

No caso específico do Espírito Santo, apenas 26,6% dos professores são concursados ou possuem estabilidade, enquanto 73,2% estão contratados de forma temporária, segundo os dados apresentados. O estado ocupa, assim, a quarta posição entre aqueles com menor proporção de docentes efetivos. Em contraste, estados como Rio de Janeiro (98,9%), Bahia (95,3%) e Amazonas (84,6%) apresentam altos índices de professores concursados, o que, à primeira vista, sugere mais estabilidade nas redes de ensino, favorecendo a consolidação de projetos pedagógicos de médio e longo prazo.

Esse contexto de instabilidade contratual pode comprometer negativamente o trabalho pedagógico desenvolvido nas escolas, refletindo diretamente nos resultados de avaliações em larga escala, como o Paebes. A contratação massiva por tempo determinado tende a gerar alta rotatividade

docente, dificultando o acompanhamento contínuo das aprendizagens e a efetiva implementação de políticas educacionais. Como apontam Soares e Riani (2024), a rotatividade constante prejudica o vínculo entre professores e comunidade escolar, afetando o clima institucional e comprometendo a construção de uma cultura pedagógica sólida.

Além disso, professores temporários costumam ter menor acesso à formação continuada, a cargos de gestão e a incentivos financeiros, o que pode impactar diretamente em sua motivação e engajamento com o processo de ensino-aprendizagem (Soares; Riani, 2024).

Quando se relacionam esses índices de contratação com os dados de desempenho apresentados no gráfico 1, surgem indícios (Ginzburg, 1989) que sugerem uma possível relação entre a baixa rotatividade docente e os bons resultados obtidos pela SRE de Afonso Cláudio. Ainda que essa região possa contar com um número elevado de contratações temporárias, o reduzido número de estabelecimentos escolares nas áreas do interior favorece a permanência dos mesmos docentes nas mesmas instituições ao longo do tempo, fortalecendo vínculos pedagógicos e institucionais.

Por outro lado, superintendências como a de Cariacica, que possivelmente concentram maior número de escolas e de professores transitando entre unidades, enfrentam mais desafios para garantir estabilidade nas equipes pedagógicas. Essa fragmentação pode impactar negativamente a continuidade dos processos educativos e, conseqüentemente, os resultados de aprendizagem.

Considerações finais

A análise dos dados do Paebes entre os anos de 2021 e 2024 revelou um panorama do desempenho dos estudantes capixabas, bem como os desafios estruturais e pedagógicos que permeiam a rede estadual de ensino. Embora algumas superintendências, como a de Afonso Cláudio e Comendadora Jurema Moretz Sohn, tenham apresentado desempenhos significativamente superiores, esse destaque parece estar associado a uma combinação de fatores, como organização da rede, menor número

de escolas, condições socioeconômicas mais estáveis e, possivelmente, menor rotatividade docente.

Por outro lado, regiões como Cariacica permaneceram com resultados aquém do esperado, mesmo não apresentando, em princípio, vulnerabilidades econômicas graves. Isso aponta para a necessidade de se considerarem variáveis mais complexas e multidimensionais, como o envolvimento dos estudantes com a avaliação, o grau de valorização da escola em relação ao Paebes e a estabilidade das equipes pedagógicas. Em alguns contextos, a responsabilização recai fortemente sobre os docentes, o que pode gerar tensões, desmotivação e distorções no uso pedagógico da avaliação.

Além disso, os mapas de distribuição de desempenho revelaram que boa parte dos municípios ainda se encontra no nível Abaixo do Básico, indicando que, embora haja avanços pontuais, ainda não se observa uma transformação substancial do panorama educacional. Os resultados médios, mesmo em crescimento, ainda ocultam desigualdades profundas entre as regiões.

Esse cenário demanda um olhar mais atento para políticas públicas estruturantes que integrem: a) a valorização docente (com ênfase na estabilidade profissional e acesso à formação continuada); b) o acompanhamento pedagógico orientado por evidências; e c) a criação de ambientes escolares que promovam o engajamento estudantil de maneira significativa. É preciso reconhecer que os desafios vão além das avaliações; elas envolvem as condições concretas de trabalho e estudo, a organização curricular e as trajetórias juvenis marcadas por demandas socioeconômicas, como a necessidade de inserção precoce no mercado de trabalho.

Diante disso, futuras investigações poderiam aprofundar a escuta de professores e estudantes para compreender as percepções sobre o Paebes e os fatores que interferem no desempenho. Além disso, recomenda-se a análise das ações formativas promovidas pela Secretaria de Educação, bem como dos efeitos da reforma do Ensino Médio na organização curricular e no engajamento dos jovens.

A avaliação em larga escala, como o Paebes, pode e deve ser um instrumento para a melhoria da educação pública. No entanto, isso só será

possível se ela for compreendida como parte de um processo mais amplo, que articule diagnóstico, ação pedagógica, justiça social e compromisso político com o direito à aprendizagem de todos.

Referências

AFONSO, A. J. Políticas avaliativas e accountability em educação: subsídios para um debate ibero-americano. **Revista de Ciências da Educação**, Lisboa, v. 9, p. 57–69, 2009. Disponível em: <https://publicacoes.fcc.org.br/eaee/article/view/9562>. Acesso em: 26 nov. 2025.

BLOCH, M. **Apologia da história**: ou o ofício de historiador. Rio de Janeiro: Jorge Zahar, 2001.

CARRASCO, A.; GUNTER, H. M. O “privado” na privatização das escolas: o caso do Chile. **Revista Revisão Educacional**, [S. l.], v. 71, n. 1, p. 67–80, 2019. DOI: <https://doi.org/10.1080/00131911.2019.1522035>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00131911.2019.1522035>. Acesso em: 26 nov. 2025.

CHIZZOTTI, A. **Pesquisa qualitativa em Ciências Humanas e Sociais**. 6 ed. Petrópolis: Vozes, 2014.

CRONBACH, L. J. **Essentials of Psychological Testing**. 2. ed. New York: Harper & Row, 1963.

DUTRA, K.; NEGREIROS, H. L.; SANTOS, W. Os exames standardizados estaduais e a constituição de políticas educacionais para a avaliação: uma análise comparada do Paebes (Brasil) e do Exeims BC (México). **Revista Arquivos Analíticos de Políticas Educativas**, [S. l.], v. 32, n. 27, 2024. DOI: <https://doi.org/10.14507/epaa.32.8454>. Disponível em: <https://epaa.asu.edu/index.php/epaa/article/view/8454>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 064–R, de 24 de maio de 2017**. Institui diretrizes para o Programa de Avaliação da Educação Básica do Espírito Santo (Paebes). Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2017. Disponível em: <https://avaliacaoeducacional.com/wp-content/uploads/2017/08/portaria-n-064-r-dio-25-05-2017.pdf>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 168–R, de 23 de dezembro de 2020**. Define parâmetros para bonificação docente com base em resultados do Paebes. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2020. Disponível em: <https://sedu.es.gov.br/Media/sedu/Portarias%20e%20Editais/Portaria%20N%C2%BA%20168-R%20DO%2028-12-2020.pdf>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 138-R, de 23 de junho de 2022.** Dispõe sobre a matriz de referência das avaliações externas aplicadas no Espírito Santo. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2022a. Disponível em: <https://sedu.es.gov.br/Media/sedu/pdf%20e%20Arquivos/Portaria%20004%20Comit%C3%AA%20de%20Monitoramento%20de%20Resultados.pdf>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 139-R, de 23 de junho de 2022.** Regulamenta os critérios de análise e uso dos resultados do Paebes. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2022b. Disponível em: <https://ioes.dio.es.gov.br/portal/edicoes/download/7310>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 181-R, de 09 de agosto de 2022.** Estabelece critérios de responsabilização pedagógica a partir do desempenho nas avaliações em larga escala. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2022c. Disponível em: <https://sedu.es.gov.br/Media/sedu/pdf%20e%20Arquivos/POR-TARIA%20N%20181-R,%2030%20de%20junho%20de%202025.pdf>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 003-R, de 13 de janeiro de 2023.** Estabelece diretrizes para a avaliação da aprendizagem na Rede Estadual de Ensino. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2023a. Disponível em: <https://sedu.es.gov.br/Media/sedu/pdf%20e%20Arquivos/EDITAL%20SEDU%20N%C2%BA%2005-2023%20-%20CHE.pdf>. Acesso em: 26 nov. 2025.

ESPÍRITO SANTO. **Portaria nº 021-R, de 30 de janeiro de 2023.** Define normas complementares para o funcionamento das unidades escolares da Rede Estadual. Vitória, ES: Diário Oficial do Estado do Espírito Santo, 2023b. Disponível em: [https://sedu.es.gov.br/Media/sedu/Portarias%20e%20Editais/POR-TARIA%20N%C2%BA%20021-R.2023%20\(DOES%20de%2001-02-2023\)-2.pdf](https://sedu.es.gov.br/Media/sedu/Portarias%20e%20Editais/POR-TARIA%20N%C2%BA%20021-R.2023%20(DOES%20de%2001-02-2023)-2.pdf). Acesso em: 26 nov. 2025.

FALABELLA, A. El poder de los números en educación: subjetivización, resistencia y nuevos horizontes. In: LEIVA, M. V. G.; PASQUAL, M. S. (ed.). **Políticas de evaluación y accountability en América Latina**. Valparaíso: Ediciones Altazor, 2023. p. 59-72.

FERRER, F. J. **La educación comparada actual**. Barcelona: Ariel Educación, 2002.

GINZBURG, C. **Mitos, emblemas, sinais: morfologia e história**. São Paulo: Companhia das Letras, 1989.

LIBÂNIO, J. C.; FREITAS, R. M. da M. **Políticas educacionais neoliberais e escola pública**: uma qualidade restrita de educação escolar. Goiânia: Editora Espaço Acadêmico. 2018

LIMA, R.; FASSARELLA, L. Programa de avaliação da educação básica do Espírito Santo: repercussões nas escolas e no ensino de matemática. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 30, n. 117, p. 1089–1111, out./dez. 2022. DOI: <https://doi.org/10.1590/S0104-40362022002902892>. Disponível em: <https://www.scielo.br/j/ensaio/a/7nQyRkfr9nbb3vd53LtttYR/?lang=pt>. Acesso em: 26 nov. 2025.

LOPES, I. P. *et al.* Políticas de formação continuada ou teaching to the test? Uma análise comparada das ações desenvolvidas pelas Secretarias de Educação dos Estados do Espírito Santo e da Baja California. **Revista Ponto de Vista**, [S. l.], v. 14, n. 2, p. 01–21, 2025. DOI: 10.47328/rpv.v14i2.19250. Disponível em: <https://periodicos.ufv.br/RPV/article/view/19250>. Acesso em: 26 nov. 2025.

LOPES, I. P.; STIEG, R.; SANTOS, W. Evaluaciones educativas a gran escala en los estados de Espírito Santo y Baja California: especificidades y aproximaciones entre Paebes y Exeims–BC. **Revista de Estudios y Experiencias en Educación**, [S. l.], v. 23, p. 124–141, 2024. DOI: <https://doi.org/10.21703/rexe.v23i52.2349>. Disponível em: <https://www.rexe.cl/index.php/rexe/article/view/2349>. Acesso em: 26 nov. 2025.

LOUREIRO, N. L. *et al.* Avaliação de matemática no ensino médio do Espírito Santo: um estudo comparativo entre Saeb. **Revista Sergipana de Matemática e Educação Matemática**, [S. l.], v. 9, n. 3, p. 284–303, 2024. <https://doi.org/10.34179/revisem.v9i3.20694>. Disponível em: <https://periodicos.ufs.br/ReviSe/article/view/20694>. Acesso em: 26 nov. 2025.

MARQUES, R. *et al.* A Evolução do Enem no Brasil: da avaliação diagnóstica ao exame de alto impacto (1998–2018). **Revista Espaço Pedagógico**, Passo Fundo, v. 31, p. 1–19, 2024. DOI: <https://doi.org/10.5335/rep.v31i15304>. Disponível em: https://www.researchgate.net/publication/381108054_A_evolucao_do_Enem_no_Brasil_da_avaliacao_diagnostica_ao_exame_de_alto_impacto_1998-2018. Acesso em: 26 nov. 2025.

MARQUES, R.; STIEG, R.; SANTOS, W. Exames standardizados: análise dos modelos e das teorias na produção acadêmica. **Revista Meta: Avaliação**, [S. l.], v. 12, n. 34, p. 1–27, 2020. DOI: <http://dx.doi.org/10.22347/2175-2753v12i34.2342>. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/2>. Acesso em: 26 nov. 2025.

MATTEI, T. F.; DA CUNHA, M. S. Impacto do trabalho infantojuvenil no desempenho escolar para o Brasil urbano. **Estudos em Avaliação Educacional**, São Paulo, v. 32, p. 1-27, 2021. DOI: <https://doi.org/10.18222/ea.v32.7159>. Disponível em: <https://publicacoes.fcc.org.br/ea/article/view/7159>. Acesso em: 26 nov. 2025.

OLIVEIRA, M. B.; CANUTO, M. B.; LEMOS, A. S. O global education reform movement (GERM) e a BNCC: performatividade para formação e práticas docentes. **Revista Teias**, [S. l.], v. 24, n. 74, p. 10-22, jul./set. 2023. DOI: <https://doi.org/10.12957/teias.2023.76015>. Disponível em: <https://www.e-publicacoes.uerj.br/revistateias/article/view/76015>. Acesso em: 26 nov. 2025.

ORFIELD, G.; WALD, J. Testing, testing: the high-stakes testing mania hurts poor and minority students the most. **The Nation**, [S. l.], v. 270, n. 22, p. 38-40, 2000. Disponível em: <https://thenation.s3.amazonaws.com/pdf/testingtesting2000.pdf>. Acesso em: 26 nov. 2025.

PASQUALI, L. **Psicometria**: teoria dos testes na psicologia e na educação. Petrópolis: Vozes, 2017.

ROCHA JUNIOR, L. D. U. **Calibração e validação da versão brasileira do banco de itens relações sociais do Patient Related Outcomes Measures Information System (PROMIS®) pediátrico**. 2018. 58 f. Tese (Doutorado em Ciências da Saúde) - Universidade Federal de Uberlândia, Uberlândia, 2018. Disponível em: <https://www.repositorio.ufu.br/handle/123456789/24422>. Acesso em: 26 de nov. 2025.

RUIZ MARTINS, M. A.; MENDES SANT'ANA, I. A recuperação da aprendizagem no pós-pandemia: medida fundamental para a garantia do direito à educação. **Revista Fragmentos de Cultura** - Revista Interdisciplinar de Ciências Humanas, Goiânia, v. 33, n. 3, p. 550-563, 2024. DOI: <https://doi.org/10.18224/frag.v33i3.13734>. Disponível em: <https://seer.pucgoias.edu.br/index.php/fragmentos/article/view/13734>. Acesso em: 26 nov. 2025.

SANTOS, M. G. *et al.* Análise comparativa de matrizes de avaliação: Enem, Paebes, Planea e Exeims-BC. **Estudos em Avaliação Educacional**, São Paulo, v. 35, p. 1-31, 2024. DOI: <http://dx.doi.org/10.18222/ea.v35.10997>. Disponível em: <https://publicacoes.fcc.org.br/ea/article/view/10997>. Acesso em: 26 nov. 2025.

SANTOS, W.; SOARES, D. J. M.; SOARES, T. E. A.; STIEG, R. Paebes TRI e seus usos em escolas de ensino médio da Secretaria de Estado da Educação do Espírito Santo: qual o lugar do trabalho dos pedagogos? *In*: PONTES JUNIOR, J. A. F.; OSTI, A.; VIDAL E. M. (ed.). **Avaliações estaduais**: nascimento, vida e protagonismo. Brasília: ANPAE, 2021. p. 156-181.

SILVEIRA, D. T.; CÓRDOVA, F. P. A pesquisa científica. *In*: GERHARDT, T. E.; SILVEIRA, D. T. (orgs.). **Métodos de pesquisa**. Porto Alegre: Editora da UFRGS, 2009. p. 31–42.

SOARES, D. J. M.; SOARES, T. E. A.; STIEG, R.; SANTOS, W. Relación entre la titulación del profesorado y la calidad de las instituciones de la Red Federal de Educación Profesional, Científica y Tecnológica brasileña. **Revista de Estudios y Experiencias en Educación**, [S. l.], v. 22, n. 48, p. 207–221, 2023. DOI: <http://dx.doi.org/10.21703/0718-5162.v22.n48.2023.012>. Disponível em: <https://www.rexe.cl/index.php/rexe/article/view/2013>. Acesso em: 26 nov. 2025.

SOARES, D.; SOARES, T.; SANTOS, W. Análise do indicador de desenvolvimento das escolas estaduais do Espírito Santo. **Educação & Sociedade**, Campinas, v. 43, p. 1–20, 2022. DOI: <https://doi.org/10.1590/ES.256722>. Disponível em: <https://www.scielo.br/j/es/a/tdg67Pr5JbQzqfch7nDHbjv/?lang=pt>. Acesso em: 26 nov. 2025.

SOARES, L. A.; RIANI, J. L. R. Efeitos da designação dos professores na educação de Minas Gerais. **Pretexto**, [S. l.], v. 25, n. 2, p. 46–63, abr./jun. 2024. DOI: <https://doi.org/10.21714/pretexto.v25i2.9586>. Disponível em: <https://revista.fumec.br/index.php/pretexto/article/view/9586>. Acesso em: 26 nov. 2025.

WERLE, F. O. C. Políticas de avaliação em larga escala na educação básica: do controle de resultados à intervenção nos processos de operacionalização do ensino. **Ensaio: Avaliação Política Pública Educacional**, Rio de Janeiro, v. 19, n. 73, p. 769–792, out./dez. 2011. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0104-40362011000400003&lng=pt&nrm=iso. Acesso em: 26 nov. 2025.

ZLUHAN, M. R.; SANTOS, A.; CORRÊA, S. de S. M. dos. As juventudes e o Ensino Médio: Perspectivas de futuro. **Revista online de Política e Gestão Educacional**, Araraquara, v. 28, n. 00, p. 1–17, 2024. DOI: <https://doi.org/10.22633/rpge.v28i00.19448>. Disponível em: <https://periodicos.fclar.unesp.br/rpge/article/view/19448>. Acesso em: 26 nov. 2025.

MINICURRÍCULO DOS AUTORES E DAS AUTORAS

ADRIANA BAUER – Pós-doutorado no *Institute of Education*, Londres. Doutora em Educação pela Faculdade de Educação da Universidade de São Paulo. É pesquisadora sênior da Fundação Carlos Chagas e docente do Departamento de Metodologia do Ensino e Educação Comparada da Faculdade de Educação da USP.

ALEXANDRE JALOTO – Doutor em Psicologia pela Universidade São Francisco e mestre em Educação em Ciências e Saúde pela Universidade Federal do Rio de Janeiro. É pesquisador do Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), onde coordena a equipe de Psicometria do Sistema de Avaliação da Educação Básica (Saeb).

ANTONI VERGER – Doutor em Sociologia e professor catedrático da Universidade Autônoma de Barcelona, onde atua como pesquisador ICREA. Coordena o mestrado europeu *Education Policies for Global Development* (GLOBED) e é editor do *World Yearbook of Education* e do *Journal of Education Policy*.

BREYNNER RICARDO DE OLIVEIRA – Doutor em Educação pela Universidade Federal de Minas Gerais. É professor associado da Universidade Federal de Ouro Preto e docente dos Programas de Pós-Graduação em

Educação e em Economia Aplicada. Integra redes nacionais e internacionais de pesquisa em monitoramento, avaliação e implementação de políticas públicas educacionais.

CRISTIANE MACHADO – Doutora em Educação pela Universidade de São Paulo. Livre-docente pela Universidade Estadual de Campinas. É professora associada na Faculdade de Educação da Unicamp, onde lidera o Laboratório de Gestão Educacional. Atua no Programa de Pós-Graduação em Educação e no Mestrado Profissional em Educação Escolar da mesma instituição.

DAIANE ANSALONI MENDES – Mestre em Educação e graduada em Pedagogia pela Universidade Federal de Ouro Preto. É coordenadora pedagógica em instituição da rede particular de ensino e atuou na Secretaria Municipal de Educação de Mariana/MG, no Programa de Educação em Tempo Integral. Foi tutora de curso de aperfeiçoamento na UFMG e é uma das idealizadoras do projeto *Pensa Educação*.

DANIEL ABUD SEABRA MATOS – Pós-doutorado na Universidade da Beira Interior e na Universidade São Francisco. Doutor e mestre em Educação pela Universidade Federal de Minas Gerais, com sanduíche na *University of Florida*. Professor associado do Depto. de Educação da Universidade Federal de Ouro Preto e docente do Programa de Pós-Graduação em Educação. Coordena o Núcleo de Avaliação Educacional.

DENILSON JUNIO MARQUES SOARES – Doutor em Educação pela UFES e mestre em Estatística Aplicada e Biometria pela Universidade Federal de Viçosa. Professor do Instituto Federal de Minas Gerais e colaborador no Programa de Pós-Graduação em Educação da UFMG. Membro da Comissão de Análise de Dados de Testes e Questionários da Educação Básica em Larga Escala do INEP. Lidera o Grupo de Pesquisa *Análise e Modelagem de Dados Educacionais*.

EDGAR QUILABERT – Pesquisador de pós-doutorado na Universidade Autônoma de Barcelona. Integra o grupo de pesquisa *Globalization, Education and Social Policies* e o Instituto de Governo e Políticas Públicas. Suas investigações abordam políticas educacionais globais, governança, privatização e desigualdades na educação.

FLAVIA PEREIRA XAVIER – Pós-doutorado em Educação pela Universidade Estadual de Campinas. Doutora e mestra em Sociologia pela Universidade Federal de Minas Gerais (UFMG). É professora associada da Faculdade de Educação da UFMG e líder do Núcleo de Pesquisas em Desigualdades Escolares (NuPeDE). Atua no curso de Pedagogia e no Programa de Pós-Graduação em Educação.

FRANCISCO AUGUSTO DA COSTA GARCIA – Mestre em Engenharia Elétrica pela Universidade Estadual de Campinas, graduado em Pedagogia e em Engenharia Elétrica pela Universidade de Brasília. Desenvolveu pesquisas nas áreas de Avaliação Educacional, Ensino de Ciências e Engenharia, com ênfase em sistemas ópticos, rádio definido por *software* e desenvolvimento de equipamentos de testes.

GERARD FERRER-ESTEBAN – Professor dos Estudos de Psicologia e Ciências da Educação da *Universitat Oberta de Catalunya*, onde coordena o Laboratório de Educação Social. Colabora com os grupos GEPS e GLANCE da Universidade Autônoma de Barcelona. Desenvolve pesquisas sobre políticas educacionais, segregação escolar, desigualdades, educação comunitária e estratégias de ensino.

GIRLENE RIBEIRO DE JESUS – Doutora e mestra em Psicologia pela Universidade de Brasília e graduada em Psicologia pela Universidade Federal da Paraíba. É professora do Departamento de Política, Gestão e Avaliação da Educação e do Programa de Pós-Graduação em Educação da UnB. Lidera o Grupo de Pesquisa *Política, Gestão e Avaliação da Educação Básica e Superior* e é coautora do teste não verbal de inteligência SON-R 2½-7.

GIZELLE ALVES XAVIER DOS SANTOS – Mestra em Educação, com ênfase em Políticas Públicas e Gestão da Educação, pela Universidade de Brasília. Especialista em Gestão Escolar pela UnB. Graduada em Letras – Português e Inglês pela Faculdade da Terra de Brasília. É professora da Secretaria de Educação do Distrito Federal e integrante da linha de pesquisa em Políticas Públicas e Gestão da Educação da Faculdade de Educação da UnB.

GUSTAVO HENRIQUE MARTINS – Pós-doutorando e doutor em Psicologia pela Universidade São Francisco, onde também é docente no curso de Psicologia. É pesquisador do Catvante Laboratório de Testagem Compu-

tadorizada e integrante da Comissão de Análise de Dados de Testes e Questionários da Educação Básica em Larga Escala do INEP. É vice-presidente da ABRAOPC.

JOYCE SOARES RODRIGUES PETRUS – Doutoranda em Educação pela Universidade Federal de Minas Gerais, mestra em Educação pela Universidade Estadual de Minas Gerais e bacharel em Estatística. Foi Superintendente de Avaliação e Diretora de Informações da Secretaria de Estado de Educação de Minas Gerais. Possui vasta experiência em análise de dados, avaliação, indicadores e políticas educacionais.

LARA NUNES BARBOSA – Licencianda em Educação Física pela Universidade Federal do Espírito Santo. É bolsista da Fundação de Amparo à Pesquisa do Espírito Santo no Programa Institucional de Iniciação Científica da UFES e membra do Instituto de Pesquisa em Educação e Educação Física (Proteoria).

LUCRECIA RODRIGO – Doutora em Sociologia da Educação pela Universidade Complutense de Madri e licenciada em Ciências da Educação pela Universidade de Buenos Aires. É pesquisadora do CONICET, com sede na Universidade Nacional do Oeste, e professora da Universidade Nacional da Patagônia Austral. Atua também como pesquisadora no Instituto de Investigações em Ciências da Educação da UBA.

LUÍS ANTÔNIO FAJARDO PONTES – Doutor e mestre em Educação e licenciado em Física pela Universidade Federal de Juiz de Fora. Realizou especializações em Métodos Quantitativos de Pesquisa Social na Universidade de Michigan (ICPSR e ISR). É Analista de Avaliação Sênior do Centro de Políticas Públicas e Avaliação da Educação e professor do Programa de Pós-Graduação em Gestão e Avaliação da Educação Pública.

MAURO C. MOSCHETTI – Professor Serra Húnter na Universidade Autônoma de Barcelona e pesquisador do grupo GEPS. Sua pesquisa aborda políticas educacionais comparadas, com foco em privatização, parcerias público-privadas e governança global da educação. Tem investigado esses processos na América Latina e em outros contextos do Sul Global.

PAULO HENRIQUE ARCAS – Doutor e mestre em Educação pela Universidade de São Paulo. É professor do Departamento de Gestão Educacional, Teorias e Práticas de Ensino da Universidade Federal de Lavras e do Programa de Pós-Graduação em Educação (Mestrado Profissional). Atuou como consultor do Instituto Internacional de *Planeamiento de la Educación* (IIPE/UNESCO), em Buenos Aires, entre 2010 e 2017.

RONILDO STIEG – Doutor em Educação Física pela Universidade Federal do Espírito Santo, com doutorado sanduíche na *Universidad de Valladolid* (Espanha). Mestre e licenciado em Educação Física pela UFES. Foi bolsista de Pós-Doutorado da FAPES junto ao Programa de Pós-Graduação em Educação da UFES e é membro do Instituto de Pesquisa em Educação e Educação Física (Proteoria).

TALITA EMIDIO ANDRADE SOARES – Doutoranda e mestra em Educação pela Universidade Federal do Espírito Santo, com estágios de pesquisa na *Universidad de Santiago de Chile* e na *Universitat de València*. É licenciada em Matemática pela Universidade Federal de Viçosa e professora EBTT do Instituto Federal de Minas Gerais.

TATIANE GONÇALVES MORAES – Doutoranda em Educação pela Universidade Federal de Juiz de Fora, mestra em Gestão e Avaliação da Educação Pública e graduada em Matemática pela mesma instituição. É coordenadora de Avaliação e Dados Educacionais na ADE Educação, professora de Matemática da Faculdade Ensino e membro da Comissão de Análise de Dados de Testes e Questionários da Educação Básica em Larga Escala do INEP/MEC.

WAGNER DOS SANTOS – Doutor em Educação pela UFES, onde é professor e diretor de Pós-Graduação. Atua como pró-reitor de Pesquisa e Pós-Graduação em exercício e é professor visitante na *Universidad de Santiago de Chile*. Foi coordenador do Programa de Pós-Graduação em Educação da UFES (2020–2024) e é líder do Instituto de Pesquisa em Educação e Educação Física (Proteoria). Bolsista de Produtividade em Pesquisa do CNPq.

Copyright © 2026 Encontrografia Editora.

Todos os direitos reservados. É proibida a reprodução parcial ou total desta obra sem a expressa autorização da editora.

Este volume inaugura a coleção *Diálogos sobre Avaliação*, reunindo especialistas do Brasil e do exterior para discutirem, de forma qualificada, os sistemas de avaliação externa, seus fundamentos e seus impactos nas políticas educacionais. Os capítulos examinam o Saeb, o Ideb, as avaliações estaduais, a atuação de organismos internacionais e os tensionamentos próprios da *accountability*, articulando evidências empíricas, análises comparadas e reflexão teórica. Oferece uma leitura atualizada, crítica e propositiva sobre qualidade, equidade, governança e justiça avaliativa, contribuindo para o fortalecimento de políticas públicas orientadas por responsabilidade social e pelo compromisso com a melhoria das aprendizagens e para a garantia do direito à educação.



GOVERNO DO ESTADO
DO ESPÍRITO SANTO
*Secretaria da Ciência, Tecnologia,
Inovação e Educação Profissional*



encontrografia

encontrografia.com
www.facebook.com/Encontrografia-Editora
www.instagram.com/encontrografiaeditora
www.twitter.com/encontrografia