

DOSSIÊ

Qualidade, aprendizagem e avaliação sistêmica: discursos dos organismos internacionais para os países latino-americanos

**Avaliações em larga escala na América Latina:
problematizando a qualidade da educação escolar*****Large-scale assessments in Latin America:
Problematizing the quality of school education***

Cristiane Machado^a
crimacha@unicamp.br

Ocimar Munhoz Alavarse^b
ocimar@usp.br

Raíssa de Oliveira Chappaz^c
raissa.chappaz@gmail.com

RESUMO

O objetivo deste artigo é caracterizar as avaliações em larga escala presentes na América Latina, das quais o Brasil é participante, problematizando a noção de qualidade que elas imprimem para a educação escolar. Trata-se de um estudo documental que tomou por base as informações disponíveis no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). São descritas e analisadas as avaliações: Estudo Regional Comparativo e Explicativo (ERCE); Estudo Internacional de Progresso em Leitura (PIRLS); Programa Internacional de Avaliação de Estudantes (Programme for International Student Assessment – PISA) e Estudo Internacional de Tendências em Matemática e Ciências (Trends in International Mathematics and Science Study – TIMSS). Conclui-se que a qualidade da educação escolar tem centralidade nessas avaliações, traduzida, em geral, em escores obtidos pelo desempenho de estudantes em testes padronizados, conformando uma noção de qualidade quase que exclusivamente associada a resultados das aprendizagens escolares, ainda que questionários contextuais acompanhem essas aferições. Ademais, a caracterização das iniciativas de avaliação em larga escala presentes na América Latina revela o interesse cada vez mais robusto de organismos internacionais em criar instrumentos capazes de aferir a aprendizagem de estudantes, possibilitando a comparação da qualidade da educação escolar entre e dentro dos sistemas educativos.

Palavras-chave: Avaliação em Larga Escala. América Latina. Qualidade da Educação Escolar. Avaliações Internacionais.

^a Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brasil.

^b Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil.

^c Instituto Federal de São Paulo (IFSP), São Paulo, São Paulo, Brasil.

ABSTRACT

This article aims to characterize the large-scale assessments implemented in Latin America, of which Brazil is a participant, problematizing the quality that such assessment tools inflict on school education. This is a documentary research based on information available on the website of the National Institute of Studies and Educational Research Anísio Teixeira (INEP). The following assessments are described and analyzed: Comparative and Explanatory Regional Study (ERCE); Progress in Reading International Study (PIRLS); Program for International Student Assessment (PISA); and Trends in International Mathematics and Science Study (TIMSS). We conclude that the quality of school education is central to these assessments, generally translated into scores obtained by students' performance on standardized tests, forming a notion of quality almost exclusively associated with the results of school learning, even though contextual questionnaires accompany these assessments. Furthermore, the characterization of large-scale assessment initiatives present in Latin America reveals the increasingly robust interest of international organizations in creating instruments capable of measuring student learning, enabling the comparison of the quality of school education between and within educational systems.

Keywords: Large-Scale Assessment. Latin America. Quality of School Education. International Assessments.

Introdução

Na América Latina, diversas iniciativas internacionais de avaliação em larga escala ganharam adesão dos países do continente desde a década de 1990. Ainda que com diferenciações entre elas, elas refletem o movimento de consolidação desse formato de avaliação no cenário das políticas educacionais como expressão de um fenômeno em escala mundial, conforme apontam, entre vários estudos, Maroy e Voisin (2013), Verger, Fontdevila e Parcerisa (2019) e Verger e Normand (2015).

Nesse contexto, as avaliações internacionais em larga escala passaram a ser fontes de referência na busca por evidências acerca do aprendizado dos estudantes da América Latina, com informações utilizadas para comparações relativas a parâmetros educacionais, incluindo associações com o desenvolvimento desses países. Esse processo foi desencadeado, com diferenciações entre e intra as nações do continente, com base no argumento do potencial que os resultados, sobretudo de testes padronizados, teriam para orientar políticas educacionais, podendo direcionar intervenções específicas nas dificuldades evidenciadas nesses resultados e nos dados de contexto, obtidos, via de regra, em questionários aplicados conjuntamente.

Acompanhando e robustecendo a tendência que teve como marco a Conferência de Jomtien, em 1990, voltada à valorização da educação escolar para impulsionar o desenvolvimento econômico dos países e à criação de instrumentos capazes de aferir a efetividade dessa educação, a melhoria da qualidade é, em geral, a justificativa e, ao mesmo tempo, a promessa das iniciativas de avaliação em larga escala. É nesse contexto que essas iniciativas eclodem mundialmente, com nuances e especificidades próprias dos locais, culturas e organismos responsáveis, mas também, com um eixo comum que relaciona testes padronizados de estudantes como expressão da qualidade da educação.

Nesse artigo, com foco na Educação Básica, procurou-se caracterizar e esmiuçar quatro avaliações em larga escala presentes na América Latina e, simultaneamente, no Brasil, problematizando a noção de qualidade que elas imprimem para a educação escolar.

O estudo, apoiado na pesquisa documental, teve como base a consulta aos documentos disponíveis no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e examinou, principalmente, as explanações relativas às características e à noção de qualidade sobre as avaliações: Estudo Regional Comparativo e Explicativo (ERCE); Estudo Internacional de Progresso em Leitura (PIRLS); Programa Internacional de Avaliação de Estudantes (*Programme for International Student Assessment – PISA*) e Estudo Internacional de Tendências em Matemática e Ciências (*Trends in International Mathematics and Science Study – TIMSS*). As informações obtidas foram sistematizadas e categorizadas de forma a permitir uma análise comparativa das características gerais dessas quatro avaliações em larga escala em relação ao ano de início das aplicações, periodicidade e quantidade de países participantes; ao início da participação do Brasil; às séries/anos avaliados; à abrangência quantitativa de estudantes; aos instrumentos aplicados e finalidades.

É oportuno pontuar que, além de participar dessas quatro avaliações internacionais, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) é responsável desde 1990 pela formulação, aplicação e sistematização do Sistema de Avaliação da Educação Básica (SAEB), avaliação amostral e nacional em larga escala que, em 2005, se desdobrou em uma avaliação censitária com o nome de Prova Brasil (Machado; Alavarse, 2014).

Este artigo está organizado em três partes, além desta introdução. Inicialmente, são descritas e analisadas as avaliações em larga escala que circulam pela América Latina, das quais o Brasil faz parte; em seguida são feitas problematizações acerca da noção de qualidade que estruturam essas avaliações e, por fim, são tecidas reflexões finais como forma de indução a novos questionamentos sobre diferentes aspectos da relação da avaliação em larga escala com a qualidade da educação escolar.

Iniciativas de avaliação em larga escala na América Latina: características e abrangência

Por meio de consulta no site do INEP¹, órgão brasileiro responsável pelo planejamento, organização e operacionalização das iniciativas internacionais no território nacional, foram identificadas quatro avaliações presentes na América Latina com o recorte da Educação Básica das quais o Brasil participa: o Estudo Regional Comparativo e Explicativo (ERCE)²; o Estudo Internacional de Progresso em Leitura (PIRLS); o Programa Internacional de Avaliação de Estudantes (*Programme for International Student Assessment – PISA*); e o Estudo Internacional de Tendências em Matemática e Ciências (*Trends in International Mathematics and Science Study – TIMSS*).

Publicizando suas atribuições em relação a essas iniciativas internacionais de avaliação, o INEP desempenha ações no planejamento e na operacionalização, o que é destacado no texto de

¹ As informações de caracterização e descrição das iniciativas internacionais de avaliação em larga escala citadas entre aspas neste tópico do artigo são provenientes do INEP, bem como as que compõem os quadros apresentados (Brasil, s.d.).

² O ERCE passou por uma mudança em sua sigla, de modo que até 2019 era incorporado ao seu nome uma nomenclatura que indicava o número da edição – PERCE, SERCE e TERCE –, mas a partir da quarta edição o Estudo passou a ser identificado pela sigla ERCE seguida do ano da aplicação.

apresentação no site do INEP. No caso do PISA e do PIRLS, são especificadas as ações realizadas, sendo elas: representar o Brasil perante a entidade responsável pela iniciativa de avaliação internacional em larga escala; coordenar as ações de tradução dos instrumentos aplicados; coordenar a aplicação dos instrumentos nas escolas participantes e a coleta e codificação das respostas dos participantes; analisar ou contribuir com a análise dos resultados; e elaborar o relatório nacional. Além disso, no caso do PIRLS, o INEP colabora com a seleção de amostra dos participantes da aplicação.

As demais etapas e ações que envolvem a elaboração e aplicação de uma avaliação internacional em larga escala no país são realizadas pelas entidades internacionais responsáveis por cada uma das iniciativas. No caso do ERCE a entidade é o Laboratório Latino-americano de Avaliação da Qualidade da Educação (LLECE), ligado à Oficina Regional de Educação para América Latina e Caribe (OREALC/UNESCO Santiago); do PISA, é a Organização para a Cooperação e Desenvolvimento Econômico (OCDE); do PIRLS e do TIMSS, é a Associação Internacional para a Avaliação do Desempenho Educacional (IEA), identificada como “cooperativa internacional de instituições nacionais de pesquisa, acadêmicos e analistas”.

O ERCE, o PIRLS, o PISA e o TIMSS têm diferentes definições que as identificam, delimitando também diferentes objetivos declarados. Considerando a ordem cronológica em que foram formuladas, o TIMSS existe desde 1995 e é definido como um “estudo internacional de tendências” que tem por objetivo “fornece[r] dados sobre o desempenho” nas áreas do conhecimento de Matemática e Ciências”. O ERCE, formulado em 1997 e identificado como uma “avaliação internacional direcionada aos países da América Latina e Caribe”, tem por objetivo declarado “monitora[r] os avanços na aprendizagem dos estudantes da região”.

Já na década de 2000, foram desenvolvidos o PISA e o PIRLS. O PISA, que teve sua primeira aplicação já em 2000, é identificado como um “estudo comparativo internacional” que tem por objetivo “revelar se os sistemas educacionais estão se tornando mais ou menos eficientes em preparar os estudantes para continuar seus estudos ou para o mercado de trabalho”, buscando, assim, “traduzir os resultados da avaliação em milhares de fragmentos de dados que, quando reunidos, revelam a representação de um sistema educacional eficiente”, de modo a “incentivar os países a aprenderem com as experiências uns dos outros na construção de sistemas educacionais mais justos e inclusivos”. O PIRLS, que surgiu em 2001, é identificado apenas como “uma iniciativa” realizada por uma cooperativa internacional que trabalha para “avaliar, entender e melhorar a educação em todo o mundo”, tendo como objetivo declarado “analisar tendências de compreensão leitora, além de coletar informações sobre os contextos de aprendizagem, para caracterizar o processo de leitura dos estudantes avaliados nos países que participam do estudo”.

Assim, com focos e abrangências distintas e considerando o potencial e a difusão dessas iniciativas na América Latina, faz-se importante caracterizá-las também em relação a outros aspectos, como suas finalidades, periodicidade, países participantes, séries/anos avaliados, abrangência, instrumentos e o que é medido, a fim de apreender diferenças e pontos de aproximação. Para tanto, a base documental adotada como referência tem como recorte o site e os documentos oficiais disponibilizados pelo INEP, que é a autarquia responsável pelos processos que envolvem a realização dessas avaliações internacionais em larga escala no território brasileiro.

O Quadro 1 abrange informações do ERCE, PIRLS, PISA e TIMSS no que tange ao ano de início de aplicação, abrangência de países participantes, participação do Brasil, séries/anos avaliados, abrangência no Brasil e instrumentos aplicados.

Quadro 1: Características gerais das iniciativas de avaliação em larga escala internacionais: ERCE, PIRLS, PISA e TIMSS

	TIMSS	ERCE	PISA	PIRLS
Ano de início	1995	1997	2000	2001
Periodicidade	A cada 4 anos	1997; 2006; 2013; 2019	A cada 3 anos	A cada 5 anos
Países participantes	“Cerca de 70 países usam os dados de tendências [...] mais países se unem ao estudo a cada ciclo de avaliação”.	1997: 13 países 2006: 16 países 2013: 15 países 2019: 18 países	“Desde sua primeira edição, em 2000, o número de países e economias participantes tem aumentado a cada ciclo”.	2021: 61 países
Participação do Brasil	Desde 2023	Desde 1997	Desde 2000	Desde 2021
Séries/anos avaliados	4º e 8º anos do EF	1997: 4º e 5º anos do EF 2006: séries escolares equivalentes ao 4º e 7º ano do EF 2013: idem a 2006 2019: idem a 2006	“Estudantes na faixa etária dos 15 anos, idade em que se pressupõe o término da escolaridade básica obrigatória na maioria dos países.”	4º ano do EF
Abrangência*		1997: 55 mil alunos 2006: 196 mil alunos 2013: 195.752 alunos	2000: 200 mil alunos 2003: 250 mil alunos (Brasil: 4.452) 2006: 400 mil alunos (Brasil: 9.345) 2009: 470 mil alunos (Brasil: 20.127) 2018: 600 mil alunos (Brasil: 10.691)	2021: aproximadamente 400 mil alunos
Instrumentos aplicados	Testes cognitivos e questionários contextuais	Testes cognitivos e questionários contextuais	Testes cognitivos e questionários contextuais	Testes cognitivos e questionários contextuais

Fonte: Sistematização feitas pelos autores a partir de informações em Brasil (s.d.).

* Os valores totais correspondem ao número total de estudantes nos países participantes.

Embora não tenham sido identificados dados quantitativos quanto à abrangência de países envolvidos e de estudantes e escolas em todas as edições realizadas, o Quadro 1 evidencia iniciativas consolidadas, com um histórico de aplicações envolvendo um número significativo de países. No caso da participação do Brasil, em duas iniciativas – ERCE e PISA –, o Brasil participou de todas as avaliações até o presente momento.

No caso do TIMSS, ainda que esta seja a mais antiga iniciativa de avaliação internacional com recorte da Educação Básica, o Brasil começou a participar apenas na última aplicação, realizada em 2023. O mesmo ocorreu com o PIRLS, que nesse recorte é a mais recente iniciativa formulada, da qual o Brasil participou após 20 anos de sua existência. Cabe ponderar que o TIMSS e o PIRLS possuem focos mais específicos em relação ao objeto de avaliação, como “conteúdos e domínios cognitivos em matemática e ciências” e “habilidades de leitura”, respectivamente, o que pode indicar uma

prioridade dada às outras iniciativas com objetos de avaliação mais abrangentes e, por isso, a adesão mais tardia a essas avaliações.

Em relação às séries/anos avaliadas/os, é possível identificar que as iniciativas priorizam os anos finais do Ensino Fundamental I e do II. No caso do PISA, o Programa está mais focado em quem concluiu a segunda etapa do Ensino Fundamental, uma vez que seu recorte são estudantes de 15 anos que, pela idade, já devem estar no Ensino Médio. No que diz respeito à abrangência de países participantes e de estudantes e escolas envolvidos em cada aplicação, não foram encontrados dados de todos os anos. Analisando os dados do ERCE, iniciativa que tem as informações mais completas divulgadas no site do INEP, observa-se que, de modo geral, houve de fato um crescimento quantitativo no número de países e estudantes participando de 1997 a 2013. Os demais afirmam também existir um crescimento no número de países, o que não é possível validar pelos dados disponíveis no site do INEP.

Além disso, identifica-se como uma característica presente em todas essas iniciativas a aplicação tanto de testes, nomeados em alguns casos, como no ERCE e PIRLS, como “testes cognitivos”, quanto a aplicação de questionários contextuais, que são direcionados a diferentes atores da comunidade escolar e buscam identificar dados de contexto, apreendendo fatores que estão associados e impactam a aprendizagem dos estudantes.

Desse modo, todas as iniciativas de avaliação aqui abordadas aplicam esses questionários em estudantes, famílias, professores e diretores, e apenas o PISA destaca como opcional os questionários distribuídos aos pais/responsáveis e professores. No que tange aos temas abordados nos questionários, evidenciados no Quadro 2, o PISA e o PIRLS destacam a existência de matrizes específicas de referência. De modo geral, exemplificam como fatores abordados os aspectos de contextos de aprendizagem e acesso às oportunidades educacionais relacionados a três âmbitos: o estudante, a escola e sala de aula e o ambiente familiar. O PIRLS abrange, além desses fatores supracitados, o âmbito nacional.

Quadro 2: Aspectos abordados nos questionários contextuais

TIMSS	Fatores escolares associados à aprendizagem Fatores de sala de aula associados à aprendizagem Fatores domésticos associados à aprendizagem
ERCE	Sentido de pertencimento à escola dos estudantes Organização do ensino Nível socioeconômico da família Violência no bairro de moradia Envolvimento dos pais nas atividades de aprendizagem no lar Participação dos pais na escola
PISA	Abordagens da aprendizagem Ambientes de aprendizagem Histórico familiar dos estudantes
PIRLS	Características dos estudantes Contexto escolar Contexto de sala de aula Contexto doméstico Contexto nacional

Fonte: Sistematização feitas pelos autores a partir de informações em Brasil (s.d.).

Quanto às finalidades estabelecidas para cada uma das iniciativas de avaliação internacional em larga escala e o que elas medem, o Quadro 3 sintetiza tais informações. Destaca-se que duas das iniciativas – TIMSS e ERCE – abrangem o monitoramento como uma finalidade, em que a primeira delimita “monitorar a eficácia de seus sistemas educacionais”, enquanto o segundo destaca a qualidade apontando o monitoramento da qualidade da educação como meio para ter informações e, assim, orientar a tomada de decisões. Já o PISA e o PIRLS são mais abrangentes e contemplam mais de uma ação como finalidade, de modo que o PISA demarca a finalidade de “avaliar” conhecimentos e habilidades dos estudantes e de “aprender” com políticas e práticas de outros países para assim “formular” políticas e programas que estejam focados em melhorar a qualidade da educação e a equidade dos resultados. O PIRLS abrange como finalidade, por sua vez, “analisar tendências” relacionadas ao que é seu objeto de avaliação e “coletar informações” relacionadas ao contexto de aprendizagem. Cabe pontuar que em nenhum dos casos em que a qualidade aparece mencionada como finalidade das iniciativas é possível identificar o que compõe tal qualidade, sem mencionar, por exemplo, fatores associados à aprendizagem, que são coletados nos questionários contextuais aplicados em todas as iniciativas.

Quadro 3: Finalidades e objetos de avaliação das iniciativas internacionais de avaliação em larga escala

	Finalidade	Objeto de Avaliação
TIMSS	“[...] países usam os dados de tendências do TIMSS para monitorar a eficácia de seus sistemas educacionais em um contexto global [...]”.	“O TIMSS avalia os principais conteúdos e domínios cognitivos em matemática e ciências a serem testados no 4º e 8º ano. As matrizes de competência (em inglês, <i>framework</i>) do TIMSS estão organizadas em torno de duas dimensões: i) dimensão de conteúdo, especificando os domínios de assunto a serem avaliados; e ii) dimensão cognitiva, especificando os processos de pensamento a serem avaliados. Entre os domínios de conteúdo estão álgebra e geometria, em matemática; e biologia e química, em ciências, enquanto os domínios cognitivos de conhecer, aplicar e raciocinar descrevem as habilidades de pensamento que se espera que os estudantes usem à medida que se envolvem com o conteúdo de matemática e ciências”.
ERCE	“O ERCE serve como uma ferramenta para monitorar a qualidade da educação na região da América Latina e Caribe, fornecendo informações para o debate e orientando a tomada de decisões nessa área”.	“Os estudantes são avaliados em leitura, escrita, matemática e ciências.” “[...] considerando os objetivos de aprendizagem que são comuns na região da América Latina e Caribe, segundo as bases curriculares de cada país”.
PISA	“[...] cada país avalie os conhecimentos e as habilidades de seus estudantes em comparação com os de outros países, aprenda com as políticas e práticas aplicadas em outros lugares e formule suas políticas e programas educacionais visando à melhora da qualidade e da equidade dos resultados de aprendizagem”.	“O PISA avalia três domínios – leitura, matemática e ciências – em todas as edições ou ciclos”.
PIRLS	“[...] analisar tendências de compreensão leitora, além de coletar informações sobre os contextos de aprendizagem, para caracterizar o processo de leitura dos estudantes avaliados nos países que participam do estudo”.	“O PIRLS avalia habilidades de leitura dos estudantes [...]”.

Fonte: Sistematização feitas pelos autores a partir de informações em Brasil (s.d.).

Em relação ao que se apresenta como objeto de avaliação, as iniciativas identificam de modo geral habilidades, conteúdos e domínios cognitivos que são comuns aos países participantes, estando a leitura presente em todos. Para além do PIRLS, que tem um recorte único, os demais abrangem também Matemática e Ciências. Em 2019, o ERCE foi aplicado em estudantes do 4º e 7º anos do Ensino Fundamental, abrangendo itens de resposta construída tanto em Matemática quanto em Ciências Naturais. Entre as iniciativas analisadas, cabe ponderar que o TIMSS é o que especifica de forma mais detalhada os conteúdos que são objeto de avaliação em sua apresentação inicial no site do INEP.

A caracterização das iniciativas de avaliação em larga escala presentes na América Latina, em relação aos termos aqui analisados, demonstra variadas semelhanças e poucas diferenciações entre elas. Mas, acima de tudo, desvela o interesse cada vez mais robusto de organismos internacionais em criar instrumentos capazes de aferir a aprendizagem de estudantes e os possíveis fatores associados a ela, com vistas a possibilitar a comparação entre e dentro dos sistemas educativos.

A qualidade da educação escolar e as avaliações em larga escala

O debate em torno da qualidade da educação escolar, com maior visibilidade nas últimas décadas, adquiriu centralidade na proposição das políticas educacionais de diferentes territórios, concepções e matizes, sobretudo pela associação de sua conceituação com resultados de testes padronizados, os instrumentos mais característicos das avaliações em larga escala. Este é um panorama recorrente em escala mundial, como analisam Adams, Acedo e Popa (2012), observado após a Conferência de Jomtien, em 1990, que imprimiu os consensos de que a educação escolar é condição indispensável para o desenvolvimento econômico e que a conquista da qualidade deve ser uma meta. Ademais, esse cenário trouxe à baila a noção de qualidade como algo mensurável.

Reconhecidamente polissêmico e ambíguo (Risopatron, 1991, p. 15), atravessado por historicidades, com a propriedade de se alterar “no tempo e no espaço” (Dourado; Oliveira, 2009, p. 203), expressando significados de acordo com “interesses e valores específicos” de seus proponentes (Azevedo, 1994, p. 456), o conceito qualidade requer definição para operacionalizar seu alcance pelas políticas educacionais.

É incontornável, como fazem vários autores, salientar a dificuldade em delimitar o que seria qualidade da educação escolar, ainda mais diante de uma avassaladora tendência internacional apoiada em organismos internacionais, como o Banco Mundial, que estabelecem uma associação biunívoca entre qualidade e resultados de testes padronizados, ou, como destaca Broadfoot (2007), pela imposição de uma agenda de qualidade que procura extrair desses resultados um conjunto mais amplo de informações do que estes podem fornecer.

Harvey e Green (1993), abordando essa problemática desde a Educação Superior, indicaram que a qualidade da educação tem uma dimensão que é privilegiada a partir dos interesses que se tem sobre essa educação, podendo recair uma ênfase em resultados ou processos, com diferenciações adicionais sobre esses elementos, conotando uma maior ou menor amplitude. Assim, a discussão sobre a qualidade reveste-se inexoravelmente de uma tensão política, entendida como um jogo de forças com argumentos variados, tal como suas conseqüências.

Embora não de forma exaustiva, as iniciativas de avaliações em larga escala que circulam no Brasil com respaldo oficial e retratadas neste artigo são configurações desse quadro internacional que vão conformando uma visão de qualidade, mesmo que sem uma apresentação mais acabada, que precisa ser problematizada.

De todo modo, como já sublinhado, esses resultados são pertinentes ao processo escolar e, por isso, uma posição de refutá-los *in totum* não seria uma estratégia que contribua com o debate e a formulação de alternativas. Assim, os termos do diálogo entre avaliação interna e avaliação externa, como formulado por Nevo (1998), parece apontar para um horizonte mais denso.

O uso indiscriminado de qualidade da educação e qualidade do ensino como análogos concorre para aprofundar as controvérsias sobre o conceito de qualidade. Silva (2008, p. 81) salienta que são inúmeros os casos de usos dessas expressões como equivalentes, mas ressalva que “não é impertinente a distinção entre educação e ensino”.

Dourado (2007, p. 11) define qualidade da educação como “resultado de uma construção de sujeitos engajados pedagógica, técnica e politicamente no processo educativo”, na qual é importante sopesar “as condições objetivas de ensino, as desigualdades socioeconômicas e culturais dos alunos, a desvalorização profissional e a possibilidade limitada de atualização permanente dos profissionais da educação”. Já para Chaves (2009, p. 9), “cuando hablamos de la educación como derecho estamos entendiendo que la calidad es uno de los atributos del derecho, una de sus condiciones esenciales” e, portanto, qualidade da educação é uma característica do direito à educação. Já a qualidade do ensino está circunscrita aos efeitos do processo pedagógico, configurando-se em um aspecto da qualidade da educação, que é importante, mas não o único. Nesse sentido, qualidade da educação é um atributo que qualifica contextos mais amplos nos quais a aprendizagem se efetiva, enquanto a qualidade do ensino se refere diretamente a essa aprendizagem.

Comumente, o sentido dado ao conceito de qualidade tem sido associado à medida, obtida por meio da implementação de avaliações em larga escala. Explica Casassus (2007, p. 43) que a medição se tornou necessária para dar visibilidade aos aspectos qualitativos da educação, traduzidos como “sucesso acadêmico” com base em notas de testes e provas padronizadas. Gil (2021, p. 205) destaca, como um desdobramento da posição de Casassus (2007), que “frequentemente, o que se considera como aferição da qualidade do ensino é o exame das capacidades do alunado”. Entretanto, sabe-se que a medida é um aspecto importante, mas é um aspecto da avaliação. Dessa maneira, cabe indagar em que medida a metodologia privilegiadamente utilizada para aferir a qualidade do ensino – os testes padronizados – não estaria induzindo as escolas a uma percepção reducionista de qualidade do ensino, sem associar outros fatores relevantes aos resultados escolares, como contexto social e econômico (Nevo, 1998).

No entanto, em documento do INEP (Brasil, 2019, p. 22), encontramos uma tentativa de reconsiderar essa ênfase no desempenho de estudantes ao delimitar outros elementos conceituais, ao indicar que “[...] a discussão sobre a qualidade da educação [...] deve abranger os fatores contextuais [...] e os resultados de desempenho dos alunos nos testes cognitivos usualmente difundidos”.

Nesse mesmo documento (Brasil, 2019, p. 24) tem-se a uma proposição que, embora fugindo ao escopo deste texto no que tange às perspectivas de se estabelecer um enfrentamento à

tendência assinalada anteriormente, aponta para a necessidade de se incorporar outros elementos no tratamento conceitual de qualidade, no sentido de:

A partir das discussões realizadas e dos estudos produzidos no âmbito do INEP, em síntese, propõe-se um sistema de avaliação que tenha como objetivo medir a qualidade da educação a partir de sete Eixos: Equidade, Direitos Humanos e Cidadania, Ensino-Aprendizagem, Investimento, Atendimento Escolar, Gestão e Profissionais Docentes.

Em face dessa iniciativa, ainda em estágio embrionário, até pelo título do documento, que salienta ser uma “versão preliminar”, devemos demarcar que no Brasil, embora a Constituição Federal de 1988, em seu art. 206, inciso VII e a Lei de Diretrizes e Bases da Educação Nacional (LDB) – Lei n.º 9.394/96, no art. 3º, inciso IX, estabeleçam que o ensino será ministrado com base no princípio de garantia de padrão de qualidade, a primeira definição oficial para o conceito foi a criação do Índice de Desenvolvimento da Educação Básica (IDEB), divulgado juntamente com o Plano de Metas Compromisso Todos pela Educação, por meio do Decreto nº 6.074, de 24 de abril de 2007, com o objetivo de ser um “um indicador de qualidade educacional”, além de possibilitar o “monitoramento permanente e medição do progresso dos programas em relação às metas e resultados fixados” na educação brasileira, de acordo com seu idealizador Fernandes (2007, p. 6).

O IDEB torna público, a cada dois anos, notas em uma escala de 0 a 10 para estados, municípios e escolas com base nas taxas de aprovação e nas proficiências dos estudantes na Prova Brasil. Nesse sentido, seu princípio é o “de que qualidade da educação pressupõe que o aluno aprenda e passe de ano” (Franco, Alves e Bonamino, 2007, p. 991). Examinando os reflexos da temática, Machado e Alavarse (2014, p. 422) sustentam que:

Embora a concepção de qualidade associada ao IDEB seja um tanto reducionista, por não contemplar aspectos relevantes do processo pedagógico, é possível considerar algumas potencialidades no IDEB por conta de duas características: por facilitar uma apreensão, mesmo que parcial, da realidade educacional brasileira, aí destacadas suas escolas, e, sobretudo, por articular dois elementos que há muito tempo parecem ser antagônicos: o aumento da aprovação e o aumento do desempenho.

A definição de qualidade é fortemente apoiada na avaliação em larga escala – no caso brasileiro, na Prova Brasil –, evidenciando um movimento de conformação do sentido de qualidade por meio de testes e provas padronizadas. Nas proposições de avaliações externas por meio de agências multilaterais, a definição de qualidade é propugnada por essas agências que são responsáveis por importantes avaliações internacionais.

Se o IDEB avança ao incorporar o fluxo escolar, um elemento que foi uma das preocupações e do significado da ideia de qualidade nos anos 1980, como ponderam Oliveira e Araújo (2005, p. 8), é preciso salientar que no cenário internacional, mesmo quando do uso de questionários contextuais para se captar fatores associados, a preocupação com os resultados em testes padronizados como balizador da qualidade segue sendo a tônica, como se nota em documento oficial (Brasil, 2020, p. 15) sobre o PISA, a principal avaliação em larga escala mundial, que

[...] oferece informações sobre o desempenho dos estudantes vinculado a dados sobre seus *backgrounds* e suas atitudes em relação à aprendizagem e também aos principais fatores

que moldam sua aprendizagem, dentro e fora da escola. Os resultados permitem que cada país avalie os conhecimentos e as habilidades de seus próprios estudantes, em comparação com os de outros países; aprenda com as políticas e práticas aplicadas em outros lugares; e formule suas políticas e programas educacionais visando à melhora da qualidade e da equidade dos resultados de aprendizagem.

Isso é reforçado em um trecho de um documento veiculado no portal do INEP (UNESCO, 2020, p. 64) ao fazer referência aos resultados do ERCE 2019 relativos ao Brasil, pois eles seriam “orientaciones para mejorar la calidad de los aprendizajes de los estudiantes brasileños”. Também se encontra em Mullis e Martin (2021, p. 41) a ideia de que os resultados de testes, no caso do PIRLS, sinalizam “a alta qualidade do desempenho em leitura”.

Depreende-se, assim, a centralidade da qualidade apoiada no desempenho de estudantes em testes padronizados como principal objetivo da avaliação em larga escala, o que vai conformando uma noção de qualidade associada quase que exclusivamente a resultados das aprendizagens escolares. Ainda que questionários contextuais acompanhem essas aferições, os documentos expressam essa tônica.

Reflexões finais

Neste artigo, procuramos descrever, caracterizar e analisar quatro avaliações em larga escala presentes na América Latina, das quais o Brasil é participante - ERCE, PIRLS, PISA e TIMSS - e problematizar a noção de qualidade que elas imprimem para a educação escolar.

O estudo permitiu observar simetrias entre tais iniciativas, principalmente em relação aos objetivos, como o diagnóstico obtido por meio de testes padronizados para possíveis ações de melhoria da qualidade da educação escolar, mas também em relação ao público-alvo, conteúdo avaliado e instrumentos utilizados.

Ainda que nos documentos não seja perceptível uma definição clara e objetiva de qualidade, o que seria absolutamente relevante tendo em vista a polissemia do conceito, conforme já indicado, é possível inferir que a qualidade propugnada pelos organismos responsáveis por essas avaliações se assenta quase que estritamente no desempenho dos estudantes. Mesmo que todas essas avaliações apliquem também, questionários contextuais que captam informações diversas que podem ser analisadas como fatores associados e explicativos do desempenho dos estudantes, há uma contundente ênfase nos resultados dos testes padronizados, que refletem a qualidade do ensino, aspecto importante, mas não único da qualidade da educação.

De alguma forma, a definição de qualidade incorporada pelo IDEB traz aspectos interessantes, embora aponte várias restrições quanto aos procedimentos de seu cálculo. Primeiramente, a qualidade no IDEB não é exatamente igual a resultados nos testes do SAEB, tampouco se associa completamente à ideia que fundamentou as políticas de ciclos e promoção automática da década de 1990, pois é uma combinação que procura evocar discussões evidenciadas no interior das escolas: como fazer os alunos aprenderem e “passarem de ano”? O IDEB tem ainda em sua formulação uma noção de avaliação ipsativa, isto é, cada escola seria comparada e avaliada em relação a ela mesma

mediante as metas estabelecidas. Adicionalmente, pode-se contestar que os únicos resultados considerados nos cálculos do IDEB sejam aqueles referentes à leitura e matemática, o que poderia ser interpretado como uma “redução curricular”; no entanto, de um lado, seriam conteúdos transversais a todas as disciplinas escolares e, de outro, são conteúdos que muito antes das avaliações externas já mereciam uma predominância sobre outros.

Nesse sentido, advoga-se que, mesmo não sendo o desempenho de estudantes em testes padronizados “a” qualidade da educação escolar, esta não se alcança sem aquele. Ademais, na tentativa de um entendimento que contemple a complexidade do conceito de qualidade da educação, buscando evitar o equívoco reducionista que identifica desempenho estudantil como seu sinônimo, defende-se qualidade não como excelência, mas como características que constituem e estruturam os sistemas educativos.

A literatura é eloquente em tecer críticas, em alguns casos severas, em relação ao desenho majoritário das avaliações em larga escala que privilegiam relevo à aferição do desempenho dos estudantes, conforme já sistematizado por Bauer, Alavarse e Oliveira (2015). Contudo, alertam os autores que “o papel que essas avaliações efetivamente assumirão no conjunto da política educacional é determinado pelos usos dos seus resultados”, sendo oportuna “a análise de como a literatura trata os usos dos resultados das avaliações em larga escala” (Bauer; Alavarse; Oliveira, 2015, p. 1376).

Dentre as críticas presentes na literatura que envolvem o papel de destaque das avaliações em larga escala, a aferição do desempenho dos estudantes e o desenvolvimento de ações de intervenção em aspectos específicos dos processos de ensino-aprendizagem, sublinha-se o currículo, com potencial de gerar consequências diretas na melhoria da qualidade da oferta da educação.

Como problematizado em Chappaz (2015), há, ainda que muitas vezes não seja ponderada, uma diferença fundamental existente entre a matriz de referência de uma iniciativa de avaliação em larga escala e a matriz curricular de uma rede de ensino – o que tem implicações diretas sobre a ideia de que as avaliações em larga escala têm impacto sobre o currículo das escolas no sentido de reduzi-los. Por isso, faz-se necessário refletir que, considerando uma iniciativa de avaliação em larga escala bem elaborada – do ponto de vista de critérios técnicos, metodológicos e pedagógicos –, os melhores resultados dos estudantes nas provas estariam refletindo um maior domínio sobre as habilidades e competências presentes nas matrizes de avaliação. Sendo esta, por sua vez, um recorte de uma matriz curricular, melhores resultados representariam um avanço na aprendizagem dos estudantes, podendo ser um efeito positivo das avaliações sobre qualidade da educação. Tal aspecto evidencia e corrobora, assim, a perspectiva de qualidade ponderada nesse texto, de modo que os resultados de avaliações em larga escala não são o todo da qualidade, mas parte importante do que a compõe.

A experiência brasileira de formulação e uso do IDEB, o principal indicador de qualidade educacional, com circulação e aderência em diferentes esferas – escolas, secretarias de educação, governos e mídia –, é uma referência fundamental para o debate sobre qualidade, pois se assenta num movimento internacional que identifica as avaliações em larga escala e seus resultados como instrumento significativo no desenho de estratégias e no desenvolvimento de políticas que visam alcançar objetivos de melhoria da eficiência e da qualidade da educação. Além disso, o IDEB, embora

não contemple todos os aspectos que compõem e impactam no processo pedagógico, avança na concepção de qualidade ao associar taxas de aprovação ao resultado de estudantes na avaliação em larga escala, indicando uma concepção de qualidade que valoriza, além do desempenho, a frequência às aulas e a permanência na escola.

Por fim, apesar da ênfase, em geral, das avaliações em larga escala quase que exclusivamente no resultado dos estudantes em testes padronizados, o reconhecimento da complexidade e amplitude do conceito de qualidade da educação e o direito de todos a uma educação de qualidade se impõem como desafio urgente das políticas educacionais.

Referências

ADAMS, Don; ACEDO, Clementina; POPA, Simona. Quality education: challenges and policy choices. In: ADAMS, Don; ACEDO, Clementina; POPA, Simona. (Ed.). *Quality and Qualities: Tensions in Education Reforms*. Rotterdam: Sense; Paris: UNESCO: International Bureau of Education, 2012. (Comparative and International Education: a Diversity of Voices, v. 16). p. 197-208.

AZEVEDO, Janete. A temática da qualidade e a política educacional no Brasil. *Educação & Sociedade*, Campinas, n. 49, especial, p. 449-467, 1994.

BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela. Avaliação em larga escala: uma sistematização do debate. *Educação & Pesquisa*, v. 41, n. especial, p. 1.367-1.382, 2015.
<https://doi.org/10.1590/S1517-9702201508144607>

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Brasil no PISA 2018*. Brasília: INEP, 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Diretoria de Avaliação da Educação Básica. *Sistema de Avaliação da Educação Básica: documentos de referência: versão preliminar*. Brasília: INEP, 2019.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP. *Avaliações e Exames Educacionais*, [s.d.].
<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais>

BROADFOOT, Patricia. *An Introduction to Assessment*. New York: Continuum, 2007.

CASASSUS, Juan. *A escola e a desigualdade*. Brasília: Líber Livro, 2007.

CHAPPAZ, Raíssa de Oliveira. *A Prova São Paulo e as tensões das avaliações externas: diálogos com o currículo oficial da rede municipal de ensino de São Paulo*. Dissertação (Mestrado em Educação) - Faculdade de Educação, Universidade de São Paulo, São Paulo, 2015.

CHAVES, Orlando. La cuestión de la calidad de la educación. *Boletín del Foro Latinoamericano de Políticas Educativas*, n. 26, p. 1-11, 2009.

DOURADO, Luiz Fernando (Coord.). *A qualidade da educação: conceitos e definições*. Brasília: INEP, 2007.

DOURADO, Luiz Fernandes; OLIVEIRA, João Ferreira de. A qualidade da educação: perspectivas e desafios. *Cad. Cedes*, v. 29, n. 78, p. 201-215, 2009.

FERNANDES, Reynaldo. Índice de Desenvolvimento da Educação Básica (*IDEB*). Brasília: INEP, 2007. (Série Documental. Textos para Discussão, n. 26).

FRANCO, Creso; ALVES, Fátima; BONAMINO, Alicia. Qualidade do ensino fundamental: políticas, suas possibilidades, seus limites. *Educação & Sociedade*, v. 28, n. 100, especial, p. 989-1014, 2007. <https://doi.org/10.1590/S0101-73302007000300017>

GIL, Natália de Lacerda. A quantificação da qualidade: algumas considerações sobre os índices de reprovação escolar no Brasil. *Sociologias*, ano 23, n. 56, p. 184-209, 2021. <https://doi.org/10.1590/15174522-109753>

HARVEY, Lee; GREEN, Diana. Defining quality. *Assessment & Evaluation in Higher Education*, v. 18, n. 1, p. 9-34, 1993.

MACHADO, Cristiane; ALAVARSE, Ocimar Munhoz. Qualidade das escolas: tensões e potencialidades das avaliações externas. *Educação & Realidade*, v. 39, n. 2, p. 413-436, 2014. <https://www.scielo.br/j/edreal/a/fyKb65xtFvXhMw3Hhpv9vNk/abstract/?lang=pt>

MAROY, Christian; VOISIN, Annelise. As transformações recentes das políticas de accountability na educação: desafios e incidências das ferramentas de ação política. Trad. Marisa Rossetto. *Educação & Sociedade*, v. 34, n. 124, p. 881-901, 2013. <https://doi.org/10.1590/S0101-73302013000300012>

MULLIS, Ina; MARTIN, Michael (Org.). *PIRLS 2021: fundamentos teóricos*. Trad. Patrícia Helena Freitag. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais, 2021.

NEVO, David. Avaliação por diálogos: uma contribuição possível para o aprimoramento escolar. In: TIANA, Alejandro (Coord.). *Anais do Seminário Internacional de Avaliação Educacional*, 1 a 3 de dezembro de 1997. Trad. John Stephen Morris. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais, 1998. p. 89-98.

OLIVEIRA, Romualdo Portela; ARAUJO, Gilda Cardoso. Qualidade do Ensino: uma nova dimensão da luta pelo direito à educação. *Revista Brasileira de Educação*, n. 28, p. 5-23, 2005. <https://www.scielo.br/j/rbedu/a/t64xS8jD8pz6yNFQNCk4n7L/?format=pdf&lang=pt>

RISOPATRON, Veronica Edwards. *El concepto de calidad de la educación*. Chile: UNESCO/Orealc, 1991.

SILVA, Vandrê Gomes da. *Por um sentido público da qualidade da educação*. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2008.

UNESCO. *¿Qué se espera que aprendan los estudiantes de América Latina y el Caribe?: análisis curricular del Estudio Regional Comparativo y Explicativo (ERCE 2019)*. Santiago: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO): Oficina Regional de Educación para América Latina y el Caribe (Orealc/UNESCO Santiago), 2020.

VERGER, Antoni; FONTDEVILA, Clara; PARCERISA, Luís. Reformando a governança através de instrumentos de políticas públicas: como e em que medida padrões, testes e accountability na educação se espalham pelo mundo. In: OLIVEIRA, Dalila Andrade; DUARTE, Adriana Maria Cancellia; RODRIGUES, Cibele Maria Lima (Org.). *A política educacional em contexto de desigualdade: uma análise das redes públicas de ensino da região Nordeste*. Campinas: Mercado de Letras, 2019.

VERGER, Antoni; NORMAND, Romuald. Nueva gestión pública y educación: elementos teóricos y conceptuales para el estudio de un modelo de reforma educativa global. *Educação & Sociedade*, v. 36, n. 132, p. 599-622, 2015. <https://doi.org/10.1590/ES0101-73302015152799>

CRISTIANE MACHADO

Doutora em Educação, Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; Professora Associada (Livre Docente), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brasil.

OCIMAR MUNHOZ ALAVARSE

Doutor em Educação, Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; Professor, Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil.

RAÍSSA DE OLIVEIRA CHAPPAZ

Doutora em Educação Universidade de São Paulo (USP), São Paulo, São Paulo, Brasil; Pedagoga, Instituto Federal de São Paulo (IFSP), São Paulo, São Paulo, Brasil.

CONTRIBUIÇÃO DOS AUTORES

Autor 1 – Concepção e desenho da pesquisa; análise e interpretação dos dados; redação de tópicos do texto; revisão final do artigo.

Autor 2 – Concepção e desenho da pesquisa; levantamento, processamento, análise e interpretação dos dados; revisão final do artigo.

Autor 3 – Concepção e desenho da pesquisa; construção e processamento dos dados; análise e interpretação dos dados; elaboração do texto final.

APOIO/FINANCIAMENTO

Não houve financiamento.

DISPONIBILIDADE DE DADOS DE PESQUISA

Não se aplica.

COMO CITAR ESTE ARTIGO

MACHADO, Cristiane; ALAVARSE, Ocimar Munhoz; CHAPPAZ, Raíssa de Oliveira. Avaliações em larga escala na América Latina: problematizando a qualidade da educação escolar. *Educar em Revista*, Curitiba, v. 40, e94031, 2024. <https://doi.org/10.1590/1984-0411.94031>

O presente artigo foi revisado por Fernanda Silva e Sousa. Após ter sido diagramado, foi submetido para validação do(s) autor(es) antes da publicação.

Recebido: 10/01/2024

Aprovado: 19/08/2024