

AUSTRÁLIA, BRASIL E CANADÁ: IMPACTO DAS AVALIAÇÕES NO ENSINO DE CIÊNCIAS

PAULO SÉRGIO GARCIA

XAVIER FAZIO

DEBRA PANIZZON

NELIO BIZZO

RESUMO

A cultura global de testes se expandiu rapidamente nas últimas décadas. Nessas avaliações, as áreas mais comumente testadas são as de Linguagem e Matemática (LM). O foco e as políticas de responsabilização têm trazido consequências para o ensino de Ciências. Neste estudo reúnem-se evidências do Brasil, do Canadá e da Austrália para compreender as implicações dessas políticas para o ensino de Ciências, com base em pesquisas recentemente publicadas. Os resultados mostraram que a cultura dos testes, com ênfase em LM, e as políticas de responsabilização têm gerado o descaso e a negligência com o ensino de Ciências, causado sobretudo pelos sistemas de ensino e pelos gestores escolares, que criam iniciativas, ações e projetos voltados para atingir os resultados e metas em LM.

PALAVRAS-CHAVE ENSINO DE CIÊNCIAS • TESTES • POLÍTICAS EDUCACIONAIS • RESPONSABILIZAÇÃO.

AUSTRALIA, BRASIL Y CANADÁ: IMPACTO DE LAS EVALUACIONES EN LA ENSEÑANZA DE CIENCIAS

RESUMEN

La cultura global de pruebas se expandió rápidamente en las últimas décadas. En estas evaluaciones, las áreas más comúnmente probadas son las de Lenguaje y Matemáticas (LM). El enfoque y las políticas de responsabilización han provocado consecuencias en la enseñanza de Ciencias. En este estudio se reúnen evidencias de Brasil, Canadá y Australia para comprender las implicaciones de tales políticas para la enseñanza de Ciencias, en base a investigaciones recientemente publicadas. Los resultados mostraron que la cultura de las pruebas, con énfasis en LM, así como las políticas de responsabilización, han generado desatención y negligencia en la enseñanza de Ciencias, sobre todo en función de los sistemas de enseñanza y los gestores escolares, que crean iniciativas, acciones y proyectos que se destinan a alcanzar los resultados y metas en LM.

PALABRAS CLAVE ENSEÑANZA DE CIENCIAS • PRUEBAS • POLÍTICAS EDUCACIONALES • RESPONSABILIZACIÓN.

AUSTRALIA, BRAZIL AND CANADA: THE IMPACT OF ASSESSMENT ON SCIENCE TEACHING

ABSTRACT

The global testing culture has expanded rapidly in recent decades. In these assessments, the most commonly tested areas are those of Language and Mathematics (LM). The focus and accountability policies have brought consequences for the teaching of the sciences. Based on recently published research, the present study gathers evidence collected in Brazil, Canada and Australia in order to understand the implications of these policies for the teaching of the sciences. The results showed that both the culture of testing, with its emphasis on LM, and accountability policies have generated a disregard and neglect of science teaching. This situation is caused mainly by the education systems and the school managers who create initiatives, actions and projects aimed at achieving LM results and targets.

KEYWORDS SCIENCE TEACHING • TESTS • EDUCATIONAL POLICIES • ACCOUNTABILITY.

INTRODUÇÃO

Com a importância da educação crescendo globalmente, o debate internacional tem mudado da questão de fornecer acesso à educação para a de garantir a eficiência e a equidade nos resultados educacionais. Para certificar que o processo educacional seja de alta qualidade, no contexto da expansão global do ensino básico universal, políticas de responsabilização sob a forma de testes acadêmicos foram lançadas em quase todas as partes do mundo (SMITH, 2014).

Nesse sentido, nas últimas décadas tem acontecido uma grande expansão, em nível global, da cultura dos testes. Um contexto em que estudantes, em várias localidades, regiões e países, têm sido expostos a testes padronizados de avaliações em larga escala (SONGER; RUIZ-PRIMO, 2012).

Essa cultura dos testes tem transpassado vários aspectos da educação, influenciando questões como: financiamento, currículo escolar, gestão da escola, responsabilização dos atores e crenças e práticas dos professores, aulas dos alunos, envolvimento dos pais, entre outras. É necessário afirmar também que essa tendência do uso dos testes não se

limita aos países industrializados, uma vez que aqueles que discutem e realizam as políticas educacionais insistem que a implantação desses sistemas de testes (nacionais, estaduais e locais) é uma estratégia importante, talvez a chave para melhorar a qualidade educacional (CHAPMAN; SNYDER, 2000).

Os resultados dessas avaliações de larga escala (ALE), dos testes, têm induzido muitos gestores, professores, pais e mídia à percepção e ao entendimento de que os resultados são sinônimos de qualidade educacional, como se a avaliação dos alunos em algumas áreas de ensino (Linguagem, Matemática) pudesse demonstrar e garantir um panorama de excelência escolar.

A utilização de testes no campo educacional não se constitui em uma novidade das últimas décadas. Os testes têm sido usados como estratégia para orientar o ensino e a avaliação e, sobretudo, para identificar e acompanhar o rendimento dos alunos já há muito tempo. No entanto, a partir da segunda metade do século XX, mudanças aconteceram no tipo de teste aplicado aos estudantes, deslocando, pelo menos de forma parcial, a responsabilidade pelos resultados dos alunos para as escolas e para os professores (SMITH, 2014). Tal situação ocasionou um processo de maior responsabilização (*accountability*) daqueles que atuam nas escolas, sobretudo professores e gestores, um fato que começou a acontecer, primeiramente, nos Estados Unidos e no Reino Unido (DORN, 2007) e depois se espalhou pelo mundo. A partir da década de 1980, de acordo com informações da rede Eurydice (EUROPEAN COMMISSION, 2009), quase todos os países europeus já tinham adotado algumas políticas e testes censitários nacionais.

Essa expansão pode ser vista entre os anos de 1995 a 2006. Nesse período, o número de países no mundo que participou de algum tipo de programa de teste anual mais que dobrou, passando de 28 para 67, de acordo com dados de Benavot e Tanner (2007). Na América Latina, atualmente 16 nações apresentam algum tipo de programa de avaliação de larga escala. Países como o Chile, México, Colômbia e Brasil avaliam seus alunos regularmente, no entanto há variações na forma de uso dos resultados. No Brasil, com a utilização do

Sistema de Avaliação da Educação Básica (Saeb) e da Prova Brasil, desde 1990, o número de testes aplicados em jovens brasileiros chega a, aproximadamente, 19 milhões (HORTA NETO, 2013).

Sem negar o valor da avaliação para o trabalho pedagógico, pois permite aos sistemas e às redes de ensino possibilidades de utilização – definir padrões e expectativas para o aprendizado dos alunos, orientar o trabalho das escolas e dos docentes, gerar dados para os professores para a tomada de decisões sobre o progresso dos jovens, entre outras (BAUER; ALAVARSE; OLIVEIRA, 2015) –, há hoje grande discussão sobre as limitações e as consequências dessa cultura dos testes: a forma como são construídos e validados, a orientação e a legitimação da política educacional que, nesse caso, desconsidera o processo educacional e responsabiliza a escola e os professores pelos resultados escolares, entre outros. Como sinaliza Freitas (2013), as avaliações direcionam políticas de responsabilização, sobretudo quando integram políticas de alto impacto (*high stakes tests*), tal como o pagamento de bônus aos professores. Além disso, esse processo de responsabilização do docente e das escolas pode, em alguns casos, levar a medidas punitivas e injustas, como a transferência ou a perda do emprego de um profissional (gestor ou professor), produzindo injustiças associadas às premiações das melhores escolas (BAUER, 2012; FREITAS, 2013).

O contexto da cultura dos testes tem induzido os sistemas educativos à busca por melhores resultados em vez de melhora da qualidade e da equidade educativa (IAIES, 2003, p. 18), um processo que tem interferido na autonomia dos professores e na forma como os conteúdos têm sido ensinados. Em muitos casos, os professores atuam preparando os alunos para os exames, utilizando metodologias para atingir maiores rendimentos nas ALE. Ou seja, a fim de conseguir resultados significativos, os docentes posicionam o conteúdo e as formas de ensino sobre a malha de conhecimentos que é cobrada nos exames, deixando de lado outros assuntos, habilidades e competências que são essenciais para a formação dos alunos (MADAUS; RUSSEL; HIGGINS, 2009; HYPÓLITO, 2013; SANTOS, 2013). Um processo claro de empobrecimento

e afunilamento curricular (BAUER, 2013; MADDAUS; RUSSELL; HIGGINS, 2009).

Outra consequência da cultura dos testes é que essas ALE possibilitam a criação de *rankings* que induzem a competição entre as escolas e, nesse caso, podem deslocar o foco do aprendizado para melhoria dos resultados. Essas avaliações podem também gerar um ensino que considere a difusão de “macetes” para potencializar as notas dos estudantes (SANTOS, 2013). Elas podem ainda induzir o aumento da desigualdade, desencadeando um processo no qual investe-se mais nos melhores alunos, como uma estratégia mais favorável na busca por melhores resultados, do que naqueles que apresentam dificuldades de aprendizado e podem comprometer o rendimento da escola (OLIVEIRA, 2013).

Há também, como consequência, a questão da indução de gestores, educacionais e escolares, a investir maiores recursos físicos, intelectuais e financeiros nas disciplinas-alvo dos testes padronizados, em detrimento das outras matérias. Trata-se do caso de que esses testes, com foco apenas em Linguagem e Matemática, deixam de lado conhecimentos e habilidades relacionadas a Ciências, por exemplo, o que, possivelmente, traz consequências para o seu ensino. Nesse sentido, este estudo pretende averiguar as consequências dessas políticas de testes padronizados e de responsabilização para o ensino de Ciências. Para tal, reuniram-se evidências de três realidades de países diferentes (Austrália, Brasil e Canadá), utilizando pesquisas recentemente publicadas nesses contextos. Trata-se de um estudo internacional que apresenta alguns dados empíricos.

POLÍTICAS NEOLIBERAIS, O CONTEXTO DA EDUCAÇÃO E A CULTURA DOS TESTES

Políticas neoliberais indicaram tendências na organização e na gestão de políticas sociais, em geral, e educacionais, em particular. Tais programas sociais e de educação, já existentes nos países desenvolvidos no final da década de 1970 e início de 1980, atingiram outras nações por meio dos processos de globalização e formas de recontextualização, a partir

da década de 1990 (América Latina e Brasil), iniciando um processo denominado de “paradigma da educação contábil” (LIMA, 1997, p. 43).

No campo da educação, verificou-se o surgimento de um grande número de iniciativas distintas – processo de privatização e de desregulação, descentralização e reorganização do sistema para flexibilizar a oferta educacional, indução de um mercado educacional, políticas de livre escolha, contratos de gestão, compra de materiais didático-pedagógicos do setor privado, políticas de avaliação institucional e de responsabilização, vigilância da qualidade educacional por meio do controle da eficácia e da eficiência das escolas, criação de sistema de méritos (meritocracia), novas estratégias de gestão escolar baseada em conceitos de qualidade total, reformulação do currículo, indução de lógica competitiva entre as escolas, alteração nas práticas pedagógicas dos professores, entre outros – que, de forma geral, instituíram um discurso ideológico e tecnocrático já consolidado no mundo empresarial. Tal discurso obscureceu a linguagem da pedagogia e vem colonizando a política educacional pelos imperativos da economia (BALL, 1999). Trata-se da influência da ciência econômica e da economia capitalista na educação, em que, em muitos processos, as decisões sobre as políticas educacionais têm como base a economia, e não mais a pedagogia. Essa visão mercadológica da educação se confronta com o ideal de democratização e de gestão escolar estabelecidos no Brasil no final da década de 1980, por meio do ordenamento jurídico da educação.

Nesse contexto, ocorreu a criação e a disseminação da cultura dos testes, que atenderam e atendem a grandes interesses econômicos de empresas nacionais e multinacionais que criam e aplicam os testes e também atuam na negociação e venda de materiais didáticos. Um grande mercado de consumo da educação (HAGOPIAN, 2014).

Para Arelaro (2003) e Esteban (2012), as avaliações em larga escala vêm responder às pressões para mudanças nos modos de administração e de controle das redes de ensino, estando aliadas a um “novo” modelo de gestão educacional pública. Tais pressões, oriundas de organismos multilate-

rais, visam a impor uma agenda de educação para os países subdesenvolvidos. As ALE apresentam, portanto, um papel político, que tem relevância na análise de suas finalidades e de seus efeitos sobre os sistemas educativos em que são aplicados (OZGA, 2000).

As discussões sobre a validade desses testes padronizados (*standards*) apresentam muitas controvérsias. Stecher (2002, p. 99-100) sinaliza que o efeito líquido dessas avaliações de alto impacto sobre a política e a prática ainda é duvidoso, considerando o alinhamento curricular, a alocação do tempo de aula para enfatizar os assuntos abordados nos exames, o treinamento excessivo e a trapaça.

Há também discussões sobre os aspectos técnicos dos testes (validade, adequação dos instrumentos e confiabilidade dos resultados). Há dúvidas ainda se esses testes conseguem medir, em termos de aprendizagem dos alunos, o centro daquilo para o que foram projetados e com a acuidade anunciada de seus resultados (CASASSUS, 2013). Outra sinalização diz respeito à seleção dos conteúdos e habilidades que embasam os testes e, ao mesmo tempo, a definição de níveis ou padrões de rendimento. Nesse sentido, questionam-se as diretrizes na elaboração das matrizes e a definição de parâmetros de avaliação (CASASSUS, 2013; IAIES, 2003).

Há também dúvidas quanto ao potencial dos testes com o propósito de melhorar a qualidade do ensino, pois apesar de muitos esforços realizados, sobretudo pelos países latino-americanos, os resultados mostram pouca alteração no quadro educacional da maioria dos países que utilizam as avaliações como instrumento de gestão (TEDESCO, 2003).

Mons (2009) e Ravitch (2010) sinalizam algumas consequências dessas políticas de avaliação: a tomada de decisões sobre os recursos financeiros a serem endereçados para as escolas, a bonificação de professores para estabelecimentos de ensino que atingirem os resultados, a indicação de caminhos para o gerenciamento das redes escolares, a contratação e a demissão de gestores escolares, entre outros.

Madaus, Russell e Higgns (2009) afirmam que alunos e professores, quando pressionados pelos testes, podem sofrer problemas de saúde (estresse), perder a motivação, ou criar

um sentimento negativo em relação à escola. Os mesmos autores afirmam também que essa pressão pela melhoria do rendimento escolar pode induzir os gestores a outras formas de gerenciamento dos tempos e conteúdos, focando nos resultados, e tal situação pode influenciar de maneira negativa tanto docentes quanto discentes.

Por fim, outra questão polêmica é a divulgação das notas das escolas nesses testes padronizados. Um processo que não atenta para a complexidade dos resultados, assim como para o contexto em que eles foram produzidos (BRUNNER, 2003). Dessa forma, banalizam e, o que é mais preocupante, distorcem a opinião pública. As avaliações não levam em consideração as condições de partida de cada escola, a infraestrutura, os profissionais, suas formações, entre outros.

UMA VISÃO GERAL DA REALIDADE BRASILEIRA

No Brasil, a responsabilidade por organizar e integrar as políticas educacionais é do Ministério da Educação (MEC). Trata-se de um órgão da administração federal direta, que, por meio de suas secretarias, coordena, entre outras coisas, as orientações para o currículo, a avaliação, a formação de professores relacionada à educação básica e ao ensino superior. Atua também na informação e na pesquisa educacional. Estados e municípios também estabelecem suas diretrizes para seus sistemas, baseados nas políticas e indicações do MEC.

No país, na década de 1990, foi criado o Saeb, com o objetivo de avaliar a qualidade do ensino das escolas brasileiras. A portaria n. 482, de 7 de junho de 2013, do MEC, dispõe sobre o Saeb, revelando que o sistema passa a ser composto pela Avaliação Nacional da Educação Básica (Aneb), Avaliação Nacional do Rendimento Escolar (Anresc) e Avaliação Nacional da Alfabetização (ANA).

Em 1990, ocorreu a primeira aplicação do Saeb, que contou com a participação, de forma amostral, de escolas públicas urbanas com primeira, terceira, quinta e sétima séries do ensino fundamental. Os jovens foram avaliados em Língua Portuguesa, Matemática e Ciências. Em 1995 foi implantada a metodologia

da Teoria da Resposta ao Item (TRI), permitindo a comparabilidade entre os resultados das avaliações ao longo do tempo. Nesse ano, os alunos avaliados foram aqueles das quarta e oitava séries do ensino fundamental e do terceiro ano do ensino médio, incluindo uma amostra da rede privada. No entanto, nesse ano não aconteceu a avaliação em Ciências. De 1997 até 1999, os estudantes das quarta e oitava séries foram avaliados em Língua Portuguesa, Matemática e Ciências, e os estudantes de 3º ano do ensino médio, em Língua Portuguesa, Matemática, Ciências, História e Geografia. A partir de 2001, o exame do Saeb passou a focar exclusivamente as áreas de Língua Portuguesa e Matemática. Uma configuração que se mantém até os dias atuais (HORTA NETO, 2007; FREITAS, 2007).

Em 2013, no entanto, foi sinalizado que a disciplina de Ciências passaria a fazer parte novamente do Saeb, e então foi publicado o primeiro caderno de “Inclusão de Ciências no Saeb”, um documento básico criado pelo Ministério de Educação. Ao mesmo tempo, nesse ano, alunos do ensino fundamental II¹ e Médio participariam do exame de forma experimental.

De acordo com o documento:

A proposta apresentada assume a natureza de matriz piloto a ser refletida, desenvolvida, testada em 2013 e aperfeiçoada com vistas a sua consolidação em 2015. Nesse sentido, a matriz desenvolvida refere-se apenas ao 9º ano do ensino fundamental, referenciando a aplicação piloto no final do ciclo. As matrizes para o 5º ano e para a 3ª série do ensino médio serão desenvolvidas em seguida, após a análise de resultados da aplicação piloto, de forma a servir de parâmetro para a construção e a revisão das demais matrizes do Saeb. (BRASIL, 2013, p. 12)

O documento sinalizou também que, a partir de 2015, o exame de Ciências seria introduzido no Brasil. No entanto, apesar de todo o esforço, a inclusão dessa disciplina não foi consolidada. Em 2015, as avaliações realizadas incidiram somente sobre as áreas de Português e Matemática (PM).

As disciplinas de Português e Matemática, exclusivamente, também são avaliadas em outras avaliações em larga

¹ Adotou-se essa nomenclatura “ensino fundamental II” para se referir aos anos finais do ensino fundamental: sexto, sétimo, oitavo e nono anos.

escala no Brasil. Os alunos são testados nacionalmente nessas disciplinas, no segundo ano, pela Provinha Brasil, uma avaliação diagnóstica aplicada duas vezes ao ano; no terceiro, pela Avaliação Nacional da Alfabetização, que se insere no contexto de atenção voltada à alfabetização; nos quinto e nono anos do ensino fundamental, pela Prova Brasil. Além dessas, estados e municípios têm suas próprias avaliações em PM.

Tal expansão dessas ALE ocorreu, principalmente, depois da criação do Índice de Desenvolvimento da Educação Básica (Ideb), em 2007, que combina o desempenho dos estudantes na Prova Brasil, para os municípios, e no Saeb, para os estados, acrescido ao fluxo escolar (BAUER, 2012; SOUSA, 2013; BONAMINO, 2013).

O estudo de Brooke, Cunha e Faleiros (2011) mostrou que, no país, já existiam iniciativas de ALE em 19 estados brasileiros. Nos municípios, a pesquisa de Bauer, Alavarse e Oliveira (2015), com 4.309 cidades, revelou que as cidades estão se incorporando às iniciativas federais – ANA, Provinha Brasil, Prova Brasil – e estaduais – Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saesp)² – de ALE. Paralelamente, os municípios também estão construindo suas próprias ações e programas de avaliação em Português e Matemática, “sejam esses concebidos pelas secretarias de educação, ou por empresas contratadas pela administração municipal” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 329), permitindo a realização de estudos sobre as características e os usos dos resultados dessas avaliações para formulação e implementação das políticas educacionais municipais (SOUSA; PIMENTA; MACHADO, 2012; ALAVARSE; MACHADO; BRAVO, 2013; GIMENES et al., 2013).

Se, no contexto brasileiro, as avaliações em larga escala, após 2001, estão assentadas exclusivamente nas disciplinas Português e Matemática, no contexto internacional, no entanto, o Brasil tem participado do Programa Internacional de Avaliação de Estudantes (PISA), no qual está inserida a avaliação em Ciências. Trata-se de um programa contínuo que objetiva o desenvolvimento de informações para o monitoramento de conhecimentos e de habilidades dos jovens

2 Avaliação aplicada pela Secretaria da Educação do Estado de São Paulo com a finalidade de produzir um diagnóstico da situação da escolaridade básica e, ao mesmo tempo, orientar os gestores escolares e monitorar as políticas associadas à qualidade educacional.

em vários países e em diferentes subgrupos demográficos de cada nação (ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OCDE, 2016).

O PISA é uma avaliação trienal e amostral em Leitura, Matemática e Ciências de alunos com idade entre 15 anos e 3 meses e 16 anos e 2 meses completos no início do período de aplicação. Esses jovens necessitam estar matriculados e cursando no mínimo a sétima série/ano. No último exame participaram 35 países da OCDE e mais 35 de economias parceiras, incluindo o Brasil.

Essa avaliação utiliza instrumentos, testes e questionários que resultam em indicadores de um perfil de conhecimentos e habilidades dos jovens. O programa também tem indicadores provenientes de questionários, variáveis demográficas, econômicas, sociais e educacionais e de tendências que monitoram o rendimento dos alunos e dos sistemas de educação. Em 2015, ano da última avaliação, participaram jovens de 27 estados, perfazendo um total de 23.141 alunos de 841 escolas, e, pela primeira vez, a aplicação do exame foi realizada totalmente pelo computador.

Os resultados de Ciências, nas últimas edições, não foram animadores para a realidade brasileira. O Brasil obteve os mesmos 405 pontos em 2009 e 2012, e, em 2015, apresentou uma queda para 401. Este escore médio (401) foi um valor significativamente menor do que a média dos alunos dos países membros da OCDE (493). Jovens da rede de ensino estadual conseguiram 394 pontos, e os da rede federal obtiveram o melhor desempenho (517), superando a média nacional. O estado do Espírito Santo apresentou o maior desempenho (435), e o de Alagoas, o pior (360). O desempenho médio dos meninos foi maior do que o das meninas na maioria dos estados do Brasil.

UMA VISÃO GERAL DA REALIDADE CANADENSE

O Canadá não tem um departamento federal ou um sistema nacional integrado de educação. Os ministérios de educação das respectivas províncias e territórios, criados em 1967, formaram uma organização intergovernamental

denominada de Conselho de Ministérios da Educação do Canadá (CMEC). Tal organização tem o intuito de prover liderança nas questões educacionais em todo o país e representá-lo em nível internacional, nomeadamente no que se refere à realização de comparações realizadas por meio de testes (CANADÁ, 2016a).

Cada uma de suas jurisdições dispõe de um departamento ou ministério de educação que realiza a organização, provisão e avaliação educacional nos níveis elementar e secundário. Dessa forma, cada província ou território apresenta um currículo próprio para o ensino, inclusive de Ciência e Tecnologia. Cada província tem também seu próprio sistema de avaliação (KLINGER; DELUCA; MERCHANT, 2016).

No Canadá, o Pan-Canadian Assessment Program (PCAP) (CANADÁ, 2016b) é o único teste nacional para o desempenho dos alunos de 13 anos (anteriormente àqueles de 16 anos) em Ciências. As províncias têm autonomia para analisar seu currículo e avaliar as escolas de ensino elementar (alunos de 6 a 14 anos) e secundário (de 14 a 18 anos). No país, várias políticas foram desenvolvidas, e recursos significativos são usados para promover a melhoria das escolas com base nos resultados das avaliações em larga escala.

O Canadá participa de testes de comparação internacional para estudantes das escolas elementares. No ensino de Ciências, trata-se do PISA e do Trends in International Mathematics and Science Study (TIMSS) (INTERNATIONAL ASSOCIATION FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT – IAEEA, 2012). Dessas comparações provêm dados internacionais que são utilizados para monitorar as mudanças do sistema educacional ao longo do tempo. Ao mesmo tempo é possível acompanhar os resultados, fomentar discussões sobre eles e analisar as políticas educacionais. Na província de Ontário, o foco das ALE é Linguagem e Matemática, e não há exames na disciplina de Ciências. Os alunos que participam dos testes estão no terceiro e no sexto ano (EDUCATIONAL QUALITY AND ASSESSMENT OFFICE – EQAO, 2016). As políticas de responsabilização e as ALE, com foco em LM, tiveram grande impacto no cenário educacional. Muitos recursos são usados, e políticas, desenvolvidas para promover a melhoria

com base em notas de avaliação em larga escala para as áreas de LM (VOLANTE; BEN JAAFAR, 2008).

UMA VISÃO GERAL DA REALIDADE AUSTRALIANA

A responsabilidade pela educação na Austrália é compartilhada, e recai sobre os seis Estados e os dois Territórios, resultando em oito órgãos ou departamentos educacionais diferentes que se responsabilizam, entre outras coisas, por monitorar os currículos e as avaliações dos alunos e por padronizar as normas educacionais. No país, não existe um único departamento de educação responsável por todos os programas educacionais. Trata-se de um processo descentralizado de administração educacional, que possibilita que cada jurisdição realize seus atendimentos relacionados às necessidades das crianças e dos jovens, estruturando e organizando currículos compatíveis com o contexto educacional específico (REID, 2005). No entanto, tal sistema se transformou, ao longo do tempo, em oito currículos diferentes em vigor no país. Apesar de receber o apoio de muitos educadores, pesquisadores e gestores escolares, essa proposta descentralizada carece de consistência nacional.

Na Austrália, existe somente uma avaliação em larga escala (teste), obrigatória em todo território nacional para todos os estudantes das escolas elementares. Trata-se do Programa Nacional de Avaliação: Linguagem e Matemática (National Assessment Program: Literacy and Numeracy – NAPLAN). O exame é administrado todo ano aos alunos da terceira e da quinta série (escolas elementares, na maioria dos estados e territórios), sétima e nona série (escolas secundárias na maioria dos estados e territórios), muito embora os familiares possam não permitir, ou seja, recusar que seus filhos realizem tais testes.

No país, o exame realizado pelo NAPLAN avalia os alunos em quatro domínios diferentes, relacionados à leitura, à escrita, às convenções da linguagem (soletração, gramática) e à aritmética (Matemática). No entanto, não existe na Austrália um exame nacional e obrigatório da disciplina de Ciências para todos os estudantes das escolas elementares.

A Austrália conta com o Programa Nacional de Avaliação de Alfabetização Científica (*National Assessment Program – Science Literacy – NAP-SL*) para avaliar os alunos em relação aos processos científicos. Trata-se de um exame que foi inicialmente implantado em 2003 para avaliar a alfabetização científica de todos os jovens. Até a presente data, o NAP-SL é administrado, a cada três anos e de forma amostral, a alunos da sexta série (jovens com 11 e 12 anos da escola elementar) de todo o território australiano selecionados randomicamente. No entanto, há uma grande discussão no país para que esse exame seja associado e incorporado aos testes realizados pelo NAPLAN. Tal associação permitiria que o exame passasse a ser anual e censitário.

A Austrália, assim como o Canadá, participa do PISA e do TIMSS, avaliações internacionais do conhecimento dos alunos em Matemática e Ciências. Os resultados do TIMSS de 2015 foram publicados recentemente (THOMSON; BORTOLI; UNDERWOOD, 2016), com a participação do quarto ano da escola elementar e do oitavo ano da escola secundária na Austrália.

O CONTEXTO DO ENSINO DE CIÊNCIAS

Em uma sociedade que valoriza o conhecimento científico e as tecnologias, o ensino de Ciências tem um papel fundamental para a formação de um cidadão crítico e transformador. Trata-se de um espaço privilegiado para a reflexão sobre o mundo, os fenômenos da natureza e a ação do homem. Em tal espaço, os jovens podem avaliar explicações, possibilitando o desenvolvimento de uma postura reflexiva, crítica, questionadora e investigativa. Esse ensino também favorece à compreensão de conceitos sobre as questões éticas implícitas nas relações entre Ciência, Sociedade e Tecnologia. Além disso, sua abrangência inclui a compreensão da crise ambiental, a destruição dos seres humanos, as questões do consumismo, o destino dado ao lixo industrial e doméstico, a manipulação gênica, entre outros.

Em muitos países o Ensino de Ciência, e em graus diferentes a tecnologia, é um dos elementos-chave para a escolarização. No entanto, esse ensino tem enfrentado vários

desafios, entre eles o desinteresse dos alunos por essa disciplina. Há também a questão do baixo interesse dos jovens em seguir as carreiras científicas, que, segundo pesquisadores, está associado às aulas de Ciências (AIKENHEAD, 2005; BERNARDO et al., 2008; BROK et al., 2006; GOUW; BIZZO, 2016), ou seja, à forma como as aulas são ministradas, sobretudo baseadas no livro didático e na memorização.

Além disso, a conjuntura atual pode ser ainda mais agravada em consequência da criação das políticas de avaliação em larga escala e de responsabilização da escola e dos professores, que focam, no Brasil, exclusivamente, as áreas de Linguagem e Matemática, deixando o ensino de Ciências em segundo plano.

METODOLOGIA DO ESTUDO

Este estudo, parte de um projeto internacional de pesquisadores na Austrália, no Brasil e no Canadá, buscou evidências, em três realidades distintas, para compreender as implicações de um cenário marcado por políticas de avaliação em larga escala, com o foco em Linguagem e Matemática, e de responsabilização das escolas para a disciplina de Ciências.

Em estudos comparativos internacionais na área de educação, é necessário reconhecer as diferenças políticas, culturais e históricas de cada país para que se avalie, de fato, o que pode ser analisado, compartilhado e incorporado. Phillips (2000) realizou uma discussão relevante sobre os motivos para a realização de estudos comparativos na área de educação. Para o autor, a proficuidade desses estudos se deve, entre outras coisas, à descoberta de possíveis alternativas para a organização dos sistemas, das redes de ensino e das escolas; às análises das consequências das reformas, avaliando experiências similares em vários países; à avaliação da possibilidade da promoção de cooperação entre as nações, a partir da discussão social, política e cultural sobre as similaridades e as diferenças.

A seleção dos países participantes decorre do fato de o Canadá e a Austrália serem parte da Commonwealth, uma organização internacional com 54 membros independentes.

Tais países atuam na cooperação inseridos em um quadro de valores e objetivos confirmados na declaração de Singapura, incluindo a promoção da democracia, a busca pelos direitos humanos e pela liberdade individual, o comércio livre, o multilateralismo e a paz entre as nações. Esses dois países não apresentam sistemas federal ou nacional de educação. Tal situação resulta na criação, em cada estado ou jurisdição, de um departamento de educação independente, o que não acontece no Brasil, que dispõe de um ministério da educação nacional para coordenar as diretrizes políticas, de currículo e de avaliação em todo território. Dessa forma, embora possam existir algumas semelhanças entre o Canadá e a Austrália, há também muitas diferenças, propiciando uma comparação interessante com o Brasil, que não tem o mesmo histórico de educação que os outros dois países. Além disso, a seleção de tais contextos dá continuidade a uma série de estudos realizados pelos pesquisadores nas três realidades (FAZIO; GARCIA; PANIZZON, 2008; GARCIA; FAZIO; PANIZZON, 2011; GARCIA; BIZZO; FAZIO, 2012; GARCIA; BIZZO; FAZIO, 2014), tendo como marco a formação de professores de Ciências, os desafios da formação contínua a distância para docentes que atuam na área de Ciências e, nesse caso, a avaliação educacional.

Nos três países, os pesquisadores coletaram evidências a partir de estudos e pesquisas recentemente publicadas. Da Austrália, foram utilizadas evidências do relatório *TIMSS 2015: a first look at Australia's results* (THOMSON; BORTOLI; UNDERWOOD, 2016), do Conselho Australiano de pesquisa educacional (Australian Council for Educational Research – ACER). Trata-se de uma organização que visa a promover conhecimento baseado em pesquisas para ser usado com o intuito de melhorar a aprendizagem ao longo da vida.

Do Canadá, região de Ontário, foram utilizadas evidências reunidas na pesquisa *Exploring the Current State of Grades 4 to 8 Science Education in Ontario* (DOW, 2015), um estudo com metodologia mista com mais de 200 professores de Ciências da região, e o estudo de Fazio e Karrow (2013), *Science takes back seat*, que, além de entrevistar docentes, analisou também dados do TIMSS de 2015 para substanciar as análises.

Do Brasil, da região do grande ABC, foram utilizados dados dos estudos: *Educational policies and science education in Brazil: a case study* (GARCIA; BIZZO, 2015), que analisou a percepção dos jovens quanto à importância das disciplinas de Português, Matemática e Ciências estudadas cotidianamente em um contexto marcado por políticas públicas que enfatizam exclusivamente a Linguagem e a Matemática; *Efeitos Ideb na região do grande ABC sobre o ensino de Ciências*, realizada pelo Observatório da Educação do Grande ABC (OEGABC, 2015), com análise do Ideb e suas consequências no ensino de Ciências; *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências* (OEGABC, 2016), que, por meio de um estudo de caso, analisou a percepção de 49 professores de Ciências dessa região, uma das mais ricas do país, sobre a percepção acerca das políticas atuais de avaliação. É necessário reconhecer, no caso desses estudos, que se trata de uma amostra particularizada, de uma região, e que, portanto, não pode ser tomada como possuidora de representatividade nacional.

Como amostra nacional foram analisados os dados do PISA, na área de Ciências, das últimas quatro edições (2006, 2009, 2012 e 2015), que são os resultados publicados logo depois da criação do Ideb no Brasil, a principal política nacional em relação à responsabilização das escolas e dos professores com foco apenas nos testes de Português e Matemática.

Por fim, foram utilizados e avaliados os dados do PISA, dos três países, na área de Ciências, dos anos de 2006, 2009, 2012 e 2015, buscando ampliar a análise e os entendimentos sobre o fenômeno. Considera-se que as influências das avaliações em larga escala, com foco exclusivo em Linguagem e Matemática, apresentem também implicações nos resultados dos alunos de 15 anos.

RESULTADOS E DISCUSSÃO

Os dados, coletados em cada contexto de cada país, são apresentados inicialmente relacionados às suas realidades e particularidades próprias: Austrália, Brasil e Canadá.

CONTEXTO AUSTRALIANO

No contexto australiano, as evidências são provenientes da análise dos dados do TIMSS, a partir do ano de 1995. A Tabela 1 sintetiza os resultados:

TABELA 1 - Média da Austrália em Ciências no TIMSS

	1995	2003	2007	2011	2015
Quarto ano	521	521	527	516	524
Oitavo ano	514	527	515	516	512

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os dados indicaram, na série histórica (1995-2015), que os resultados obtidos pelos jovens, em Ciências, permaneceram praticamente no mesmo patamar ao longo dos anos, considerando que as médias das pontuações não foram estatisticamente diferentes de um período para outro, na referida série ($p > 0,05$), para os quartos e oitavos anos. Uma estagnação que coincidiu com a proliferação das políticas de avaliação com o foco substantivo em Linguagem e Matemática no país.

Em 2015, os resultados dos alunos do quarto ano mostraram que, enquanto a Austrália superou, em relação ao desempenho dos estudantes em Ciências no TIMSS, 17 países, ela foi superada por outras 17 nações. Da mesma forma, no oitavo ano, o país superou, em termos de diferenças estatisticamente significantes, 20 outros, mas foi superado por outros 14 países. Tal situação ilustra também o processo de estagnação em relação ao desempenho do ensino de Ciências.

Outros estudos locais indicaram (REID, 2009; LINGARD, 2010) que o forte investimento em políticas de avaliação em Linguagem e Matemática e de responsabilização, utilizando programas de mérito, tem, entre outras coisas, afetado e reduzido os currículos àquilo que tem sido cobrado nos testes nacionais, revelando o fenômeno de afinilamento curricular presente em outras partes do mundo (BAUER, 2013; MADAUS; RUSSELL; HIGGINS, 2009). Ao mesmo tempo, essas iniciativas têm interferido na prática pedagógica dos professores e, em alguns casos, indicado um processo de uniformização de ensino.

É necessário sinalizar que os investimentos nessas políticas, com foco em LM e de responsabilização, continuam em pleno vigor na Austrália e estão sendo ainda mais ampliadas, como pode ser apreciado em um documento recente publicado pelo governo de Nova Gales do Sul: *Literacy and Numeracy Strategy 2017-2020* (NEW SOUTH WALES, s.d.).

CONTEXTO BRASILEIRO

No Brasil, as evidências coletadas e apresentadas são provenientes das análises realizadas sobre os dados do PISA dos anos de 2009, 2012 e 2015, e outras originárias particularmente da região do Grande ABC Paulista.

BRASIL NO PISA

Os dados do PISA mostraram que o país estava em 63º lugar no *ranking* de Ciências, em 2015, quatro posições abaixo da atingida em 2012. Uma posição incômoda para a nação que é a nona economia do mundo. A Tabela 2 sintetiza os dados do PISA:

TABELA 2 – Médias do Brasil em Ciências no PISA

2006	2009	2012	2015
390	405	405	401

Fonte: Programme for International Student Assessment.

Os dados do PISA revelam, apesar de uma redução na pontuação média do ano de 2015, estagnação do país no desempenho dos alunos em Ciências, considerando que

[...] não foram encontradas evidências empíricas que apontem diferenças estatisticamente significativas entre o desempenho dos estudantes brasileiros em Ciências no PISA 2015 e nas três últimas edições da avaliação. (BRASIL, 2015, p. 16)

Não se pode assumir diretamente que a ausência de avanços nos dados seja reflexo das políticas de avaliação em larga escala e de responsabilização. No entanto, pode-se constatar que a estagnação começou, coincidentemente, após a criação da principal política para a verificação da qualidade da educação básica no Brasil, o Ideb, que utiliza dados

da Prova Brasil e do Saeb, ambos incidindo apenas sobre as áreas de Português e Matemática.

A REGIÃO DO GRANDE ABC

Na região do grande ABC, as evidências foram coletadas a partir de três pesquisas realizadas na área. O primeiro estudo (GARCIA; BIZZO, 2015) mostrou a percepção de um grupo de alunos de escolas de ensino fundamental I³ em relação às disciplinas mais relevantes no contexto escolar. Esses jovens acreditavam que as matérias mais importantes eram as de Português e Matemática e secundarizavam as Ciências.

3 Adotou-se essa nomenclatura “ensino fundamental I” para se referir aos anos iniciais do ensino fundamental: primeiro, segundo, terceiro, quarto e quintos anos.

Os pesquisadores mostraram que os jovens foram influenciados na construção desse imaginário pelos professores, escola e sistema de ensino, inseridos em um contexto marcado pela cultura do teste em Português e Matemática e de responsabilização. No primeiro caso, os docentes solicitavam que os jovens estudassem mais as matérias de PM para a realização dos simulados da escola, para a prova local e para a Prova Brasil ou Saresp. Alguns professores atribuíam notas para os alunos realizarem esses testes de forma séria. No segundo, várias estratégias eram utilizadas para que os estudantes tivessem altos resultados nas avaliações de larga escala: realizavam algum tipo de prova em PM duas ou três vezes ao ano, simulando a Prova Brasil, participavam de projetos de reforço relacionados a essas disciplinas no contraturno, entre outros. Salienta-se ainda que alguns diretores apresentavam estratégias bem definidas para que os alunos com baixo rendimento intelectual não realizassem as provas (alguns ficavam na biblioteca da escola). No último caso, em relação às secretarias de educação e suas políticas educacionais, eram realizadas provas municipais para averiguação das habilidades dos alunos em PM, e os resultados eram avaliados pelos especialistas da secretaria, pelos diretores e coordenadores das escolas e depois pelos professores e os jovens. Esses departamentos endereçavam maiores recursos (compra de materiais) para as áreas de Português e Matemática. Um grande movimento que induzia nos alunos a crença de que essas matérias eram as mais importantes no contexto escolar.

O segundo estudo, na mesma região (OEGABC, 2015), com escolas de ensino fundamental II, mostrou que elas valorizavam os resultados do Ideb ou do Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp), e concentravam esforços nas áreas de PM: os professores dessas disciplinas dispunham de mais tempo com a coordenação pedagógica para reuniões, participavam mais de projetos externos (olimpíada de Matemática, por exemplo) e possuíam maiores recursos para a compra de materiais (jogos, *softwares*, entre outros). A disciplina de Ciências ficava em um segundo plano e não era valorizada. Em alguns casos, ela era encarada como uma matéria de conhecimentos secundários. Tal situação ocorria porque uma alta posição no *ranking* do Ideb indicava que a escola seria muito mais reconhecida e valorizada pelos pais e pela rede de ensino. Em alguns contextos (Santo André), os professores recebiam um bônus salarial em virtude dos bons resultados alcançados. Trata-se do projeto de meritocracia (FREITAS, 2013).

O terceiro estudo foi realizado com 49 professores de Ciências do ensino fundamental II da mesma região: *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências* (OEGABC, 2016). A pesquisa tinha o objetivo de conhecer a percepção do professor sobre as políticas de avaliação em larga escala com foco em Português e Matemática e de responsabilização difundidas na região do grande ABC.

Os resultados mostraram uma realidade preocupante em relação ao ensino de Ciências. Em geral, faltavam recursos para equipar os laboratórios e materiais para a realização de experiências, a disciplina não era valorizada pelas redes de ensino e pela gestão escolar e o ensino era baseado praticamente no uso do livro didático. Por outro lado, os maiores recursos das secretarias de educação eram endereçados aos programas e projetos nas áreas de PM. As escolas ofereciam reforço escolar, olimpíadas, jogos, simulados focando nessas duas disciplinas. Em algumas cidades, só existiam professores eventuais em PM. Havia também mais tempo destinado à formação desses professores (cursos, congressos). Os cursos, palestras e encontros, em alguns municípios

(São Caetano do Sul), focavam mais os conteúdos dessas disciplinas. O tempo de reuniões de diretores, coordenadores e docentes dessas áreas era bem maior nas escolas (São Caetano do Sul). Nenhuma dessas estratégias descritas existiam para fomentar a aprendizagem em Ciências.

Um fato a se destacar é que alguns professores de Ciências faziam piadas sobre seus colegas de PM, sinalizando que eles trabalham muito mais, e no final o bônus era recebido por todos. Em geral, esses docentes sinalizaram que essas questões começaram a acontecer depois da criação e instauração do processo de avaliação em larga escala.

Uma análise documental das escolas desses 49 professores revelou que nos últimos cinco anos mais de 55 projetos associados às disciplinas de Português e Matemática foram criados nas instituições de ensino para melhorar o rendimento dos alunos nos testes padronizados (simulado da escola, prova local, Saesp, Prova Brasil). Não houve mais de dez projetos relacionados às Ciências.

CONTEXTO CANADENSE

No contexto canadense, o estudo de Dow (2015), com professores de Ciências, mostrou que os docentes consideravam que a disciplina não era prioridade nas políticas públicas educacionais recentes, nos programas e projetos dos sistemas de ensino e nas iniciativas e ações realizadas pelas escolas e pelos gestores escolares.

Para a maioria dos professores, a disciplina de Ciências era considerada como sendo de “segunda classe”, considerando que os docentes das áreas de Linguagem e Matemática possuíam os maiores recursos pedagógicos e materiais para as aulas (laboratório de redação, *software* para Matemática), mais horas de formação (cursos, encontros), maior número de projetos extraclasse (olimpíada de Matemática), mais tempo de ensino (número maior de aulas semanais) e mais apoio dos diretores. O foco excessivo em LM, visando a melhorar o rendimento do aluno, é fortemente incentivado nos documentos oficiais recentes publicados pela Educational Quality and Assessment Office (EQAQO, 2015), uma organização que avalia o quanto o sistema de ensino público de Ontário está

desenvolvendo as habilidades de leitura, escrita e matemática dos alunos. A instituição fornece informações confiáveis e úteis que são usadas para auxiliar a melhorar o desempenho dos jovens e garantir a responsabilidade dos conselhos escolares (EQAQ, 2016).

O estudo de Fazio e Karrow (2013), na mesma região, mostrou, por um lado, a falta de recursos para o ensino de Ciências, por outro, a desmotivação dos professores para o ensino. Paralelamente, esses pesquisadores revelaram uma redução do desempenho dos alunos no TIMSS nos anos de 2003, 2007 e 2011. A Tabela 3 sintetiza os dados dos quarto e oitavo anos:

TABELA 3 - Média do Canadá em Ciências no TIMSS

	2003	2007	2011	2015
Quarto ano	540	536	528	530
Oitavo ano	533	526	521	524

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os dados revelam uma tendência de queda nos resultados dos jovens na área de Ciências a partir de 2003, em ambos os anos (quarto e oitavo). Em 2011, o desempenho diminuiu, em média, 15 pontos em relação ao ano de 2003. Quando é realizada uma análise do Índice de Referência Internacional (IRI), os dados também apontam para uma diminuição no aproveitamento dos alunos: A Tabela 4 revela os dados do IRI:

TABELA 4 - Desempenho dos jovens de acordo com o Índice de Referência Internacional

NÍVEL DE DESEMPENHO		2003	2007	2011
Avançado	Quarto ano	13%	12%	9%
	Oitavo ano	7%	7%	6%
Alto	Quarto ano	47%	45%	40%
	Oitavo ano	41%	37%	35%
Intermediário	Quarto ano	81%	79%	77%
	Oitavo ano	81%	77%	76%
Baixo	Quarto ano	96%	95%	94%
	Oitavo ano	97%	96%	96%

Fonte: Trends in International Mathematics and Science Study (elaboração dos autores).

Os resultados revelam que no IRI o percentual de alunos no nível avançado, alto e intermediário também vem diminuindo a partir de 2003. Esses dados indicam um declínio do rendimento dos jovens nos níveis, como mostra a Tabela 4.

Sinteticamente, os resultados mostram queda no aproveitamento dos jovens. Em 2011, as notas médias dos testes em Ciências (Tabela 3) eram significativamente mais baixas do que em 2003 para os alunos do quarto e oitavo anos. Essa queda do rendimento, no país, coincidiu com o crescimento das políticas de avaliação com o foco em LM, em que os principais recursos eram e são utilizados para melhorar as notas de avaliação em larga escala nessas áreas (VOLANTE; BEN JAAFAR, 2008).

O PISA NAS TRÊS REALIDADES ANALISADAS

A análise histórica dos dados do PISA⁴ dos três países avaliados é reveladora de realidades distintas. Embora a ausência de melhora nos resultados exija uma análise mais detalhada, é importante notar o comportamento do desempenho dos alunos em Ciências no período compreendido entre 2006 e 2015. A Tabela 5 ilustra os resultados:

⁴ Embora seja possível a comparação de mudanças no desempenho dos alunos ao longo do tempo em cada domínio do PISA (Leitura, Matemática e Ciências), há que se considerar que cada edição do exame tem um foco particular. Em 2006 e 2015, o foco recaiu sobre Ciências.

TABELA 5 – Média em Ciências no PISA, segundo o país

PAÍS	2006	2009	2012	2015
Austrália	527	527	521	510
Brasil	390	405	405	401
Canadá	534	529	525	528

Fonte: Programme for International Student Assessment (elaboração dos autores).

Os resultados apresentados na Tabela 5 revelam, no contexto brasileiro (BRASIL, 2015, p. 16) e também no canadense (O'GRADY et al., 2016, p. 29), uma estagnação nos dados de Ciências no período de 2006 a 2015, considerando que não existem diferenças estatisticamente significativas.

No caso australiano, que apresenta uma queda ao longo das últimas três avaliações, os dados, no entanto, apontam em outra direção, indicando que existem diferenças estatisticamente significativas no desempenho em Ciências no período entre 2006 e 2015 (THOMSON; BORTOLI; UNDERWOOD, 2015).

Embora Canadá e Austrália não tenham apresentado aumento no desempenho médio de seus estudantes em Ciências, afetado também pelas ALE que priorizam somente Língua Materna e Matemática, a situação do Brasil é claramente mais grave, uma vez que os aspectos de letramento científico contemplado no exame estão muito aquém dos resultados dos dois países, em média cerca de 20%, ou mais de 100 pontos numa escala de 0 a 700, em todas as edições do exame.

CONSIDERAÇÕES FINAIS

Este estudo reuniu e apresentou evidências de três realidades diferentes, demonstrando as consequências para o ensino de Ciências de um cenário marcado por políticas de avaliação em larga escala com foco exclusivo em Linguagem e Matemática e de responsabilização.

Entre as consequências, verificou-se que essas políticas estão induzindo à distribuição desigual dos recursos e, o que é mais preocupante, muitas vezes, dentro da mesma escola,

incluindo a questão do tempo de formação e da criação de programas e projetos educativos e pedagógicos, que estão muito mais centralizados nas disciplinas de Linguagem e Matemática.

Essas políticas estão desalinhando a questão da formação dos professores, à medida que os tempos para capacitação docente tem sido, significativamente, maiores para os docentes das áreas de Linguagem e Matemática, assim como os investimentos também têm sido mais amplos nessas áreas, deixando as outras disciplinas em um segundo plano.

Nas três realidades investigadas, uma das conclusões válidas relaciona-se à questão de que essas políticas com ênfase em LM e de responsabilização têm afetado negativamente a qualidade do ensino de Ciências. No contexto australiano, foi encontrado estagnação no desempenho dos alunos em Ciências no TIMSS, nos quartos e nonos anos, isso depois de 20 anos dessas políticas no país. Além disso, tem acontecido uma redução dos currículos e uma uniformização nas práticas pedagógicas dos professores, como sinalizado pelos pesquisadores. Há ainda a questão de que os investimentos nesses programas continuam avançando.

No Canadá e no Brasil, a disciplina de Ciências não tem sido prioridade para as autoridades e gestores, e também não tem sido valorizada pelas redes de ensino e pelos gestores escolares. Entre outras situações identificadas em ambas as realidades avaliadas, a disciplina foi considerada como de segunda classe. Há ainda a falta de recursos para as escolas e para os docentes (laboratórios e materiais) e a desmotivação dos professores para o trabalho de sala de aula.

No Canadá, houve uma redução no desempenho dos alunos no TIMSS que também se manifesta no Índice de Referência Internacional dos últimos 15 anos, coincidindo com os avanços das políticas de avaliação e de responsabilização. No Brasil, há uma estagnação no desempenho dos alunos em Ciências no PISA nas últimas edições (nos outros dois países, queda), coincidentemente após a criação da principal política de responsabilização (Ideb). Ao mesmo tempo, em uma das áreas mais ricas do país, a região do Grande ABC, essas políticas têm influenciado o imaginário, a

percepção dos alunos em relação à importância das disciplinas de Português e Matemática nas escolas e, nesse sentido, a disciplina de Ciências tem sido secundarizada.

Todas essas iniciativas e ações direcionadas com o foco em Linguagem e Matemática e essas políticas de prestação de contas têm sido endereçadas para melhorar os resultados dos alunos nos testes nas áreas de LM. No entanto, essas políticas têm gerado como consequência não intencional o descuido e o descaso com o ensino de Ciências, que tem sido negligenciado. Tal desvalorização e secundarização, considerando a relevância dessa disciplina na formação do cidadão em um contexto social que valoriza o conhecimento científico e as tecnologias, pode ampliar ainda mais o desinteresse dos alunos no ensino de Ciências, contribuindo para que os jovens sejam menos críticos, reflexivos, questionadores e transformadores da realidade em relação à crise ambiental, ao consumismo, à manipulação gênica, entre outros. Ao mesmo tempo, essas políticas podem distanciar ainda mais os jovens das carreiras científicas, o que já vem acontecendo em várias partes do mundo (AIKENHEAD, 2005; BERNARDO et al., 2008; BROK et al. 2006; GOUW; BIZZO, 2016).

Os resultados deste estudo podem ser utilizados pelas secretarias de educação e pelos gestores escolares para ampliar a compreensão sobre o fenômeno, alertar as autoridades políticas e educacionais sobre as consequências de tais políticas e fomentar o debate sobre a qualidade educacional.

REFERÊNCIAS

AIKENHEAD, Glen S. Research Into STS Science Education. *Educación Química*, Belo Horizonte, v. 16, p. 384-397, 2005.

ALAVARSE, Ocimar Munhoz; MACHADO, Cristiana; BRAVO, Maria Helena. 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.

ARELARO, Lisete Regina Gomes. Direitos sociais e política educacional: alguns ainda são mais iguais que outros. In: SILVA, S.; VIZIM, M. (Org.). *Políticas públicas: educação, tecnologias e pessoas com deficiências*. Campinas: Mercado de Letras, 2003. p. 13-36.

- BALL, Stephen J. Cidadania global, consumo e política educacional. In: SILVA, L. H. da. *A escola cidadã no contexto da globalização*. Petrópolis: Vozes, 1999. p. 127-137.
- BAUER, Adriana. Estudos sobre sistemas de avaliação educacional: um retrato em branco e preto. *Revista @mbienteeducação*, São Paulo, v. 5, p. 7-31, 2012.
- BAUER, Adriana. Limites, desafios e possibilidades das avaliações de sistemas educacionais: contribuições do ciclo de debates para as políticas de avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialava R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores*. Florianópolis: Insular, 2013. p. 281-294.
- BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela. Avaliações em larga escala: uma sistematização do debate. *Educação e Pesquisa*, São Paulo, v. 41, n. especial, p. 1367-1382, dez. 2015.
- BENAVOT, Aaron; TANNER, Erin. The growth of national learning assessments in the world, 1995-2006. Paris: Unesco, 2007. Background paper for the Education for All Global Monitoring Report 2008, Education for All by 2015: Will we make it.
- BERNARDO, Allan B. I. et al. Students' perceptions of science classes in the Philippines. *Asia Pacific Education Review*, Philippines, v. 9, n. 3, p. 285-295, ago. 2008.
- BONAMINO, Alicia C. de. Avaliação educacional no Brasil 25 anos depois: onde estamos? In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores*. Florianópolis: Insular, 2013. p. 43-60.
- BRASIL. Instituto Nacional de Estudos e Pesquisas. *Inclusão de Ciências no SAEB*: documento básico. Brasília, DF: Inep, ago. 2013.
- BRASIL. Diretoria de Avaliação da Educação Básica. *Brasil no PISA 2015*: sumário executivo. Brasília, DF, 2015. Disponível em: <http://download.inep.gov.br/acoes_internacionais/pisa/documentos/2016/pisa_brasil_2015_sumario_executivo.pdf>. Acesso em: 04 mar. 2016.
- BROK, Perry Den et al. Californian science students' perceptions of their classroom learning environments. *Educational Research and Evaluation: An International Journal on Theory and Practice*, England & Wales, London, v. 12, n. 1, p. 3-25, 2006.
- BROOKE, Nigel P.; CUNHA, Maria Amália; FALEIROS, Matheus. *A avaliação externa como instrumento da gestão educacional nos estados*: relatório final. Belo Horizonte: Game/UFGM; Fundação Victor Civita, 2011. Disponível em: <http://www.educadores.diaadia.pr.gov.br/arquivos/File/pdf/avaliacao_externa_fvc.pdf>. Acesso em: 27 dez. 2012.

BRUNNER, José Joaquín. Límites de la lectura periodística de resultados educacionales. In: IAIES, Gustavo et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIPE, 2003. p. 67-84.

CANADÁ. Council of Ministers of Education. *About CMEC*. Toronto: CMEC, 2016a. Disponível em: <<http://www.cmec.ca/11/About/index.html>>. Acesso em: 2 dez. 2017.

CANADÁ. Council of Ministers of Education. *Pan-Canadian Assessment Program (PCAP)*. 2016b. Disponível em: <[http://www.cmec.ca/302/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-\(PCAP\)/PCAP-2013/index.html](http://www.cmec.ca/302/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-(PCAP)/PCAP-2013/index.html)>. Acesso em: 2 dez. 2017.

CASASSUS, Juan. Política y metáforas: un análisis de la evaluación estandarizada en el contexto de la política educativa. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 21-46.

CHAPMAN, David; SNYDER, Conrad. Can high stakes national testing improve instruction: re-examining conventional wisdom. *International Journal of Educational Development*, Amsterdam, v. 20, p. 457-474, 2000.

DORN, Sherman. *Accountability Frankenstein: understanding and taming the monster*. Charlotte, NC: Information Age Publishing, 2007.

DOW, Allison. *Exploring the current state of grades 4 to 8 science education in Ontario*. 2015. Dissertação (Mestrado em Educação) – Faculty of Education, Brock University, Ontario, 2015. Disponível em: <<http://dr.library.brocku.ca/handle/10464/6983?show=full>>. Acesso em: 12 abril 2016.

EDUCATIONAL QUALITY AND ASSESSMENT OFFICE. *Highlights of the provincial results*. Assessments of reading, writing and Mathematics, primary division (grades 1-3) and junior division (grades 4-6). English-Language Students, 2014-2015. Ontario: OQRE, 2015. Disponível em: <<http://www.eqao.com/en/assessments/results/communication-docs/provincial-report-highlights-elementary-2015.pdf>>. Acesso em: 12 dez. 2015.

EDUCATIONAL QUALITY AND ASSESSMENT OFFICE. *Highlights of the provincial results*. Assessments of reading, writing and Mathematics, primary division (grades 1-3) and junior division (grades 4-6). English-Language Students, 2015-2016. Ontario: OQRE, April 2016. Disponível em: <<http://www.eqao.com/en/assessments/results/communication-docs/provincial-report-highlights-elementary-2016.pdf>>. Acesso em: 15 mar. 2017.

ESTEBAN, Maria Teresa. Considerações sobre a política de avaliação da alfabetização: pensando a partir do cotidiano escolar. *Revista Brasileira de Educação*, Rio de Janeiro, v. 17, n. 51, p. 573-592, 2012.

EUROPEAN COMMISSION. Eurydice Network. *National testing of pupils in Europe: objectives, organisation and use of results*. Bruxelas: Eurydice Network, 2009.

FAZIO, Xavier; GARCIA, Paulo Sergio; PANIZZON, Debra Lee. Comparative analysis of elementary science teacher preparation in Australia, Brazil, and Canada. In: INTERNATIONAL ORGANIZATION FOR SCIENCE TECHNOLOGY EDUCATION SYMPOSIUM, 18., set. 2008, Esmirna, Turquia. *Proceedings of...* Ancara: Palme & Bookshops, 2008. v. 1, p. 737-745.

FAZIO, Xavier E.; KARROW, Doug. Science takes a back seat: an unintended consequence of prioritizing literacy and numeracy achievement. *Canada Education*, Ontario, v. 56, n. 1, jun. 2013.

FREITAS, Dirce Nei Teixeira de. *A avaliação da educação básica no Brasil*. Campinas: Autores Associados, 2007.

FREITAS, Luiz Carlos de. Caminhos da avaliação de sistemas educacionais no Brasil: o embate entre a cultura da auditoria e a cultura da avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 147-176.

GARCIA, Paulo Sérgio; BIZZO, Nelio. Educational policies and science education in Brazil: a case study. In: EUROPEAN SCIENCE EDUCATION RESEARCH ASSOCIATION, 11., 2015, Helsinki. *Proceedings...* Helsinki: European Science Education Research Association, 2015.

GARCIA, Paulo Sérgio; BIZZO, Nelio; FAZIO, Xavier. Challenges faced by science teachers in distance continuing education. In: INTERNATIONAL ORGANIZATION FOR SCIENCE TECHNOLOGY EDUCATION SYMPOSIUM, 15., 2012, Hammamet, Tunis. *Proceedings...* Hammamet, Tunis: 2012. v. 1, p. 23-33.

GARCIA, Paulo Sérgio; BIZZO, Nelio; FAZIO, Xavier. Desafios da formação contínua a distância para professores de ciências. *Revista Iberoamericana de Educacion a Distancia*, Ecuador, San Cayetano Alto, v. 17, n. 2, p. 37-57, 2014.

GARCIA, Paulo Sérgio; FAZIO, Xavier; PANIZZON, Debra. Formação Inicial de professores de ciências na Austrália, Brasil e Canadá: uma análise exploratória. *Ciência e Educação*, Bauru, v. 17, p. 1-19, 2011.

GIMENES, Nelson et al. Além da Prova Brasil: investimento em sistemas próprios de avaliação externa. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 55, p. 12-32, maio/ago. 2013.

GOUW, Ana Maria; BIZZO, Nelio. A percepção dos jovens brasileiros sobre suas aulas de Ciências. *Educar em Revista*, Curitiba, n. 60, p. 277-292, abr./jun. 2016.

HAGOPIAN, Jesse (Ed.). *More than a score: the new uprising against high-stakes testing*. Chicago: Haymarket Books, 2014.

HORTA NETO, João L. Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o SAEB de 2005. *Revista Iberoamericana de Educación*, Madrid, v. 42, n. 5, p. 1-14, abr. 2007.

HORTA NETO, João L. As avaliações externas e seus efeitos sobre as políticas educacionais: uma análise comparada entre a União e os Estados de Minas Gerais e São Paulo. 2013. Tese (Doutorado em Política Social) – Instituto de Ciências Humanas, Departamento de Serviço Social, Universidade de Brasília, Brasília, 2013.

HYPÓLITO, Álvaro Moreira. A necessária meta-avaliação das políticas de avaliação. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. Florianópolis: Insular, 2013. p. 211-227.

IAIES, Gustavo. Evaluar las evaluaciones. In: IAIES, G. et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIEPE, 2003. p. 15-36.

INTERNATIONAL ASSOCIATION FOR THE EVALUATION OF EDUCATIONAL ACHIEVEMENT. *Trends in International Mathematics and Science Study (TIMSS), 2011: Ontario Report*. Toronto: Educational Quality and Accountability Office, 2012. Disponível em: <http://www.eqao.com/pdf_e/12/TIMSS_Ontario_Report_2011.pdf>. Acesso em: 07 mar. 2015.

KLINGER, Don A.; DELUCA, Chistopher; MERCHANT, Stefan. Canada: the intersection of international achievement testing and educational policy development. In: VOLANTE, L. (Ed.). *The intersection of international achievement testing and educational policy: global perspectives on large-scale reform*. Routledge: Nova York, 2016. p. 140-159.

LIMA, Licínio. O paradigma da educação contábil. Políticas educativas e perspectivas gerencialistas no ensino superior em Portugal. *Revista Brasileira de Educação*, Rio de Janeiro, n. 4, p. 43-59, 1997.

LINGARD, Bob. Policy borrowing, policy learning: testing times in Australian schooling. *Critical Studies in Education*, London, v. 51, n. 2, p. 129-147, 2010.

MADAUS, George; RUSSELL, Michael; HIGGINS, Jennifer. *The paradoxes of high stakes testing: how they affect students, their parents, teachers, principals, schools, and society*. Charlotte: Information Age, 2009.

MONS, Nathalie. Theoretical and real effects of standardised assessment. Bruxelas: Eurydice Network, 2009.

NEW SOUTH WALES. Department of Education. *Literacy and numeracy strategy 2017-2020*. Sydney: New South Wales Government, [s.d.]. Disponível em: <<https://www.det.nsw.edu.au/media/downloads/about-us/literacy-numeracy/literacy-and-numeracy-strategy.pdf>>. Acesso em: 10 fev. 2017.

OBSERVATÓRIO DA EDUCAÇÃO DO GRANDE ABC. *Efeitos IDEB na região do grande ABC sobre o Ensino de Ciências*. São Caetano do Sul: Universidade Municipal de São Caetano do Sul, 2015.

OBSERVATÓRIO DA EDUCAÇÃO DO GRANDE ABC. *Políticas de avaliação e de responsabilização no grande ABC: a percepção dos professores de Ciências*. São Caetano do Sul: Universidade Municipal de São Caetano do Sul, 2016.

O'GRADY, Kathryn et al. *Measuring up: Canadian results of the OECD PISA study. The performance of Canada's youth in science, reading, and Mathematics. 2015 First Results for Canadians Aged 15.* Toronto: Council of Ministers of Education, 2016.

OLIVEIRA, Romualdo P. de. A utilização de indicadores de qualidade na unidade escolar ou porque o IDEB é insuficiente. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores.* Florianópolis: Insular, 2013. p. 87-100.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. *PISA 2015 assessment and analytical framework: Science, Reading, Mathematics and Financial Literacy.* Paris: OECD, 2016.

OZGA, Jenny. *Investigação sobre políticas educacionais: terreno de contestação.* Porto: Porto, 2000.

PHILLIPS, David. Learning from elsewhere in education: some perennial problems revisited with reference to British interest in Germany. *Comparative Education*, London, v. 36, n. 3, p. 297-307, 2000.

RAVITCH, Diane. *The death and life of the great American school system: how testing and choice are undermining education.* New York: Basic Books, 2010.

REID, Alan. *Rethinking national curriculum collaboration: towards an Australian curriculum.* Canberra: Department of Education, Science and Training, 2005.

REID, Alan. Is this really a revolution? A critical analysis of the Rudd government's national education agenda. *Curriculum Perspectives*, Berlin, v. 29, n. 3, p. 1-13, 2009.

SANTOS, Lucíola L. A avaliação em debate. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos.* Florianópolis: Insular, 2013. p. 229-245.

SMITH, William. *National testing policies, school practices, and student outcomes: an analysis using data from PISA 2009.* Paris: OECD, 2014. (Education Working Papers).

SONGER, Nancy Butler; RUIZ-PRIMO, Maria Araceli. Assessment and science education: Our essential new priority? *Journal of Research in Science Teaching*, New Jersey, v. 49, n. 6, p. 683-690, 2012.

SOUSA, Sandra Zákia. Avaliação externa e em larga escala no âmbito do Estado brasileiro: interface de experiências estaduais e municipais de avaliação da educação básica com iniciativas do governo federal. In: BAUER, Adriana; GATTI, Bernardete A. (Org.). *Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores.* Florianópolis: Insular, 2013. p. 61-85.

SOUSA, Sandra Zákia; PIMENTA, Claudia Oliveira; MACHADO, Cristiane. Avaliação e gestão municipal da educação. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 53, p. 14-36, set./dez. 2012.

STECHEER, Brian M. Consequences of large-scale, high-stakes testing on school and classroom practice. In: HAMILTON, Laura; STECHER, Brian M.; KLEIN, Stephen P. (Ed.). *Making sense of test-based accountability in education*. Santa Monica: Rand, 2002. p. 79-100.

TEDESCO, Juan C. Prólogo. In: IAIES, Gustavo et al. *Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. Buenos Aires: Unesco/IIPE, 2003. p. 11-14.

THOMSON, Sue; BORTOLI, Lisa de; UNDERWOOD, Catherine. *TIMSS 2015: a first look at Australia's results*. Melbourne: Australian Council for Educational Research, 2016.

VOLANTE, Louis; BEN JAAFAR, Sonia. Educational assessment in Canada. *Assessment in Education: Principles, Policy & Practice*, London, v. 15, n. 2, p. 201-210, 2008.

PAULO SÉRGIO GARCIA

Professor titular da Universidade Municipal de São Caetano do Sul (USCS). Coordenador do Observatório de Educação do Grande ABC, São Caetano do Sul, São Paulo, Brasil
paulo.garcia@uscs.edu.br

XAVIER FAZIO

Professor associado da Faculty of Education da Brock University, Ontário, Canadá
xfazio@brocku.ca

DEBRA PANIZZON

Professora associada da University of South Australia (UniSA), Adelaide, Austrália
debra.panizzon@unisa.edu.au

NELIO BIZZO

Professor titular da Faculdade de Educação da Universidade de São Paulo (FE/USP), São Paulo, São Paulo, Brasil
bizzo@usp.br

Recebido em: 17 JULHO 2017

Aprovado para publicação em: 10 OUTUBRO 2017