



Currículo de matemática: um olhar sobre a Prova Brasil no estado do Rio de Janeiro¹

Maria Isabel Ramalho **Ortigão**

Faculdade de Educação da Baixada Fluminense, Universidade do Estado do Rio de Janeiro
Brasil

iortigao@uerj.br

Larissa Ribeiro **Lessa**

Faculdade de Educação da Baixada Fluminense, Universidade do Estado do Rio de Janeiro
Brasil

lari_lessa@yahoo.com.br

Resumo

Esse trabalho apresenta os resultados de uma investigação que buscou compreender o currículo de matemática das redes municipais do estado do Rio de Janeiro, a partir da avaliação nacional Prova Brasil. Trata-se, portanto, de uma pesquisa quantitativa que analisou os dados dos estudantes do quinto ano do ensino fundamental que fizeram a prova de matemática. Utilizou-se na investigação uma análise do funcionamento diferencial do item (DIF), que busca detectar itens cuja probabilidade de acertos difere entre distintos grupos. No total, foram analisados 87 itens de matemática. Os resultados evidenciaram que os itens não apresentaram DIF entre os grupos considerados. Argumenta-se que tais resultados podem estar associados tanto à melhoria nos processos teórico-metodológicos da avaliação em larga escala, como por certa “homogeneização” no currículo de Matemática das escolas públicas do estado do Rio de Janeiro, que sistematicamente vem evidenciando pouco avanço no desenvolvimento das habilidades matemáticas básicas.

Palavras chave: currículo, avaliação, Prova Brasil, DIF.

Resumen

Este trabajo presenta los resultados de una investigación que busca comprender el currículo de matemáticas de las redes municipales en el estado de Río de Janeiro, a

¹ Pesquisa realizada com apoio de: CAPES/INEP/OBEDUC e FAPERJ.

partir de la evaluación nacional brasileña *Prova Brasil*. Es, por lo tanto, un estudio de investigación cuantitativa que analiza los datos de los estudiantes del quinto grado de la escuela primaria que hizo el examen de matemáticas. Fue utilizado en la investigación un análisis de funcionamiento diferencial del ítem (DIF), que busca detectar objetos cuya probabilidad de accesos difiere entre los diferentes grupos. En total, 87 ítems fueron analizados. Los resultados mostraron que los ítems no mostraron DIF entre los grupos considerados. Se argumenta que estos resultados pueden estar asociados con una mejoría en ambos procesos teóricos y metodológicos de la evaluación a gran escala, como cierta "homogeneización" en el currículo de matemáticas de las escuelas públicas en el estado de Río de Janeiro, que ha venido mostrando consistentemente poco progreso en desarrollo de habilidades matemáticas básicas.

Palabras Clave: Currículo, Evaluación, *Prova Brasil*, DIF.

Introdução

Em diversos países, as últimas décadas do século XX foram marcadas por uma evidente preocupação com a melhoria da qualidade da educação, que terminaram por impulsionar tanto reformas educativas como a criação de sistemas de avaliação. No Brasil, em particular, essa preocupação culminou com a proposição pelo Governo Federal do Sistema de Avaliação da Educação Básica (SAEB) no final dos anos 1980 (Bonamino, 2002).

O SAEB ao longo de sua existência passou por diversas modificações (Ortigão; Sztajn, 2001). Atualmente, congrega três avaliações de larga escala: Avaliação Nacional da Educação Básica (Aneb), Avaliação Nacional do Rendimento Escolar (Anresc - também denominada Prova Brasil) e a Avaliação Nacional da Alfabetização – ANA.

Especificamente, em relação à Prova Brasil, trata-se de um sistema avaliativo cujo objetivo é o de produzir sistematicamente informações sobre o desempenho médio dos estudantes das escolas públicas e colocá-las à disposição da sociedade em geral, oferecendo dados não apenas para o Brasil e unidades da Federação, mas também, para cada município e escola participante. A coleta de informações baseia-se em teste de Matemática e Língua Portuguesa e em questionários, aplicados aos alunos, seus professores e os diretores de suas escolas (Bonamino; Souza, 2012)

Os itens que compõem os testes são construídos a partir das Matrizes de Referência para a Avaliação e cada item está associado a uma competência específica. Em outros estudos por nós conduzidos (Ortigão; Sztajn, 2001; Lessa; Ortigão, 2014) evidenciou-se estreita relação entre o que é proposto nas matrizes de avaliação e nos documentos curriculares - Parâmetros Curriculares Nacionais (Brasil, 1997).

Este texto apresenta os resultados de uma investigação conduzida com o objetivo de analisar o comportamento diferencial dos itens de matemática aplicados pela Prova Brasil, em 2007, nas diferentes redes municipais do estado do Rio de Janeiro. Especificamente, serão apresentados aqui os uma análise realizada com itens de Matemática aplicados a estudantes do 5º ano do Ensino Fundamental. No desenvolvimento da pesquisa as seguintes questões foram conduzidas:

- (a) há diferenças de competência cognitiva entre estudantes de escolas municipais situadas na capital e aqueles que estudam em escolas das outras regiões do estado do Rio de Janeiro?

(b) é possível identificar competências matemáticas que são exploradas de modo diferenciado nos currículos destas escolas?

Antes de apresentarmos a abordagem utilizada no estudo, trazemos, de modo breve, algumas considerações sobre o currículo de matemática contido nos documentos oficiais brasileiros. Na continuidade descrevemos abordagem priorizada e, por fim, discutimos os resultados obtidos.

Currículo de matemática: proposta nacional

A divulgação dos Parâmetros Curriculares Nacionais (PCN) de Matemática (Brasil, 1997) e as sucessivas avaliações de livros didáticos do Programa Nacional de Avaliação do Livro Didático (PNLD) foram decisivas para algumas modificações ocorridas nos currículos de matemática da escola básica de ensino fundamental, dentre as quais, destaca-se a ampliação das áreas de ensino. No PCN a resolução de problemas é destacada como o foco do ensino de Matemática desde os anos 1980. Além dos aspectos cognitivos – tradicionalmente valorizados – o documento refere-se também à relevância aos aspectos sociais, antropológicos e linguísticos como tendo imprimido novos rumos às discussões curriculares. Para o documento, o ensino precisa valorizar não mais o ensino propedêutico, mas, a construção de competências básicas necessárias ao cidadão. Estas ideias estão em consonância com as propostas do campo da Educação Matemática por: (1) achar que o ensino de Matemática tem produzido baixos resultados no desempenho dos alunos; (2) pelo reconhecimento de que o mundo necessita de estudantes com maiores habilidades no uso de ferramentas matemáticas; (3) pelos avanços educacionais que passaram a valorizar a aprendizagem coletiva, os conhecimentos prévios dos alunos e a construção do conhecimento pelos estudantes.

O documento PCN-Matemática organiza-se em duas partes. A primeira apresenta a caracterização da área da matemática abordando a análise da trajetória e o quadro atual do ensino da disciplina, as principais características do conhecimento matemático, o papel fundamental da matemática no ensino fundamental e a matemática enquanto formadora de cidadania. Ainda na primeira parte, trata das relações entre aprender e ensinar matemática no ensino fundamental, indica caminhos para “fazer matemática” na sala de aula, destaca os objetivos gerais para a matemática no ensino fundamental, apresenta blocos de conteúdos e traz propostas para a avaliação. A segunda parte aborda, apresenta sugestões para o trabalho em sala de aula. De um modo geral os conteúdos são divididos em quatro grandes blocos: Números e Operações, Espaço e Forma, Grandezas e Medidas e Tratamento da Informação.

Comportamento Diferencial do Item (DIF)

Esta sessão destina-se a apresentar a abordagem metodológica adotada na investigação. Em especial a análise de DIF, que visa a verificar se um item tem ou não o mesmo comportamento para indivíduos pertencentes a dois grupos distintos, mas de mesma habilidade cognitiva. Em outras palavras, ao realizar esse tipo de análise o investigador está interessado em saber se um item apresenta grau de dificuldade diferente para subgrupos da população que tem o mesmo nível de conhecimento, ou seja, se há itens mais fáceis para um grupo em detrimento de outro. É importante salientar que, aqui, a aceção de *itens mais fáceis para alunos de um grupo em relação a outro grupo* pressupõe, sempre, comparações controladas pela proficiência em Matemática, isto é, comparações entre alunos com desempenhos semelhantes no teste.

A elaboração de um item para um teste de uma avaliação em larga escala, como a Prova Brasil, deve levar em consideração vários aspectos, tais como, a idade e a escolarização do grupo

ao qual o teste será aplicado. Além disso, o item deve ser objetivo e claro, não conter termos que favoreçam um grupo em detrimento de outro, bem como, não oferecer dicas que, facilmente, conduzam o aluno à resposta.

Historicamente, a preocupação com o DIF está associada ao desejo de que se construíssem questões de teste que não fossem afetadas por características étnico-culturais dos grupos submetidos aos testes de avaliação educacional – muito ligada, portanto, às campanhas em prol da melhoria dos direitos civis dos cidadãos comuns, nos anos de 1960, nos Estados Unidos da América. Esses anos foram marcados por uma enorme preocupação com a igualdade de oportunidades, pelas críticas aos sistemas educacionais discriminadores, pelo desenvolvimento de um conceito popular e legal de ações afirmativas e pela consciência racial/étnica. Diferenças educacionais, resultantes de sistemas educacionais com muita iniquidade, passaram a ser vistas como vestígios de uma velha ordem segregadora. Assim, escores de testes, refletindo essas diferenças, foram considerados, da mesma forma, discriminadores, e passou-se a usar o termo *viés*, ao referenciá-los. Os estudos para identificação de *viés* tinham por objetivo provar que os testes ou instrumentos de medida não possuíam nenhum tipo de *viés*. Muitos pesquisadores começaram a se dedicar ao estudo sistemático das diferenças entre os grupos étnicos, com o objetivo de tentarem encontrar explicações convincentes para as grandes diferenças de rendimento, observados entre os diversos grupos étnicos e socioeconômicos, que refletiam, na realidade, disparidades nas oportunidades educacionais e se mostravam injustos, ao exigirem tarefas estranhas às culturas de algumas minorias. Tais estudos adotaram a seguinte definição de *viés*: *um item é enviesado se sujeitos de habilidades iguais, mas de culturas diferentes, não têm a mesma probabilidade de acertar o item*².

Estudos envolvendo DIF vêm ganhando força no campo da Avaliação da Educação, em especial por se constituir uma boa estratégia para a compreensão do currículo escolar. Essa abordagem metodológica possibilita identificar itens que violam um dos principais pressupostos da TRI, segundo o qual a probabilidade de acertar um item é função da proficiência do aluno (Andriola, 2006; Soares *et al*, 2005; Soares, 2005; Aguiar; Ortigão, 2012; Ortigão, 2014). Ou seja, alunos de grupos distintos com igual proficiência têm a mesma probabilidade de acertar um item. Caso isso não ocorra, podemos afirmar que algum fator que extrapola a habilidade cognitiva do aluno está tornando um item mais fácil para um dos grupos.

No âmbito da TRI, é possível dizer que um item não apresenta DIF quando a curva característica do item (CCI) é a mesma para os grupos comparados em um mesmo nível de habilidade ou proficiência medida através do item. As figuras a seguir ilustram as CCIs de dois itens. No primeiro caso, as curvas estão sobrepostas, indicando a não presença de DIF. Já no segundo caso, há presença de DIF, pois as curvas são diferentes. A análise gráfica da figura 2 evidencia que o item favoreceu a um grupo de alunos, representados pela linha verde, em detrimento de outro, pois, para todas as faixas de proficiências, a probabilidade de acertar o item é maior um dos grupos de alunos.

² Ver, por exemplo: Angoff, 1993; Linn et al, 1981; Shepard; Camilli; Ironson, 1982; Linn; Drasgow, 1987.

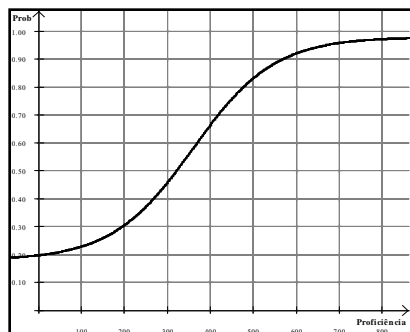


Figura 1. Curva Característica do Item (CCI) – item não apresenta DIF

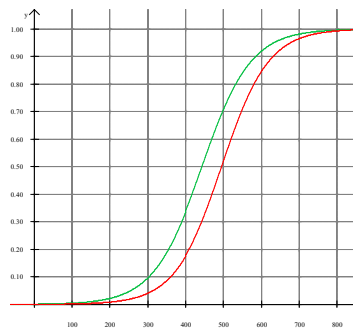


Figura 2. Curva Característica do Item (CCI) – item apresenta DIF

Embora o DIF possa significar que algum grupo particular de indivíduos esteja sendo privilegiado, em detrimento dos demais, ainda assim, a evidência estatística da diferença entre o desempenho de grupos não deve ser persuasiva quanto à decisão de excluir ou não um item do teste, pois sua análise pode ser uma ferramenta de diagnóstico do sistema educacional bastante útil. Doolittle e Cleary (1987) mostraram que o desempenho das meninas é inferior ao dos meninos, quando se trata de itens que medem habilidade matemática em Geometria e em raciocínio matemático. Tais habilidades são objetivos legítimos na Educação Matemática e os estudantes, sejam meninos ou meninas, devem saber lidar igualmente com problemas dessa natureza. Excluir tais itens de um teste, por favorecerem a um determinado grupo, torna o instrumento incompleto e é prejudicial para os que estão em desvantagem, pois tende a perpetuar a diferença.

Soares e colaboradores (2005) realizaram análise de DIF aos dados da avaliação da educação mineira - SIMAVE - para investigar diferenças de competência em Geografia dos alunos das diferentes regiões do Estado. Os autores concluem que há diferença nessas competências, em especial em itens que procuram avaliar as diferenças entre o espaço urbano e o rural (que se mostraram desfavoráveis para os alunos da região metropolitana) e também as questões ao meio ambiente (que se mostraram desfavoráveis aos alunos do interior, quando comparados aos da região metropolitana). Para eles, este fato sugere que, “*para se alcançar equidade, o conteúdo desses itens precisa ser reforçado, adequadamente, nas regiões onde o item apresentou um comportamento aquém do esperado*” (Soares *et al*, 2005, p. 99).

A análise de DIF aos itens de matemática da Prova Brasil

Nesta sessão apresentamos a análise de DIF aos itens de matemática da Prova Brasil 2007³. Para efeito de análise, quando se está comparando o desempenho de um item em dois grupos distintos, um deles denomina-se grupo de Referência (R) e o outro grupo Focal (F). Na análise aqui realizada, o grupo de referência adotado foi o correspondente à rede municipal da cidade do Rio de Janeiro, que congrega cerca de 1200 escolas. Já o grupo focal engloba as redes municipais das demais cidades do estado do Rio de Janeiro. No total foram analisados 87 itens aplicados aos alunos do 5º ano do Ensino Fundamental.

A análise de DIF foi realizada segundo o modelo desenvolvido por Mantel e Haenszel

³ Inicialmente, pretendia-se implementar análise de DIF aos dados da avaliação de 2009. No entanto, isto não foi possível devido à falta de informações necessárias na base de dados de 2009. Cabe observar ainda que a base de dados de 2011 não se encontra disponível até o momento, o que nos levou a trabalhar apenas com os dados de 2007.

(1959) que tem por base a comparação de frequências observadas nos subgrupos constituídos por indivíduos com habilidades semelhantes. Sua avaliação é feita pelo nível de significância das diferenças. O programa estatístico usado para fazer as análises foi o SisAnI, desenvolvido pelo Centro de Avaliação e Políticas Públicas – CAED (UFJF). Embora exista um bom número de procedimentos bastante sofisticados para detectar o DIF em suas distintas manifestações (Fidalgo, 1996), esses procedimentos são apropriados apenas para detectar o viés em potencial de um item. Eles não possibilitam uma explicação de suas causas, pois, de acordo com Mellenbergh (1989, p. 13) *sua interpretação deve ser o resultado de uma conjunção entre procedimentos estatísticos e análises teóricas ou julgamento de especialistas*. Na realidade, detectar o DIF não possibilita afirmar que ele é enviesado, na medida em que é necessário complementar essa informação relacionando-a com as causas ou razões de seu funcionamento diferenciado em tal grupo e se essas diferenças são ou não parte legítima do constructo em questão.

No âmbito da TRI, é possível dizer que um item não apresenta DIF quando a Curva Característica do Item (CCI) é a mesma para os grupos comparados em um mesmo nível de habilidade ou proficiência medida através do item. A figura abaixo ilustra o resultado da análise de DIF, tomando como exemplo um item aplicado ao 5º ano (Figura 3) na avaliação Prova Brasil 2007. Observa-se que os pontos sobre a curva, em ambos os casos, apresentam-se sobrepostos na mesma linha, indicando a não presença de DIF entre os dois grupos considerados.

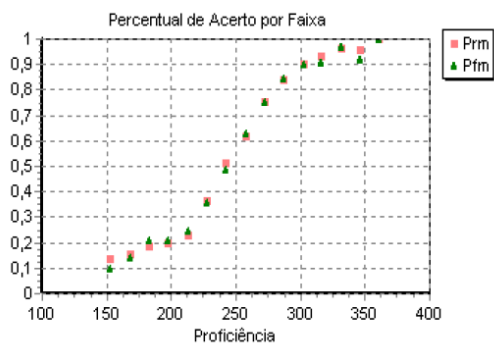


Figura 3. Resultado da análise de DIF ao item M22189 da Prova Brasil 2007 – 5º ano EF

De modo geral, a análise de DIF realizada em todos os itens mostrou resultados análogos, ou seja, a não presença de DIF nos itens da Prova Brasil 2007 entre os estudantes do 5º ano do Ensino Fundamental de escolas municipais da capital e do interior do estado do Rio de Janeiro. Duas hipóteses podem estar associadas a esse resultado: (i) melhoria nos processos teórico-metodológicos da avaliação em larga escala; (ii) certa “homogeneização” no currículo de Matemática das escolas públicas do estado do Rio de Janeiro, que pouco possibilita diferenciar as habilidades dos estudantes que estudam em escolas da capital e do interior.

Com relação à primeira hipótese, é notório o avanço teórico e o aprimoramento metodológico e técnico nos procedimentos que envolvem a avaliação em larga escala. No Ciclo de debates – Vinte e Cinco Anos de Avaliação de Sistemas Educacionais no Brasil (Bauer; Gatti; Tavares, 2013; Bauer; Gatti, 2013), promovido pela Fundação Carlos Chagas, em 2012, evidenciou bem este aspecto. Para Bauer e Tavares (2013), desde a criação do sistema brasileiro de avaliação, no início dos anos 1990, foram detectados quase trezentos estudos acadêmicos,

entre teses e dissertações, sobre as avaliações de sistemas educacionais no Brasil. Desses, há estudos teóricos sobre os procedimentos metodológicos adotados nas avaliações em larga escala, que evidenciam o aprimoramento das teorias das medidas e o avanço em relação à comparabilidade de grupos distintos. Há ainda estudos que se dedicaram a compreender as implicações das avaliações de sistemas educacionais em seus diversos aspectos – na gestão da escola e do sistema de ensino, na formação continuada e nas práticas dos professores em sala de aula, etc. Destacam-se ainda estudos que utilizam bases de dados das diversas avaliações para fazer análises de fatores associados ao desempenho dos estudantes.

Com relação à segunda hipótese, estudos anteriores vêm evidenciando o pouco avanço dos estudantes das redes públicas municipais e estaduais em relação ao desenvolvimento de habilidades matemáticas. Um deles, que envolveu uma análise detalhada nos resultados da avaliação do estado do Rio de Janeiro, por meio da análise de todos os itens de prova de Matemática, aplicados pelo Sistema de Avaliação do estado do Rio de Janeiro - SAERJ – aos alunos do terceiro, quinto, sétimo e nono anos do ensino fundamental (CAEd, 2011), realizada com o intuito de construir uma escala de habilidades em Matemática, evidenciou urgência em rever os currículos desta área de conhecimentos nas escolas públicas do estado do Rio de Janeiro: pouco avanço foi verificado no desenvolvimento das habilidades básicas.

Em outro estudo, realizado no âmbito do Laboratório de Avaliação da Educação (LAED/PUC-Rio), em parceria com a Fundação Getúlio Vargas, com os dados do Sistema Nacional de Avaliação (SAEB) de 1995 a 2009, evidenciou-se que as redes municipais e estaduais em todas as Unidades da Federação, pouco se distinguem, quando são comparados os resultados médios dos estudantes em Matemática e Língua Portuguesa.

Certamente, que este resultado precisa ser olhado com cautela. Como afirmado anteriormente, os sofisticados procedimentos utilizados para detectar DIF são apropriados apenas para detectar o viés em potencial de um item. Eles não possibilitam uma explicação de suas causas, pois sua interpretação deve ser o resultado de uma conjunção entre procedimentos estatísticos e análises teóricas ou julgamento sobre o que de fato os estudantes estão aprendendo e o que lhes estão ensinando, ou seja, sobre o currículo das escolas.

Considerações finais

As características próprias de cada município exercem influência na organização do seu sistema educacional, na forma como os currículos são organizados e na ênfase com que se explora um determinado conteúdo em sala de aula. Além disso, sabemos que os desempenhos escolares dos alunos são influenciados por seus modos de pensar, seus valores culturais e sociais, bem como por características econômicas de suas famílias. A possibilidade de conhecer essas características por meio dos itens que favorecem determinados grupos, e perceber a existência de padrões que passam despercebidos aos olhos dos especialistas que os elaboram, é, sem dúvida, a grande contribuição que a análise do DIF trás para avaliação educacional. Estudos dessa natureza evidenciam que uma boa comparabilidade das proficiências de diferentes grupos de alunos depende de que haja itens comuns aplicados a esses grupos que não apresentem DIF.

O procedimento que desenhou o percurso deste trabalho foi o de responder se Prova Brasil utiliza itens que apresentam DIF. A hipótese inicial da presença de itens com DIF na avaliação da Prova Brasil não se confirmou. Ou seja, não foi verificado itens de matemática com DIF, quando comparamos os estudantes de escolas da rede municipal da capital e os das demais redes municipais do estado do Rio de Janeiro. Os primeiros indícios deste resultado nos conduziram à

busca de argumentos que nos ajudassem a explicá-los. A literatura apontou-nos dois caminhos: o avanço teórico e o aprimoramento metodológico e técnico nos procedimentos que envolvem a avaliação em larga escala e, por outro lado, certa homogeneidade no perfil cognitivo dos estudantes da rede pública brasileira. A adoção de qualquer das duas hipóteses, certamente, deve ser vista com parcimônia e cautela. É necessário um aprofundamento da investigação para que se possa confirmá-las ou refutá-las e isso demandaria um novo mergulho nos dados e nos modelos de análise.

Concluimos, afirmando que o esforço para aquisição das ferramentas empregadas nesta pesquisa e o rigor metodológico que as acompanha foi decisivo por permitir uma análise relevante dos dados apresentados. É desejável que haja um desdobramento deste trabalho no sentido de complementá-lo, não só com os recursos da pesquisa qualitativa, mas, também, com análises mais sofisticadas, que incluam os resultados de outros estados, de modo a aumentar o número de observações.

Referências

- Aguiar, G.; Ortigão, M.I.R. (2012). Letramento em Matemática: Um estudo a partir dos dados do PISA2003. *Boletim de Educação Matemática - Bolema (UNESP. Rio Claro. Impresso)*, 26, 1-21.
- Andriola, W. B. (2006). Estudo sobre o viés de itens em testes de rendimento: uma retrospectiva. *Estudos em Avaliação Educacional*, 17, 35, set./dez.
- Angoff, W. H. (1993). Perspectives on differential item functioning. Em P. W. Holland & H. Wainer (Orgs.). *Differential Item Functioning* (pp. 3-24). Hillsdale, NJ: LEA.
- Bauer, A.; Gatti, B.A. (Org.). (2013). *Ciclo de Debates Vinte cinco anos de avaliação de sistemas educacionais no Brasil / Implicações nas redes de ensino, no currículo e na formação de professores*. V. 2. Florianópolis: Editora Insular.
- Bauer, A.; Gatti, B.A.; Tavares, M.R. (Org.). (2013). *Ciclo de Debates Vinte cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos*. V. 1. Florianópolis: Editora Insular.
- Bonamino, A. (2002). *Tempos de avaliação educacional: o SAEB, seus agentes, referências e tendências*. Rio de Janeiro: Quartet.
- Bonamino, A.; Souza, S. Z. (2012). Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo, 38(2), 373-388, abr./jun.
- Brasil. (1997). *Ministério da Educação. Parâmetros Curriculares Nacionais – Matemática / 1ª a 4ª séries*. Brasília: MEC/SEEFF.
- Doolittle, A. E.; Cleary, T.A. (1987). Gender-based differential item performance in mathematics achievement items. *Journal of Education Measurement*, 24, 157-166.
- Fidalgo, A.M. (1996). Funcionamiento diferencial de los ítems. En J. Muñiz (Org.), *Psicometría*. Madrid: Universitas.
- Ironson, G.H. (1982). Use the chi-square and latent trait approaches for detecting item bias. In: Holland, P. W. (eds). *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum associates, Publishers.
- Lessa, L. R; Ortigão, M.I.R. (2014). *Currículo de Matemática e Avaliação Nacional - e Prova Brasil em questão*. Trabalho apresentado na modalidade Comunicação Científica no VI EEMat/SBEM-RJ.
- Linn, R.L & Drasgow, F. (1987). Implications of golden Rule setting for test construction. In: Holland, P. W.; Wainer, H. (eds). *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum associates, Publishers.

- Linn, R.L.; Levine, M.V.; Hastings, C.N. & Wardrop, J.L. (1981). Item bias in a test of reading Comprehension. *Applied Psychological Measurement*, 5, 159-173.
- Mantel, N.; Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719-748.
- Mellenbergh, G.J. (1989). Item bias and item response theory. *International Journal of Educational Research*, 13, 127-143.
- Ortigão, M. I. R. (2014). *Estudo comparativo entre escolas situadas em periferias e em capitais brasileiras sobre ênfases curriculares em matemática, a partir da análise do comportamento diferencial do item (DIF) na Prova Brasil*. Relatório Técnico: Prociência. Universidade do Estado do Rio de Janeiro.
- Ortigão, M. I. R.; Sztajn, P. (2001). Dilemas para a avaliação: o caso dos conjuntos no ensino da matemática. In: Franco, C. (Org.). *Avaliação, Ciclos e Promoção na Educação*, 69-84. Porto Alegre: Artmed Editora.
- Shepard, L. A.; Camilli, G. & Averill, M. (1981). Comparison of procedures for detecting test bias with both international and external ability criteria. In: Holland, P. W.; Wainer, H. (eds.), *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum associates, Publishers.
- Soares, J. F. (2005). Qualidade e equidade na educação básica brasileira: fatos e possibilidades. In: Brock, C.; Schwartzman, S. *Os desafios da educação no Brasil* (pp. 91-118). Rio de Janeiro: Nova Fronteira.
- Soares, T. M.; Genovez, S. F. M.; Galvão, A. F. (2005). Análise do Comportamento Diferencial dos Itens de Geografia: estudo da 4ª série avaliada no Proeb/Simave 2001. *Estudos em Avaliação Educacional*, São Paulo, 16, 32, jul./dez.