

O que a TRI não nos conta? O que os itens excluídos pela TRI dizem sobre o ensino de Matemática?

Rodrigo de Souza Bortolucci – Fundação VUNESP; Guaracy Tadeu Rocha – Fundação VUNESP;
Natália Noronha Barros – Fundação VUNESP; Nayara Negrão Pereira – Fundação VUNESP;
Christiane Bellorio Gennari de Andrade Stevão – Fundação VUNESP

RESUMO

Este estudo objetivou mostrar uma análise pedagógica feita para itens de matemática descartados em dois pré-testes. Dessa análise resultou informações relevantes do grupo de respondentes dos pré-testes, principalmente suas fragilidades. Dado que não há clareza sobre o que os itens descartados representam no processo de significação do traço-latente é necessária uma reflexão quanto a possibilidade de um tratamento estatístico que valorize as informações neles contidas.

Palavras-Chave: Itens descartados; TRI; TCT; Pré-teste; Traço latente.

INTRODUÇÃO

As avaliações em larga escala que utilizam a Teoria da Resposta ao Item (TRI) para a apuração de médias de proficiência dos participantes dessas avaliações em projetos de abrangência e propósitos diversificados, se diferenciam daquelas que fazem uso da Teoria Clássica (TCT), entre outras metodologias. Isto porque as avaliações em larga escala objetivam a descrição do traço latente, ou seja, o conhecimento que o aluno demonstra possuir em determinada área do conhecimento, quando se mostra capaz de resolver tarefas propostas em uma avaliação. Para tanto, a TRI possibilita organizar e situar itens em uma escala de forma linear, a partir de um conjunto de respostas para uma coletânea de questões, chamadas de itens nesse tipo de avaliação. A partir dessa escala são definidos intervalos de proficiência, que caracterizam estágios de aprendizagem dos estudantes e indicam aqueles que estão aquém das expectativas para determinada disciplina em um dado momento da escolarização, assim como aqueles que estão avançados em relação às expectativas projetadas. Isso é possível pois a metodologia estatística da TRI permite conhecer as propriedades estatísticas dos itens caracterizadas por três parâmetros: índice de discriminação (a), índice de dificuldade (b) e probabilidade de acerto casual (c).

A importância da utilização de itens pré-testados na composição de instrumentos de avaliação educacional de larga escala é bem conhecida e apoia-se justamente na qualidade estatística dos itens selecionados para compor um instrumento que se articule aos referenciais da avaliação, que seja capaz de investigar conhecimentos, habilidades e competências com a amplitude que a avaliação requer e que ofereça, com a maior precisão possível, a medida de proficiência definida para a avaliação que se realiza. Dessa condição emerge a prática da pré-testagem, como principal garantia da qualidade dos itens que serão utilizados em um projeto de avaliação em larga escala. Ocorre que nesse processo, há itens para os quais não é possível obter parâmetros nos intervalos que definem um item de bom desempenho pela TRI. Sempre que isso acontece, os itens são descartados. As razões para o descarte são indicadas pelo relatório estatístico e correm por conta principalmente do não atendimento aos critérios definidos para os intervalos de valores que qualificam o bom desempenho no âmbito dos parâmetros *a*, *b*, *c*.

Para o presente estudo, foram utilizados dados de 2 projetos de pré-testagem de itens realizados em 2014 e em 2018. Pretendeu-se fazer uma análise quantitativa e qualitativa dos itens descartados pela metodologia TRI nesses dois eventos e se, apesar do descarte, esses itens não seriam úteis para o diagnóstico de particularidades que, na perspectiva pedagógica, permitiriam melhor compreensão da aprendizagem no sistema avaliado.

METODOLOGIA

Para o presente estudo, foram utilizados dados de dois projetos de pré-testagem de itens realizados em 2014 e em 2018. A escolha desses dois projetos apoia-se em duas razões específicas:

- o processo de 2014 é pré-testagem aplicada como parte de provas de avaliação educacional de larga escala, de natureza censitária, amplamente divulgada. O participante respondeu os itens da pré-testagem em situação idêntica à da prova;
- a pré-testagem realizada em 2018 compreendeu a aplicação de testes de Língua Portuguesa, Matemática, Ciências e Ciências da Natureza em todos os anos escolares em que essas disciplinas são obrigatórias no percurso da Educação Básica.

A realização de pré-testagem de itens é um processo necessário e importante quando se trata de avaliações educacionais de larga escala. Dentre os diferentes processos realizados para a obtenção dos parâmetros dos itens que irão compor o banco de itens da pré-testagem, o utilizado no presente

trabalho, foi o processo de equalização via itens comuns. Segundo ANDRADE et al. (2000), o uso de itens comuns entre provas distintas aplicadas a populações distintas permite que todos os itens estejam na mesma escala ao final dos processos de estimação. O processo de equalização é um dos grandes objetivos da avaliação educacional, segundo KOLEN & BRENNAN (1995), equalizar significa, tornar comparável, e ao colocarmos os itens em uma mesma escala, tornamos possível a comparação de resultados e a criação de uma escala de proficiência com os diferentes grupos. Dentre os desafios pertinentes à criação de um banco de itens está a confecção de um banco suficientemente grande para a aplicação de provas e a perda de um menor número de itens após a aplicação da prova, obedecendo os critérios estabelecidos para a garantia de boas estimativas dos parâmetros dos itens.

Nem todos os itens que compõem uma prova apresentam um comportamento estatístico esperado. Como parte do processamento, são retirados itens que não apresentam qualidade no processo de estimação, como a não convergência, valores altos dos erros das estimativas e itens com baixo poder de discriminação, ou seja, itens pelos quais não é possível discriminar indivíduos de alta ou baixa habilidade cognitiva. No pré-teste de 2014 itens foram eliminados quando o parâmetro *b* (dificuldade) apresentasse índices fora do intervalo -4,0 e 4,0 ou se o parâmetro *c* (acerto casual) superior a 0,45. Já para o pré-teste de 2018, além desses casos, os itens foram eliminados quando o parâmetro *a* (discriminação) em Língua Portuguesa e Matemática foi menor que 0,5 ou maior que 3,0 na escala normal, e menor que 0,35 ou maior que 3,0 em itens de Ciências da Natureza. Os percentuais de itens descartados pela TRI, após a aplicação dos pré-testes 2014 e 2018, por ano/série, são apresentados na tabela a seguir.

Tabela 1. Descarte de Itens					
Ano/Série Escolar	Proporção em relação ao total de itens no teste (em %)				
	Projeto Pré-Testagem 2014		Projeto Pré-Testagem 2018		
	Língua Portuguesa	Matemática	Língua Portuguesa	Matemática	Ciências / Ciências da Natureza
4º Ano EF	4,8	0,0	10,3	11,1	5,1
5º Ano EF	2,4	0,0	9,4	6,0	6,8
7º Ano EF	-	-	6,0	9,4	12,8
8º Ano EF	7,1	35,7	-	-	-
9º Ano EF	4,8	26,2	5,1	10,3	13,7
1ª, 2ª e 3ª Série EM	11,1	24,6	10,3	8,1	13,7
Total	7,5	21,1	10,3	10,6	11,3

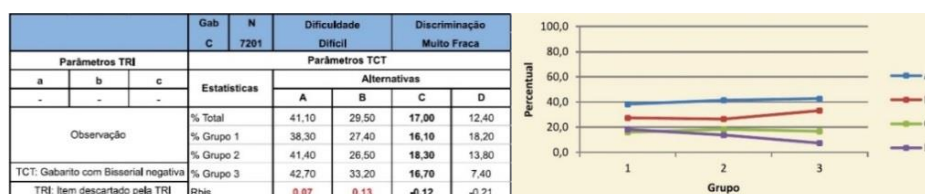
Como se pode constatar, as proporções de itens descartados, na maioria dos anos escolares são da ordem dos 10%. Uma vez identificados, os itens foram analisados a fim de detectar particularidades e informações relevantes, apesar do descarte. Na sequência, é apresentada parte da análise feita para itens de Matemática, cuja exclusão está atrelada ao parâmetro a.

DISCUSSÃO E RESULTADOS

Tendo a disposição os itens que foram excluídos do pré-teste; optou-se por analisar as propriedades desses itens segundo a TCT, além da qualidade pedagógica, uma vez que se mostravam corretamente elaborados. A partir dessa análise observou-se que muitos trazem informações relevantes mostrando principalmente as fragilidades em tarefas que o grupo de respondentes, inclusive aqueles de proficiência avançada, não se mostrou capaz de resolver. A seguir, três exemplos ilustram a análise feita.

ITEM 1 – PRÉ-TESTE 2014 – 9º ANO EF – MATEMÁTICA

O item proposto apresenta o número aproximado de glóbulos vermelho que o corpo humano possui, sendo esse número da ordem da dezena de trilhão, e cabe ao respondente indicar a forma correta de representá-lo por meio de escrita utilizando notação científica. Os dados obtidos a partir da resposta do grupo de respondentes foram os seguintes:



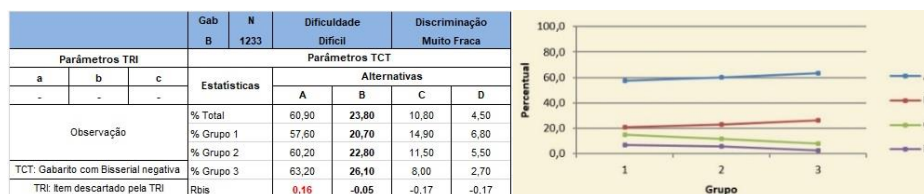
Dois distratores apresentaram coeficiente bisserial positivo, indicando que essas opções de resposta foram escolhidas por um percentual maior de estudantes que tiveram melhor desempenho na prova como um todo. Ou seja, mesmo os alunos de melhor desempenho na prova não souberam identificar a representação correta do número utilizando notação científica.

Chama a atenção o fato de que um desses distratores foi a alternativa mais assinalada, apesar de propor uma escrita em que a notação científica é composta por uma potência de 10 com expoente negativo, ou seja, representa um número menor que 1 e que, portanto, não poderia representar um número da ordem de trilhão. Já a outra alternativa traz como resposta uma notação, cujo valor

presente no expoente da potência é igual ao número de algarismos zeros presentes na escrita do número, pensamento válido somente para números com determinadas características, que aparentemente foi generalizado pelo alunado. Consequentemente, o item foi classificado como difícil para o grupo de respondentes, sendo acertado somente por 17% do público, e apresentou um índice de discriminação classificado como muito fraco. Mediante esses dados, é razoável sugerir que o conceito de notação científica não foi devidamente apropriado por esses alunos, não sendo compreendido como uma escrita que mantém relação estreita com a grandeza do número proposto.

ITEM 2 – PRÉ-TESTE 2018 – 7º ANO EF – MATEMÁTICA

O item proposto exigia a análise de dados apresentados em duas tabelas para determinar se há proporcionalidade entre os dados descritos. Na primeira tabela há a relação entre o peso de um produto e seu respectivo preço, sendo que caberia ao estudante perceber que é cobrado R\$4,00 por kg do produto. Já os dados presentes na segunda tabela mostram o peso de uma pessoa associado à sua idade, de modo que o aluno deveria atentar-se para o fato de que não há uma relação de proporcionalidade entre peso e idade, mesmo o peso apresentando um aumento constante ano a ano. Os dados obtidos para esse item foram:

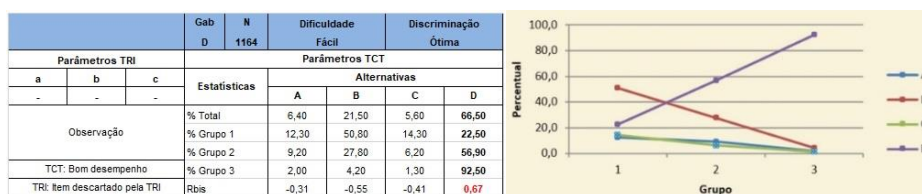


O item se mostrou difícil para o grupo de respondentes uma vez que a maioria dos estudantes optou pela alternativa incorreta que diz que as duas situações apresentam proporcionalidade. Essa alternativa foi a mais escolhida, indicando que pouco mais de 60% dos respondentes se mostraram inclinados a reconhecer grandezas como sendo proporcionais baseados somente no fato de as grandezas envolvidas aumentarem juntas, mesmo sem um fator de proporcionalidade que as relacione. Além disso, o fato do aumento observado ser constante, o que indica a existência de um padrão nesse crescimento, pode reforçar a existência de uma possível proporcionalidade para os estudantes. Isso sugere que o fator de proporcionalidade pode estar sendo entendido como um padrão de crescimento e não como uma constante que deve ser multiplicada no valor de uma das grandezas para obtenção do valor correlato da outra grandeza. Segundo a TCT, esse item apresentou

coeficiente bisserial negativo no gabarito, indicando que não há um ganho de desempenho por parte dos melhores respondentes da prova como um todo, o que reforça a proposição de uma tarefa difícil e que não foi apropriada pela maioria dos estudantes que participaram do teste. Frente a essa constatação, o professor pode atuar a fim de evitar que o conceito seja sistematicamente reproduzido de forma incorreta.

ITEM 3 – PRÉ-TESTE 2018 – 9º ANO EF - MATEMÁTICA

O item proposto solicita ao respondente a identificação do número decimal e da porcentagem equivalentes à fração $\frac{3}{5}$. Os dados relativos às respostas apresentadas foram:



O item apresentou ótima discriminação, uma vez que 92,5% dos alunos de melhor desempenho geral no teste (Grupo 3) indicaram a resposta correta, enquanto que no grupo de alunos de pior desempenho geral no teste (Grupo 1) o índice de acerto foi igual a 22,5%. Considerando todos os respondentes, dois terços optaram pela alternativa correta, caracterizando o item como fácil. Dentre os que erraram, a maioria optou pela alternativa que associava a fração $\frac{3}{5}$ ao decimal 3,5 e a porcentagem de 35%, indicando que esses alunos provavelmente inferiram que as diferentes formas de escrever um número racional exigem o emprego dos mesmos algarismos. Por fim, é importante ressaltar que os exemplos de itens discutidos não apresentam erros em sua formulação, tampouco as chamadas “pegadinhas”, mas permitem que parte dos alunos que não se mostraram capazes de concluir a tarefa corretamente encontre uma alternativa de resposta, indicando assim um equívoco em sua forma de pensar sobre o problema e, consequentemente sobre a habilidade em questão, o que é relevante para uma reflexão sobre o ensino da Matemática, mas foi descartado pela TRI.

CONCLUSÃO

A exclusão dos itens pela TRI pode estar atrelado a diversos motivos. Um parâmetro “a” muito baixo costuma indicar que um aumento da proficiência agrega pouquíssimo na probabilidade de resolver corretamente a tarefa. Já no caso do parâmetro “a” ser muito alto, a probabilidade de acerto varia

bruscamente num curto intervalo de proficiência e, com isso, a escala de proficiência praticamente fica dividida em duas partes, a dos que não devem e a dos que devem ser capazes de resolver a tarefa. Aparentemente, itens com essas características não servem para contribuir na descrição do traço latente. No entanto, muitos desses itens excluídos pela TRI indicam a construção de um conceito ou o emprego de um pensamento equivocado por parte relevante do grupo de respondentes, assim como a apropriação de uma ideia correta somente por um grupo muito restrito de alunos.

Tendo em vista a importância da valorização do erro no processo de ensino-aprendizagem, cabe pensar na importância e na relevância desses itens sob um ponto de vista pedagógico. Os agentes que atuam no processo de avaliação têm contato com esses itens, o que possibilita uma análise dos mesmos, mas que não costuma chegar ao conhecimento do público que fará uso desse material. O fato de o item não poder ser posicionado adequadamente em uma escala não deve invalidar as informações importantes que ele carrega, deixando de ser conhecido pelo público que faz uso dos resultados das avaliações. A análise pedagógica de itens descartados traz uma contribuição que justifica o título desse trabalho. Em oposição aos itens que não foram eliminados pela TRI, os descartados evidenciam déficits significativos no ensino e na construção do conhecimento daquele público que participou do teste para uma dada área de conhecimento.

Cabe uma reflexão quanto à possibilidade de esses itens receberem um tratamento estatístico que valorize suas potencialidades, uma vez que parece não haver clareza em relação ao que esses itens descartados representam no processo de significação do traço latente. A simples exclusão ao longo do processo certamente não parece ser a solução para tanto.

REFERÊNCIAS BIBLIOGRÁFICAS

- SOARES, T.M.; SILVA, W. Eficácia dos processos de linkagem na avaliação educacional em larga escala. **Est. Aval. Educ.**, São Paulo, V. 21, n.45, p. 191-211, jan/abr, 2010.
- ANDRADE, D.; TAVARES, H.; VALLE, R. C. **Teoria da Resposta ao Item: Conceitos e Aplicações**. São Paulo: ABE, 2000.
- KOLEN, M. J.; BRENNAM, R. L. **Test Equating – Methods and Practices**. New York. Springer, 1995.
- ZIMOWSKI, M.; MURAKI, E.; MISLEVY, R.; BOCK, D. **Software Bilog-MG 3 for windows**. Scientific Software International, Inc. 2003.