

ANÁLISE DE TESTES INTEGRATIVOS (PIRLS E PISA): CONTRIBUIÇÕES À AVALIAÇÃO EM LARGA ESCALA

Rosângela Veiga Júlio Ferreira

CAEd/UFJF
Juiz de Fora, MG, Brasil
rosangelaveiga.ferreira@ufjf.br

Begma Tavares Barbosa

CAEd/UFJF
Juiz de Fora, MG, Brasil
begmatb@gmail.com

Hilda Aparecida Linhares da
Silva

UFJF 3
Juiz de Fora, MG, Brasil
hildaaparecidalinhares@gmail.com

Resumo

Este trabalho apresenta os resultados de um estudo sobre a natureza dos chamados testes integrativos, realizado no âmbito da Pesquisa em Avaliação em Língua Portuguesa, desenvolvida pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd)¹. Realizou-se uma análise comparativa entre dois modelos de testes integrativos – o PIRLS e o PISA – e destes com as avaliações em larga escala convencionais, com o intuito de compreender de que modo e a partir de que critérios tais testes aferem as habilidades leitoras. De um lado, um resultado importante do estudo diz respeito à centralidade das habilidades (ou processos cognitivos) de localização e inferência de informação. Por outro, esses novos modelos de avaliação afetam, em alguma medida, nossa compreensão de determinadas habilidades e dos critérios para avaliá-las. Tomamos aqui a habilidade de “localização de informação” para demonstrar como isso se dá.

Palavras-chave: Avaliação em larga escala; Língua Portuguesa; Testes integrativos; PIRLS; PISA.

¹ As autoras agradecem à Fundação CAEd - Centro de Políticas Públicas e Avaliação da Educação pelo financiamento da pesquisa.

INTRODUÇÃO

Este texto tem por objetivo apresentar resultados de uma análise que buscou compreender a natureza de diferentes modelos de testes integrativos². Para essa análise, partimos dos achados da Pesquisa em Avaliação em Língua Portuguesa, realizada no âmbito do CAEd (2016-2022), que propôs critérios de complexidade para os textos suporte e para as tarefas de leitura relativas a cada uma das habilidades da Matriz de Referência³. O estudo que apresentamos amplia a compreensão das avaliações de leitura em larga escala convencionais, procedendo uma análise inicial de dois modelos de testes internacional.

Sua metodologia apoia-se em estudo teórico realizado em três passos: i) estudo sobre a natureza de dois modelos de testes integrativos internacionais, identificando suas principais características; ii) estudo comparativo dos testes integrativos (matrizes de referência, tipos de tarefas e habilidades (ou “processos cognitivos”⁴); iii) cruzamento dos resultados desse estudo comparativo com os construtos da Pesquisa em Avaliação (2016/2022), relativos a dados de avaliação externa convencional.

Dos resultados obtidos, trazemos aqui um recorte que considera a análise da habilidade de **localização de informação**, para demonstrar de que modo o remodelamento da avaliação proposto pelos testes PIRLS e PISA altera, em alguma medida, o modo como aferimos o desenvolvimento de algumas habilidades de leitura.

2. TESTE INTEGRATIVO: principais características do objeto de análise

Um teste integrativo caracteriza-se por integrar textos e itens (ou tarefas de leitura). Esses testes podem envolver itens de resposta selecionada e itens de resposta construída. Os modelos analisados envolvem textos suportes bem mais longos (principalmente o PIRLS) e tarefas dirigidas a mais de um texto (modelo PISA). Os itens que compõem os testes são organizados de modo a mobilizar, inicialmente, porções menores do texto, o que vai se ampliando e, ao final, requerem do estudante a análise tanto de porções maiores quanto do sentido global do texto. Outro aspecto interessante diz respeito à forma de organização da sequência de tarefas, que são ordenadas considerando a sequência de informações presentes

² Chamamos de teste integrativo àquele que integra textos e tarefas. São modelos de testes que, de forma diferente dos testes mais “convencionais” de avaliação em larga escala, propõem tarefas (itens) que avaliam a leitura de textos em diálogo e sequências de tarefas dirigidas a um mesmo texto.

³ De base quali-quantitativa, a pesquisa analisou cerca de 3500 itens já aplicados em avaliações externas distribuídos em uma escala de proficiência, o que permitiu chegar a resultados importantes relativos à progressão de dificuldade de tarefas de leitura. As habilidades contempladas são aquelas descritas em uma Matriz de Referência específica oferecida pelo CAED.

no texto e tendem a caminhar de tarefas mais simples para mais complexas.

Um modelo de teste integrativo desloca-se da ideia de avaliar um conjunto de habilidades em diferentes textos, com um texto suporte para cada item, como observamos nas avaliações externas convencionas. Com a integração de itens dirigidos a um mesmo texto (ou a um conjunto de textos, como no PISA) é possível, por exemplo, obter informações sobre a competência de leitura de determinados gêneros textuais com nível de dificuldade apropriado a uma etapa de escolarização.

2.1 O modelo PIRLS

Esse modelo de teste integrativo considera que a compreensão leitora implica o desenvolvimento das várias habilidades necessárias para o entendimento de um texto narrativo e de um texto informativo. Trata-se de um modelo que requer dos estudantes do 4º ano fluência leitora e engajamento, seguindo o percurso de leitura que as tarefas de avaliação propõem. No PIRLS, o foco central são os textos e seus propósitos comunicativos. Assim sendo, as habilidades estão “a serviço” dos processos de compreensão – matriz de referência (Brasil, 2023a; 2023b) . Portanto, é um modelo de teste que difere do que vem sendo realizado desde a década de 1990, em nosso país⁵.

As principais características do PIRLS são: i) é formado por textos narrativos e informativos com diferentes níveis de dificuldade - fácil, médio e difícil; ii) o caderno do teste contém um texto que considera cada um dos propósitos comunicativos agrupados em função do nível de dificuldade; iii) é constituído por textos de nível 2, de acordo com critérios da Pesquisa em Avaliação, mas de longa extensão⁶; iv) o teste avalia 4 processos cognitivos, descritos em sua matriz de referência; v) é formado pela combinação de itens de resposta selecionada e de resposta construída, em média 16 ao todo; vi) os itens são apresentados considerando a ordenação das informações nos textos.

2.2 O modelo PISA 2018

O Pisa é um exame comparativo internacional (avaliação amostral)⁷ que procura

⁴ As matrizes dos dois testes integrativos analisados estruturam-se em “processos cognitivos”.

⁵ O PIRLS considera, predominantemente, critérios pedagógicos, a partir dos quais as tarefas de leitura são formuladas, com base no conhecimento sobre o texto e sobre a língua, partindo de porções menores do texto para se chegar à mobilização do sentido global do texto. O critério estatístico, nos parece, não é o que guia a proposição do teste integrativo, sendo este considerado central na montagem dos testes em larga escala em nosso país.

⁶ Sobre os níveis do texto, ver Barbosa; Micarello e Ferreira (2020; 2023).

⁷ Realizado a cada três anos pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE, 2019), o teste avalia três domínios: Leitura, Matemática e Ciências, elegendo-se, a cada edição, um domínio

avaliar se os alunos são capazes de, além de reproduzir conhecimentos, aplicar o que aprenderam em situações não familiares, dentro e fora da escola. Dentre as competências a serem desenvolvidas para uma participação ativa nas modernas sociedades tecnológicas, a leitura no meio digital é, sem dúvida, das mais relevantes. As tarefas propostas pelo PISA estão muito voltadas para a avaliação dessa competência, levando-nos a refletir sobre um conjunto de habilidades e estratégias de leitura específicas, diferentes das avaliadas pelas avaliações convencionais.

São características formais do teste PISA: i) é constituído por textos do universo digital - textos multimodais que conjugam diferentes semioses, ou seja, hipertextos; ii) os textos são mais longos - com aproximadamente de 35 linhas; iii) exige do leitor o estabelecimento de diálogo entre textos; iv) é formado por itens com formatos mais variáveis e enunciados mais longos; v) os itens avaliam mais de uma habilidade; vi) vários itens são dirigidos a um único texto ou textos; vi) requer do leitor que considera a interação entre os itens.

3 PONTOS DE CONVERGÊNCIA ENTRE AS MATRIZES DE AVALIAÇÃO (PIRLS/PISA/AC⁸)

A matriz de referência do PIRLS organiza-se em torno de quatro “processos de compreensão”, a saber: i) localização e recuperação de informações explícitas; ii) inferências diretas; iii) interpretação e articulação ideias e informações; iv) avaliação e análise crítica do conteúdo e dos elementos textuais.

A Matriz de Letramento em Leitura do PISA 2018⁹ organiza-se a partir de três grandes “processos cognitivos”, representando, desse modo, como se dá o processamento do texto. São eles: i) localizar informações; ii) compreender o texto; iii) avaliar e refletir.

O primeiro aspecto a ser observado é que ambas as matrizes se estruturam a partir das duas principais habilidades de leitura (ou “processos cognitivos”): a localização de informação e a inferência de informação. A Pesquisa em Avaliação (2016-2022) dedicou-se detidamente à investigação dessas duas habilidades, dada sua relevância para a aferição da competência leitora, buscando apontar para níveis de complexidade em que elas se realizam, bem como para critérios que as distinguem.

Observa-se muita convergência entre as matrizes do PIRLS e do PISA (e dessas com

como principal, o que implica em um número maior de questões voltadas para sua avaliação. Neste estudo, analisamos o PISA 2018, cuja Matriz de Referência sofreu modificações em relação às anteriores.

⁸ Com a sigla AC queremos fazer referência às Avaliações em Larga Escala “clássicas” ou “convencionais”.

as das avaliações em larga escala convencionais) em se tratando da descrição desses dois processos ou habilidades. Podemos dizer que são processos focados principalmente na interação do leitor com o texto, com as sinalizações textuais, enquanto os demais processos ou habilidades descritos nas matrizes consideram também outros aspectos da produção de sentido na leitura, como os que envolvem a atuação de conhecimento prévio: conhecimento de mundo, textual, linguístico.

Considerando os limites deste resumo, apresentamos, a seguir, a título de exemplo, os resultados relativos à análise de como a habilidade de localizar informação é tratada nos modelos de testes aqui estudados.

3.1 A habilidade de localização de informação

A habilidade de localizar uma informação é bastante fácil. Mesmo leitores com pouca proficiência em leitura, ainda iniciantes, são capazes de localizar informações explícitas, sobretudo quando as tarefas envolvem textos curtos ou em parcelas menores de texto.

A Pesquisa em Avaliação (2016-2022) distingue três procedimentos de localização de informação: i) o recurso à repetição (correspondência total entre o gabarito e a informação do texto), ii) o recurso à paráfrase (correspondência parcial); iii) o procedimento de referenciação por anáfora direta. A partir desses procedimentos, distinguem-se níveis de complexidade de tarefas de localização de informação explícita.

No PIRLS, a utilização de textos bem mais extensos que aqueles das avaliações convencionais, redimensiona a habilidade, tornando-a mais complexa, pois o estudante terá mais trabalho para buscar no texto o local onde a informação está expressa. Nos itens analisados pela Pesquisa em Avaliação, os textos mais longos têm, em média, vinte e cinco linhas, ocupando, normalmente três terços de uma página. O PIRLS seleciona textos de aproximadamente cinco páginas. O tamanho do texto¹⁰, neste caso, é fator determinante de dificuldade da tarefa de localização de informação. Tarefas que tenham como suporte textos curtos e envolvendo o procedimento da repetição são bastante apropriadas para avaliação das séries iniciais. O PIRLS, tendo o propósito de avaliar estudantes do 4º ano do Ensino Fundamental, complexifica a tarefa de localizar informação ao introduzir textos mais longos envolvendo o recurso à paráfrase.

A propósito do recurso à paráfrase, há uma distinção importante a ser feita. A

⁹ A matriz PISA 2018, com a qual trabalhamos aqui, reformulou as matrizes anteriores.

¹⁰ Na Pesquisa em Avaliação outros fatores, que não o tamanho do texto, se apresentaram como mais determinantes na avaliação da complexidade de um texto ou de uma tarefa de leitura, como, por exemplo, o tamanho do período, sobretudo se considerarmos a avaliação das séries iniciais (Barbosa, Micarello e Ferreira, 2020; 2023).

Pesquisa em Avaliação, ao descrever a habilidade de localização de informação, identifica o recurso à “paráfrase simples”, que se caracteriza por pequenas alterações lexicais e sintáticas na estrutura de uma frase (inversão da ordem canônica, substituição por um sinônimo, por exemplo). No PIRLS, as paráfrases são mais complexas, envolvendo uma reescrita bastante modificada do texto.

No PISA, a habilidade de localização de informação se complexifica ainda mais, quando está em jogo a leitura de mais de um texto, aspecto caracterizador do exame. Como descrito em sua matriz, o teste avalia a habilidade de acessar uma informação pontual dentro de um texto e – habilidade mais complexa – a de localizar o texto ou a parcela de um texto onde a informação possa ser encontrada, no caso de testes com textos em diálogo.

CONCLUSÕES: CONTRIBUIÇÕES DO ESTUDO À AVALIAÇÃO EM LARGA ESCALA

Os estudos realizados até o presente momento sobre os testes integrativos PISA e PIRLS já oferecem algumas contribuições à avaliação em larga escala. Considerados os resultados da Pesquisa em Avaliação, que vêm impactando a formulação de itens e testes, uma primeira contribuição é a ratificação da centralidade das duas habilidades fundamentais de leitura- a localização de informação e a produção de inferência. Ratificam-se também categorias de análise importantes que permitem identificar/distinguir tais habilidades (ou processos cognitivos): a repetição, a quase repetição, a paráfrase.

Por outro lado, o estudo de testes integrativos aponta para possibilidades de avaliação externa mais próximas do currículo escolar e das práticas de sala de aula, permitindo explorar, por exemplo, textos maiores e produzir um conhecimento mais profundo sobre como os estudantes constroem sentido de gêneros distintos de texto. Tais testes, sobretudo o PISA, apontam também para a importância de que as avaliações iluminem nossa compreensão sobre um tipo de leitura e de leitor dos quais ainda sabemos pouco e que nos desafia: o leitor do meio digital.

REFERÊNCIAS

BARBOSA, Begma Tavares; SILVA, Hilda Linhares da ; FERREIRA, Rosângela Veiga Júlio. **Avaliações em larga escala de Língua Portuguesa**: uma pesquisa sobre a complexidade dos textos que dão suporte a itens que avaliam a leitura. Pesquisa e Debate em Educação, [S. l.], v. 10, n. 1, p. 1064–1081, 2020.

BARBOSA, Begma Tavares; SILVA, Hilda Linhares da; FERREIRA, Rosângela Veiga. A definição dos níveis de complexidade das tarefas de leitura: relações entre currículo e avaliação em larga escala. In: Hilda Aparecida Linhares da Silva; Liamara Scortegagna.

(Org.). **Pesquisa aplicada a instrumento de avaliação educacional**: desenvolvimentos e perspectivas com base na experiência do CAEd/UFJF. 1ed. Juiz de Fora: Editora UFJF/ Gestão e Avaliação educacional, 2023, v. 1, p. 11-32.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Brasil no PISA 2018** – versão preliminar. Brasília, DF: Inep, 2019.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Brasil no PIRLS 2021*: análise dos resultados da compreensão leitora dos estudantes do 4º ano do ensino fundamental. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2023a.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **PIRLS 2021 Resultados Internacionais em Leitura**: texto do relatório. Brasília, DF: Inep, 2023b.

OCDE- produzido por ETS, núcleo A. **Released Field Trial and Main Survey New Reading Items**. OCDE: out.2019.