

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE EDUCAÇÃO

THAYS CRISTINA DOS SANTOS RODRIGUES

**Teste Adaptativo Informatizado como instrumento para a avaliação da aprendizagem
em leitura: o teste em teste**

São Paulo
2022

THAYS CRISTINA DOS SANTOS RODRIGUES

**Teste Adaptativo Informatizado como instrumento para a avaliação da aprendizagem
em leitura: o teste em teste**

Trabalho Complementar de Curso apresentado
à Faculdade de Educação da Universidade de
São Paulo para a conclusão do Curso de
Licenciatura Plena em Pedagogia.

Área de Concentração: Educação

Orientador: Prof.º Dr.º Ocimar Munhoz
Alavarse

São Paulo
2022

RODRIGUES, T. C. S. Teste Adaptativo Informatizado como instrumento para a avaliação da aprendizagem em leitura: o teste em teste. Trabalho Complementar de Curso (Licenciatura Plena em Pedagogia) – Faculdade de Educação, Universidade de São Paulo. São Paulo, 2022.

Aprovado em:

Banca Examinadora

Prof. Dr.

Instituição:

Julgamento:

Prof. Dr.

Instituição:

Julgamento:

Prof. Dr.

Instituição:

Julgamento

AGRADECIMENTOS

À Denise Fernandes Gonçalves, Sonia Borges dos Santos Rodrigues, Sergio Ricardo Rodrigues, Sergio Ricardo Rodrigues Júnior, Gabriela de Almeida Alvarez, Erika Fujimura e Aline Cristina S. de Oliveira pelo suporte, crença, dedicação e carinho, além de todo apoio dado ao longo de todos esses anos na graduação da Universidade de São Paulo.

Ao Prof.º Dr.º Ocimar Munhoz Alavarse por todos os ensinamentos, conselhos e direcionamentos cuidadosamente dados, incluindo também todo seu entusiasmo ao longo da pesquisa.

À Faculdade de Educação da Universidade de São Paulo e, somada a ela, a todas as professoras que dedicam seu tempo integral e/ou parcial para a formação de graduandas capazes de debater, questionar e pesquisar, sobretudo, também, por todas as oportunidades fornecidas ao longo da minha formação.

À Coordenação de Aperfeiçoamento de Pessoal em Nível Superior (Capes) pela oportunidade de bolsa dentro do Programa Residência Pedagógica, onde o TAI-Leitura foi debatido e apresentado a mim enquanto um instrumento avaliativo capaz de ultrapassar certas limitações, seja em âmbito conceitual, seja em âmbito pessoal.

Aos amigos que formaram uma rede de apoio ao longo de toda a minha vida, dando conselhos sempre que precisei.

Meus mais sinceros agradecimentos a todas, todos e todes que participaram direta ou indiretamente dessa caminhada.

RODRIGUES, T. C. S. Teste Adaptativo Informatizado como instrumento para a avaliação da aprendizagem em leitura: o teste em teste. Trabalho Complementar de Curso (Licenciatura Plena em Pedagogia) – Faculdade de Educação, Universidade de São Paulo. São Paulo, 2022.

Resumo:

Este Trabalho Complementar de Curso apresenta os resultados de uma pesquisa que visou problematizar a aplicação de um Teste Adaptativo Informatizado, denominado TAI-Leitura, concebido para a medição da proficiência da competência leitora de alunas dos anos iniciais do Ensino Fundamental. Para tanto, foram problematizados os conceitos de alfabetização e letramento; estudados os principais conceitos do campo da avaliação educacional; bem como estudada a Provinha Brasil, instrumento avaliativo que se propõe a avaliar o nível de proficiência em leitura de crianças a partir do 2º ano do Ensino Fundamental. Além disso, o trabalho retoma um Teste Adaptativo Informatizado, idealizado previamente pelo Grupo de Estudos e Pesquisas em Avaliação Educacional (Gepave) da Faculdade de Educação da Universidade de São Paulo (Feusp), desenvolvido em 2016, o TAI-Provinha Brasil. Também foram apresentadas as principais discussões sobre medição e avaliação, a fim de que fossem erguidos subsídios para a pesquisa proposta. Com o intuito de compreender a funcionalidade, limitações e potencialidades do TAI-Leitura, instrumento formulado e idealizado para ser um teste de caráter adaptativo pelo Gepave, a pesquisa se deu em duas escolas parceiras, uma pública e outra privada. Dentre os dados obtidos, foi possível perceber que o instrumento é funcional, coerente ao alocar as respondentes em níveis diferentes de proficiência de acordo com a literatura a esse respeito e capaz de captar percepções similares às das professoras de ambas as escolas. Ademais, apesar de serem necessários aprimoramentos em relação ao fornecimento de dados às docentes e à equipe de gestão escolar, ampliação do banco de itens, revisão de determinados itens problematizados ao longo da aplicação do teste, e adequação aos demais sistemas operacionais e suportes eletrônicos, o TAI-Leitura possui o potencial de ser um instrumento capaz de apoiar as professoras ao longo do processo de avaliação da competência em leitura de suas alunas.

Palavras-chave: Avaliação Educacional; Avaliação da Aprendizagem; Teste Adaptativo Informatizado; Tecnologia Educacional; Proficiência em Leitura.

Lista de Gráficos

| | | |
|-----------|---|----|
| 1 | TAI-Leitura. Resultados por Nível e ano. Escola Ciclo. Aplicação 1. 2021.... | 47 |
| 2 | TAI-Leitura. Resultado por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 1. 2021..... | 50 |
| 3 | TAI-Leitura. Resultados por Nível e ano. Escola Ciclo. Aplicação 2. 2021.... | 52 |
| 4 | TAI-Leitura. Resultados por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 2. 2021..... | 53 |
| 5 | Resultado por Nível e ano. Escola Maurîtânia. Aplicação 1. 2021..... | 57 |
| 6 | Resultado por Nível de V_{\min} e ano. Escola Maurîtânia. Aplicação 1. 2021..... | 58 |
| 7 | TAI-Leitura. Nível atribuído e Nível Theta. Escola Maurîtânia. 2021..... | 62 |
| 8 | TAI-Leitura. Nível atribuído e Nível V_{\min} . Escola Maurîtânia. 2021..... | 63 |
| 9 | TAI-Leitura. Nível atribuído e Nível Theta. Escola Ciclo. 2021..... | 64 |
| 10 | TAI-Leitura. Nível atribuído e Nível V_{\min} . Escola Ciclo. 2021..... | 65 |

Lista de Tabelas

| | | |
|---|---|----|
| 1 | Resultado por Nível e Ano. Escola Ciclo. Aplicação 1. 2021..... | 47 |
| 2 | TAI-Leitura. Resultado por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 1. 2021. | 49 |
| 3 | Resultado por Nível e Ano. Escola Ciclo. Aplicação 2. 2021..... | 51 |
| 4 | Tratamento pelo V_{\min} do Theta. Aplicação 2. 2021..... | 52 |
| 5 | Theta médio, tempo médio e variação média das duas aplicações. Escola Ciclo. 2021..... | 54 |
| 6 | Resultado por Nível e ano. Escola Mauritânia. Aplicação 1. 2021..... | 56 |
| 7 | Resultado por Nível de V_{\min} e ano. Escola Mauritânia. Aplicação 1. 2021..... | 57 |
| 8 | Theta médio, tempo médio e variação média dos anos escolares. Escola Mauritânia. 2021..... | 59 |
| 9 | Theta médio, tempo médio e variação média por ano e turma. Escola Mauritânia..... | 60 |

Lista de Quadros

| | | |
|----------|--|----|
| 1 | Número de acertos, nível de proficiência e competências das respondentes. Provinha Brasil – Teste 2. 2016..... | 26 |
| 2 | Distribuição das matrículas por turma. Escola Ciclo. 2021..... | 43 |
| 3 | Distribuição das matrículas por turma. Escola Maurîtânia. 2021..... | 44 |
| 4 | TAI-Leitura. Níveis de proficiência..... | 47 |

Lista de Siglas

| | |
|-------------|--|
| API | <i>Application Programming Interface</i> |
| BNCC | Base Nacional Comum Curricular |
| Capes | Coordenação de Aperfeiçoamento de Pessoal em Nível Superior |
| Enem | Exame Nacional do Ensino Médio |
| Feusp | Faculdade de Educação da Universidade de São Paulo |
| Gepave | Grupo de Estudos e Pesquisas em Avaliação Educacional |
| Inep | Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira |
| PB | Provinha Brasil |
| PRP | Programa de Residência Pedagógica |
| SME-SP | Secretaria Municipal de Educação de São Paulo |
| TAI | Teste Adaptativo Informatizado |
| TAI-Leitura | Teste Adaptativo Informatizado – Leitura |
| TAI-PB | Teste Adaptativo Informatizado – Provinha Brasil |
| TBC | Teste Baseado em Computador |
| TCT | Teoria Clássica dos Testes |
| TRI | Teoria de Resposta ao Item |

Lista de Símbolos

| | |
|---------------------------|-----------------|
| Θ_i | Theta Médio |
| Lim_{inf} | Limite Inferior |
| Lim_{sup} | Limite Superior |
| V_{min} | Valor mínimo |
| V_{max} | Valor máximo |

SUMÁRIO

| | |
|--|-----------|
| 1. INTRODUÇÃO | 11 |
| 2. A IMPORTÂNCIA DA LEITURA..... | 15 |
| 2.1. A IMPORTÂNCIA DO ENSINO DA LEITURA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL | 15 |
| 3. AVALIAÇÃO DO APRENDIZADO DA LEITURA COMO PONTO DE APOIO PARA SEU ENSINO..... | 19 |
| 3.1 DESAFIOS PARA AVALIAR O APRENDIZADO DA LEITURA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL..... | 21 |
| 3.2 OS INSTRUMENTOS DE AVALIAÇÃO E OS ITENS DE MÚLTIPLA ESCOLHA NO APRENDIZADO DE LEITURA | 22 |
| 3.3 O CASO DA PROVINHA BRASIL: POTENCIALIDADES E LIMITAÇÕES..... | 23 |
| 4. MEDIR E AVALIAR | 30 |
| 4.1 PROFICIÊNCIA E NÍVEL DE PROFICIÊNCIA COM INTERPRETAÇÃO PEDAGÓGICA..... | 32 |
| 4.2 TESTES EM PAPEL E DIGITAIS: POTENCIALIDADES E LIMITAÇÕES..... | 34 |
| 4.3 TESTE ADAPTATIVO INFORMATIZADO (TAI) COMO UMA ALTERNATIVA E SUAS CARACTERÍSTICAS GERAIS | 36 |
| 4.4 O TAI-LEITURA DO GEPAVE | 38 |
| 5. O TESTE EM TESTE: APLICAÇÃO DO TAI-LEITURA EM DUAS ESCOLAS.... | 41 |
| 5.1 CARACTERÍSTICAS DAS ESCOLAS, DAS PARTICIPANTES E AS CIRCUNSTÂNCIAS DA APLICAÇÃO | 42 |
| 5.2 OS RESULTADOS DAS APLICAÇÕES..... | 44 |
| 5.2.1. <i>Escola Ciclo</i> | 44 |
| 5.2.2. <i>Escola Maurîtânia</i> | 54 |
| 5.3 AS PERCEPÇÕES DAS PROFESSORAS E OS RESULTADOS | 60 |
| 6. CONSIDERAÇÕES FINAIS..... | 67 |
| REFERÊNCIAS | 72 |
| ANEXO | |

1. INTRODUÇÃO

Uma das partes fundamentais do trabalho de professores é a tomada de decisão com base nos julgamentos que realiza sobre as produções de suas alunas¹ no âmbito da avaliação da aprendizagem, com os impactos no percurso escolar das mesmas, sem desconsiderar outros impactos de ordem subjetiva que acabam por influenciar, também, em seu engajamento na vida escolar, incluindo a autoimagem sobre, especificamente, suas capacidades. Sabe-se ainda que procedimentos e instrumentos avaliativos podem condicionar os resultados do aprendizado, sobretudo quando apresentam limitações que não garantem informações necessárias e adequadas.

Nesse sentido, o estudo e a análise de instrumentos avaliativos é um meio de se compreender a avaliação da aprendizagem dos sujeitos para quem se propõe. Em igual medida, respaldado pela relevância da avaliação da competência em leitura, preconizada pela literatura (Cf. FREIRE, 1989; KRUG, 2015; SOARES, 2009), o presente estudo elenca a avaliação em leitura como sua temática, posto a sua relevância social, histórica e cultural para os sujeitos da educação e sua formação escolar.

Igualmente, reconhecendo a importância da leitura em âmbito social e, sobretudo, educacional, o presente trabalho tem por objeto analisar a avaliação da aprendizagem em leitura por meio de um instrumento avaliativo digital: o TAI-Leitura. Como demarcado em Alavarse et al. (2018), Catalani (2019) e Travitzki et al. (2020), o TAI-Leitura, desenvolvido no âmbito do Grupo de Estudos e Pesquisas em Avaliação Educacional (Gepave), da Faculdade de Educação da USP (Feusp), foi elaborado para estimar a proficiência – o domínio – em leitura, e projetado para ser aplicado nos anos iniciais do Ensino Fundamental.

Em seu cerne, problematizamos o uso de um Teste Adaptativo Informatizado (TAI) por professoras como um recurso para avaliar o aprendizado de leitura ao longo dos anos iniciais do Ensino Fundamental. Como qualquer instrumento avaliativo, o TAI se inscreve no processo de levantamento de informações relativas às aprendizagens, tanto para as docentes, quanto para a escola e seus sujeitos, contribuindo para o (re)planejamento pedagógico e, especialmente, para definir a avaliação desse aprendizado.

Portanto, buscamos analisar a funcionalidade (benefícios, dificuldades e desafios) colocados ao longo da aplicação do TAI-Leitura que foi baseado na Provinha Brasil (BRASIL, 2016a;

¹ Neste trabalho, optamos por definir os sujeitos citados no gênero feminino por se tratarem de crianças e pessoas, sendo ambos substantivos femininos.

2016b), conforme exposto por Alavarse e Melo (2013) ao discorrerem sobre esse tipo de teste, voltado para a competência de leitura. Buscamos dar continuidade aos trabalhos dos autores citados², através de certos questionamentos: o que a aplicação apresenta? É possível afirmar sua funcionalidade? Quais dados fornece? De que modo esses dados refletem o cotidiano observável pelas professoras em sala de aula? Como as professoras se posicionaram diante do TAI-Leitura? Mas, o grande desafio da pesquisa, da qual deriva este Trabalho de Conclusão de Curso, foi responder à questão: o TAI-Leitura funciona como um recurso para estimar a proficiência em leitura?

É importante sublinhar que o TAI-Leitura, até nossa pesquisa, não tinha sido aplicado efetivamente em escolas com alunas dos anos iniciais do Ensino Fundamental, sendo ele um desdobramento de uma versão anterior, o TAI-Provinha Brasil (TAI-PB), abordado em estudos prévios³, onde os códigos foram reescritos, sendo imprescindível verificar sua funcionalidade. Também destacamos que o TAI-Leitura ainda deve merecer aperfeiçoamentos e os resultados da pesquisa podem contribuir nesse sentido.

Assim, para respaldo conceitual quanto à compreensão da relevância da competência à qual o objeto a ser estimado e avaliado está vinculado, com suas implicações para a educação escolar em geral e para seus sujeitos, buscamos em autores como Freire (1989), Krug (2015) e Soares (2009) compreender os conceitos de alfabetização e letramento, além dos principais debates em relação a sua avaliação. Ademais, por ser uma pesquisa no campo da avaliação educacional, é importante elucidar conceitos caros à compreensão da análise dos dados coletados em campo, bem como de sua apropriada significação.

Em seguida, o capítulo 2 pretende percorrer alguns entendimentos e debates sobre instrumentos de avaliação em alfabetização e letramento, atribuindo especial lugar aos debates e desafios acerca da avaliação das competências e habilidades da disciplina de Língua Portuguesa (Cf. MICOTTI, 2007, 2013; SOARES, 2009). O capítulo busca elucidar a importância da competência em leitura como parte fundamental da formação dos sujeitos, trazendo à tona a sua relevância social, histórica e cultural. Propõe a problematização dos conceitos de alfabetização e letramento, especificamente por serem dois conceitos que se complementam, embora possuam competências específicas, a fim de que a competência em leitura seja compreendida, enfim, como o objeto a ser estimado pelo TAI-Leitura. Entendemos que alfabetização é a capacidade de decodificar e interpretar os grafemas e fonemas, transcrevê-los e utilizá-los na

² Sendo Alavarse et al. (2015), Catalani (2019) e Travitzki et al. (2020).

³Cf. Alavarse et al. (2018), Catalani (2019) e Travitzki et al. (2020).

formação de palavras; ao passo em que o letramento é a capacidade de contextualização das palavras na (re)elaboração de textos, ao longo da leitura e da utilização da língua. Consideramos que são conceitos complementares ao processo escolarização, mais acentuadamente nos anos iniciais do Ensino Fundamental.

Ainda, *pari passu*, a relevância em se compreender o que é avaliação, quais seus aspectos, características e principais discussões trouxe-nos o desafio de sintetizar os conceitos que fundamentam o campo teórico e epistemológico em que se coloca nossa pesquisa. Portanto, sem a pretensão de exaurir o tema, no capítulo 3, recuperamos elementos conceituais da literatura em avaliação para elucidar e clarificar as definições, ainda que aproximadas, do que é entendido por avaliação, medição, instrumentos avaliativos, matriz de referência, níveis de proficiência, testes etc., para que sejam subsidiadas as análises dos dados coletados em campo. Trata do debate da elaboração de instrumentos avaliativos que deem conta de medir a proficiência em leitura; tendo a leitura enquanto objeto principal de medição, isso nos leva a salientar que a leitura não deve se confundir com a escrita. Por fim, o capítulo apresenta um instrumento, a Provinha Brasil (PB), como uma possibilidade de medir e avaliar a proficiência em leitura, considerando que nesse instrumento está alicerçado o TAI-Leitura.

A discussão sobre avaliação e medição leva-nos a percorrer os principais aspectos para seu entendimento. Para tanto, no capítulo 4, a relação entre avaliação e medição é discutida, demonstrando que a medição é relevante para o processo avaliativo, ainda que esta possa ser um dos elementos do processo avaliativo. Ainda são igualmente apresentados e discutidos os conceitos de proficiência e nível de proficiência com interpretação pedagógica. São problematizados os instrumentos informatizados para a avaliação em leitura, a fim de que algumas limitações de instrumentos em papel e lápis sejam superadas.

Ademais, o capítulo busca discutir diferenças entre os tipos de testes existentes, tal como a elaboração de um Teste Baseado em Computador (TBC) e as alterações para torná-lo um Teste Adaptativo Informatizado, cuja formulação gerou o TAI-Leitura. De modo que o TAI-Leitura é apresentado enquanto um recurso informatizado, de caráter adaptativo, para a avaliação da proficiência em leitura, sendo discorrido seu histórico e características inerentes à sua (re)elaboração. É importante salientar que a adaptabilidade de um TAI consiste em fornecer novos itens ao respondente de acordo com a resposta ao item anterior, realizando a medição da proficiência leitora ao longo da realização do teste, no caso do TAI-Leitura e do TAI-PB.

Para problematizar a funcionalidade do instrumento avaliativo elaborado pelo Gepave, fez-se necessária a ida a duas escolas, denominadas Escola Ciclo e Escola Maurítânia, a primeira uma

escola privada e a outra, pública, a fim de realizar a testagem do TAI-Leitura com alunas dos anos iniciais do Ensino Fundamental. No capítulo 5, o estudo traz dados que elucidam os resultados apreendidos na aplicação do teste, as impressões de seu contexto (apontamentos das professoras, das crianças etc.), caminhando para as considerações finais.

Ao fim, no capítulo 6, são encaminhados alguns apontamentos acerca do que se pôde observar em relação aos principais desafios da aplicação do instrumento estudado, para que o seu devido estudo e sua (re)elaboração sejam enriquecidos. Dentre os quais estão a necessidade de aprimoramento de apresentação dos dados obtidos com a medição às professoras e, talvez, às respondentes; o acesso do TAI-Leitura por determinados suportes eletrônicos; ampliação do banco de itens, a fim de que a taxa de exposição de item não seja considerável; revisão de itens apontados enquanto confusos pelas respondentes ao longo da observação na pesquisa à campo; e a contínua discussão da matriz de referência e demais conceitos norteadores que fundamentam o TAI-Leitura com as professoras e a comunidade, principais interessados ao longo do processo de avaliação. São apontadas também algumas linhas de pesquisas possíveis, que fomentariam a amplitude que o TAI-Leitura possui enquanto instrumento avaliativo. Finda-se o estudo com a interpretação de que o TAI-Leitura se apresenta enquanto um instrumento avaliativo importante e rico para subsidiar a medição da competência em leitura das crianças ao longo dos primeiros anos do Ensino Fundamental.

2. A IMPORTÂNCIA DA LEITURA

Neste capítulo, busca-se elucidar a relevância da leitura ao longo do processo de alfabetização e letramento em âmbito educacional. Inspiramo-nos em Freire (1989) ao levarmos em consideração que esta supera a mera decodificação do fonema, do grafema, da letra, da sílaba, da palavra, do parágrafo e outros tantos conceitos linguísticos, caros ao ensino da Língua Portuguesa. Embora a proficiência em leitura seja o objeto avaliativo a ser centralizado neste trabalho, reiteramos que a morfologia e demais aspectos da Língua Portuguesa não devem ser dispensados para quaisquer fins, até mesmo pelas relações que estabelecem com o processo de leitura.

Atemo-nos à leitura ser uma das partes do processo de alfabetização e letramento e, na esfera educacional, ser imprescindível para as crianças⁴ e professoras alcançarem as metas preconizadas pela legislação educacional brasileira. Portanto, lançamos olhares a certos conceitos e caminhos trilhados na formação das crianças ao longo de sua escolarização, a fim de compreendê-los.

2.1. A IMPORTÂNCIA DO ENSINO DA LEITURA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL

Quando Freire (1989, p. 9) reitera que “a leitura do mundo precede a leitura da palavra”, igualmente reafirma que a existência da pessoa no mundo e sua relação com os signos têm um papel social, histórico e político, em um processo dinâmico inerente à alfabetização e ao letramento. Social porque, afinal, é através da palavra e a sua relação com o meio que a pessoa, seja ela criança ou adulta, se faz existir, interagindo com outras, incluindo-se em sua comunidade, moldando a sua realidade de acordo com as necessidades individuais e coletivas, para além de suas capacidades e possibilidades. Histórico porque diz respeito à memória e à apreensão da construção da palavra ao longo do tempo. E político porque é através da palavra, de seu entendimento e das formas como as pessoas concebem sua realidade, na retórica e no discurso, que as comunidades se organizam politicamente, adquirindo espaço, promovendo direitos e deveres, e buscando alcançar os objetivos éticos comuns. Enfim, o sujeito

⁴ Referimo-nos as crianças, em passagens deste trabalho, pelo grupo etário das alunas participantes da realização do estudo nos anos iniciais.

contextualiza-se, adquire sentidos com e do mundo antes de conhecer a palavra escrita, antes de desenhá-la e realizá-la.

Assim, quando a proficiência em leitura, fruto de medição por via de um teste como é o caso do TAI-Leitura, se coloca como objeto da avaliação dentro do campo da Língua Portuguesa, sobretudo por sua importância social inerente ao mundo, deve ser entendida como um fenômeno educacional que necessita de instrumentos próprios e adequados para sua apreensão, como aqueles compostos por itens de múltipla escolha, de cujas respostas processadas tem-se um resultado que é, justamente, essa medida da proficiência.

Soares (2009, p. 68), ao abordar o conceito de letramento, traz à luz uma discussão relevante, mas que se torna complexa à medida em que é tecida: discriminar os conceitos de leitura e de escrita dos quais derivam os objetos a serem avaliados e medidos, sem confundir um com o outro. Compreende, portanto, a leitura enquanto parte do processo de letramento, ressaltando a complementariedade do processo de ler e escrever, enfatizando que a

[...] *leitura*, do ponto de vista da dimensão individual de letramento (a leitura como uma “tecnologia”), é um conjunto de habilidades linguísticas e psicológicas, que se estendem desde a habilidade de decodificar palavras escritas até a capacidade de compreender textos escritos. Essas categorias não se opõem, complementam-se; **a leitura é um processo de relacionar símbolos escritos a unidades de som e é também o processo de construir uma interpretação de textos escritos.** (Grifos nossos)

Em concordância, ressaltamos que a leitura, ao integrar o processo de letramento e consequentemente da alfabetização, deve ser um objeto disciplinar relevante para as crianças e demais sujeitos da educação. E, assim reconhecida, deve ser avaliada, uma vez que o processo de avaliação escolar, ressaltadas as devidas proporções, tem como desafio alocar o objeto avaliado dentro de uma escala de proficiência, capaz de fazer aproximações sobre o seu nível de entendimento a respeito desse objeto disciplinar. Isso, na perspectiva de sucesso escolar de todas as estudantes, pode ser, a considerar resultados insatisfatórios, uma referência para se retomar ou desencadear o ensino para que atinjam patamares satisfatórios de aprendizado da leitura.

Respalado por diversos documentos norteadores, os currículos de ensino e aprendizagem em alfabetização e letramento sempre visaram o debate dos objetos que a ela são intrínsecos (BRASIL, 1997, 2008). Conforme o estudo feito por Santos e Silva (2016) acerca da Provinha Brasil e seus pressupostos, as diretrizes caminham para a construção de noções que visem a correlação entre alfabetização e letramento, em uma relação de dependência entre ambos.

Na Base Nacional Comum Curricular (BNCC) (BRASIL, 2018, p. 89) o ensino e a aprendizagem em Língua Portuguesa estão divididos em eixos, sendo eles: oralidade, análise linguística/semiótica, leitura/escuta e produção de textos, em um movimento progressivo de estratégias que viabilizem sempre ampliar os conhecimentos e habilidades adquiridas previamente. Indica-se, portanto, que

[...] no Ensino Fundamental – Anos Iniciais, no eixo **Oralidade**, aprofundam-se o conhecimento e o uso da língua oral, as características de interações discursivas e as estratégias de fala e escuta em intercâmbios orais; no eixo **Análise Linguística/Semiótica**, sistematiza-se a alfabetização, particularmente nos dois primeiros anos, e desenvolvem-se, ao longo dos três anos seguintes, a observação das regularidades e a análise do funcionamento da língua e de outras linguagens e seus efeitos nos discursos; no eixo **Leitura/Escuta**, amplia-se o letramento, por meio da progressiva incorporação de estratégias de leitura em textos de nível de complexidade crescente, assim como no eixo **Produção de Textos**, pela progressiva incorporação de estratégias de produção de textos de diferentes gêneros textuais.

Com isso, é possível perceber o ensino e a aprendizagem focados na alfabetização em um processo contínuo, que leva a estudante a progredir em sua aprendizagem na área de Língua Portuguesa. A alfabetização, isto é, a decodificação da letra, é entendida enquanto um processo de *ortografização*. O letramento em Língua Portuguesa vincula-se, na BNCC, a diversos tipos de letramentos, relacionado ao conceito de *multiletramentos*, onde estão alocados o letramento midiático e o letramento nas demais linguagens existentes no mundo contemporâneo. Pouco se aprofunda a respeito da leitura em si, sendo dada maior atenção à aspectos fonológicos da língua, o que indica atrito na conceptualização dos fundamentos que norteiam o trabalho docente e o afastamento desta em relação aos documentos anteriores (CAVARSAN; MIGUEL, 2020), como a própria BNCC (Brasil, 2018).

Desse modo, ficam subentendidos os aspectos que Soares (2009) previamente apontou sobre a dicotômica relação entre uma visão individual dos objetos do processo de aquisição de linguagem, pela qual ficam separados os processos concernentes ao trabalho pedagógico da alfabetização e letramento. Ainda percebemos que há um reforço, por parte de documentos normativos, da confusão conceitual que as professoras possuem ao apoiarem-se em tais diretrizes para a realização da avaliação de suas alunas. Isto é, a alfabetização e o letramento desassociados, sendo o ensino da correspondência grafema-fonema desvinculados dos processos de leitura, intrínsecos ao letramento, ao longo das aulas de Língua Portuguesa. E o presente estudo aponta tal confusão, como o caso da professora A.⁵

⁵ Verificar subseção 5.3, p. 65, do presente estudo.

Soares (2009) ainda nos provoca ao frisar as dificuldades acerca dos instrumentos e processos avaliativos que tocam o letramento, trazendo à tona a relevância da devida conceptualização do objeto, a definição dos instrumentos, bem como o cuidado em sua realização. Em concordância, ao criticar a formação de leitores na escola, Krug (2015) aponta a preocupação sobre a valorização de aspectos formais do texto ao longo da escolarização, colocando o papel da escola (e seus sujeitos) enquanto mediadora principal entre as leitoras e aquilo que é lido, para além do formal.

De modo que, ao longo da escolarização, sendo a leitura imprescindível para o seu sucesso nas demais disciplinas escolares, os percursos das crianças – com seus sucessos ou não –, são devidamente analisados e trabalhados como parte dos desafios que as docentes e a escola devem inerentemente lidar. Ainda, como engrenagem desse processo, está a mediação das professoras, que são as agentes responsáveis por promover a leitura em busca de alunas leitoras (KRUG, 2015), atuando enquanto mediadoras. Preocupa-nos compreender como esses aspectos conversam entre si ao longo do processo de avaliação.

Desse modo, como fazer para avaliar a proficiência da competência em leitura de crianças nos anos precoces da escolarização sem, contudo, confundir os objetos a serem avaliados? Sem haver atrito quanto aos conceitos fundamentais da alfabetização e do letramento, onde a competência em leitura está situada? Quais instrumentos se mostram favoráveis a apreensão dessa competência em leitura enquanto objeto a ser avaliado?

A seguir, buscamos elencar os principais debates a respeito da avaliação do aprendizado da leitura.

3. AVALIAÇÃO DO APRENDIZADO DA LEITURA COMO PONTO DE APOIO PARA SEU ENSINO

Agregando-se às questões básicas e conceituais, ainda tendo em mente a inerente função de avaliadora que as professoras possuem em seu cotidiano, faz-se imprescindível compreender o que a literatura aponta a respeito dos processos de avaliação em leitura, aspecto que se encontra no seio do percurso trilhado em alfabetização e letramento ao longo do ensino fundamental – em especial nos anos iniciais.

Avaliação é um processo no qual é (são) elencado(s) um (ou mais) objeto(s) a ser(em) avaliado(s), bem como instrumentos que sirvam a este propósito, a fim de fornecer(em) informações a serem analisadas e utilizadas para tomadas de decisão. Para que seja realizada, portanto, se faz necessária a existência de uma causa, uma finalidade, um sujeito, um objeto, uma metodologia e um usuário. Comumente, a avaliação, em âmbito escolar, é envolvida em uma dicotômica relação entre avaliação interna e externa, também compreendida respectivamente como formativa e somativa, no senso comum do professorado.

A função formativa da avaliação é um processo que se dá no cotidiano escolar que visa reorganizar o processo de ensino em função dos resultados da avaliação para melhorar a aprendizagem. Contudo, se a avaliação é aplicada através de instrumentos sem o devido rigor em sua construção, especialmente se seus itens não estão plenamente articulados com os conteúdos que embasam o objeto de avaliação, com critérios muito subjetivos, os resultados da avaliação não garantem, apesar de ter intenções, um uso formativo, pois “podem priorizar aspectos que são alheios ao conhecimento que se quer descrever” (CATALANI, 2019, p. 40). Então, para que se consolide a função formativa e que sejam tecidos subsídios para tomadas de decisões do planejamento pedagógico de determinado sujeito da educação, a avaliação precisa ser válida e fidedigna.

Quanto à função somativa, ela ocorre quando temos uma síntese, como finalização dos processos de ensino, especialmente pelo fato de que não há mais tempo para incidir sobre o processo de ensino. Nesses termos, a função somativa assume ares de um “balanço pedagógico” educacional de determinado período, descrevendo o desempenho dos sujeitos.

A avaliação realizada em sala de aula, denominada por alguns autores, como avaliação interna, tem como pressuposto a avaliação com função formativa, mas pode estar sujeita, por características que historicamente foram se estabelecendo no cotidiano escolar, à falta de necessária padronização e de transparência, de tal modo que, muitas vezes, seus resultados não

se constituem em pontos de apoio para uma intervenção pedagógica para incidir na aprendizagem das estudantes.

Já uma avaliação externa, apesar das controvérsias e resistências apresentadas por docentes, por estarem relacionadas a instrumentos padronizados (testes, provas), aplicados em larga escala, podem fornecer dados que contribuiriam para o entendimento sobre as aprendizagens numa escola. Assim, apontamos, em concordância com Catalani (2019, p. 39), que

[...] nem toda avaliação interna é realizada para orientar a ação do(a) professor(a), deixando de ter o caráter formativo. Além disso, é possível que os resultados de uma avaliação externa proporcionem informação útil para replanejamento do professor(a).

Portanto, uma avaliação configura-se enquanto um processo que exige uma epistemologia específica a sua construção, com tratamentos metodológicos próprios, atribuindo a ela um caráter qualitativo, ainda que, a depender da instrumentalização adotada, essa qualificação possa se apoiar em quantificações. Mobiliza conceitos que, para fins de entendimento do presente estudo, fomentam o trabalho avaliativo educacional. Uma avaliação educacional bem fundamentada, por exemplo, de acordo com Travitzki et al. (2020, p. 69) necessita de

[...] uma matriz de avaliação, na qual se estabelece o que será objeto de avaliação e que serve de guia para a elaboração dos itens e constituição do teste; uma escala de resultados, quando se adota procedimentos de medida, que permita aquilatar diferenças de desempenho; um critério para efetuar o julgamento dos resultados; um conjunto de itens com boa cobertura da matriz e parametrizados pela escala de resultados; uma interpretação pedagógica para que esses resultados não sejam apenas valores numéricos e facultem ações pedagógicas em decorrências dos mesmos; e a fundamentação teórica documentada de todos esses elementos.

Embora seja um processo – um conjunto de elementos –, é comum que a avaliação seja entendida apenas enquanto *a prova, o teste*, um equívoco corriqueiro, de senso comum, que restringe o processo a um elemento desse processo, ou seja, ao instrumento para levantamento de informações sobre o aprendizado ao qual se relaciona a prova ou o teste.

Usualmente, o que se pode desencadear com os resultados de uma avaliação em larga escala, baseada em testes padronizados, é a função somativa, pois quando esses resultados chegam à escola, as estudantes às quais estão associados esses resultados não se encontram mais na instituição, impedindo, portanto, um uso formativo dos mesmos.

Quando o debate se volta para a avaliação escolar, o grande epicentro da polêmica sobre a sua existência material na escola, seja ela formativa ou somativa, interna ou externa, é permeada pela valorização de *rankings* (CASASSUS, 2009) e políticas de responsabilização e

accountability (FERNANDES; GREMAUD, 2020), gerando assim, ruídos no que deveria ser entendido enquanto um processo comum à escola e à educação.

Com base em Alavarse (2013), ao discorrer a respeito do paradoxo docente e a constante necessidade de se estudar os conceitos do campo da avaliação, a fim de que as docentes sejam capazes de atenuar a ausência de formação específica em avaliação educacional e da aprendizagem na formação superior em Pedagogia e demais Licenciaturas, buscamos pressupostos teóricos que favorecem análise, leitura e debate *a priori*, ao longo e *a posteriori* da construção da presente pesquisa.

Nesse sentido, o presente trabalho foca nos instrumentos avaliativos porque visa analisar a funcionalidade, os possíveis benefícios e os problemas de um software elaborado para fornecer informações às docentes e suas estudantes sobre o progresso da aprendizagem em leitura.

Para tanto, esta seção busca discutir os principais desafios e debates a respeito da avaliação do aprendizado de leitura nos anos iniciais do Ensino Fundamental (Cf. MICOTTI, 2007, SOARES, 2009, p. 83-89); em seguida, expor questões epistemológicas e metodológicas, em especial, a questão dos itens de múltipla escolha na construção de instrumentos avaliativos; para depois elucidar o caso da Provinha Brasil (BRASIL, 2016a, 2017b) como uma prova padronizada que caminha para a avaliação da competência em leitura das crianças em seus primeiros anos de escolarização básica, abordando suas limitações e potencialidades (ALAVARSE et al., 2018; CATALANI, 2019; CORRÊA; DORNELLES, 2010; MORAIS; LEAL, 2013; TRAVITZKI et al., 2020).

3.1 DESAFIOS PARA AVALIAR O APRENDIZADO DA LEITURA NOS ANOS INICIAIS DO ENSINO FUNDAMENTAL

Avaliar, embora seja parte inerente do papel docente no cotidiano escolar, mostra-se enquanto um desafio – principalmente quando estudos indicam a ausência de disciplinas obrigatórias nos cursos de educação superior de Pedagogia e demais Licenciaturas (ALAVARSE, 2013) – para as professoras, escolas e crianças. De forma que o ato de avaliar, ao longo da docência, pode se tornar uma das polêmicas da educação, enfatizada nos anos iniciais.

Colocando avaliação em leitura no centro da problematização, é importante citarmos a complexidade de serem construídos instrumentos avaliativos que agradem a “gregos e troianos”: que busquem aferir a proficiência das crianças e fornecer informações às professoras

ao longo do ano letivo, e que forneçam dados mais precisos e menos subjetivos, ainda que, em relação ao trabalho individual das professoras, a subjetividade seja um traço inerente. Mas, traço que, no que tange à avaliação, deve ser controlado pelo fato de poder levar a resultados enviesados, sem refletir o efetivo nível de aprendizado.

Ao pensar o entrelaçamento de noções que ocorrem com a avaliação do aprendizado de letramento, buscamos em Soares (2009) algumas reflexões. A autora se preocupa com três aspectos, sendo um deles a elaboração de boas avaliações, que discriminem os conceitos a serem levados em conta ao longo de seu processo (construção dos instrumentos e demais materiais, aplicação, análise de resultados etc.) sem, contudo, restringi-los (Op. cit., p. 86).

Aponta, ainda, a difícil tarefa de se abordar o letramento enquanto um objeto a ser avaliado devido à subjetividade de sua conceptualização: é um conceito que pode ser entendido em uma dimensão individual, cuja avaliação seria feita sem parâmetros mais amplos; ou um conceito que pode ser visto sob uma ótica social, submetido ao entendimento ideológico, político e social de determinada comunidade, agregando ou abolindo novos significados no decurso do tempo.

Neste caso, compreendemos a necessidade do desenvolvimento de um instrumento avaliativo que se preocupe com a competência de leitura, tratando na dimensão individual, que sirva ao cotidiano escolar e às professoras, tal como a suas alunas, e que, acima de tudo, forneça informações que aproximem-se ao máximo do que pode ser a proficiência da criança. Assim, buscamos testes que se colocassem a esse pressuposto, encontrando na Provinha Brasil um instrumento que atende as especificidades de uma avaliação externa e bem fundamentada.

3.2 OS INSTRUMENTOS DE AVALIAÇÃO E OS ITENS DE MÚLTIPLA ESCOLHA NO APRENDIZADO DE LEITURA

A preocupação com a apreensão da leitura em um instrumento avaliativo leva ao questionamento de como pode ser feita a sua construção. Devemos ter em mente a complexidade exposta por Soares (2009) acerca do processo de aquisição de linguagem, em especial no processo de escolarização, que engloba a alfabetização e o letramento – com especial atenção aos anos iniciais. Partindo do princípio de que a leitura é imprescindível e estabelecida a sua relevância ao longo do aprendizado das crianças, captar as dificuldades desde os primórdios do Ensino Fundamental se firma como um dos caminhos para sua superação, considerando as implicações em seu ensino com vistas ao sucesso escolar.

Um instrumento avaliativo é todo recurso que visa levantar informações, associadas a um sujeito ao longo de um processo educacional, em relação a um determinado conjunto de conteúdos, como é a leitura. Conforme Micotti (2007, p. 39) explicita, as avaliações devem ser entendidas como formas de “informar a professores e alunos a respeito de seus acertos e dos aspectos de seu desempenho nos quais há necessidade de mudar a atuação”. Em síntese: a avaliação é um processo imprescindível para a análise e interpretação pedagógica dos resultados do processo educacional, seja ela voltada para objetos de Língua Portuguesa, seja para quaisquer outras.

Pensando o desafio de se conceber práticas pedagógicas que visem a apreensão do processo de aquisição de leitura e escrita, a autora (Op. cit., p. 44) indica que:

O enfoque interativo dos vários aspectos da aquisição da escrita propõe à prática pedagógica a necessidade de desenvolver o ensino de modo tal que contemple a abordagem pelos alunos dos vários tipos de textos em situações reais de comunicação, de modo a desenvolver a compreensão e os procedimentos intelectuais que entram em jogo nessa aquisição, sem perder de vista o trabalho com os elementos constitutivos da escrita em suas relações com a língua oral. Este contexto teórico indica novas orientações para o ensino.

Dessa forma, aponta que deve haver uma reavaliação constante das atividades didáticas, a fim de organizá-las para o apoio interativo entre as crianças e os escritos, isto é, a língua. O desafio consiste em elaborar métodos para trazer maior independência na relação entre criança e o objeto, sobretudo em aspectos que remetem à medição e ao julgamento do que é preconizado como habilidades a serem adquiridas, por parte das crianças, ao longo da escolarização.

Assim, forma-se o desafio de construir instrumentos avaliativos que estejam em concordância com o que é realizado no cotidiano escolar, através do trabalho docente e do que é vivenciado, passos que nos levam a debater tanto sobre esses instrumentos, quanto a formação docente, inicial e continuada, nesse sentido. Um dos instrumentos que focalizam a avaliação da proficiência em leitura no início do Ensino Fundamental é a Provinha Brasil.

3.3 O CASO DA PROVINHA BRASIL: POTENCIALIDADES E LIMITAÇÕES

Implementada pelo Ministério da Educação em 2008, sob responsabilidade do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), a Provinha Brasil (PB)⁶

⁶Denominada como Avaliação da Alfabetização Infantil – Provinha Brasil (BRASIL, 2016a, p. 4) na época de sua implementação em 2008.

caracteriza-se como um instrumento cujo objetivo é fornecer informações que visem auxiliar nos processos de avaliação da alfabetização e letramento, e alfabetização matemática, tal como possibilitar um “conhecimento imediato dos resultados da turma” (BRASIL, 2016a, p. 4), capazes de fornecer subsídios para a tomada de decisão pedagógica das docentes em sala de aula, bem como as gestoras e a comunidade escolar. A PB foi disponibilizada pelo Inep até o ano de 2016, tendo sido descontinuada na sequência.

A Provinha Brasil (Op. cit., p. 6) foi apresentada tendo como objetivos:

- avaliar o nível de alfabetização dos educandos nos anos iniciais do ensino fundamental
- oferecer às redes e aos professores e gestores de ensino um resultado da qualidade da alfabetização, prevenindo o diagnóstico tardio das dificuldades de aprendizagem
- concorrer para a melhoria da qualidade de ensino e redução das desigualdades, em consonância com as metas e políticas estabelecidas pelas diretrizes da educação nacional.

Nesse sentido, ao ter como respaldo o Plano de Metas – Compromisso Todos pela Educação (BRASIL, 2007) em que a meta era a alfabetização das crianças até os 8 anos de idade, sendo que a aferição dos resultados se daria por meio de um instrumento avaliativo periódico e específico, a Provinha Brasil é elaborada e instituída. Uma das justificativas para a faixa etária indicada é a de que as crianças teriam vivenciado, pelo menos, um ano letivo ao ser aplicada no início do 2º ano do Ensino Fundamental. É importante frisar que a adoção da PB nas escolas era de cunho opcional, ficando a critério das escolas e de sua comunidade, sem haver retorno dos resultados para o Inep.

A organização de sua aplicação, inicialmente, se daria da seguinte forma: uma primeira aplicação no início do ano letivo, com os resultados organizados pelas professoras e repassados para a gestão escolar; e, no fim do ano letivo, seguindo a mesma organização de coleta e tratamento de dados. Desta forma, seria possível apreender os momentos iniciais e finais das crianças no processo de alfabetização e letramento do ano letivo vigente. Portanto, a Provinha Brasil tinha o intuito de ser uma avaliação diagnóstica (BRASIL, 2016a, p. 7).

O teste da PB era composto por itens de múltipla escolha relacionados a uma matriz de referência com descritores acerca das habilidades preconizadas para a alfabetização e letramento. Abrangia apenas a leitura, com exceção do primeiro ano de aplicação, no qual havia itens de escrita, pois, especialmente por questões metodológicas e organizacionais, não eram avaliados os processos de escrita das crianças respondentes (BRASIL, 2016a), apoiando-se em

três eixos: apropriação do sistema de escrita; leitura; compreensão e valorização da cultura escrita (Op. cit., p. 9).

De acordo com os resultados dos testes, as respondentes eram alocadas em uma escala com níveis de proficiência com suas respectivas interpretações pedagógicas, evidenciando o que dominavam as estudantes ali alocadas, subsidiando o trabalho pedagógico acerca dos domínios e dificuldades que suas alunas possuem.

Logo, a PB caracterizava-se enquanto uma avaliação fundamentada, uma vez que possuía características metodológicas próprias de uma avaliação em larga escala (ALAVARSE; CATALANI, 2016, TRAVITZKI et al., 2020), pois cumpria com os pressupostos desse tipo de avaliação, chamando-nos a atenção por voltar-se especificamente para a leitura, sendo um instrumento avaliativo importante para o processo pedagógico.

No entanto, o que se pode observar nos estudos realizados a seu respeito (ALAVARSE et al., 2018; CATALANI, 2019; TRAVITZKI et al., 2020) é que, embora seja construída de modo coerente, com itens que contribuíam para avaliar a proficiência em leitura, com matriz de referência e orientações acerca de sua aplicação e análise de resultados, a PB possuía algumas limitações que diziam respeito ao seu financiamento e logística, aplicação e organização dos resultados. Os argumentos utilizados destacavam:

- **Financiamento e logística:** por se tratar de uma avaliação em larga escala e padronizada, a PB dependia de impressão, organização e entrega para fosse realizada⁷.
- **Aplicação:** Sendo aplicada pelas professoras, demandava, em média, duas horas, o que gerava desconforto para as professoras devido às limitações sobre o gerenciamento do tempo e da organização de sala de aula, pois, como as crianças tinham, em geral, níveis de proficiência variáveis, desestimulava aquelas cuja proficiência leitora se encontra nos extremos – inferior ou superior – da escala⁸.

⁷ É importante apontarmos que, uma das motivações para a descontinuidade da PB, para além de financeira, foi a elaboração de uma Base Nacional Comum Curricular, conforme uma nota divulgada pelo MEC (TRAVITZKI et al., 2020, p. 69).

⁸ Isto porque, por se tratar de um teste impresso fixo e de realização, pela aplicação conduzida pela professora, linear, a Provinha Brasil apresenta todas as 21 questões para todas as respondentes. As crianças que se encontram nos níveis iniciais de proficiência não encontram dificuldades para a realização das primeiras questões, mas frustram-se com as questões finais; em contrapartida, as crianças que possuem proficiência mais avançada, ficam desestimuladas com os itens iniciais, bem como com o tempo que as demais colegas levam para respondê-los. (CATALANI, 2019, p. 53-54).

- **Organização dos resultados:** a sistematização primária dos dados por parte das professoras era um trabalho feito de forma manual que, além de demandar muito tempo, era passível de erros⁹.

Levando em conta a interpretação pedagógica dos dados obtidos através da aplicação a PB, o kit da avaliação indicava as competências que as respondentes possuem de acordo com o número de acertos nos itens elaborados.

No Quadro 1 estão sintetizadas as informações disponibilizadas pelo Inep (BRASIL, 2016a, p. 29) para a interpretação dos resultados, intimamente relacionada à matriz de referência, cujo cerne é apontar os objetos a serem avaliados.

Quadro 1 – Número de acertos, nível de proficiência e competências das respondentes.

Provinha Brasil – Teste 2. 2016

| Nº de Acertos | Nível de Proficiência | Competências |
|-----------------|-----------------------|--|
| 3 acertos | 1 | <ul style="list-style-type: none"> • Diferenciar letras de outros sinais gráficos; • Identificar letra ou sequência de letras do alfabeto lida pelo aplicador; |
| 4 – 9 acertos | 2 | <ul style="list-style-type: none"> • Reconhecer palavras de formação silábica canônica escritas de diferentes formas; • Estabelecer relação entre grafemas e fonemas, identificando, por exemplo, a letra ou a sílaba inicial de uma palavra; • Ler palavras formadas por sílabas canônicas e não canônicas |
| 10 – 14 acertos | 3 | <ul style="list-style-type: none"> • Identificar o número de sílabas em uma palavra; • Ler frases de sintaxe simples com o apoio de imagens ou ditadas pelo aplicador; • Identificar informação explícita de fácil localização em textos curtos com o apoio da leitura pelo aplicador ou pela leitura individual; • Inferir informações em textos curtos de gêneros usuais, pela leitura individual e com o apoio em linguagem não verbal; • Reconhecer o assunto do texto com o apoio do título ou de conteúdo informacional trivial, com base nas características gráficas do gênero, pela leitura individual ou com o auxílio da leitura pelo aplicador; • Reconhecer a finalidade de textos de gêneros usuais (receita, bilhete, curiosidades, cartaz) com base nas características gráficas desses e na leitura individual. |

⁹Ver Alavarse e Catalani (2016).

| | | |
|--------------------|---|---|
| 15 – 16 acertos | 4 | <ul style="list-style-type: none"> • Identificar informação explícita não trivial em textos curtos ou médios, com o apoio da leitura pelo aplicador ou com base em leitura individual; • Reconhecer a finalidade de um texto a partir de leitura individual, sem o apoio das características gráficas do gênero ou explorando seu conteúdo informacional; • Reconhecer o assunto de textos curtos e médios lidos individualmente sem o apoio das características gráficas do gênero; • Inferir informações não triviais em textos curtos pela leitura individual e apoio nas características do gênero; • Relacionar um nome a seu referente anterior em textos curtos e médios. |
| 17 – 20 acertos | 5 | <ul style="list-style-type: none"> • Reconhecer o assunto de um texto longo com base no título, a partir de leitura individual; • Reconhecer o assunto de textos médios por meio de inferências com forte base no conteúdo informacional, a partir de leitura individual; • Identificar informação explícita não trivial, por vezes secundária, em um texto curto ou médio, com base em leitura individual; • Inferir informação não trivial em textos médios com base em leitura individual ou com o apoio de leitura pelo aplicador; • Reconhecer a finalidade de um texto de construção complexa lido silenciosamente com o apoio de suporte. |

Fonte: Brasil (2016a)

Os dados do Quadro 1 nos levam a perceber a epistemologia acerca dos conceitos de alfabetização e letramento que eram considerados na PB. De modo semelhante, também aponta as restrições da prova em relação à oralidade e escrita (BRASIL, 2016a), um dado relevante ao se pensar a avaliação das competências de Língua Portuguesa preconizadas pelos documentos orientadores aos quais se referenciava (BRASIL, 1997, 2008).

A esse respeito, os estudos de Micotti (2013) e Moraes, Leal e Pessoa (2013) acabam por indicar limitações da PB e de sua utilização nas escolas. Frisa-se, nos artigos, o peso da matriz de referência e sua relação com o trabalho pedagógico das professoras em sala de aula, somado, ainda, aos entendimentos conceituais no campo da alfabetização e letramento.

Moraes, Leal e Pessoa (2013) entendem que a PB era um importante instrumento avaliativo para a formulação de um referencial quanto ao que se deve avaliar no processo de alfabetização e letramento. Entendem-na como um instrumento com potencial para enfatizar políticas de formação continuada voltadas para a discussão da avaliação educacional, bem como pressupostos epistemológicos na aquisição da língua portuguesa e materna. Possibilitaria, portanto, compreender em que pé estaria o aprendizado das crianças que eram submetidas a ela, relacionando os resultados às práticas pedagógicas, a fim de auxiliar o (re)planejamento das atividades e propostas didáticas.

No entanto, indicavam algumas limitações quanto à organização da PB, sendo elas: a reavaliação qualitativa, a cada edição, do banco de itens, a fim de que fossem exploradas as características dos mesmos e suas limitações e adequações; indicavam a necessidade de discussão do currículo e da alfabetização, em especial a formulação da matriz de referência, uma vez que apontam como imprescindível a existência da avaliação da habilidade de escrita e ponderavam como injustificável a sua exclusão da PB nas edições mais recentes; apontavam a necessidade de se pensar os meios estatísticos de apuração de resultado, para que sejam epistemologicamente críveis e confiáveis; e ainda apontavam a necessidade de diálogo com as professoras, na formulação da PB em si, já que, de acordo com o observável, não são as professoras que exercem seu direito de estar à frente da aplicação da PB, bem como da análise de seus resultados.

Em consonância ao exposto, Micotti (2013), ao relacionar as principais concepções metodológicas no campo da alfabetização e letramento, enfatiza a interação e integração dos processos de avaliação no trabalho docente. Seu principal apontamento é referente à matriz pedagógica da PB, especialmente no tocante aos objetos que são enfatizados e aqueles que são excluídos do processo de avaliação deste instrumento. Sua preocupação é a ausência da avaliação do processo textual de escrita das crianças submetidas à prova, encaminhando para a necessidade de se existirem estudos à inclusão desta à PB, de forma a contribuir para a avaliação do desempenho na produção de textos – algo que considera ser um dos pontos mais fracos do ensino brasileiro (Op. cit., p. 193).

Assim, ambos os artigos caminham para a discussão de formas de avaliar que estejam em consonância com os debates e as práticas das professoras em sala de aula, pensando a PB enquanto um instrumento avaliativo, não como a totalidade de um processo de avaliação.

Embora se deva ter presente os debates em relação às limitações expostas, chama-nos especial atenção as orientações sobre a análise dos resultados (BRASIL, 2016a, p. 41), sua divulgação e utilização (p. 42) e a socialização à comunidade escolar (idem). As orientações concernentes à orientação dos dados, quando da interpretação das respostas, apontam a análise das respostas incorretas nos itens do teste. Isto porque cada uma das alternativas dos itens da PB demonstra algum raciocínio realizado pela respondente, caminhando para maior compreensão das dificuldades que a criança possui. Assim, serve de aporte ao planejamento do trabalho docente.

Quanto à divulgação e do uso dos dados, o material propõe uma frente de trabalho específica e dentre as indicações feitas, encontra-se a discussão entre as professoras e seus grupos de trabalho, a comunicação efetiva com as turmas a respeito dos dados coletados, tal como o

trabalho coletivo em prol do reconhecimento dos erros e das estratégias na correção dos mesmos; o replanejamento das ações pedagógicas; e a comunicação aberta com a comunidade escolar (famílias das respondentes), a fim de garantir o trabalho conjunto, através de (sugeridas) estratégias para a devida orientação acerca dos resultados.

Interessa-nos compreender as diversas dimensões do processo de avaliação da PB, uma vez que está em consonância com o que a literatura diz respeito aos processos avaliativos externos e a aplicação de testes que se figurem a apoiar as docentes. Sobretudo, no que tange ao estudo da funcionalidade de um instrumento avaliativo que nela se baseia, é imprescindível compreender a organicidade do processo pressuposto pela instituição ao aderir à Provinha Brasil, já que a pesquisa se deu na medição da proficiência de alunos de duas escolas.

4. MEDIR E AVALIAR

Admitindo-se que há algo a ser ensinado e para ser aprendido no processo escolar, a problematização do presente trabalho gira em torno da avaliação da aprendizagem, o que, por sua vez, se desdobra em outras questões: o que é avaliar ou a avaliação? Quais instrumentos podem ser utilizados para se captar as diferenças de conhecimento entre os sujeitos? Como avaliar as diferenças de desempenho entre os anos escolares? Qual seria a relação entre medir e avaliar?

Inicia-se, então, com a problematização em torno da avaliação. Se afirmamos que a avaliação é um processo, com levantamento, processamento e análise das respostas dadas às tarefas avaliativas – o desempenho – vinculadas a objetos educacionais, a fim de que sejam subsidiadas as tomadas de decisões no âmbito da gestão escolar e/ou na sala de aula, podemos igualmente afirmar que a avaliação é um processo no qual um juízo de valor a respeito da aprendizagem de uma estudante, sobre uma competência, é tecido após a coleta de dados por meio de instrumentos avaliativos. Desse modo, a avaliação pode estar atrelada à ideia de medição e à de julgamento, sendo este uma consideração a partir de um valor que se configura como critério ao qual o resultado da medição é comparado.

Quando falamos de avaliação educacional, alguns pressupostos epistemológicos e de ordem de senso comum tornam-se barreiras ao processo, conforme já explicitado no capítulo anterior. Frisa-se os apontamentos já realizados acerca dos métodos e estudos particulares ao campo, que levam a elaboração de formas de se medir um objeto não-observável, tal qual a aprendizagem. Historicamente, a avaliação está atrelada à educação, uma vez que para planejar uma prática metodológica, didática e educacional se faz necessária a compreensão sobre o quanto os sujeitos a ela submetidos compreendem de um conteúdo que teria sido objeto de ensino. Portanto, a medição em um processo avaliativo é necessária para o dimensionamento do que teria sido apropriado, apesar das dificuldades inerentes ao processo de medida de algo que não pode ser medido diretamente¹⁰. E para que seja feita de forma coerente, alguns raciocínios estatísticos e de ordem organizacional foram movidos e estudados, a fim de que o juízo de valor obtido fosse fundamentado e justificado.

Uma das metodologias adotadas em algumas avaliações, para o tratamento das respostas aos itens de seus instrumentos com vistas a estimar a proficiência dos respondentes, é a Teoria de

¹⁰ Isto porque o aprendizado é um traço latente, interno ao aprendiz, do qual o que se pode fazer é inferir – uma suposição – em função das respostas que apresenta às tarefas avaliativas.

Resposta ao Item (TRI), apontada como “uma forma para se descrever ou medir uma variável não observável (ou não mensurável) por mecanismos convencionais, como a aprendizagem” (BRASIL, 2012, p. 11), através dos resultados obtidos nos itens de testes elaborados. Concorrendo a isto, se estabelece como uma metodologia de grandes avaliações padronizadas, externas e em larga escala, tal como o Exame Nacional do Ensino Médio (Enem). Entendemos a TRI enquanto uma metodologia que gera resultados numa métrica capaz de favorecer a “interpretação pedagógica desses resultados, o que, por sua vez, possibilita que orientações curriculares e didáticas sejam produzidas” (ALAVARSE; MELO, 2013, p. 106).

Outra metodologia empregada no tratamento de respostas a testes com vistas à aferição do desempenho cognitivo e de conhecimentos é a Teoria Clássica dos Testes (TCT) que, no entanto, fica restrita ao total de acertos do teste para o grupo de respondentes desse teste, sendo este total de acertos uma síntese do que se pretende avaliar. Tanto a TRI quanto a TCT são metodologias empregadas visando, de algum modo e com diferenças entre elas, medir o conhecimento – domínio – sobre determinado(s) objeto(s) educacional(ais) através da performance do sujeito em tarefas – os itens – a ele submetidas, sendo consolidadas como métodos psicométricos de controle de coleta de dados, buscando consolidar precisão e confiabilidade do processo de avaliação que se apoia em medidas educacionais.

Os testes, com o tratamento das respostas e um resultado disso decorrente, permitem um juízo de valor através da análise de seus resultados, de modo que a sua funcionalidade, dentro do campo da educação, tem como propósito o suporte à avaliação educacional. Assim, devem se atrelar a elaboração de uma matriz de referência, a uma escala e níveis de proficiência, subsidiando os objetivos qualitativos elencados pelo grupo docente, em consonância com as diretrizes educacionais democraticamente instituídas enquanto parâmetros aos quais a qualidade educacional deve se basear.

A fim de compreender tais aspectos que levam ao juízo de valor, em especial por ser ele o que guiará, sem com ele se confundir, a tomada de decisão pedagógica do docente no cotidiano escolar e seu trabalho, buscamos discutir aspectos que dizem respeito a interpretação pedagógica baseada na análise da proficiência das estudantes, bem como os estudos sobre os testes e seus suportes, explorando estudos sobre a elaboração de testes computadorizados e adaptativos.

4.1 PROFICIÊNCIA E NÍVEL DE PROFICIÊNCIA COM INTERPRETAÇÃO PEDAGÓGICA

Uma matriz de referência é um documento que se propõe a estabelecer as habilidades e os conhecimentos em relação aos quais um objeto será avaliado por meio de instrumentos. Assim sendo, uma avaliação bem estruturada deve possuir uma matriz de referência com descritores – isto é, a descrição das habilidades organizadas no documento – claros a quem a utilizará para a análise das respostas dos respondentes. É com base nessas habilidades que os itens são elaborados de tal modo que as respostas a eles, sobretudo pelo tratamento dado a elas, possam gerar um resultado com interpretação pedagógica, isto é, que aludem a essas habilidades e conhecimentos.

Conforme apontado no material da PB (BRASIL, 2012, p. 14) consta que esses descritores “traduz[em] as habilidades que são avaliadas e que refletem a associação entre os conteúdos curriculares e as operações mentais realizadas pela criança ao responder à questão”. Chama-nos atenção, no entanto, que a matriz de referência de uma avaliação não é sinônima de currículo, embora deva, por processos apropriados, nele encontrar ressonância, isto é, aquilo que será base para a construção dos instrumentos deve integrar, por formas mediadas, o currículo, sem necessariamente esgotá-lo. Está, portanto, intimamente relacionada aos itens – questões – de um teste e à escala de proficiência, já que a matriz de referência servirá de aporte, via a elaboração dos itens, para a organização dos níveis de proficiência estimados. Podemos entender um item de um instrumento como “uma unidade básica de um instrumento de coleta de dados [...]” (BRASIL, 2006, apud BRASIL, 2012, p. 31).

A respeito do nível de proficiência, podemos realizá-lo enquanto “ponto[s] do processo de aprendizagem em que os estudantes se encontram no momento da aplicação [...] e que podem ser usados como referência para o planejamento do ensino e da aprendizagem” (BRASIL, 2016a, p. 23). Para a definição de níveis de proficiência com a alocação de estudantes em cada um dos níveis de uma escala de proficiência, existem tratamentos estatísticos que extrapolam os objetivos deste trabalho.

Os itens de um teste, que devem ser elaborados com rigorosa metodologia¹¹, podem ser categorizados como itens de resposta construída, itens de múltipla escolha ou itens de atuação; sendo que a condição para serem objetivos é dada por uma resposta correta, assim considerada por uma dada comunidade de elaboradores desses itens ou de profissionais vinculados aos seus

¹¹ Ver BRASIL (2010).

conteúdos, como é o caso dos professores. Os itens de múltipla escolha, que são os itens utilizados no TAI-Leitura, “permitem ao participante do teste escolher a resposta entre várias alternativas, das quais apenas uma é correta” (BRADFIELD; MOREDOCK, 1964, apud BRASIL, 2012, p. 8). Assim, verificam “tanto comportamentos simples, de memorização ou reconhecimento, como comportamentos mais complexos, envolvendo compreensão, aplicação, análise, síntese e avaliação” (idem). Mostram-se necessários e indicados para avaliações em larga escala devido a sua rapidez de processamento de respostas, além de serem menos vulneráveis a erros de julgamento.

Os itens e suas respostas são responsáveis, após tratamento psicométrico, pela aferição da estimativa de proficiência dos respondentes dentro de uma escala de resultados que pode estar segmentada em níveis de proficiência com suas respectivas interpretações pedagógicas. A proficiência estimada, por sua vez, quando comparada a um critério leva ao seu julgamento, conforme supracitado. Ou seja, serve de suporte à análise da alocação da proficiência estimada da criança dentro de uma escala nivelada, já que nesta estão apontadas as habilidades que, aproximadamente, a criança domina, levando-se em conta os níveis e seus descritores.

A interpretação pedagógica ganha destaque no presente estudo uma vez que é ela que contribui para a tomada de decisão, por parte das professoras, no processo pedagógico da escolarização. A partir da interpretação pedagógica é possível obter “uma descrição quanto aos domínios cognitivos dos(as) estudantes cujas proficiências aí se localizam [...]” (CATALANI, 2019, p. 47) dentro de uma escala de proficiência. Isto é, a interpretação pedagógica pode guiar o planejamento ou replanejamento da prática pedagógica das professoras, com o objetivo de lidar com as dificuldades dos sujeitos da educação, bem como com o avanço dos objetos preconizados para seu efetivo sucesso escolar.

Ainda sobre a escala de proficiência, é importante frisar que uma escala de qualidade deve ser construída com o suporte de diretrizes educacionais, tal como com o constante debate e discussão com as professoras e demais especialistas educacionais. Estes, por sua vez, interpretam aquilo que deve ser considerado enquanto nível satisfatório para as crianças e demais sujeitos. Essa interpretação facilita a determinação de juízos de valor para o desempenho dos estudantes, algo que deixa entendida a necessidade de construção de instrumentos precisos e construídos com qualidade para a aferição da proficiência.

Deste modo, um instrumento avaliativo deve ser funcional para as professoras e suas estudantes, servindo aos propósitos dos trabalhos pedagógicos, articulando-se com as necessidades e determinações educacionais. A realização de procedimentos metodológicos através de testes

psicométricos não leva apenas ao ranqueamento, conforme já explicitado, e muito menos a demais políticas educacionais; leva também ao entendimento da síntese do que foi realizado, cognitivamente, até então, a depender do modo como são utilizados. A respeito da utilização dos dados medidos em um processo avaliativo, Catalani (2019, p. 47-48) expõe que

A [tomada de] decisão envolve ponderar as informações que a medida oferece com outros aspectos do contexto educacional, quais sejam, histórico-sociais, materiais, culturais, entre outros, para ensejar medidas pedagógicas que possam fazer com que cada estudante, estando abaixo do que era esperado, possa avançar em sua aprendizagem no objeto em questão.

Ressaltamos, ainda, a relevância da consideração de que a medição da proficiência de uma estudante é uma aproximação – uma estimativa – quanto ao domínio dos conteúdos do teste, baseados numa matriz, que deve ser considerada enquanto compreensão do objeto a ser avaliado. Isto porque o erro de medida existe em toda a qualquer metodologia de estimativa de proficiência, e deve ser considerada para a tomada de decisão, conforme exposto por Alavarse e Melo (2013, p. 105).

A seguir, expomos a discussão sobre os diferentes tipos de testes psicométricos e seus suportes.

4.2 TESTES EM PAPEL E DIGITAIS: POTENCIALIDADES E LIMITAÇÕES

Compreende-se um teste enquanto uma prova, um questionário, constituído de tarefas propostas aos respondentes do teste, chamadas de itens, que podem ser de resposta construída ou de múltipla escolha. Essas tarefas podem ser na forma de pergunta, de completar ou de executar alguma operação. É, portanto, um instrumento que busca coletar a compreensão da respondente em relação a conteúdos cujas respostas, após o devido tratamento, constituirão o objeto de avaliação. Suas respostas podem ser sintetizadas e sistematizadas posteriormente, a fim de se chegar a uma medição daquilo que se quer compreender enquanto aprendizagem do objeto educacional¹². É entendido enquanto um instrumento que, ao ser aplicado, é capaz de fornecer dados a respeito da proficiência das respondentes, na verdade, uma estimativa da proficiência que será julgada, de acordo com os critérios determinados.

A aplicação de teste é comumente associada à ideia de um questionário impresso em um papel, cujas respostas devem ser escritas ou marcadas, seja na própria questão, seja em uma folha de

¹²Lembramos, no entanto, que a apreensão da aprendizagem é apenas uma aproximação do que foi compreendido pelos sujeitos da educação, conforme já explorado no presente trabalho, pois trata-se de um objeto não observável, imaterial, restando aos instrumentos o desafio de se aproximar o máximo possível da realidade daquilo que é subjetivo.

resposta à parte. Tradicionalmente, sobretudo quando se trata de avaliações em larga escala, o suporte utilizado é o papel, mas isso tem sido repensado nos últimos anos com o advento da tecnologia digital, como podemos ver nas aplicações do Enem digital. Portanto, o quadro mais frequente são provas impressas, distribuídas entre as mesas das crianças, jovens e mesmo adultos, seja para um vestibular, seja para ver o que o sujeito compreendeu sobre determinado objeto.

Chama-nos atenção, no entanto, as limitações que um teste em papel, respondido à lápis ou caneta, apresentam. Conforme exposto por Catalani (2019) em relação a disponibilização dos testes da PB, a aplicação de uma avaliação desse gênero demanda a impressão, o que significa afirmar que é necessário a compra dos materiais e demais necessidades para que seja efetivamente distribuída às escolas. Colocamos em evidência o gasto de verba com tal movimentação, uma das discussões mais presentes em âmbito educacional nos últimos tempos.

Alguns estudos acerca de testes informatizados, realizados através de computadores e demais aparelhos eletrônicos que acessem à internet ou que consigam se colocar à disposição da elaboração de aplicativos e demais recursos tecnológicos, aparecem enquanto possíveis meios de se repensar as formas com as quais as aplicações dessas provas têm se dado. Guardadas as polêmicas envolvendo a sua realização, o Enem (2020) buscou se consolidar enquanto um exame que lança mão de alguma tecnologia para a sua realização, para citar um exemplo.

Entretanto, o suporte digital também oferece desafios passíveis de estudos para a sua devida realização, seja o investimento público na democratização de acesso ao suporte eletrônico (*tablets*, computadores, internet e demais meios e mídias) e sua efetiva utilização, ou mesmo a segurança e confiabilidade em relação as respostas realizadas pelas respondentes ao longo deste teste.

Portanto, ao pensar a PB enquanto um material com potencial para a avaliação da competência em leitura, adaptá-la para um recurso informatizado foi um dos desafios preconizados pelo Gepave. Para tal, o estudo de tal tecnologia se fez necessário e, à luz das reflexões sobre o recurso, buscamos elucidar alguns entendimentos por e sobre um teste informatizado, isto é, um teste digital.

4.3 TESTE ADAPTATIVO INFORMATIZADO (TAI) COMO UMA ALTERNATIVA E SUAS CARACTERÍSTICAS GERAIS

Um teste adaptativo, em geral, caracteriza-se como um instrumento avaliativo que busca aferir a proficiência através da adaptação dos itens disponibilizados para as estudantes de acordo com suas respostas aos itens anteriores, o que lhe empresta a adjetivação de teste adaptativo que, a princípio, pode ser aplicado em papel com a apresentação de itens em papel à respondente. Um TAI é um teste que tem como suporte para a sua aplicação os recursos tecnológicos informatizados (computadores, *tablets* e/ou *smartphones*) (ALAVARSE; MELO, 2013, p. 106). Especificamente a respeito de ser adaptativo, podemos realizá-lo enquanto um teste cujos itens se adaptam as respostas fornecidas pelo respondente, de tal modo que somente quando o teste foi concluído é que se pode saber quais e quantos itens o compuseram, diferentemente dos testes fixos, em papel ou digitais, que estão prontos antes de seus itens serem respondidos.

Quanto à sequência de respostas aos itens do teste, o TAI-Leitura é linear, pois o respondente só recebe novo item após responder ao item anterior, sem poder retroceder aos itens respondidos ou avançar sem responder a um item apresentado. A PB, em sua versão em papel, é um exemplo de teste linear na perspectiva de quem o aplica, pois oferece itens de forma linear – contínua –, com um item sendo apresentado – ou indicado para ser respondido – na sequência do anterior. Ainda, a PB é um teste fixo – com 20 itens para quaisquer estudantes.

Então, a PB foi configurada como um teste fixo – linear e não adaptativo. Essa última característica indica que, independentemente da resposta dada pela respondente ao item anterior, não há a consideração do que foi respondido para definir o item que será colocado a seguir. O resultado da aplicação da PB depende de uma análise final, um trabalho que demanda a leitura e alocação da respondente em uma escala de acordo com seus acertos, utilizando maior tempo e maior organização para a sistematização dos resultados *a posteriori*.

Entretanto, um teste informatizado não significa que seja adaptativo. Um Teste Baseado em Computador (TBC), embora seja realizado através de meios eletrônicos e informatizados, com algumas vantagens em relação a um teste em papel, sobretudo no processamento das respostas, não necessariamente adapta a apresentação dos itens, como num TAI, em função das respostas aos itens. No TBC, essencialmente, a interação com o recurso eletrônico facilita a captação rápida dos dados do teste, de forma que há a obtenção de resultados quase que instantaneamente – a depender da conexão, por exemplo, ou mesmo do aparelho utilizado.

Contudo, como uma característica comum, por ser um teste informatizado, tanto um TBC quanto um TAI favorecem novas possibilidades de elaboração de seus itens. Alavarse e Catalani (2016, p. 40), ao lançar olhares para as vantagens da informatização dos testes, apontam algumas possibilidades:

a) ampliação dos modos de apresentar os contextos e/ou objetos auxiliares/suportes na reflexão proposta pelo item e que mobilizam uma resposta do respondente, podendo integrar movimento e som às figuras, gráficos, textos e desenhos já utilizados nos testes de lápis e papel; e b) um incremento das operações cognitivas solicitadas, suplantando as possibilidades de expressar escolhas, descrições, identificações, comparações, relacionamentos, análises e avaliações em torno dos fatos, fenômenos ou linguagens, objetos da aferição pretendida.

Assim, os itens em meio digital podem apresentar potencialidades em relação àqueles elaborados para apresentação em papel, uma vez que incorporaram recursos que o papel não permite, como, por exemplo, som e imagem animada.

Sendo o TAI um teste informatizado, ele proporciona maior rapidez na obtenção de respostas e oferece suporte mais preciso no processo pedagógico de avaliação por parte das docentes, já que os dados coletados são processados de forma mais rápida (TRAVITZKI et al., 2020, p. 527) através de suportes computacionais e informacionais. É apontado como um instrumento com potencial para a aferição do nível de proficiência de determinado objeto de conhecimento. Os autores descrevem um teste adaptativo do seguinte modo:

No início do teste, a cada respondente é apresentado um item, o qual pode, por exemplo, ser selecionado aleatoriamente ou escolhido de acordo com uma condição pré-determinada. Posteriormente, durante a realização do teste, ele se adapta a cada indivíduo, dependendo de seu desempenho (representado por seus acertos e erros), principalmente através da seleção de itens mais difíceis ou mais fáceis, e tem o objetivo tanto de otimizar a precisão da estimativa de proficiência de cada indivíduo quanto de reduzir o tempo de aplicação, em contraste com os testes em papel ou apresentados em computador, considerados lineares, que desde o início estão prontos com uma determinada quantidade de itens, mesmo que apresentados sequencialmente. Essa é a essência de um teste adaptativo, que pode apresentar diversas variações e níveis de complexidade. Quando esse tipo de teste é aplicado em formato digital, mediante processos informatizados, denomina-se Teste Adaptativo Informatizado.

Ademais, vista como uma possibilidade para sanar alguns dos problemas previamente expostos, um TAI pode ser entendido como um meio de viabilizar a diminuição de erros. Alavarse e Melo (2013, p. 107) elencam que

[...] com a utilização da TRI e o desenvolvimento de um algoritmo informatizado, a quantidade de itens que será utilizada na estimativa da proficiência e o intervalo de tempo para a divulgação dos resultados serão significativamente menores do que os necessários quando a avaliação é realizada no formato lápis e papel.

Em relação ao desempenho das respondentes, Alavarse e Mello (2013, p. 107) enfatizam a hipótese de que “o desgaste físico e mental do respondente será diminuído, reduzindo assim os erros provocados pelo chamado ‘efeito-cansaço’, que expressa o fato de [...] exibir menor proficiência do que efetivamente tem quando se aumenta o número de itens a responder”. Destacam, ainda, o TAI enquanto uma ferramenta com potencial para contribuir na construção e adaptação das provas aplicadas em larga escala.

Alavarse e Catalani (2016, p. 39) apontam algumas das vantagens que um TAI pode proporcionar se em comparação com os testes em papel:

[...] adaptar o teste aos conhecimentos do respondente; possibilitar integração das ferramentas tecnológicas na elaboração das questões (itens) da prova; reduzir o tempo de aplicação do teste; fornecer acesso imediato, após o término da aplicação, aos resultados da análise; conferir maior precisão às estimativas da proficiência; e possibilitar a interpretação pedagógica dos resultados.

Entretanto, embora existam hipóteses que elevem o TAI enquanto um instrumento avaliativo de inúmeras potencialidades, para que seja construído, se faz necessário investimento em recursos tecnológicos, em pessoal capacitado (um grupo interdisciplinar que lance mão de conhecimentos necessários para a sua estruturação), além das características intrínsecas a uma avaliação fundamentada, em especial no que tange ao banco de itens.

4.4 O TAI-LEITURA DO GEPAVE

Em desdobramento de apontamentos de Alavarse e Melo (2013), foi elaborado, por meio de um trabalho interdisciplinar no Gepave, um Teste Adaptativo Informatizado baseado na Provinha Brasil (ALAVARSE; CATALANI, 2016) e, posteriormente, uma nova versão, denominada TAI-Leitura, instrumento que o presente trabalho tem como objeto de estudo. Trata-se de um recurso com uma série de itens para alocar o respondente, ao final do teste, dentro de uma escala de proficiência, mas, é importante frisar que tal escala, assim como os itens destinados ao teste, foram baseados na PB (BRASIL, 2016a; 2016b).

Catalani (2019) investigou a construção do TAI-Provinha Brasil desenvolvido para avaliar a aprendizagem em leitura de crianças pertencentes aos anos iniciais do Ensino Fundamental, que se alinhou com as ponderações realizadas por Alavarse e Melo (2013) a respeito da criação de um *software* capaz de fornecer itens adaptados aos respondentes de acordo com a sua resposta,

processando e fornecendo os dados de forma mais prática. Alavarse e Catalani (2016, p. 36) indicam alguns pressupostos quanto ao que o TAI-PB permitia, a saber:

- a) flexibilizar a aplicação do teste sempre que for necessário o diagnóstico da leitura;
- b) evitar a dispersão dos alunos no momento da aplicação do teste;
- c) proporcionar o uso do instrumento como diagnóstico rápido, facilitando a identificação das intervenções.

Como principais características, o TAI-PB foi um recurso com capacidade de automatizar a leitura das comandas dos itens, disponibilizar os itens de acordo com o desenvolvimento da respondente ao teste e, além disso, captar e organizar os dados adquiridos, diminuindo o tempo para a organização dos resultados. Ademais, por ser realizado em computador, com o uso de fones de ouvido e com os aparelhos eletrônicos, a heterogeneidade da proficiência das crianças foi facilmente gerenciada uma vez que o processo da resposta ao teste é autônomo e independe da leitura dos itens e indicações de realização por parte da professora.

O TAI-PB foi elaborado em parceria com a Secretaria Municipal de Educação de São Paulo (SME-SP), mas não houve continuidade depois de 2016 e uma nova versão do teste foi construída, dessa vez de forma autônoma pelas estudantes e docentes da Faculdade de Educação, mantida na nuvem da Universidade de São Paulo, o TAI-Leitura.

Deve ser ressaltado que o TAI-PB foi aplicado em 15 escolas da RME-SP no ano de 2016, onde participaram cerca de 80 professoras e 2.000 estudantes, conforme relatado em Alavarse et al. (2018). Iniciou-se como um TBC de 20 itens, elaborado via *web* a fim de que fosse de fácil acesso, a depender apenas de internet. Nele foram transportados os itens para os meios eletrônicos, sem adaptações. Como as comandas dos itens que precisavam ser lidas pela professora, conforme instrução da aplicação em papel, foi introduzida no TBC da PB a leitura automatizada de texto, sendo que a repetição se dava o tanto de vezes permitidas pela PB, que foi possível através do recurso *Application Programming Interface* (API). Ademais, no TBC da PB não foi permitido que as crianças voltassem os itens já respondidos conforme avançavam.

Para comparar a dificuldade dos itens da PB, originalmente em papel, com a dos mesmos itens em meio digital, uma parcela dos estudantes respondeu a PB em papel e outra, ambas selecionadas de modo aleatório, respondeu no TBC. A aplicação do TBC foi realizada em sala de aula, com as crianças em suas mesas, em *tablets* fornecidos pela própria escola, com fones de ouvido. Foram aplicadas 524 provas de TBC, cujo tempo médio foi de 25 minutos (ALAVARSE et al., 2018, p. 74), sendo a PB impressa realizada em média de 1 hora.

Em relação ao TBC que antecedeu a versão TAI-PB, Alavarse et al. (2018, p. 72) apontam que a sua elaboração foi necessária para que fossem atendidos os pressupostos psicométricos, objetivando

[...] observar se as características psicométricas como índice de dificuldade/facilidade, índice de discriminação e o percentual de acerto ou erro ao acaso dos itens, obtidos na aplicação em lápis e papel se mantinham quando esses itens eram disponibilizados no meio digital, um tablet.

O passo seguinte, assumindo-se que a dificuldade dos itens em meio digital muito próxima às dos itens em papel, foi a construção do TAI-PB. Para tanto, a equipe do Gepave integrou o algoritmo responsável pela aferição da proficiência ao TBC. O algoritmo apresenta 3 componentes, conforme Travitzki et al. (2020, p. 75), sendo: “1) método de estimação de proficiência; 2) critério para seleção de itens; e 3) critério para finalizar o teste”, visando dar dinamismo ao longo das respostas fornecidas pelas crianças. O primeiro item estaria alocado no meio da escala, fornecendo outro item ao longo da escala enquanto aloca a proficiência. Quando atinge o critério de parada, isto é, quando a proficiência está posicionada no intervalo que caracteriza determinado nível de proficiência a fim de se adquirir certeza (a saber intervalo de confiança), o teste é encerrado. A apresentação de itens é repetida várias vezes antes de atingir o critério de parada.

É importante salientar que sempre há o erro de medida, algo que deve ser levado em consideração na interpretação dos dados. Travitzki et al. (2020, p. 77) apontam que o TAI parte da lógica de oferecimento de item para se ter maior certeza da proficiência e menor erro de medida, pressupondo-se maior precisão. Nesse sentido, o estudo de Alavarse et al. (2018) destaca que “embora seja inerente a toda e qualquer medida, sua diminuição pode trazer consequências relevantes para o julgamento que está na base do processo de avaliação realizado pelo professor”.

Como mencionado, com o término do projeto do TAI-PB com a SME-SP, foi necessária a elaboração de um novo *software* do TAI-Leitura pelo Gepave. Conforme supracitada, a segunda fase do TAI-Leitura foi hospedada na nuvem da Universidade de São Paulo e apresenta características similares às do recurso anterior. Especialmente por sua nova elaboração, o presente estudo visa problematizar sua aplicação, a fim de compreender sua funcionalidade pelas crianças e professoras. De forma que, a seguir, apresentamos a contextualização da pesquisa.

5. O TESTE EM TESTE: APLICAÇÃO DO TAI-LEITURA EM DUAS ESCOLAS

Como objetivo central deste trabalho, procuramos investigar a utilização do Teste Adaptativo Informatizado elaborado pelo Gepave, baseado nos pressupostos da Provinha Brasil (BRASIL, 2016a; 2016b), no que tange à competência de leitura, denominado TAI-Leitura. A presente seção se propõe a apresentar o contexto em que se deu a ida à campo para a coleta de dados e análise do material elaborado. Em primeira instância, são caracterizadas as escolas participantes da pesquisa; em seguida, são apresentadas e analisadas as observações e dados coletados nas aplicações do TAI-Leitura; e, ao fim, as observações coletadas em reuniões com as professoras que lecionam nas turmas do Ensino Fundamental.

Como parte da preparação do trabalho de campo, nas duas escolas participantes as professoras dos anos iniciais do Ensino Fundamental receberam material impresso com o conteúdo da Provinha Brasil Teste 2-2016, tanto pelo fato do TAI-Leitura estar baseado neste instrumento quanto, sobretudo, para que elas tivessem acesso à fundamentação teórica dessa avaliação da proficiência em leitura e ao tratamento das respostas, pelo qual os resultados são apresentados em uma escala de cinco níveis com a respectiva interpretação pedagógica. Algumas apontaram conhecer a Provinha Brasil, enquanto outras não tiveram contato com o teste.

O funcionamento do TAI-Leitura enquanto um software proposto a realizar a medição e alocação de respondentes em uma escala de proficiência possui detalhamentos que fogem a proposta do presente trabalho. Neste sentido, conforme exposto pelo pesquisador do Gepave, Thiago Costa, o Anexo I propõe a discussão de seu funcionamento e o detalhamento desses aspectos.

Visamos questionar se as hipóteses preconizadas nos estudos anteriores seriam mantidas na nova programação do *software* – em especial sendo o TAI-Leitura a segunda versão do TAI-PB – em relação às otimizações de tempo em comparação à aplicação da PB, obtenção e validade dos dados, além de compreender aspectos que dizem respeito a fundamental parceria com as escolas que se dispuseram a aplicar o teste (Cf. ALAVARSE et al., 2018; CATALANI, 2019; TRAVITZKI et al., 2020).

5.1 CARACTERÍSTICAS DAS ESCOLAS, DAS PARTICIPANTES E AS CIRCUNSTÂNCIAS DA APLICAÇÃO

A seguir, são elencadas a caracterização das participantes e os resultados da aplicação do TAI-Leitura em duas escolas. Conforme já citado de forma introdutória, o estudo se deu em duas escolas, denominadas Ciclo e Mauritânia, ambas localizadas na cidade de São Paulo.

A aplicação do teste na Escola Ciclo foi realizada em duas etapas ao longo de quatro meses, com intervalo de dois meses entre as aplicações. Foram usados recursos tecnológicos e uma sala multiuso, pois a aplicação do TAI-Leitura foi feita por notebooks em vez de *tablets* e/ou iPads¹³. É uma escola privada localizada na zona sul de São Paulo, possui sistema apostilado próprio, e das 16 professoras que fazem parte da equipe docente, 5 lecionam Língua Portuguesa. A escola possui 79 estudantes matriculadas, conforme o Quadro 2.

Quadro 2 – Distribuição das matrículas por turma. Escola Ciclo. 2021

| Turma | Alunos |
|--------------|---------------|
| 1A | 12 |
| 1B | 12 |
| 2A | 9 |
| 3A | 11 |
| 3B | 11 |
| 4A | 12 |
| 5A | 12 |
| Total | 79 |

Fonte: Elaboração própria.

A respeito dos dispositivos tecnológicos: a escola dispõe de 7 iPads, 7 fones de ouvido e 30 notebooks com sistema operacional Windows 10. Foram utilizados apenas os fones de ouvido e os notebooks, haja vista que o instrumento avaliativo não funcionou no sistema operacional iOS do iPad. O teste foi realizado em uma sala denominada Multiuso, onde foi possível espalhar cerca de 12 computadores para cada turma. Desses, apenas 7 fizeram a prova e, conforme cada criança fosse terminando, o fone de ouvido era repassado para a outra. O modelo de aplicação se repetiu nas duas vezes, tanto em meados de outubro quanto em meados de dezembro.

¹³ Verificar Conclusão para maiores informações sobre apontamentos técnicos acerca da não utilização dos iPads e *tablets*.

A segunda escola, denominada Escola Maurîtânia, é pública e está localizada na zona leste paulistana. Possui 11 professoras atuantes nos anos iniciais do Ensino Fundamental. Conta com 360 estudantes matriculadas nesses anos iniciais, divididas em 11 turmas, conforme consta no Quadro 3.

Na Escola Maurîtânia, o TAI-Leitura foi aplicado por meio de 30 notebooks, dos quais todos possuem o sistema operacional Windows 10, com acesso à internet cabeada fornecida e mantida pela própria escola. Por se tratar de aparelhos públicos, os notebooks possuem restrição de acesso a alguns sites, sendo necessária a solicitação de liberação do endereço eletrônico para a instância regional da rede à qual pertence a escola. A aplicação do TAI-Leitura na Escola Maurîtânia ocorreu no Laboratório de Informática, localizado no segundo andar do prédio. Nesta sala foi possível encaminhar as crianças de forma confortável, uma vez que existiam computadores suficientes para acomodar o número de crianças presentes nos dias dos testes.

Quadro 3 – Distribuição das matrículas por turma. Escola Maurîtânia. 2021

| Turma/Ano | Matrículas |
|------------------|-------------------|
| 1A | 34 |
| 1B | 30 |
| 1C | 36 |
| 1º Ano | 100 |
| 2A | 34 |
| 2B | 30 |
| 2º Ano | 64 |
| 3A | 30 |
| 3B | 27 |
| 3º Ano | 57 |
| 4A | 32 |
| 4B | 36 |
| 4º Ano | 68 |
| 5A | 34 |
| 5B | 37 |
| 5º Ano | 71 |
| Total | 360 |

Fonte: Elaboração própria.

5.2 OS RESULTADOS DAS APLICAÇÕES

Na Escola Ciclo foram realizadas duas aplicações, com um intervalo de dois meses entre as duas, enquanto na Escola Mauritânia foi possível realizar a aplicação apenas uma vez – em especial por estar com a maioria de seus alunos em modo híbrido/remoto devido ao quadro pandêmico e de distanciamento social em decorrência do vírus Sars-Cov-2, causador da Covid-19.

Na Escola Ciclo, os testes foram aplicados após uma breve conversa sobre o que fariam. Foi introduzida, brevemente, a ideia do teste, apontando que aqueles que não fizessem o teste imediatamente por falta de fones de ouvido o suficiente para abranger a todas as crianças ao mesmo tempo, realizariam outra atividade nos computadores, e, conforme os fones de ouvido fossem disponibilizados, iniciariam o teste. O procedimento foi repetido em todas as 7 turmas da Escola Ciclo nas duas aplicações feitas.

5.2.1. Escola Ciclo

A primeira aplicação se deu em uma semana com todas as 7 turmas, nas quais não houve ausências. Antes de ser iniciada a etapa de interação entre as crianças e o TAI-Leitura, e de levá-las à sala multiuso, foi-lhes explicado o que ocorreria: fariam uma atividade em uma sala à parte, com notebooks e o uso da internet. As crianças se animaram ante a perspectiva de fazer uma atividade fora da sala de aula e com o uso de aparatos tecnológicos diferentes do iPad, que já costumam usar em outras aulas.

Ao longo da primeira semana de aplicação, as crianças fizeram apontamentos acerca da atividade: perceberam a extensão das provas entre elas e a diferença da quantidade de questões, bem como os níveis de proficiência; demonstraram preocupação com a nota daquela prova; apontaram o interesse em fazer uma atividade legal como aquela, porque envolvia o uso de computadores – uma das crianças ficou incomodada com a duração do teste, pois para ela “a prova durou tempo demais”; para outra criança a atividade foi chata e, quando questionada sobre o motivo, afirmou que “provas são chatas”.

Durante as atividades, ocorreram problemas com alguns notebooks antes dos testes serem iniciados, em especial com o *hardware* de som, de forma que o fone de ouvido não funcionava bem. Nesses casos, as crianças foram realocadas para notebooks que demonstraram bom funcionamento e iniciaram seus testes. Outra observação foi a de que as crianças menores demoraram mais tempo para fazer a atividade por terem dificuldades em manusear os recursos

do teclado (as setas para movimentar a tela para cima e para baixo a fim de ver as alternativas dos itens), e não sabiam como utilizar o site e o painel tátil dos notebooks.

Igualmente, constatou-se ter um problema de item, uma vez que a imagem ou estava muito pequena, ou o nome de um objeto podia ser confundido com um substantivo sinônimo (caneca com xícara, por exemplo). Outra criança, da turma do 2ºA, precisou de auxílio para entender um dos itens usados como exemplo, interpretando-o como o desenho de escada e não a figura de uma estante. Assim sendo, foi necessária uma intervenção para compreender qual o problema ocorrido, constatando-se, ao fim, que era um problema similar ao de outras turmas.

É importante frisar que a primeira aplicação ocorreu com duas crianças ainda em ensino remoto, não sendo possível alocá-las em um nível na escala de proficiência na primeira aplicação do TAI-Leitura por não terem acesso aos dispositivos e a realização da aplicação do teste. Realizaram o TAI-Leitura na segunda aplicação, com as suas respectivas turmas devido a obrigatoriedade da volta a frequência escolar presencial. Em igual importância, cita-se o acompanhamento ao longo da atividade por parte de uma das professoras de Língua Portuguesa, que trabalha com a turma do 2º ano. Sua observação se restringiu à um dos alunos vendo o comportamento do *software*, sem optar por responder ao teste devido ao tempo hábil para fazê-lo.

Após um intervalo de quase dois meses, foi realizada uma segunda aplicação do TAI-Leitura para as turmas dos anos iniciais do Ensino Fundamental da Escola Ciclo. A abordagem da turma ocorreu do mesmo modo que da primeira vez, onde foi explicada, de forma resumida, a atividade a ser realizada em uma sala a parte nos computadores com acesso à internet. Nesta semana, no entanto, houve a adição de duas crianças que antes estavam em ensino remoto e que passaram ao ensino presencial – uma da turma do 2º ano A e uma da turma do 3º ano A. Das duas crianças que nunca fizeram, uma teve problemas para acessar a prova devido ao acesso bloqueado, sendo necessário retirá-la de uma aula posterior, ainda na mesma semana, para que fizesse a atividade.

As crianças notaram a similaridade com a atividade anterior e muitas comentaram que gostaram de realizá-la novamente por ser menos chata que uma prova em sala de aula. Também apontaram que a atividade era, na verdade, bem fácil e legal. Nesta aplicação, diferentemente das outras, em uma das turmas a professora solicitou realizar o teste, percebendo a dinâmica e a funcionalidade do instrumento avaliativo. A docente, que leciona Língua Portuguesa nas turmas dos 3º anos A e B, acompanhou a atividade com dois alunos de ambas as salas,

observando as respostas dadas e vendo como o *software* se comporta de acordo com as alternativas selecionadas.

Foram apontadas dificuldades similares com as questões previamente citadas. No entanto, as crianças menores demonstraram menos dificuldades no manuseio das máquinas. Similarmente, os notebooks e fones de ouvido que apresentaram problemas na primeira aplicação não foram utilizados, a fim de que não trouxessem ônus às crianças durante a prova. Aponta-se também a ausência de alguns alunos dos anos avançados do ensino fundamental, em especial dos 4º e 5º anos devido à finalização do ano letivo estar bem próxima à segunda semana de aplicação do TAI-Leitura, em meados de dezembro.

Ao total, foram aplicados 74 testes para as crianças, sendo consideradas as respostas das crianças que conseguiram realizar o TAI-Leitura nas duas aplicações. Conforme pudemos observar, os resultados apontam diferenças de proficiência de acordo com o nível de forma crescente, sendo do nível 1 para aquelas matriculadas no 1º ano do Ensino Fundamental e nível 5 para aquelas matriculadas no 5º ano do Ensino Fundamental. Tal dado é relevante pois demonstra que o instrumento é capaz de captar as diferenças de aprendizagens nos anos escolares dentre os sujeitos, diferenças essas que são apontadas tanto pela literatura quanto pelas professoras.

Em relação aos resultados das aplicações, especificamente o que tange à alocação das respondentes em um nível de proficiência, é importante frisar que o *software* está fundamentado na escala elaborada pela PB, de 0 a 1000. Para efeito do resultado, a proficiência de cada estudante é alocada numa escala de 5 níveis, numerados de 1 a 5, com a interpretação pedagógica para cada um deles. No Quadro 4 constam os valores dos intervalos de cada nível.

Quadro 4 – TAI-Leitura. Níveis de proficiência

| TAI-Leitura. Níveis de proficiência | | | | |
|-------------------------------------|--------------------|-------------------|--------------------|-----------|
| Nível | Lim _{inf} | Θ_i | Lim _{sup} | Intervalo |
| Nível 1 | | $\Theta_i <$ | 325 | |
| Nível 2 | 325 | $\leq \Theta_i <$ | 425 | 100 |
| Nível 3 | 425 | $\leq \Theta_i <$ | 525 | 100 |
| Nível 4 | 525 | $\leq \Theta_i <$ | 595 | 70 |
| Nível 5 | 595 | $\leq \Theta_i$ | | |
| Θ_i : Theta do respondente | | | | |

Fonte: Elaboração própria.

Portanto, o Θ_i corresponde ao resultado obtido pela respondente ao final do teste, posteriormente alocado no nível dentro do qual se encontra em função do limite inferior (Lim_{inf}) e limite superior (Lim_{sup}) de cada nível. O nível 1 corresponde ao intervalo entre 0 e 325; o nível 2 ao intervalo entre 325 e 425; o nível 3 ao intervalo entre 425 e 525; o nível 4 ao intervalo entre 525 e 595; e o nível 5 ao intervalo entre 595 e 1000. De forma que, conforme a resposta aos itens do teste, a proficiência da respondente é medida e alocada dentro desses intervalos.

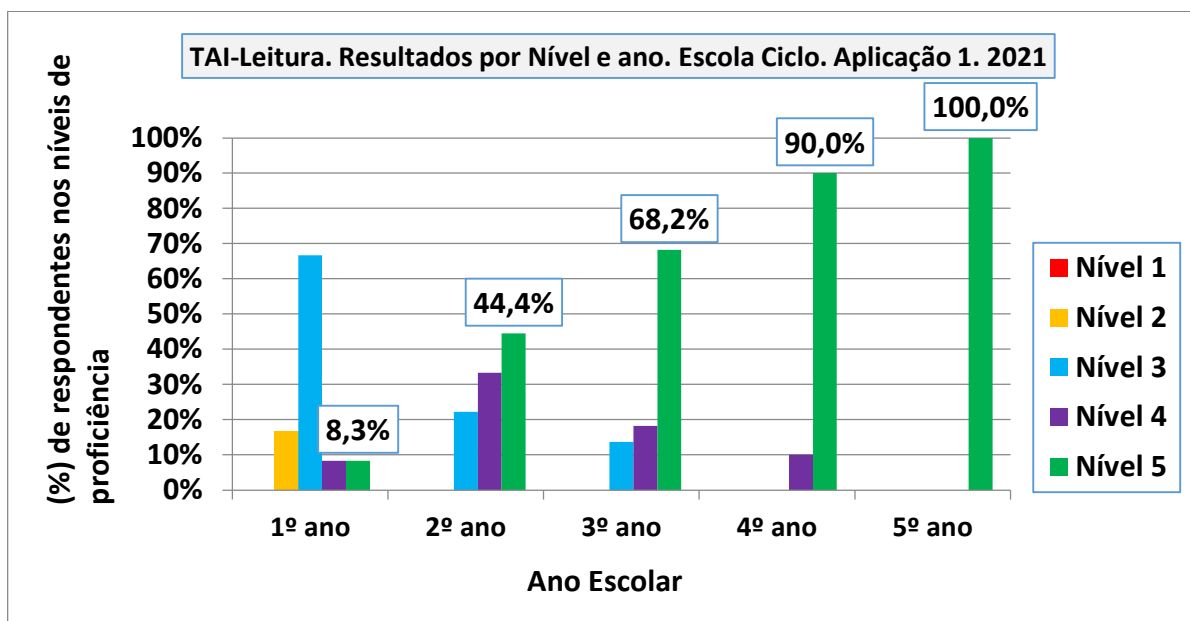
O Quadro 5 nos mostra a quantidade de estudantes alocadas nos níveis de proficiência, separadas por ano escolar. Como se pode observar, os resultados da primeira aplicação indicam que a quantidade de alunos da Escola Ciclo aumenta ao longo dos níveis de proficiência conforme os anos avançam. Nesse sentido, pressupõe-se um avanço na compreensão das competências da Língua Portuguesa por parte das crianças respondentes. Os dados vão ao encontro dos pressupostos pedagógicos expostos na escala de proficiência e interpretação pedagógica da PB. Assim, esses resultados apontam que o TAI-Leitura funcionou como esperado.

Tabela 1 – Resultado por Nível e Ano. Escola Ciclo. Aplicação 1. 2021

| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|-----------|-----------|----------|-----------|-----------|----------|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 4 | 4 | 0 | 0 | 0 | 0 |
| 3 | 21 | 16 | 2 | 3 | 0 | 0 |
| 4 | 10 | 2 | 3 | 4 | 1 | 0 |
| 5 | 39 | 2 | 4 | 15 | 9 | 9 |
| Total | 74 | 24 | 9 | 22 | 10 | 9 |

Fonte: Elaboração própria.

Em igual medida, o Gráfico 1 traz a leitura da Tabela 1 em relação às porcentagens dos grupos em seus respectivos níveis. O que nos interessa nesse gráfico é a crescente porcentagem de crianças no nível 5 conforme a evolução no decorrer dos anos, tal como pode ser observado ao longo da passagem do 1º ao 5º ano, saindo de 8,3% e chegando a 100% respectivamente. Assim, o TAI-Leitura demonstra a porcentagem das proficiências de um mesmo grupo, indicando às professoras os objetos compreendidos por determinada turma de acordo com a alocação das respondentes em determinado nível, facilitando o trabalho pedagógico quando a escala de proficiência e os descritores do nível de proficiência são utilizados para a (re)elaboração de atividades voltadas para a competência em leitura de determinado grupo.

Gráfico 1 – TAI-Leitura. Resultados por Nível e ano. Escola Ciclo. Aplicação 1. 2021

Fonte: Elaboração própria.

De modo que, para fins de exemplificação, é possível compreender que na primeira aplicação do TAI-Leitura, as crianças da turma do 4º ano submetidas ao teste estão majoritariamente alocadas no nível 5 de proficiência, sendo 90% do total de crianças. Pautando-nos nas indicações pedagógicas da PB (BRASIL, 2016a), significaria compreender que a média das crianças é capaz de:

- Reconhecer o assunto de um texto longo com base no título, a partir de leitura individual;
- Reconhecer o assunto de textos médios por meio de inferências com forte base no conteúdo informacional, a partir de leitura individual;
- Identificar informação explícita não trivial, por vezes secundária, em um texto curto ou médio, com base em leitura individual;
- Inferir informação não trivial em textos médios com base em leitura individual ou com o apoio de leitura pelo aplicador;
- Reconhecer a finalidade de um texto de construção complexa lido silenciosamente com o apoio de suporte.

Entretanto, o erro de medida deve ser considerado para a compreensão da aprendizagem das crianças. A esse respeito, o que o presente estudo considerou como erro de medida é a variação do que é estimado pelo recurso dentro de um intervalo, estabelecendo um Valor máximo (V_{\max}) e um Valor mínimo (V_{\min}), nos extremos do intervalo no centro do qual está a proficiência estimada ao final da aplicação do TAI-Leitura, denominado intervalo de confiança. Sendo este inerente ao processo de medição, o TAI-Leitura fornece uma estimativa de alocação considerando o erro de medida em relação à proficiência dos sujeitos submetidos ao teste. Nesse

sentido, o V_{\min} da proficiência é considerado como um dado relevante, em especial por indicar a proficiência mínima dos respondentes na escala¹⁴. A Tabela 2 apresenta os níveis das crianças de acordo com a proficiência tendo em conta o erro para menos.

Tabela 2 – TAI-Leitura. Resultado por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 1. 2021

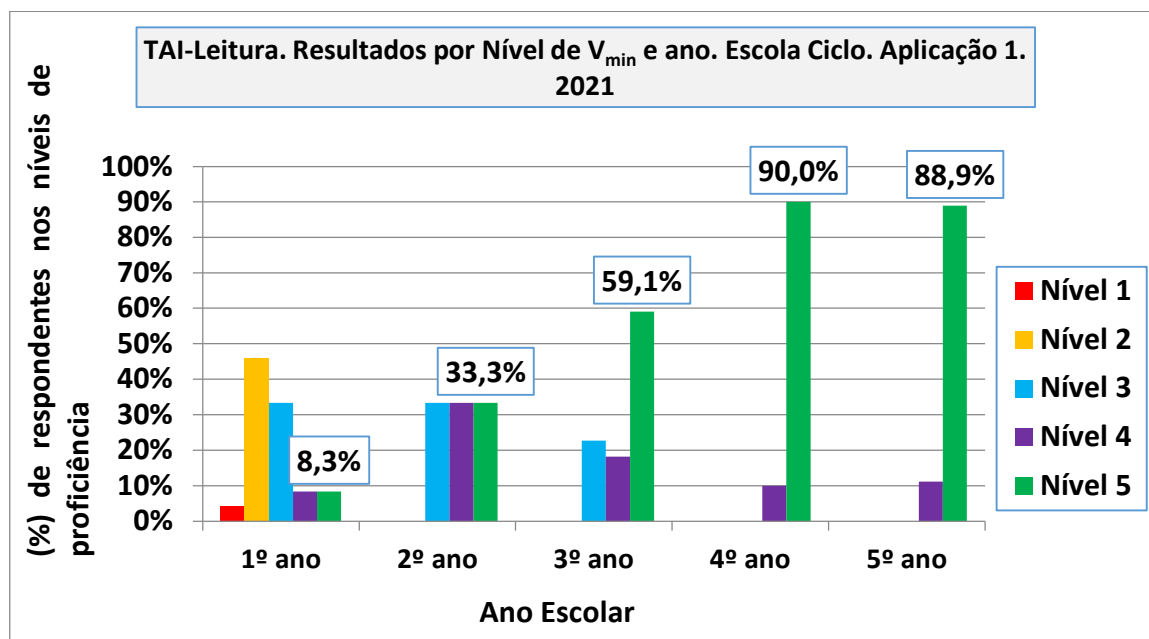
| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|-----------|-----------|----------|-----------|-----------|----------|
| 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 11 | 11 | 0 | 0 | 0 | 0 |
| 3 | 16 | 8 | 3 | 5 | 0 | 0 |
| 4 | 11 | 2 | 3 | 4 | 1 | 1 |
| 5 | 35 | 2 | 3 | 13 | 9 | 8 |
| Total | 74 | 24 | 9 | 22 | 10 | 9 |

Fonte: Elaboração própria.

Portanto, podemos observar que o V_{\min} fornece a perspectiva de que existem alunos em um nível menor que o nível médio de proficiência da turma. Por exemplo, no 1º ano os dados indicavam que quatro alunos estavam alocados no nível 2, mas considerando o V_{\min} , aumentam as crianças que estão no nível 2 e uma criança apresenta o nível 1 de proficiência. . Nesse sentido, pode orientar a docente em decisões pedagógicas. De forma a fazer seu planejamento de aulas e atividades levando em conta as competências que seguramente as respondentes possuem, trabalhando pedagogicamente as habilidades e competências que precisam ser reforçadas (no caso, aquelas de nível 2 de acordo com a escala de nível de proficiência e sua respectiva interpretação pedagógica).

De igual forma, o Gráfico 2, referente a Tabela 2, apresenta a alocação das crianças que realizaram o teste, considerando a margem de erro do V_{\min} . O que deve ser observado é que a linha crescente do nível de proficiência se mantém, indicando coerentemente a evolução no aprendizado em leitura ao longo dos anos. Excetua-se a pequena queda no 5º ano de Ensino Fundamental que apresenta uma criança cujo nível pode estar alocado no nível 4, muito próximo ao nível 5 – pela quantidade de alunos matriculados no respectivo ano, uma criança representa 11,1% da turma.

¹⁴ A este respeito, é importante frisar que foi utilizado um intervalo de confiança de 95%, de modo que a percentagem de incerteza é de 5% dos casos.

Gráfico 2 – TAI-Leitura. Resultado por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 1. 2021

Fonte: Elaboração própria.

Nesse sentido, o respondente ao teste pode estar mais próximo do início do nível 4, ainda dentro do nível 5 na escala de proficiência, permitindo que o planejamento pedagógico leve em conta os pressupostos do que a respondente em questão conseguiu compreender em sua aprendizagem e quais competências precisam ser trabalhadas pela docente.

Levando em consideração o V_{\min} , portanto, pode-se obter um indicativo de trabalho pedagógico mais seguro, uma vez que leva em consideração um possível resultado em um momento estatístico específico do que a respondente pode apresentar, evitando um trabalho pedagógico acima do que essa respondente poderia acessar. Tal gráfico é importante pois é capaz de indicar aos professores quais conceitos devem ser trabalhados para que os níveis de proficiência das crianças respondentes evoluam.

De forma que, se selecionarmos o 3º ano para fins de exemplificação, podemos compreender que mais da metade das crianças, isto é, 13 crianças, possuem níveis de proficiência esperados para turmas de 3º ano em diante, isto é, o nível 5. Além disso, 22,7% do total de crianças da turma em questão demonstra conhecimentos compatíveis com o nível 3, enquanto outra porcentagem (18,2%) apresentam nível 4 de proficiência. Reiteramos aqui que o nível esperado pela Provinha Brasil, a partir do 2º ano dos anos iniciais do ensino fundamental, é o nível 5 de proficiência.

Em relação a segunda aplicação, realizada num intervalo de tempo de quase dois meses, buscou-se compreender se o teste seria capaz de captar algum tipo de mudança no nível de proficiência das respondentes.

Tabela 3 – Resultado por Nível e Ano. Escola Ciclo. Aplicação 2. 2021

| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|-----------|-----------|----------|-----------|-----------|----------|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 | 0 | 0 |
| 3 | 17 | 16 | 1 | 0 | 0 | 0 |
| 4 | 11 | 5 | 2 | 4 | 0 | 0 |
| 5 | 45 | 2 | 6 | 18 | 10 | 9 |
| Total | 74 | 24 | 9 | 22 | 10 | 9 |

Fonte: Elaboração própria.

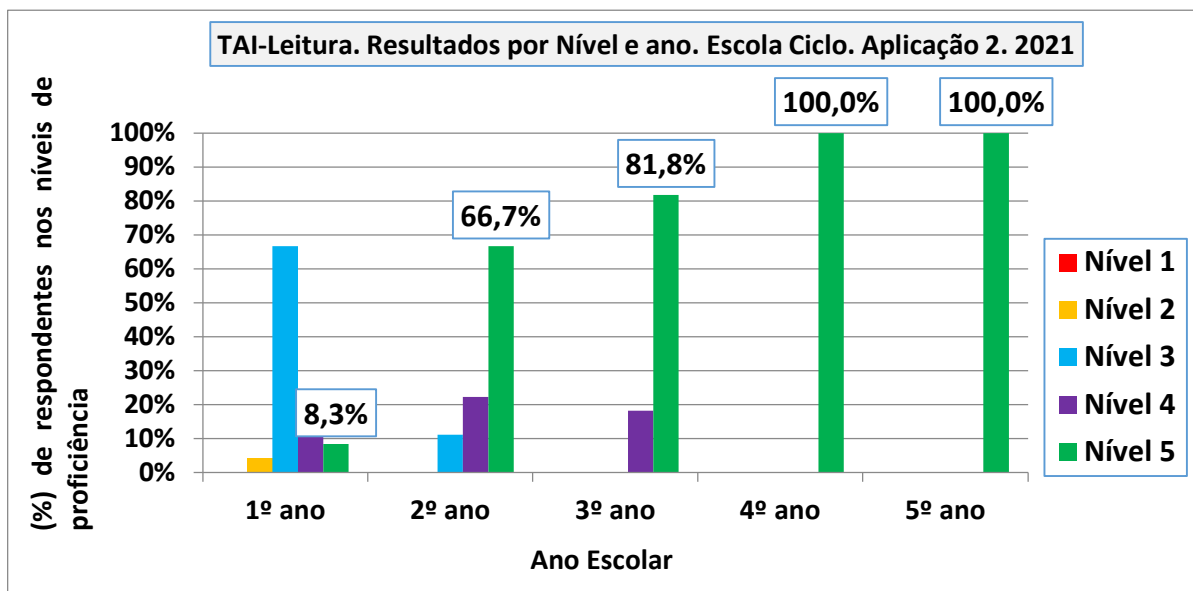
Levando-se em conta o total de respondentes alocadas nos cinco níveis de proficiência da PB, é possível notar uma evolução na proficiência dos anos escolares da Escola Ciclo, o que aponta para uma funcionalidade que vai ao encontro do que se pressupõe ao longo do trabalho docente, isto é, o avanço na proficiência conforme a passagem do tempo e da aprendizagem. Assim, nota-se que no 4º ano, onde antes havia uma criança alocada no nível 4 de proficiência, na segunda aplicação já não se acusa, por meio do Θ_i , a alocação de nenhuma das respondentes em tal nível, passando de um total de 9 crianças no nível 5, na primeira aplicação conforme a Tabela 2, para 10 respondentes alocadas no nível 5, de acordo com os dados demonstrados pela Tabela 3.

De forma semelhante, o Gráfico 3 demonstra o crescimento dos níveis de proficiência em dois meses, guardadas as proporções, em detrimento de um menor percentual de crianças com o Θ_i alocado nos menores níveis de proficiência. Pode-se notar o crescimento do nível de proficiência das respondentes, em especial dos primeiros anos do Ensino Fundamental – Anos Iniciais, como o aumento de mais de 20% no nível 5 na turma do 2º ano em relação ao total de alunos.

Frisa-se aqui que este crescimento está atrelado às características da PB, cujo medição restringe-se às competências elencadas para o nível 5. Isto é, não havendo outros níveis e demais competências de leitura, o teto da medição alcança o nível 5, como é possível observar com a turma do 5º ano, cuja alteração não é tão expressiva se em contraposição às turmas dos 1º, 2º e

3º anos. Se houvesse maiores níveis na escala de proficiência, seria provável assumir que o crescimento da alocação da estimativa da proficiência das crianças respondentes seria expressivo.

Gráfico 3 – TAI-Leitura. Resultados por Nível e ano. Escola Ciclo. Aplicação 2. 2021



Fonte: Elaboração própria.

A Tabela 4 apresenta o número de alunos alocados, levando-se em conta o V_{\min} . É possível observar que a quantidade de crianças diminui conforme há uma evolução ao longo dos anos escolares. Entretanto, como o V_{\min} leva em conta a média de proficiência para menos, considerado um erro de medida (inerente a qualquer medição), é importante frisar a alocação de crianças em níveis inferiores, o que pode indicar que estas estão mais próximas do nível de proficiência abaixo daquele que a proficiência média (o Theta) apresenta – como é o caso da quantidade de alunos do nível 2 do 1º ano, que passa de uma para cinco.

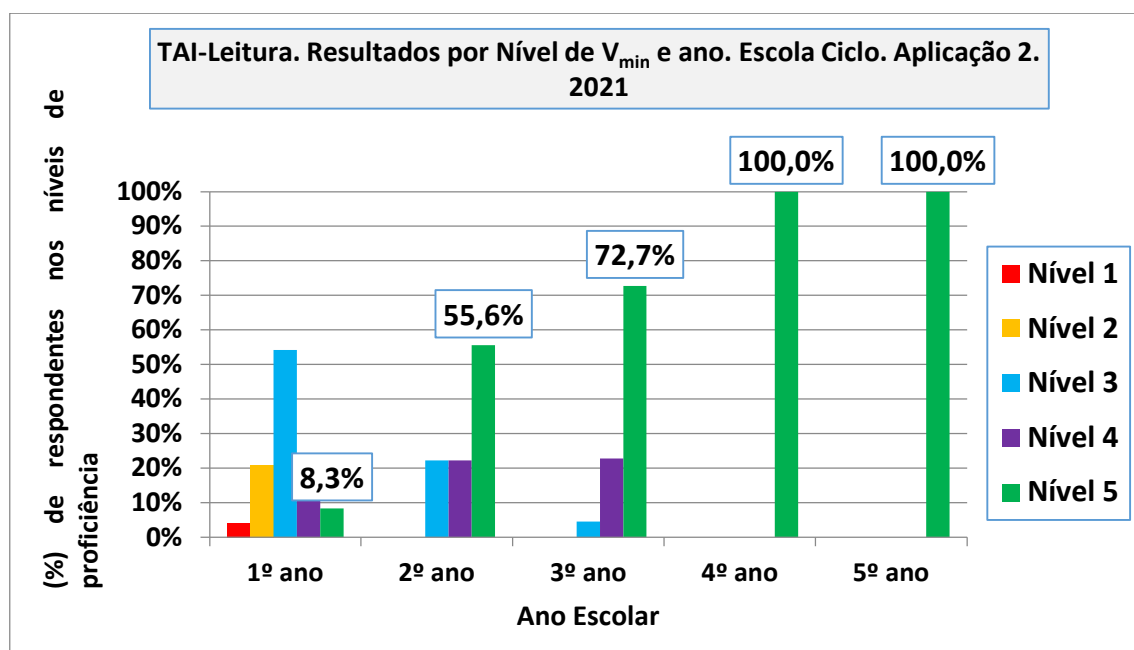
Tabela 4 – Tratamento pelo V_{\min} do Theta. Aplicação 2. 2021

| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|-----------|-----------|----------|-----------|-----------|----------|
| 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 5 | 5 | 0 | 0 | 0 | 0 |
| 3 | 16 | 13 | 2 | 1 | 0 | 0 |
| 4 | 10 | 3 | 2 | 5 | 0 | 0 |
| 5 | 42 | 2 | 5 | 16 | 10 | 9 |
| Total | 74 | 24 | 9 | 22 | 10 | 9 |

Fonte: Elaboração própria.

Ao colocarmos em primeiro plano o V_{\min} , o avanço na escala de proficiência na segunda aplicação é igualmente possível perceber que o mesmo fenômeno se faz presente: há uma evolução no nível de proficiência das turmas: no 4º ano já não existem alunos cujo V_{\min} esteja alocado próximo à fronteira de dois níveis (nível 4 e 5), ao passo em que no 3º ano o nível 3 demonstra uma queda, passando de 22,2% em relação às respondentes do 2º ano para 4,5% das respondentes do 3º ano em detrimento de um aumento nos níveis 4 e 5 de proficiência.

Gráfico 4 – TAI-Leitura. Resultados por Nível de V_{\min} e ano. Escola Ciclo. Aplicação 2. 2021



Fonte: Elaboração própria.

Tais dados colaboram para a compreensão de que o instrumento consegue captar a evolução de proficiência em leitura de crianças em um recorte de tempo maior do que dois meses.

Outro dado relevante é o tempo de aplicação do TAI-Leitura. Os dados apontam uma média de 12 minutos de duração na primeira aplicação e 11 minutos na segunda aplicação para as crianças dos anos iniciais do Ensino Fundamental. Conforme a Tabela 5, onde tempo de aplicação é discriminado por turma com sua variação, podemos constatar que o tempo médio de aplicação apresenta uma queda: conforme os anos escolares avançam em direção aos anos finais do Ensino Fundamental, menor é o tempo gasto com a aplicação do TAI-Leitura. Ainda: a segunda

aplicação demonstra certa otimização de tempo em algumas turmas, com uma diferença de cerca de 02min e 30s da primeira aplicação.

Tabela 5 – Theta médio, tempo médio e variação média das duas aplicações. Escola Ciclo.

2021

| | Aplicação 1 | | Aplicação 2 | | Variação | |
|--------|-------------|---------------|-------------|---------------|-------------|---------------|
| | Theta médio | Duração média | Theta médio | Duração média | Theta médio | Duração média |
| 1º ano | 473,8 | 00:21:33 | 500,4 | 00:20:40 | 26,6 | 00:00:53 |
| 2º ano | 583,4 | 00:17:39 | 617,1 | 00:14:49 | 33,7 | 00:02:50 |
| 3º ano | 619,4 | 00:11:04 | 654,8 | 00:10:12 | 35,5 | 00:00:52 |
| 4º ano | 655,8 | 00:08:49 | 683,3 | 00:06:24 | 27,6 | 00:02:25 |
| 5º ano | 666,9 | 00:07:02 | 682,4 | 00:06:31 | 15,5 | 00:00:31 |

Fonte: Elaboração própria.

Além de apresentar a diminuição do tempo de aplicação do teste, os dados também apresentam o aumento da média de proficiência estimada, (o Theta) das respondentes, corroborando para a ideia de que o *software* é capaz de discriminar a evolução da proficiência das crianças respondentes. Também apontamos para a diminuição do tempo ao longo dos anos escolares, bem como o aumento do Theta. Contrariamente ao que se observa na Escola Maurîtânia, na Escola Ciclo não há uma curvatura: o tempo é menor no 1º ano, aumenta no 3º ano, para diminuir conforme avança até o 5º ano.

Considerando um bom funcionamento do TAI-Leitura, o estudo analisa, na subseção seguinte, os dados obtidos e observados na Escola Maurîtânia, em que apenas uma aplicação foi realizada em decorrência do distanciamento social.

5.2.2. Escola Maurîtânia

A aplicação do TAI-Leitura, na Escola Maurîtânia, se deu em duas semanas, para a qual houve preparação: a gestão realizou testes nos *tablets* disponibilizados a fim de que as atividades ocorressem corretamente quando levadas à sala. No entanto, em decorrência da aplicação da avaliação de sua rede, o acesso a endereços eletrônicos externos e não autorizados foi bloqueado nos dispositivos das escolas paulistanas. De forma que, ao realizar a primeira aplicação, foi

necessário deslocar as crianças para o laboratório de informática, turma por turma, utilizando como aparato eletrônico os *notebooks* previamente descritos neste trabalho.

Inicialmente, houve uma reunião de apresentação do TAI-Leitura às professoras em um dia do horário coletivo em um sábado¹⁵. Em seguida, em uma semana subsequente, a primeira turma do período da manhã foi encaminhada para que fossem feitos os primeiros testes. A turma do 1ºC pôde interagir com o TAI-Leitura após uma explicação fornecida pela professora a respeito das atividades que fariam fora da sala. Aos poucos os alunos foram dirigidos para o laboratório de informática, fazendo o seu *log-in* e iniciando a atividade.

Acerca da interação: as crianças acharam a atividade diferente e legal. Notou-se a heterogeneidade em relação aos conhecimentos dos aparatos tecnológicos, em especial porque algumas tiveram dificuldades para manusear *mouse* e teclado, enquanto outras conseguiram ultrapassar as questões sem delongas. Chamou nossa atenção a empolgação das alunas em relação a ouvir uma voz do computador, especificamente uma das crianças que buscou interagir oralmente com o *API*, pedindo que a moça¹⁶ repetisse o enunciado.

De forma similar à Escola Ciclo, as crianças notaram as diferenças das quantidades de itens que apareciam, em especial aquelas que sentaram próximas. Em determinado caso, foi necessária a intervenção para que a aluna finalizasse a atividade, pois encontrava dificuldades em compreender o que era pedido. Pareceu-nos que, assim que se sentiu apoiada com a presença de uma professora ao seu lado, conseguiu prosseguir com as questões de modo independente. Tanto na Escola Ciclo quanto na Escola Mauritânia a recepção à atividade foi satisfatória, com as crianças aprovando a realização de uma atividade no computador e com entusiasmo, pois em nenhum dos casos houve reclamações sobre a realização de uma prova ou teste.

Apontamos problemas similares à Escola Ciclo no que tange a determinados itens: um item que pedia que a criança assinalasse a alternativa em que todas as palavras das figuras mostradas fossem iniciadas com a sílaba CA, cuja figura de uma caneca era lida como xícara; um item em que a imagem de um armário, no qual a criança deveria escolher a palavra que nomeia o objeto, era entendido como uma escada e um item em que o informativo da doença da Dengue aparece de forma muito diminuta, a criança não consegue ler as informações mostradas de forma efetiva.

Outro dado relevante a respeito da aplicação da Escola Mauritânia foi a ampliação do tempo de aplicação, embora ainda de um intervalo em que a proficiência não aponta muitas alterações,

¹⁵ Ver subseção 5.3

¹⁶ A voz do *API* é um recurso baseado na voz da plataforma de tradução da Google.

devido à ausência de muitas crianças. A aplicação foi continuada ao longo dos outros dias, pela direção da escola, conforme as crianças fossem regressando para o ensino presencial obrigatório¹⁷.

Em relação à participação geral, pode-se dizer que das 360 crianças, 200 realizaram o TAI-Leitura em apenas uma aplicação. As informações obtidas demonstram um crescimento de proficiência ao longo dos anos escolares, em uma repetência daquilo que foi observado na Escola Ciclo, guardadas as proporções¹⁸. Conforme a Tabela 6, nota-se a crescente presença de crianças do 1º ano do Ensino Fundamental no nível 1 de proficiência, algo esperado, e a diminuição gradual em detrimento do aumento de crianças presentes no nível 5, *pari passu* o avanço dos anos escolares dos anos iniciais do Ensino Fundamental.

Tabela 6 – Resultado por Nível e ano. Escola Maurîtânia. Aplicação 1. 2021

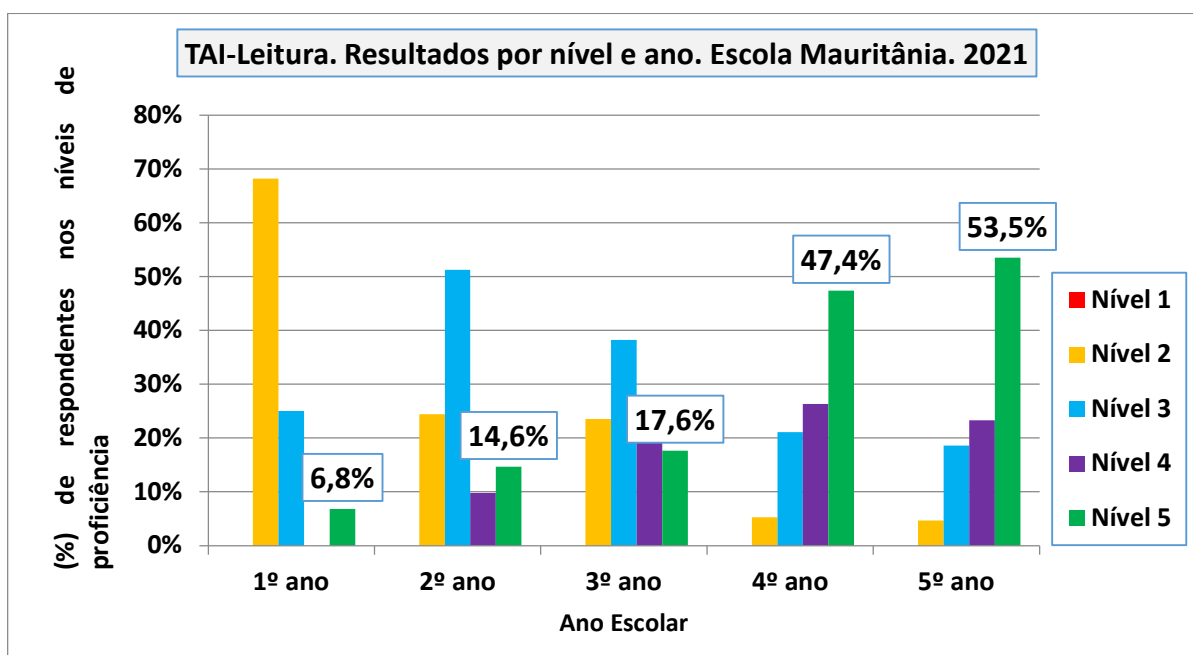
| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|------------|-----------|-----------|-----------|-----------|-----------|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 52 | 30 | 10 | 8 | 2 | 2 |
| 3 | 61 | 11 | 21 | 13 | 8 | 8 |
| 4 | 31 | 0 | 4 | 7 | 10 | 10 |
| 5 | 56 | 3 | 6 | 6 | 18 | 23 |
| Total | 200 | 44 | 41 | 34 | 38 | 43 |

Fonte: Elaboração própria.

O Gráfico 5 mostra a proporção das respondentes ao longo dos níveis de proficiência, levando em consideração o nível médio obtido na estimativa da proficiência em leitura. Percebe-se que, tal qual na aplicação da Escola Ciclo, na Escola Maurîtânia foi possível apreender diferentes níveis de proficiência ao longo das turmas de 1º a 5º ano, reforçando a capacidade de captura da aprendizagem. De forma semelhante, embora com suas particularidades, o gráfico mostra o crescimento do nível 5 desde o 1º ano até o 5º ano do Ensino Fundamental, em detrimento da diminuição do nível 2, o menor nível de proficiência captado pelo Θ_i apontado.

¹⁷ Isto porque, dentro do contexto em que se passou a pesquisa, parte da turma ainda poderia se manter em ensino remoto, sendo a obrigatoriedade retomada poucos dias depois do início das aplicações.

¹⁸ Desconsideram-se as diferenças em relação ao total de alunos por turma, a desigualdade socioeconômica e o capital cultural, cuja problematização compete a outro tipo de pesquisa, indicada ao final deste trabalho.

Gráfico 5 – Resultado por Nível e ano. Escola Maurîtânia. Aplicação 1. 2021

Fonte: Elaboração própria.

Levando-se em consideração as turmas do 4º ano, por exemplo, em comparação com as turmas do 5º ano, é possível perceber o crescimento do nível 5. Do outro lado do gráfico, no entanto, vemos o 1º ano com uma grande porcentagem das respondentes alocadas no nível 2 em contraposição ao nível 5, cujo percentual é de quase 7%. Assim, portanto, o TAI-Leitura consegue apreender situações de aprendizagem preconizadas pelas docentes e pela literatura.

Quando levado em conta o V_{\min} , conforme demonstrado na Tabela 7, percebemos a diferença na alocação de crianças ao longo dos níveis, em especial nos níveis 5 do 2º ano do Ensino Fundamental, se em comparação com o resultado da mesma turma, por nível, sem considerar o V_{\min} .

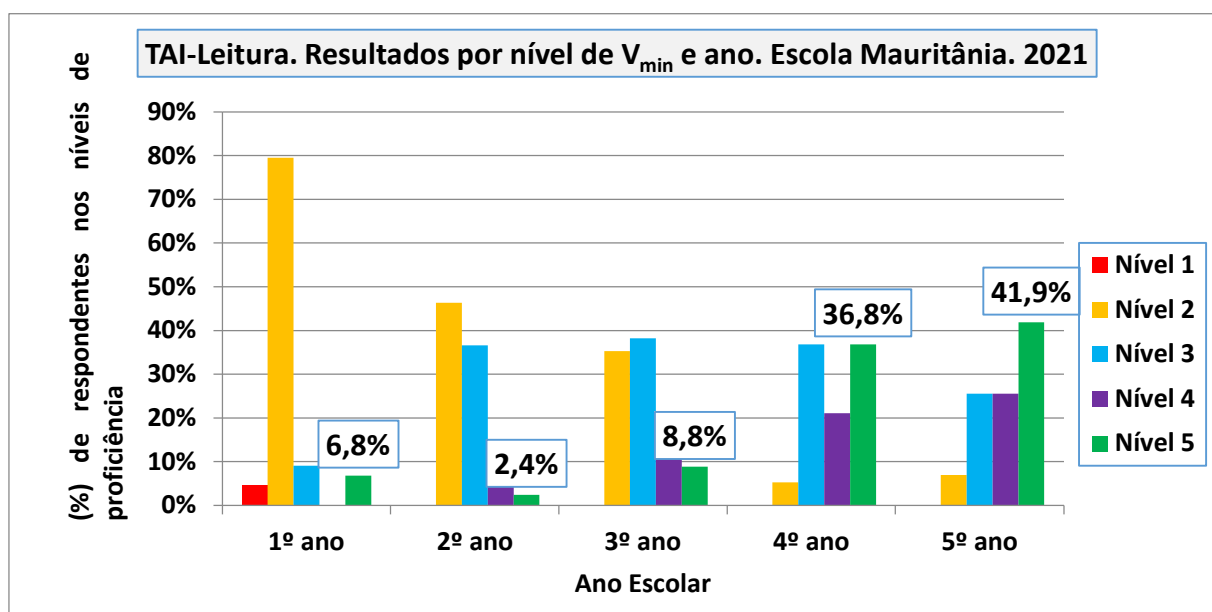
Tabela 7 – Resultado por Nível de V_{\min} e ano. Escola Maurîtânia. Aplicação 1. 2021

| Nível | Total | 1º ano | 2º ano | 3º ano | 4º ano | 5º ano |
|--------------|------------|-----------|-----------|-----------|-----------|-----------|
| 1 | 2 | 2 | 0 | 0 | 0 | 0 |
| 2 | 71 | 35 | 19 | 12 | 2 | 3 |
| 3 | 57 | 4 | 15 | 13 | 14 | 11 |
| 4 | 31 | 0 | 6 | 6 | 8 | 11 |
| 5 | 39 | 3 | 1 | 3 | 14 | 18 |
| Total | 200 | 44 | 41 | 34 | 38 | 43 |

Fonte: Elaboração própria.

De forma semelhante o Gráfico 6 demonstra essa curva, onde o nível 1 de proficiência diminui ao passo em que o nível 2 cresce. Contrariamente, o nível 3, por exemplo, aparece em diminuta porcentagem, estabilizando-se num intervalo entre 36% e 38% entre os 2º, 3º e 4º anos, para enfim cair no 5º ano. O nível 4 de proficiência aparece a partir, apenas, do 2º ano do Ensino Fundamental, crescendo até o 5º ano.

Gráfico 6 – Resultado por Nível de V_{\min} e ano. Escola Maurîtânia. Aplicação 1. 2021



Fonte: Elaboração própria.

Tais resultados apontam que o TAI-Leitura demonstra a diferença na aprendizagem das crianças, mas que, além disso, consegue alocar a estimativa de proficiência em leitura de acordo com o intervalo dos níveis de proficiência elaborados a partir da escala da PB. Com isso, ao discriminar a aproximação a níveis inferiores e considerado o erro de medida para menos (V_{\min}), o que foi percebido pelo instrumento na Escola Ciclo, em ambas as aplicações, igualmente ocorre com a Escola Maurîtânia: o crescimento nos níveis maiores é perceptível conforme os anos escolares avançam, opostamente diminuindo ao passo em que os níveis inferiores aumentam nos anos primários dos anos iniciais do Ensino Fundamental. Dessa forma, notamos a coerência mantida pelo TAI-Leitura em ambas as escolas em que foi aplicado. Em relação ao tempo de aplicação do TAI-Leitura, os dados do Quadro 12 indicam uma média de 19 minutos de aplicação dentre as turmas, com a discriminação da duração média por turma.

Tabela 8 – Theta médio, tempo médio e variação média dos anos escolares. Escola Maurítânia. 2021

| | Theta (Θ_i) médio | V _{min} médio | Tempo médio no teste |
|--------|-------------------------------|---------------------------|-------------------------|
| 1º ano | 398,7 | 361,7 | 00:17:49 |
| 2º ano | 471,9 | 436,5 | 00:21:17 |
| 3º ano | 499,5 | 462,8 | 00:23:27 |
| 4º ano | 585,7 | 543,4 | 00:21:00 |
| 5º ano | 589,8 | 546,2 | 00:14:13 |

Fonte: Elaboração própria.

Os dados da Tabela 8 revelam, portanto, um tempo máximo de 23 minutos de aplicação e um mínimo de 14 minutos. É importante notar que o tempo sofre uma curvatura, uma vez que é menor no 1º ano, tendo seu ápice no 3º ano e diminuindo consideravelmente no 5º ano. Isto porque o 1º e 5º anos são turmas extremas, cuja pressuposição é de níveis menores e maiores de suas respondentes, respectivamente, devido ao tempo de escolarização e o avanço da aprendizagem preconizado ao longo desta.

A Tabela 9 traz com maior detalhe o tempo médio de duração das aplicações, tal como o Theta médio e a média de V_{min}, por turma, por ano.

É importante ressaltar a coerência dos dados em relação aos anos em que o TAI-Leitura foi aplicado. As turmas de um mesmo ano escolar mostram pouca alteração em relação à diferença de tempo médio, V_{min} e o Θ_i médio, de forma que o software consegue discriminar semelhanças entre os grupos, ano por ano. Igualmente chamamos atenção para o fato de o tempo médio ser maior no terceiro ano, diminuindo no 4º e 5º ano. Isso demonstra que foi necessário maior tempo para responder ao teste devido ao maior número de itens apresentados às respondentes, uma vez que o Theta médio está alocado nos níveis intermediários da escala de proficiência. Nesse sentido, em turmas intermediárias dos anos iniciais do Ensino Fundamental, é esperado maior tempo de realização do TAI-Leitura, posto que, devido à evolução pedagógica as crianças podem estar alocadas no meio da escala de proficiência.

Tabela 9 –Theta médio, tempo médio e variação média por ano e turma. Escola Maurîtânia.

2021

| Turma/Ano | Tempo médio | DP Tempo | Média Vmin | Média Theta |
|------------------|--------------------|-----------------|-------------------|--------------------|
| 1A | 00:18:52 | 00:07:32 | 397,97 | 435,96 |
| 1B | 00:17:47 | 00:10:13 | 350,50 | 384,85 |
| 1C | 00:17:10 | 00:07:03 | 346,30 | 384,52 |
| 1º Ano | 00:17:49 | 00:08:08 | 361,73 | 398,65 |
| 2A | 00:19:02 | 00:10:35 | 420,41 | 456,11 |
| 2B | 00:23:52 | 00:13:26 | 455,20 | 490,11 |
| 2º Ano | 00:21:17 | 00:12:05 | 436,53 | 471,87 |
| 3A | 00:21:54 | 00:09:37 | 473,01 | 511,37 |
| 3B | 00:25:24 | 00:09:47 | 449,97 | 484,50 |
| 3º Ano | 00:23:27 | 00:09:42 | 462,85 | 499,52 |
| 4A | 00:21:07 | 00:09:22 | 527,40 | 568,93 |
| 4B | 00:20:54 | 00:09:47 | 557,83 | 600,77 |
| 4º Ano | 00:21:00 | 00:09:28 | 543,41 | 585,69 |
| 5A | 00:15:04 | 00:07:10 | 569,39 | 611,91 |
| 5B | 00:13:29 | 00:06:45 | 526,11 | 570,49 |
| 5º Ano | 00:14:13 | 00:06:55 | 546,24 | 589,76 |

Fonte: Elaboração própria.

A seguir, buscamos problematizar se o instrumento consegue captar uma proficiência que seja similar à observada pelas professoras, em especial quando levam em consideração suas experiências pedagógicas dentre suas turmas.

5.3 AS PERCEPÇÕES DAS PROFESSORAS E OS RESULTADOS

A participação das docentes ao longo do processo mostrou-se produtiva em diversos aspectos. Em ambas as escolas foi possível realizar conversas a respeito da avaliação educacional, sendo que esta etapa se mostrou mais organizada e produtiva na Escola Maurîtânia. Tal diferença se deu devido ao fato de que nessa escola foi realizada uma reunião de apresentação e preparação de aplicação do TAI-Leitura, na qual o grupo de pesquisadores do Gepave, somados a estagiários do Programa de Residência Pedagógica (PRP), puderam acompanhar as discussões preliminares com as docentes.

Ao todo, 11 professoras participaram da reunião de apresentação. Nesta foram apresentados aspectos iniciais sobre o projeto, discutidas as noções de avaliação educacional que fundamentam a pesquisa, e explicitados os objetos a serem avaliados ao longo das coletas de dados via TAI-Leitura. Igualmente, após o seminário de apresentação, as professoras tiveram a

oportunidade de realizar um teste no software, a fim de compreenderem como que o teste se comporta nas mais variadas hipóteses. Além disso, foi pedido que as docentes alocassem suas estudantes de acordo com a escala de cinco níveis PB, a mesma utilizada para os resultados do TAI-Leitura, tendo em mente o aprendizado observado no cotidiano escolar em relação à leitura.

Em contrapartida, na Escola Ciclo as conversas se deram de forma individual com as professoras das cinco turmas, onde a aceitação para participar do projeto foi firmada com a coordenação e corpo docente. Foi introduzido TAI-Leitura e entregue os kits impressos da PB a fim de que pudessem verificar a matriz de referência em leitura e demais particularidades da avaliação – em especial os testes comumente entregues para as crianças e as professoras. Igualmente foi pedido que as professoras alocassem suas alunas de acordo com a matriz de referência da PB em relação ao cotidiano escolar, mas a entrega dessa avaliação se deu em diferentes momentos dos últimos meses de 2021.

A análise dos resultados dos testes caminha para a percepção de que considerar o V_{\min} é mais relevante para o trabalho pedagógico, quando aliado aos materiais que discriminam aquilo que é entendido enquanto a proficiência de respectivo nível, somado ao olhar das professoras que atuam nas salas de aula. Isto porque o *software* demonstra que é plenamente possível que o intervalo da proficiência estimada de uma criança esteja entre dois níveis de proficiência – no intervalo em torno da proficiência estimada (Θ_i) –, algo dificilmente captado na prova em papel ou por um TBC que não utilizam a TRI. É importante ter-se em mente que os intervalos são importantes para a medição da proficiência, já que os itens são apresentados em relação aos valores de corte de cada nível de proficiência da escala da PB.

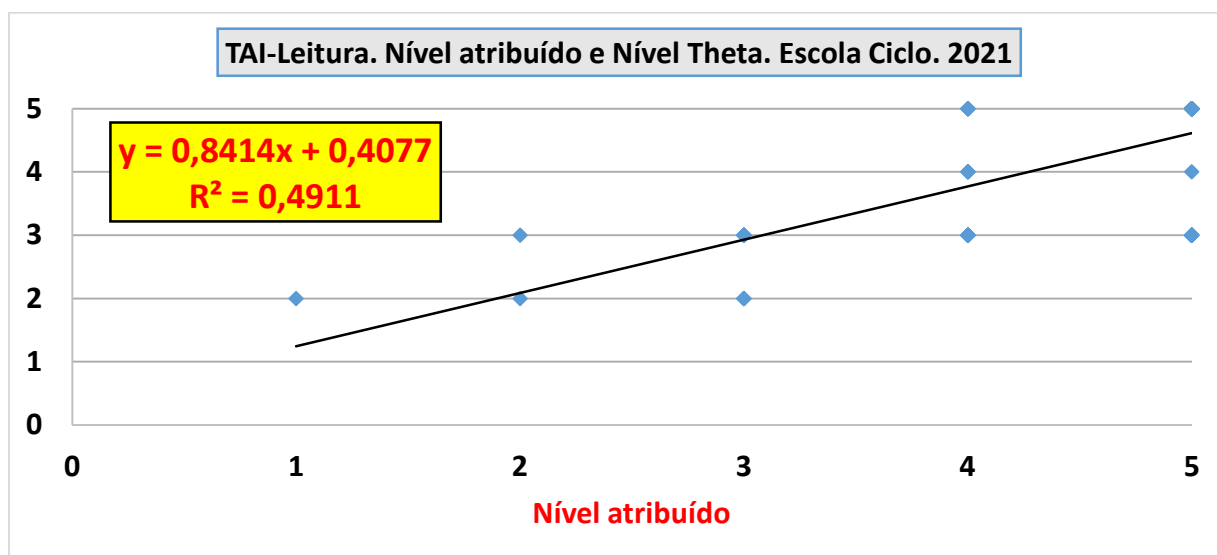
Especial atenção deve ser dada para a atribuição de níveis do TAI-Leitura, principalmente, quando levamos em consideração possíveis níveis atribuídos pelas próprias professoras às crianças respondentes. O estudo apresenta, então, gráficos que buscam analisar comparativamente as alocações da proficiência estimada das respondentes pelas professoras e pelo *software*, de modo a captar possíveis semelhanças nas percepções entre ambos. Esses gráficos apresentam os quadrados azuis, cujo propósito é representa as alunas, plotados com os valores atribuídos pelas educadoras no eixo do x e os do TAI-Leitura no eixo y, existindo superposições.

Em relação à Escola Ciclo, quando comparadas as percepções das professoras e os resultados das aplicações das crianças, percebe-se a existência de uma coerência entre a percepção das

professoras e o instrumento avaliativo, sendo que o TAI-Leitura trilha o mesmo caminho do olhar pedagógico das professoras.

O Gráfico 7 busca analisar a concordância entre a alocação das professoras em relação a alocação da proficiência estimada do *software*, demonstrando que as professoras da Escola Ciclo atribuem às suas alunas níveis de proficiência similares ao próprio instrumento, após concordarem com a interpretação pedagógica da matriz de referência da escala de proficiência da PB. A reta demonstra a concordância uma vez que é ascendente, o que caminha para a crescente associação entre o TAI-Leitura e a alocação das professoras na escala de proficiência a qual o *software* se baseia.

Gráfico 7 – TAI-Leitura. Nível atribuído e Nível Theta. Escola Ciclo. 2021



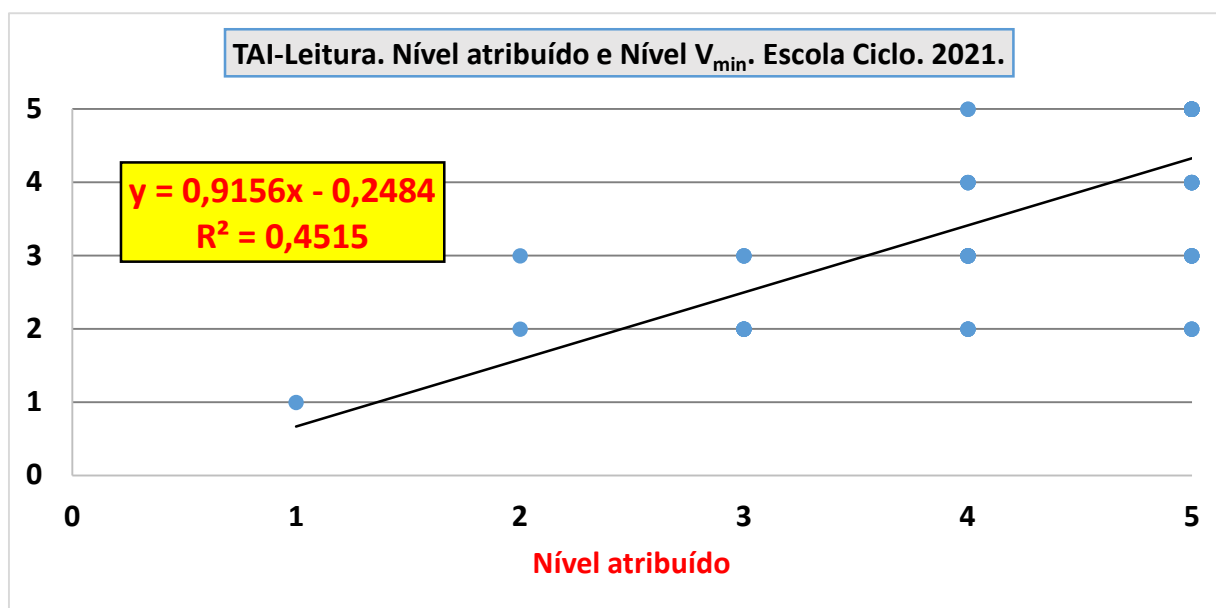
Fonte: Elaboração própria.

Em relação ao V_{\min} , conforme exposto no Gráfico 8, a linha de tendência comporta-se de forma similar: a linha de tendência demonstra a concordância entre o TAI-Leitura e a atribuição de níveis, embora a taxa de concordância no gráfico 8 seja maior que a exposta no gráfico 7, que diz respeito a análise entre a relação da alocação da estimativa média da proficiência das crianças respondentes, a percepção das professoras e a alocação por parte do *software*. Nesse sentido, portanto, apontamos que a similaridade da percepção pedagógica da estimativa de proficiência das professoras e o que é apontado pelo instrumento avaliativo (TAI-Leitura) é coerente e caminha para a similaridade, um dado positivo.

A linha de tendência comporta-se de forma linear e o posicionamento dos pontos está próximo a esta, tanto com a média Theta quanto com o V_{\min} , o que indica que o TAI-Leitura e as

professoras apontam as mesmas alocações em relação ao nível de proficiência de grande parte das estudantes respondentes. Frisamos ainda que existem discordâncias, as quais indicam linhas de pesquisas que visem compreender tais comportamentos.

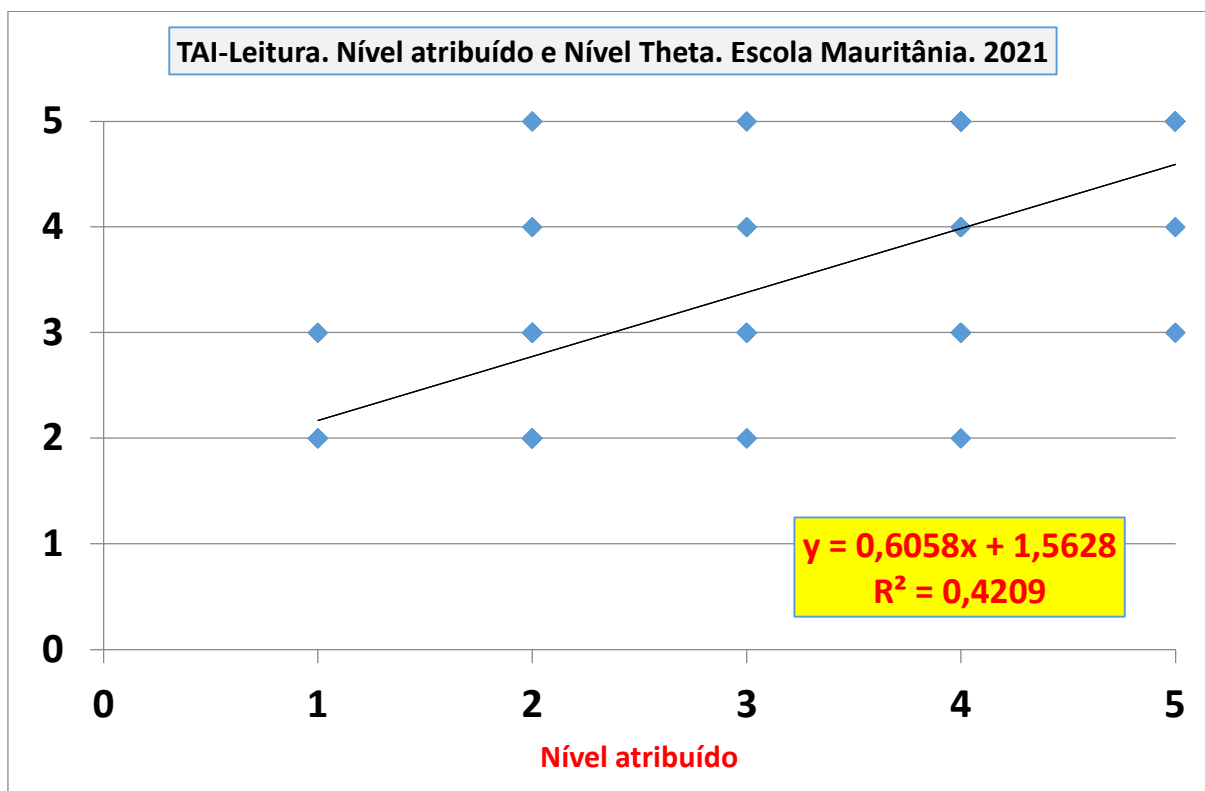
Gráfico 8 – TAI-Leitura. Nível atribuído e Nível V_{\min} . Escola Ciclo. 2021



Fonte: Elaboração própria.

Portanto, ao observarmos os Gráficos 7 e 8, podemos perceber que há maior concordância entre a atribuição de professoras e a alocação do nível de proficiência pelo TAI-Leitura, ainda que com diferença quando se considera o V_{\min} , mais latente na Escola Ciclo do que na Escola Maurîtânia, guardadas as proporções em relação a quantidade de estudantes respondentes, dentre demais aspectos conforme se pode observar a seguir.

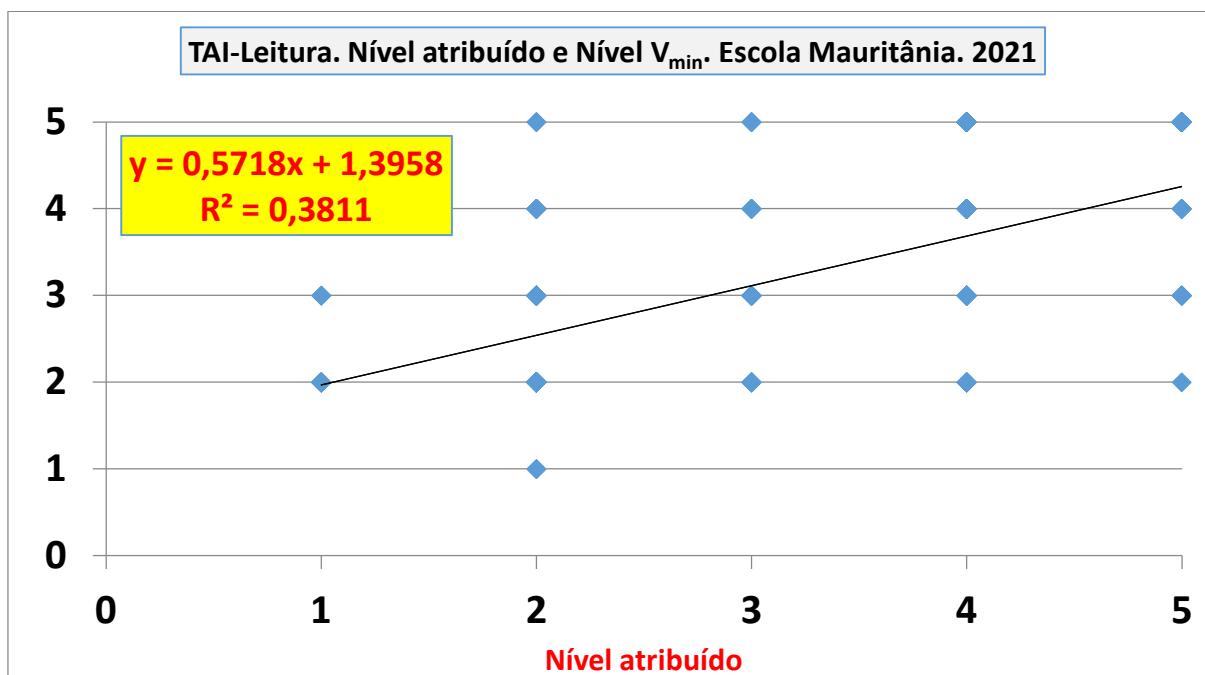
O Gráfico 9 traz os dados relativos à Escola Maurîtânia. Sobre ele observamos que a linha de tendência indica que a percepção pedagógica das professoras da em relação às alunas que responderam ao TAI-Leitura se aproxima dos resultados do teste, embora com vários casos discrepantes, que podem ser decorrentes ou de consideração de fatores que não são exatamente de aprendizado em leitura ou de avaliação da capacidade de leitura com limitações, entre outras possíveis explicações que a continuidade de pesquisas pode revelar.

Gráfico 9 – TAI-Leitura. Nível atribuído e Nível Theta. Escola Maurîtânia. 2021

Fonte: Elaboração própria.

Os pontos azuis são as alocações ao longo dos níveis. O eixo X diz respeito aos níveis atribuídos pelas professoras; o eixo Y são as medições realizadas pelo instrumento. As atribuições feitas pelas professoras foram solicitadas no primeiro encontro, onde o TAI-Leitura foi introduzido, bem como os materiais da PB foram discutidos. Conforme pode-se perceber, há uma concordância com os níveis dados, descritos na matriz de referência e levados em conta pelas docentes, e aquilo que é medido pelo TAI-Leitura.

O Gráfico 8 faz uma comparação com os níveis atribuídos pelas docentes e o nível V_{\min} , sendo este calculado levando em conta a proficiência média (Theta) menos o erro de medida, conforme já citado.

Gráfico 10 – TAI-Leitura. Nível atribuído e Nível V_{\min} . Escola Maurîtânia. 2021

Fonte: Elaboração própria.

Observando-se os Gráficos 9 e 10, nota-se que a discrepância foi um pouco maior quando se considerou o V_{\min} , ressaltando ainda que novas pesquisas podem explorar tais diferenças. É importante frisar que tais gráficos dizem respeito a todas as turmas dos anos iniciais do Ensino Fundamental. Assim, é possível notar a concordância entre a percepção pedagógica das docentes da Escola Maurîtânia e a medição realizada pelo TAI-Leitura.

Ao analisarmos ambas as escolas, temos um quadro similar entre elas no que diz respeito ao comportamento das atribuições de níveis de proficiência às crianças das turmas respondentes em relação as alocações feitas pelas professoras. Tal indicativo diz-nos que o TAI-Leitura se comportou de forma similar tanto em uma escola privada quanto em uma escola pública.

Vale destacar que em uma das reuniões da pesquisa, ao serem problematizados e discutidos os dados obtidos na aplicação às turmas da Escola Maurîtânia, uma das professoras – denominada professora A. –, relatou que os objetos e conceitos que levava em consideração para avaliar as suas crianças diziam respeito às competências de escrita e não necessariamente de leitura. Tal dado é importante, pois coloca em evidência aspectos levados em conta quando as professoras avaliam a leitura, algo preconizado pela literatura e previamente discutido no presente trabalho.

Deste modo, a percepção docente de ambas as escolas campos concordam e caminham no mesmo sentido no que diz respeito às atribuições de níveis de proficiência e os resultados do

TAI-Leitura. Ressaltamos, portanto, que a função formativa advinda da medição realizada pelo TAI-Leitura, por parte das professoras, pode ser realizada, posto que os resultados não divergem de suas próprias observações.

6. CONSIDERAÇÕES FINAIS

Na finalização deste trabalho, retomamos alguns elementos conceituais que subsidiaram a pesquisa e destacamos alguns elementos que podem contribuir tanto na problematização do TAI-Leitura quanto para seu aperfeiçoamento enquanto um instrumento de avaliação da aprendizagem em leitura a ser utilizado em sala de aula, especialmente por professoras no início do Ensino Fundamental.

Em relação ao diálogo entre avaliações internas e externas, conforme discutido por Nevo (1998), concordamos com a necessidade dessa relação mais dialógica entre tais avaliações e entre seus aspectos formativos e somativos, a fim de enriquecer o planejamento educacional das instituições escolares.

No caso da pesquisa, pode-se considerar que esse diálogo se manifesta na possibilidade de utilização de um instrumento com características típicas de uma avaliação em larga escala, com a demanda testes e interpretações padronizadas, mas que pode ser apropriado para o trabalho docente e a avaliação da aprendizagem conduzida em sala de aula, mesmo com a necessidade de aperfeiçoamentos no TAI-Leitura.

Ao levarmos em conta o paradoxo docente, como salientado por Alavarse (2013), como um fomento à problematização de instrumentos avaliativos e seu uso no cotidiano escolar, o presente estudo traz dados que envolvem a avaliação educacional, tanto quanto se encaminham para a utilização de um instrumento informatizado e adaptativo, problematizando-o, sobretudo pelo fato de que sua utilização pode levar professoras a refletirem sobre a avaliação da aprendizagem que conduzem, destacando a importância de se garantir que avaliação da aprendizagem seja da aprendizagem, no caso, da capacidade de leitura.

Ao tratar os desafios de avaliar a capacidade de leitura nos anos iniciais do Ensino Fundamental, a pesquisa apresenta argumentos que problematizam a construção de instrumentos avaliativos para o objeto supracitado. Soares (2009) traz pertinentes apontamentos acerca da avaliação do letramento e da alfabetização, atribuindo um lugar de questionamento aos resultados dos processos avaliativos, reiterando a preocupação sobre os primórdios da elaboração destes. Acreditamos que a preocupação é pertinente, ainda mais se somada ao desconhecimento de como os instrumentos são idealizados, construídos, aplicados e de quais as possibilidades de utilização de seus resultados; destacando que essas possibilidades devem estar presentes no início do processo, em sua etapa de idealização.

A preocupação em avaliar o rendimento gramatical das crianças ao longo do processo de alfabetização e letramento em detrimento de aspectos semânticos e de compreensão do que é lido, coloca a avaliação da leitura em um local de ambiguidade. Pois existem práticas que podem fomentar a valorização de aspectos decodificáveis da língua, relacionada à leitura, mas que não deve restringi-la a tais aspectos. O que a pesquisa demonstra é o que já é preconizado por Soares (2009) quando aponta a complexidade da concepção de conceitos que discriminem o que deve ser entendido por leitura.

A exemplo do que se expõe, apontamos o depoimento da professora A., que, ao avaliar a proficiência de leitura de sua turma, confundiu-se com aspetos relacionados à escrita, atribuindo-lhes um conceito avaliativo que não condiz com a discriminação da matriz de referência em leitura. De forma similar, com vistas às diferenças entre os conceitos de alfabetização e letramento ao longo de sua avaliação, Corrêa e Dornelles (2010), ao estudarem o ensino de Língua Portuguesa e as práticas didático-pedagógicas aplicadas em sala de aula, apontam uma visão aditiva de língua que está relacionada com uma perspectiva ultrapassada de ensino de língua, em especial para a alfabetização. Indicam que as práticas observadas em sala de aula, promovidas pelas professoras em atividades, como também na aplicação da PB, focam a tecnicidade da alfabetização, nas quais é dada maior ênfase ao ensino de gramática em detrimento de práticas focadas na leitura e no texto. As autoras entendem que a avaliação proposta pela Provinha Brasil ajuda a manter a valorização da gramática devido ao modo como ela é organizada, em especial sua matriz de referência, o que compreendemos ser um olhar demasiadamente raso em relação aos documentos orientadores da PB. Isto pois os conceitos abordados no estudo não são compatíveis com aqueles trabalhados ao longo dos documentos de apresentação da PB, fundamentais para sua compreensão e aplicação.

Tal compreensão a respeito da avaliação de Língua Portuguesa, em especial ao longo da alfabetização e letramento, diz respeito a um desentendimento de aspectos da competência em leitura que não deve ser confundida com a competência escrita tão somente, visto que possui aspectos e processos específicos. Assim, portanto, pensar a PB enquanto uma avaliação capaz de instituir uma visão aditiva de língua é similar ao raciocínio da professora A.; indica-nos que a matriz de referência da PB pode não ter sido devidamente lida e apropriada pelas autoras, e os conceitos levados em consideração no processo de avaliação dos sujeitos podem fugir daqueles preconizados enquanto objetos da competência em leitura, tanto da PB quanto de demais documentos – tal como a própria BNCC aponta (BRASIL, 2018, p. 89).

O entendimento de que a avaliação é um processo, e que a PB serve enquanto instrumento avaliativo de modo a fazer parte deste, em vez de ser entendida enquanto a totalidade de uma avaliação, parece-nos ser um exemplo da necessidade de se discutir a avaliação educacional e instrumentos avaliativos desde a formação básica do professorado, estendendo-se para a formação continuada. Para se evitar apreciações sem o devido aprofundamento, particularmente quanto ao que é preconizado na matriz de referência da PB e com relação aos níveis de proficiência da competência em leitura, entendemos que há a necessidade de se sistematizar a discussão desses tópicos entre as professoras, posto que a comum avaliação da competência em escrita pode ser confundida com a leitura, embora sejam esses processos diferentes, como pôde ser observado na Escola Maurîtânia ao longo de nossa pesquisa através da fala da professora A.

Quando Soares (2009) traz à luz a discriminação das habilidades a serem testadas em uma prova e/ou teste, concordamos com a sua preocupação. Neste caso, uma avaliação deve trazer consigo materiais que sirvam de aporte a sua idealização, os conceitos que fundamentam os objetos a serem avaliados e, acima de tudo, servir enquanto instrumento de apoio ao trabalho docente. O Teste Adaptativo Informatizado em Leitura, construído pelo Gepave, tem como proposição este desafio: oferecer aporte na avaliação de leitura para a tomada de decisões no cotidiano escolar.

A concordância das professoras em ambas as escolas ao que foi sistematizado pela PB enquanto matriz referência é apontada nos gráficos de 7 a 10, em que são alocados os níveis obtidos tanto pela média Theta quanto pelo V_{\min} , e apresenta similaridade ao que foi observado pelas docentes em sala de aula em relação às suas alunas. Nesse sentido, percebemos que o TAI-Leitura serve enquanto um instrumento para a medição da proficiência em leitura, em especial por ser capaz de realizá-la seguindo as competências elencadas em sua matriz.

A pesquisa demonstrou o potencial do TAI-Leitura, uma vez que sua aplicação não está na mera realização do teste em sala de aula, mas visa o apoio à formação continuada das professoras em noções de avaliação escolar, voltando-se para a leitura como um (não o todo) dos objetos inerentes à alfabetização e letramento. Apoiado na Provinha Brasil (BRASIL, 2016a; 2016b), o TAI-Leitura utiliza os itens da prova, tal como parte de sua matriz de referência, restringindo-se aos aportes tecnológicos que o caracterizam enquanto um teste eletrônico, computacional e informatizado.

Nesse sentido, estudos a respeito do que se compreende por leitura, em matrizes de referência – isto é, as habilidades discriminadas e propostas pelos itens dos instrumentos avaliativos – são uma proposição de trabalho futuro. Outro aspecto apontado no estudo e que diz respeito à observação ao longo das aplicações foi a taxa de exposição de item, ocorrida em decorrência

da quantidade de itens do banco do TAI-Leitura. Assim, outra indicação seria a ampliação dos itens. Essa ampliação também se relaciona a precisão da aferição da proficiência das crianças, uma vez que quanto maior for o banco de itens, maior pode ser a precisão em relação a alocação da respondente na escala de proficiência.

O que observamos ao longo da pesquisa foi a aprovação por parte das docentes de duas escolas em relação à utilização do instrumento. O TAI-Leitura oferece saída de dados mais rápidos, é estimulante para as crianças e de fácil utilização. Além disso, a devolutiva às professoras fez com que o processo de pesquisa fosse bem aceito e nossa percepção foi a de que o TAI-Leitura surgiu enquanto uma ferramenta válida para a utilização do corpo docente – uma preocupação por parte da equipe, posto que a premissa sempre foi a de oferecer subsídios para a avaliação em leitura e não a substituição do professorado.

Tal dado encontra-se em concordância com o que pode ser observado em relação aos níveis atribuídos às alunas, pelas docentes, nas duas escolas e os níveis mensurados pelo TAI-Leitura com baixa taxa de erro. A concordância de ambos os dados aponta para um bom funcionamento do *software*. Somadas a isso, estão as próprias aplicações que, conforme apontado previamente pelos comentários das crianças das duas escolas, mostraram-se positivas.

Ainda, quando apresentados os dados obtidos nas aplicações, é possível notar que o recurso é capaz de captar diferenciações dentre as respondentes, levando em conta a coerência do desenvolvimento da aprendizagem em leitura, onde a premissa é a de que há um avanço, guardadas as proporções, ao longo dos anos escolares. Desta forma, o TAI-Leitura é um recurso que está em sintonia com o que é debatido pela literatura em relação à avaliação, especialmente a avaliação da proficiência em leitura.

Dentre os apontamentos citados e demais pontos positivos observados, ressaltando-se que um dos principais desafios de qualquer processo de medição é que o resultado reflita no nível de aprendizado de quem tem a sua proficiência estimada, uma das vantagens do TAI-Leitura é a possibilidade de ser utilizado várias vezes em pouco tempo. Assim, pode fornecer resultados que, sendo semelhantes, podem ser considerados como mais próximos do nível de aprendizado da respondente. Apontamos essa enquanto uma possibilidade de pesquisa dentro do próprio instrumento e em escolas parceiras, pois a compreensão do avanço do aprendizado em leitura ao longo da escala de proficiência e a sua captura pelo instrumento, permitiria verificar se o próprio TAI-Leitura seria capaz de perceber alguma mudança em determinado período de tempo.

Embora seja um instrumento rico e caro ao trabalho docente, em especial pela facilidade de aplicação e demais pontos previamente abordados, o TAI-Leitura ainda precisa ser aperfeiçoado em relação ao banco de itens, saída de dados e de resultados, e acessibilidade. Em relação ao banco de itens, é importante frisar que o banco atual possui menos de 40 itens, o que diz respeito à taxa de exposição destes e, além disso, à precisão da alocação das respondentes na escala de proficiência – afinal, quanto maior a variedade de itens, menor é a taxa de erro e mais precisa é a medição. Outro importante apontamento que pôde ser observado em relação ao TAI-Leitura é a utilização de alguns de seus itens, cuja aplicação se mostrou complicada, como foi o caso do item da xícara/caneca, do armário/escada e da imagem da Dengue, que é pequena.

Sobre a saída de dados e resultados, é importante apontar que os dados são visíveis para as professoras, coordenadoras e demais gestoras das escolas. No entanto, é necessário realizar um tratamento dos dados antes de analisá-los, o que significa maior tempo dedicado à sua organização. Desta forma, as saídas dos dados poderiam ser otimizadas, a fim de que o tratamento dos dados fosse menos trabalhoso. Outro detalhe é a possibilidade de impressão e de publicação dos resultados às respondentes, posto que o processo de avaliação é importante não apenas para as docentes e a escola, como também aos seus sujeitos.

Ainda a respeito de aperfeiçoamentos ao TAI-Leitura, apontamos a necessidade de se pensar o acesso em determinados aparelhos, já que a funcionalidade em suportes eletrônicos que utilizam o sistema operacional iOS demonstrou problemas para a execução do recurso. Além disso, certos navegadores também mostraram problemas na utilização do TAI-Leitura, em especial o Mozilla Firefox. Tais aperfeiçoamentos podem enriquecer o TAI-Leitura, fazendo de si um instrumento com o potencial para servir aos propósitos do processo de avaliação em leitura.

Concluimos, portanto, que as condições para avaliação do TAI-Leitura foram satisfeitas, ainda que aperfeiçoamentos no próprio teste, aprofundamentos teóricos sobre avaliação da proficiência em leitura e discussões de validação podem e devem ser feitos. De modo que o TAI-Leitura se mostra enquanto um instrumento capaz de fornecer subsídios para o processo de avaliação da competência em leitura de crianças nos anos iniciais do Ensino Fundamental. Ademais, fornece novos desdobramentos para seu aperfeiçoamento, novos questionamentos e desafios que interessam tanto aos seus pesquisadores, quanto à própria comunidade.

REFERÊNCIAS

- ALAVARSE, Ocimar Munhoz. Desafios da avaliação educacional: ensino e aprendizagem como objetos de avaliação para a igualdade de resultados. **Cadernos Cenpec**, São Paulo, v. 3, n. 1, p. 135-153, jun. 2013.
- ALAVARSE, Ocimar Munhoz; CATALANI, Érica Toledo. Alfabetização e TIC: os testes adaptativos informatizados (TAI) como recurso. In: COMITÊ GESTOR DA INTERNET NO BRASIL – CGI.br. **Pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras**: TIC Educação 2015. São Paulo: CGI.br, 2016. p. 35-44.
- ALAVARSE, Ocimar M.; CATALANI, Érica M. T.; MENEGETTI, Douglas de R.; TRAVITZKI, Rodrigo. Teste Adaptativo Informatizado Como Recurso Tecnológico para Alfabetização Inicial. **Revista Sistemas, Cibernética e Informática**, v. 15, n. 3, p. 68-78, 2018.
- ALAVARSE, Ocimar Munhoz; MELO, Wolney Candido de. Avaliação educacional e testes adaptativos informatizados (TAI): desafios presentes e futuros. In: COMITÊ GESTOR DA INTERNET NO BRASIL – CGI.br. **Pesquisa sobre o uso das tecnologias de informação e comunicação no Brasil**: TIC Educação 2012. Coord. Alexandre F. Barbosa. São Paulo: CGI.br, 2013. p. 103-112.
- BRASIL. Parâmetros Curriculares Nacionais: língua portuguesa. Brasília: MEC, 1997.
- _____. Pró-Letramento: Programa de Formação Continuada de Professores dos Anos/Séries Iniciais do Ensino Fundamental: alfabetização e linguagem. ed. rev. e ampl. Secretaria de Educação Básica. Brasília: Ministério da Educação, Secretaria de Educação Básica, 2008.
- _____. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Guia de elaboração e revisão de itens**. Brasília, 2010. v. 1
- _____. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Guia de elaboração de itens**: Provinha Brasil. Brasília: MEC/Inep, 2012.
- _____. Ministério da Educação. **Provinha Brasil**: avaliando a alfabetização: guia de apresentação, correção e interpretação dos resultados: Teste 1. Brasília: Inep, 2016a.
- _____. Ministério da Educação. **Provinha Brasil**: avaliando a alfabetização: guia de apresentação, correção e interpretação dos resultados: Teste 2. Brasília: Inep, 2016b.
- _____. Ministério da Educação. Secretaria de Educação Básica. **Base Nacional Comum Curricular**. Brasília: MEC: SEB, dez. 2018.
- CASASSUS, Juan. Uma nota crítica sobre a avaliação estandardizada: a perda de qualidade e a segmentação social. Tradução de Gabriela Lourenço. **Sísifo: Revista de Ciências da Educação**, n. 9, p. 71-78, maio/ago. 2009.
- CATALANI, Érica Maria Toledo. **Teste Adaptativo Informatizado da Provinha Brasil**: a construção de um instrumento de apoio para professores(as) e gestores(as) de escolas. Tese (Doutorado em Educação). Faculdade de Educação, Universidade de São Paulo. São Paulo, 2019.
- CAVARSAN, Cássia Helena Guillen; MIGUEL, Maria Elisabeth Blanck. A alfabetização nos Parâmetros Curriculares Nacionais (PCN) e na Base Nacional Comum Curricular (BNCC): o que mudou de 1997 a 2017. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 101, n. 259, p. 567-582, set./dez. 2020.

CORRÊA, Jocielle; DORNELLES, Clara. Leitura, alfabetização e Provinha Brasil - um olhar sobre o objeto do ensino de português em séries iniciais. IX Encontro do CELSUL, 2010, Palhoça, SC. Anais do IX Encontro do CELSUL, Palhoça: Universidade do Sul de Santa Catarina, p. 1-12. FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. Avaliação educacional em larga escala e *accountability*: uma breve análise da experiência brasileira. **Revista Pesquisa e Debate em Educação**, Juiz de Fora, v. 10, n. 1, p. 1103-1137, jan./jun. 2020.

FREIRE, Paulo. A Importância do ato de ler. In: _____. **A importância do ato de ler: em três artigos que se completam**. 23. ed. São Paulo: Cortez, 1989. p. 9-15.

KRUG, Flavia Susana. A importância da leitura na formação do leitor. **Revista de Educação do Ideau**, v. 10, n. 22, p. 1-13, jul./dez. 2015.

MICOTTI, Maria Cecília de Oliveira. O ensino e o aprendizado da escrita: novos olhares. **Educação: Teoria e Prática**, v. 16, n. 28, p. 29-46, jan./jul. 2007.

_____. A avaliação do ensino e do aprendizado de língua portuguesa nas séries iniciais da escola fundamental. In: BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Avaliações da educação básica em debate**: ensino e matrizes curriculares de referência das avaliações em larga escala. Brasília: Inep, 2013. p. 175-194.

MORAIS, Artur Gomes de; LEAL, Telma Ferraz; PESSOA, Ana Cláudia Rodrigues Gonçalves. O ensino da Língua Portuguesa no Ciclo de Alfabetização e sua avaliação pela Provinha Brasil. In: Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Avaliações da educação básica em debate**: ensino e matrizes curriculares de referência das avaliações em larga escala. Brasília: Inep, 2013. p. 153-174.

NEVO, David. Avaliação por diálogos: uma contribuição possível para o aprimoramento escolar. In: TIANA, Alejandro (Coord.). **Anais do Seminário Internacional de Avaliação Educacional, 1 a 3 de dezembro de 1997**. Tradução de John Stephen Morris. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), 1998. p. 89-97

SANTOS, Solange; SILVA, Leilane Ramos. Leitura, textos oficiais e Provinha Brasil: reflexões sobre um mesmo tema. **Interfaces Científicas-Educação**, Aracaju, v. 4, n. 2, p. 97-106, fev. 2016.

SOARES, Magda. **Letramento**: um tema em três gêneros. 3. ed. Belo Horizonte: Autêntica, 2009.

TRAVITZKI, Rodrigo et al. Teste Adaptativo Informatizado da Provinha Brasil - Leitura: Resultados e Perspectivas. **Estudos em Avaliação Educacional**, São Paulo, v. 31, n. 78, set./dez. 2020.

ANEXO I - As principais características psicométricas do TAI-Leitura

Thiago Fernando Costa

Pesquisador

Grupo de Estudos e Pesquisas em Avaliação Educacional (Gepave)

O TAI-Leitura, como qualquer outro TAI, tem como principal objetivo selecionar itens, de um Banco de Itens, a serem apresentados aos respondentes de acordo com as respostas dadas anteriormente, cujo item inicial pode ser aleatório ou condicionado e a finalização da aplicação do teste também deverá atender alguma condição.

Para Olea e Ponsoda (2002, p. 5), “[u]m Teste Adaptativo Informatizado (TAI) é uma prova construída para fins de avaliação psicológica ou educativa, cujos itens se apresentam e são respondidos mediante um computador, sendo sua característica fundamental a adaptação ao nível de competência progressivo que manifeste a pessoa”. Nestes termos, diferentemente dos testes tradicionais nos quais um conjunto de itens já se encontra definido a priori, independentemente dos respondentes, um TAI é um teste que se adapta ao respondente, administrado por algoritmos informatizado e desta forma, é possível chegar a resultados com maior confiabilidade – resultados mais precisos – e menor quantidade de itens apresentados – menor tempo de aplicação. Como consequência, um TAI deve facilitar e melhorar a tarefa avaliativa, por exemplo, de um professor.

Nesse sentido, o TAI-Leitura, por estar baseado nos itens e na escala da Provinha Brasil, consegue apresentar com maior facilidade essa relação entre o teste e a competência manifesta do estudante, sendo possível verificar a progressão do conhecimento – proficiência – em leitura em um período curto entre duas aplicações. Entretanto, o exposto como características principais precisa de um maior detalhamento dos elementos centrais para a constituição do TAI-Leitura, sendo assim, a seguir, vamos descrever tais elementos.

Principais elementos de um Teste Adaptativo Informatizado (TAI)

Existem alguns elementos básicos para a constituição do TAI-Leitura, tais como: Banco de Itens; algoritmos de início do teste, escolha do próximo item e parada do teste; estrutura de *back-end* e *front-end* do software utilizado para apresentar o TAI.

Banco de Itens e a Teoria de Resposta ao Item

O Banco de Itens (BI), pode-se dizer, é análogo ao acervo de uma biblioteca, visto que pode propiciar informações relativas à magnitude da proficiência que se pretende aferir, com variabilidade de itens sobre o objeto da medição e que posteriormente se poderá avaliar. Essa posição de centralidade do BI em um TAI revela-se desafiadora para os seus administradores, tendo em vista as especificidades psicométricas e estatísticas que a constituição de um BI exige, ambas dependentes da realização de um pré-teste. De acordo com Magis, Yan e van Davier (2017, p. 39),

Uma vez que o pré-teste é feito, psicometristas e os estatísticos precisam calibrar os itens usando a Teoria Clássica dos Testes (TCT) ou a Teoria de Resposta ao Item (TRI), e realizar uma revisão estatística para a qualidade dos itens, usando a TCT ou critérios TRI. Quaisquer itens que não atendem às especificações de qualidade devido a falhas ou ambiguidades, baixa correlação com o escore total observado, ou com valores extremamente baixos ou alto nível de dificuldade, devem ser eliminados. Assim, apenas itens qualificados e selecionados podem ser adicionados a um Banco de Itens. O banco de itens também será periodicamente avaliado / reavaliado por seu tamanho, especificações e equilíbrio de conteúdo.

Os resultados obtidos no pré-teste, por meio da análise citada, concedem ao BI a robustez que o TAI necessita. Entre as duas teorias citadas, a TRI tem sido amplamente utilizada para a constituição dos BI, principalmente pela possibilidade de comparabilidade das aplicações ao longo do tempo. Sendo assim, vamos apresentar algumas características dessa teoria.

Teoria de Resposta ao Item (TRI)

A TRI, em seu cerne, indica a possibilidade de se apresentar itens e respondentes na mesma escala de medida, algo que desencadeia a composição de uma escala de proficiência, como também da interpretação pedagógica dessa mesma escala. A TRI tem em sua constituição um conjunto de modelos matemáticos que buscam representar a probabilidade de um respondente em acertar um item, dada sua quantidade relativa ao “traço latente” ou à “proficiência em face ao que se está perguntando”.

É frequente o uso dos modelos logísticos unidimensionais para itens dicotômicos ou “dicotomizados” (que foram “corrigidos” como certo ou errado) e que se diferenciam pelo número de parâmetros que cada modelo tem; dessa maneira, existem os modelos logísticos de 1, 2 e 3 parâmetros, o que conforme Andrade, Tavares e Valle (2000, p. 9) podemos definir, respectivamente, como:

- (i) somente a dificuldade do item;
- (ii) a dificuldade e a discriminação;
- (iii) a dificuldade, a discriminação e a probabilidade de resposta correta dada por indivíduos de baixa proficiência.

No Brasil, após a introdução, em meados da década de 1990, da TRI nas análises das avaliações nacionais, o modelo logístico de 3 parâmetros ficou sendo o mais utilizado, ainda que, na Provinha Brasil, o modelo de Rasch, onde temos um único parâmetro, acabou sendo o escolhido, pois ele possibilita relacionar mais facilmente a quantidade de acertos e medida do traço latente. A seguir, apresentamos algumas características a respeito desses modelos.

Modelo logístico unidimensional de Rasch

O Modelo de Rasch é definido por:

$$P(\theta_j) = \frac{1}{1 + e^{-D(\theta_j - b_i)'}}$$

com $i = 1, 2, \dots, I$, e $j = 1, 2, \dots, n$, onde:

U_{ij} = variável dicotômica que assume os valores 1, quando o respondente j responde corretamente ao item i , e 0 quando o respondente j responde incorretamente ao item i .

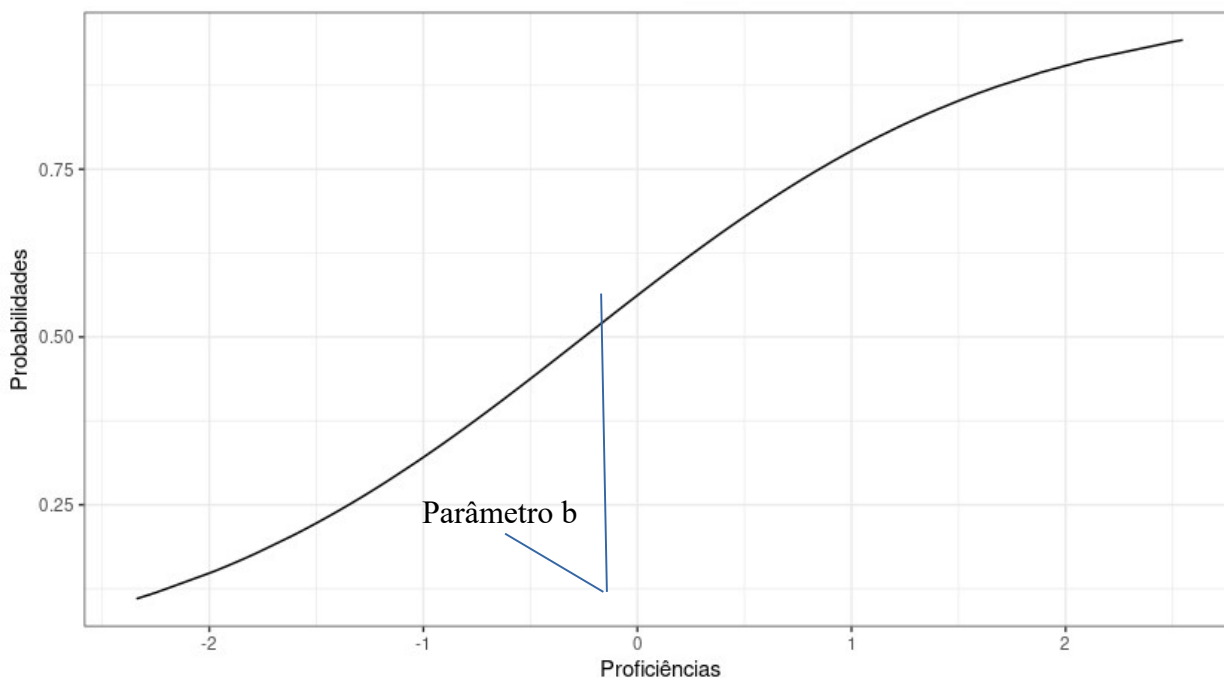
θ_j = representa o traço latente (habilidade) do j -ésimo estudante.

$P(\theta_j)$ = probabilidade do j -ésimo estudante responder corretamente um item i , dado o seu traço latente.

b_i = parâmetro que representa a “dificuldade” do item medido na mesma escala do traço latente.

D = fator de escala constante e igual a 1. Utiliza-se 1,7 quando se quer que a função logística forneça valores semelhantes a função da ogiva normal.

O modelo proposto representa as probabilidades de estudantes responderem corretamente um determinado item, essa relação é evidenciada graficamente pela Curva Característica do Item (CCI), como pode ser observado no Gráfico 1.

Gráfico 1 – Curva Característica de um Item (CCI)

Fonte: Elaboração própria

As Curvas Características dos Itens (CCI) são a melhor representação do princípio que rege os modelos da TRI que, segundo Muniz (1997, p. 19), “assumem que existe uma relação funcional entre os valores da variável que mede os itens e a probabilidade de acertar estes, denominando a função CCI”. Em outras palavras, a probabilidade de os respondentes acertarem um item está intrinsecamente ligada à medida de seus “traços latentes”.

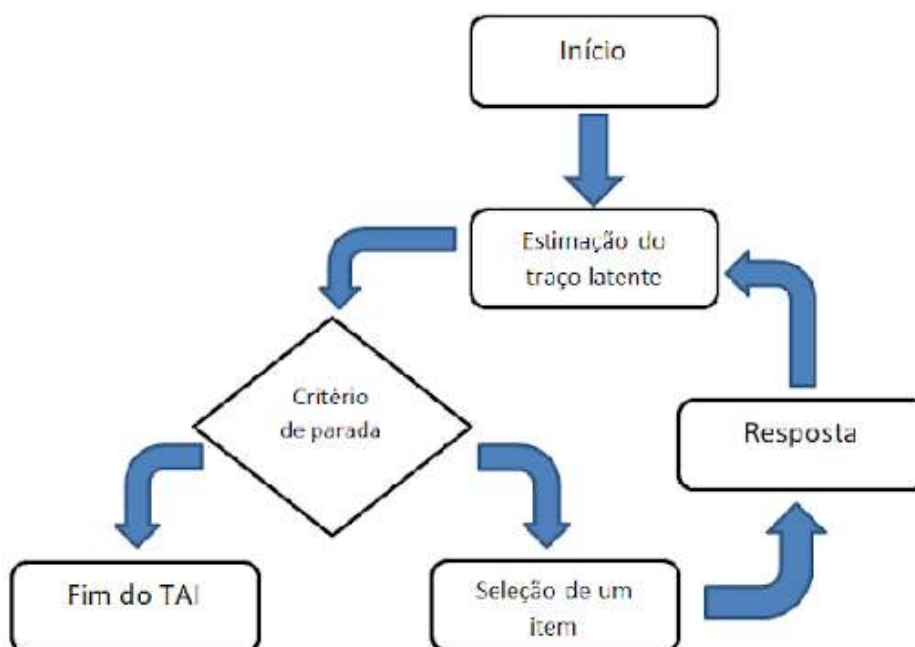
A escala do traço latente (proficiência) é arbitrária, onde, segundo Andrade, Tavares e Valle (2000, p. 10), “o importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude”. O parâmetro “b” utiliza a mesma unidade de medida do traço latente; na realidade o “b” é análogo à própria medida do traço latente, uma vez que o mesmo representa o quanto de “proficiência” o estudante precisa para ter maior probabilidade de acertar um item, dessa maneira, quanto maior o parâmetro “b” mais difícil é considerado o item e o inverso também é verdadeiro.

No que convém ao TAI-Leitura, tanto o BI como, por consequência, os parâmetros dos itens são de grande importância, principalmente o parâmetro “b”, em razão de ser por meio dele que se dá a escolha dos itens que serão administrados. Desta maneira, os algoritmos que compõem o TAI se tornam essenciais.

Os algoritmos do Teste Adaptativo Informatizado

O TAI-Leitura é composto por três algoritmos principais, que são: início do teste ou partida; seleção do próximo item; e parada do teste. Com essa disposição, o TAI-Leitura se constitui como um teste que se adapta ao respondente, assim, como já exposto, traz maior precisão e rapidez para com a medida do traço latente. Dessa maneira, a estrutura do TAI-Leitura pode ser representada segundo o que Barrada (2012, p. 290) apresenta em seu diagrama, como consta na Figura 1.

Figura 1 – Diagrama de fluxo de um TAI



Fonte: Adaptado de Barrada (2012)

De acordo com essa estrutura, a fluidez de um TAI depende de três pontos cruciais responsáveis por esse fluxo adaptativo, “Início”, “Critério de parada” e “Seleção de um item”. Vamos abordar cada um desses pontos separadamente, com o intuito de compreendermos não somente cada um dos algoritmos do TAI-Leitura, mas as possibilidades que qualquer TAI possa vir a ter.

Algoritmo de partida

Existem diversas possibilidades de se iniciar um TAI, sendo que o formato digital potencializa a criatividade, tornando viável a utilização de diversas informações do estudante previamente, como, por exemplo, o Nível Socioeconômico (NSE) ou as notas atribuídas pelos professores. Importante frisar que o primeiro item, que compõe o processo de início do TAI-Leitura, foi fixado no centro da escala, no caso, entorno de 500 pontos na escala; logo, temos o mesmo item sendo apresentado todas as vezes, algo que, potencialmente, pode ser uma restrição ao teste, pois o item passa a interferir no processo de avaliação, na medida que os estudantes conhecem a resposta pelas repetidas vezes em que foi apresentado, incluindo, eventuais, comentários entre os estudantes. Isso coloca o desafio de se implementar um controle para a quantidade de exposições do item, para evitar o denominado “viés de exposição do item”, ou seja, a excessiva exposição do item enviesaria as respostas dos estudantes.

Algoritmo de seleção de um item

O algoritmo de seleção de um item é fundamental para o funcionamento do TAI, pois é com base nesse procedimento que o teste pode ter um fluxo de apresentação de itens para o estudante respondê-los.

O algoritmo de seleção de um item é o “sistema nervoso” de um TAI; por meio dele se faz a constituição do teste que o estudante respondeu ao final da aplicação, como também permite a aproximação desse conjunto de itens à proficiência do estudante, garantindo o fluxo entre apresentação e coleta da informação.

Existem diversos algoritmos que merecem destaque, porém no TAI-Leitura o critério de Máxima Informação de Fisher foi o escolhido para ser utilizado por conta ampla utilização em outros TAI, além de ser um modelo estatístico amplamente testado e provada a sua eficácia.

Algoritmo de parada do Teste Adaptativo Informatizado

O algoritmo de parada está embasado no princípio de que existe um passo a passo para se definir, com relação ao ponto da escala que a estimativa de proficiência do estudante que está realizando o teste se encontra, se o teste deve ser concluído. Nesse sentido, esse algoritmo tem papel central na estimativa de proficiência do estudante. Existem muitas formas de se fazer um teste parar, desde a estabilização do erro, que não quer dizer “sem existir erro”, até o processo,

como é feito no TAI-Leitura, de estabilização de um intervalo de confiança dentro de um intervalo de interpretação da escala da Provinha Brasil.

Os usos do TAI também devem ser observados para a escolha do algoritmo de parada. Dependendo da finalidade do teste, um único algoritmo pode ser utilizado, contudo não é raro verificar a junção de algoritmos, como, por exemplo, encontramos no TAI presente em Alavarse; Catalani, Meneghetti e Travitzki (2017, p. 5), com um critério misto, a saber:

a) número de itens do teste (mínimo de 8 e máximo de 20 itens); b) limite permitido de incerteza (Erro Padrão menor que 35) na estimação da proficiência; e c) confiabilidade do nível de proficiência, tendo sido considerado confiável quando o intervalo de confiança da estimação da proficiência corresponde a um único nível de proficiência, dentre os cinco níveis de proficiência da Provinha Brasil (PB).

No Brasil, é comum para a estimativa da proficiência, utilizar o método de Esperança a Posteriori (EAP) que, ao contrário da estimação por máxima verossimilhança, não apresenta grandes inconsistências para vetores totalmente corretos ou incorretos (quando o respondente acerta ou erra todos os itens). Os métodos para estimar as proficiências são: Máxima Verossimilhança – MV; Esperança a Posteriori – EAP; Moda a Posteriori – MAP; e Máxima Verossimilhança Conjunta – MVC. Importante salientar que realizada uma escolha de método para estimar as proficiências, qualquer modificação em tal critério afeta diretamente a medida, como também, a consistência da série histórica.

Referências

- ALAVARSE, Ocimar Munhoz; CATALANI, Érica Toledo; MENEGHETTI, Douglas de Rizzo; TRAVITZKI, Rodrigo. Teste adaptativo informatizado como recurso tecnológico para alfabetização inicial. In: MEMORIAS de la Séptima Conferencia Iberoamericana de Complejidad y Cibernética (CICIC 2017). Winter Garden, FL: International Institute of Informatics and Systems (IIIS), March 2017. p. 165-169.
- ANDRADE, Dalton Francisco de; TAVARES, Héilton Ribeiro; VALLE, Raquel da Cunha. **Teoria da Resposta ao Item: conceitos e aplicações**. São Paulo: Associação Brasileira de Estatística, 2000.
- CATALANI, Érica Maria Toledo. **Teste Adaptativo Informatizado da Provinha Brasil: a construção de um instrumento de apoio para professores(as) e gestores(as) de escolas**. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo. São Paulo, 2019.
- MAGIS, David; YAN, Duanli; von DAVIER, Alina A. **Computerized Adaptive and Multistage Testing with R: Using Packages catRandmstR**. [S.l.]: Springer International Publishing, 2017.
- MUÑIZ, José. **Introducción a la Teoría de Respuesta a los Ítems**. Madrid: Pirámide, 1997.

OLEA Julio, PONSODA, Vicente; PRIETO, Geraldo. **Tests informatizados**: fundamentos y aplicaciones. Madrid: Pirámide, 1999.