

# Teste Adaptativo Informatizado Como Recurso Tecnológico para Alfabetização Inicial

**Ocimar M. ALAVARSE**

Faculdade de Educação, Universidade de São Paulo  
São Paulo, São Paulo, 05508-040, Brasil

**Érica M. T. CATALANI**

Faculdade de Educação, Universidade de São Paulo  
São Paulo, São Paulo, 05508-040, Brasil

**Douglas de R. MENEGHETTI**

Faculdade de Educação, Universidade de São Paulo  
São Paulo, São Paulo, 05508-040, Brasil

**Rodrigo TRAVITZKI**

Faculdade de Educação, Universidade de São Paulo  
São Paulo, São Paulo, 05508-040, Brasil

## RESUMO

O presente artigo é uma adaptação de Alavarse et al (2017) e aponta parte dos resultados da pesquisa em andamento, realizada pelo Grupo de Estudos e Pesquisas em Avaliação Educacional (Gepave), da Faculdade de Educação da Universidade de São Paulo (Feusp) em parceria com órgãos públicos. O estudo delineou a construção do Teste Adaptativo Informatizado (TAI) para a Provinha Brasil. A Provinha é um teste padronizado, disponibilizado na versão impressa pelo órgão federal e respondido por crianças brasileiras em processo de alfabetização. Teoricamente, os TAI têm características que superam limites dos testes impressos, entre elas, captação eletrônica das respostas, permitindo apuração automática das pontuações e apresentação de um teste diferente para cada nível de proficiência do respondente. O desenvolvimento do TAI pressupõe o envolvimento de especialistas de diferentes áreas, configurando uma atuação interdisciplinar e articulada no uso das Tecnologias da Informação e Comunicação (TIC). Incorporou-se as TIC no processo de avaliação educacional e consequentemente no de aprendizagem. O TAI desenvolvido da PB possibilitou: testes com menor número de questões comparado ao teste impresso; sequência de questões mais ajustadas aos domínios dos alunos; e algoritmo de seleção que permitiu maior precisão na alocação dos respondentes nos níveis de proficiência, contribuindo para melhorar e

intervenção pedagógica, objetivo central da avaliação.

**Palavras-chave:** TAI, Avaliação da aprendizagem, Alfabetização, Tecnologia da informação e comunicação, Plataforma de avaliação educacional.

## 1. INTRODUÇÃO

A alfabetização é um dos principais objetos de ensino nos anos iniciais do ensino fundamental e um dos objetivos mais salientes da educação no mundo inteiro, pois é essencial para democratizar o acesso a uma das mais importantes competências de nossa cultura, conforme destacou Williams (2000). Contudo, no Brasil, vários indicadores apontam que a plena alfabetização parece tratar-se de uma meta longe de estar totalmente atingida, quer quando nos reportamos às taxas de aprovação nos anos iniciais, fruto de decisões de professores, muito focadas nessa competência, quer quando nos fixamos nos resultados de avaliações externas em larga escala tais como a Avaliação Nacional da Alfabetização (ANA) e a Prova Brasil, aplicadas para alunos do 3º e 5º anos, respectivamente, que revelam baixas proficiências para a imensa maioria das crianças matriculadas. Ademais, constata-se a própria dificuldade em avaliar a alfabetização na sala de aula, sobretudo, quando se considera a conceituação de avaliação defendida por Luka Mujika e Santiago

Ettxeberria (2009), que a concebem como julgamento de algo visando sua melhoria, ou por Nevo (1998), que considera o diálogo entre avaliações externas e internas. Esses problemas foram fundamentais para a proposta de desenvolvimento de um Teste Adaptativo Informatizado (TAI) para a Província Brasil (PB) em leitura, conduzida pelo Gepave, envolvendo 15 escolas da Rede Municipal de Educação de São Paulo (RMESP) e aproximadamente 2000 alunos, 80 professores e 20 pesquisadores durante o ano de 2016. Essas escolas foram escolhidas a partir de dois critérios: o primeiro foi o interesse em participar do estudo, manifestado primeiramente pelos gestores e o segundo, de natureza mais técnica, deveu-se à existência de alunos que apresentassem grande variabilidade nos resultados da aprendizagem em leitura, observados em aplicações anteriores da PB nessas unidades educacionais.

A PB na versão impressa, de responsabilidade do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), autarquia do Ministério da Educação brasileiro, é oferecida anualmente desde 2008 para ser aplicada aos estudantes do segundo ano do ensino fundamental. Cabe ressaltar que sua distribuição foi suspensa no segundo semestre de 2016, devido a “restrições financeiras”, de acordo nota publicada pelo Inep, na qual foi dito que a prova não seria aplicada até que a publicação das Bases Curriculares Nacionais, em fase de elaboração em 2017 e 2018. Nessa prova, para atender ao propósito de ser um diagnóstico da progressão da competência leitora, são fornecidos aos professores dois cadernos de prova diferentes durante o ano letivo, sendo um para ser aplicado no início e outro no final de cada ano letivo. A aplicação da PB é uma decisão de cada escola ou mesmo de cada professor, atendendo à sua formulação original, caracterizando-a como instrumento para auxiliar os professores e tornando-a expressão de uma política pública que, oficialmente, proclama a necessidade de instituir instrumentos e procedimentos a serviço da democratização do conhecimento.

O principal objetivo da PB é permitir o diagnóstico rápido e mais preciso dos problemas no processo de aprendizagem e possibilitar instrumento padronizado que auxilie em intervenções potencializadoras para que o processo inicial de

alfabetização seja concretizado até o final do terceiro ano do ensino fundamental, conforme expresso nos documentos que tratam da PB, mas que podem ser alterados tendo em vista que a Base Curricular especifica período diferente. Os instrumentos formulados para os testes da PB são elaborados segundo as normas de elaboração de medidas educacionais em larga escala, contando com cuidados que visam proporcionar uma aferição mais precisa das competências leitoras, no entanto, ainda se constata os limites convencionais da apresentação linear de um conjunto de itens, próprios da administração de testes em lápis e papel, denominação para versões impressas.

Assim, a escolha da PB para o desenvolvimento do TAI, além das premissas pedagógicas e político-educacionais destacadas na seção a seguir, se pautou no fato dela possuir todos os elementos necessários para uma avaliação educacional: 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.

Mencionados os aspectos da PB que favoreceram a construção do TAI, outra consideração essencial na pesquisa era de que esse TAI pudesse ser disseminado entre professores dos anos iniciais do ensino fundamental, consistindo em verdadeiro instrumento de enfrentamento do desafio da alfabetização. O envolvimento dos (as) professores(as) alfabetizadores(as) em encontros formativos durante a produção do TAI problematizou os vários aspectos da PB e da avaliação, especialmente porque não se queria difundir a ideia de desenvolver um dispositivo que dispensasse a participação de professoras em sua aplicação e, principalmente, no entendimento dos resultados e seus desdobramentos. Por isso, foi interessante proporcionar o amplo debate sobre a PB, suas características, fundamentos e implicações na etapa denominada de formação em avaliação da

alfabetização com professores, cujos detalhes não serão objeto desse artigo, mas que igualmente constituíram etapa da pesquisa.

O debate sobre as contribuições do TAI da PB em leitura seriam elementos adicionais nos processos de formação dos professores e gestores no âmbito da pesquisa e contribuiria para o processo de incorporação das TIC, com o diferencial de ultrapassar o contexto de desenvolvimento da aula. Concordamos com artigos publicados nesta revista, como o de González Mariño et al (2018) e de Vargas (2018), que defendem a incorporação das TIC no desenvolvimento da aula, contudo, este estudo propõe a incorporação das TIC no processo de avaliação educacional, cuja principal função é a de permitir a identificação dos avanços e dificuldades na aprendizagem, possibilitando intervenção pedagógica satisfatória.

Os pesquisadores envolvidos na pesquisa, concentraram seus esforços para verificar: *Quais são as limitações da versão impressa da PB que o TAI consegue superar, considerando a participação de professores no seu desenvolvimento?*

Considerando a questão proposta, a discussão com professores e gestores sobre a incorporação da TIC no processo de avaliação, certamente agregou nova perspectiva para inclusão das TIC no âmbito educacional.

Na próxima seção salienta-se a importância da alfabetização e da avaliação educacional no contexto do projeto. Na seção 3, destacam-se o desenvolvimento do TAI para a PB em leitura e os desafios da sua implementação nas escolas, identificando os limites tecnológicos e a atuação articulada e interdisciplinar de pesquisadores das áreas de alfabetização, avaliação, psicométrica, desenvolvimento de softwares e programação. Na seção 4, denominada - discussões finais -, apresentam-se parte das análises dos resultados.

## **2. AVALIAÇÃO, ALFABETIZAÇÃO E A FORMAÇÃO DE PROFESSORES NO CONTEXTO DO PROJETO**

Um dos produtos culturais mais notáveis se materializa na escrita e nas práticas de leitura, adensando social e politicamente, portanto, as

tarefas e objetivos da escolarização, particularmente, com vistas à alfabetização.

Para romper com algumas simplificações (Cf. Soares, 2016) no que tange ao seu ensino e para dar o devido peso à questão da leitura e escrita, um projeto pedagógico consequente implica, entre outros elementos, cuidar dos procedimentos avaliativos, considerando suas potencialidades formativas. Se a própria concepção de alfabetização é fator que pode restringir sua consecução (Cf. Kleiman, 1995), não menos verdade é constatar que a maneira como se avalia os domínios da leitura e escrita pode se constituir entrave em sua apropriação, particularmente, para se evitar que a escola se transforme numa agência social seletiva desde os anos iniciais, marcando-se pelo fracasso de seus alunos, precocemente considerados como incapazes de leitura autônoma. Para isso, as atividades escolares devem ser pautadas pelo objetivo de inclusão social quanto aos domínios cognitivos e, em decorrência, a proposta de avaliação deve dar à leitura importância política e cultural, por ser uma competência indispensável para a apropriação de vários conteúdos escolares, ademais de suas implicações sociais.

No entanto, demarcar o objetivo de autonomia leitora das crianças não significa desconsiderar os apontamentos de Street (1993) sobre os riscos de considerar a alfabetização de modo mecânico e acrítico, sem levar em conta o contexto e as expectativas dos alunos. Essa visão reducionista nos levaria ao que se denomina modelo autônomo de alfabetização, com consequências que não poderiam passar despercebidas pelo sistema escolar.

Se o rompimento com a seletividade e as práticas fragmentadas não é um objetivo que se resolva no campo da avaliação, nela aparecem elementos que podem, pelo debate e novos procedimentos, incorporar estratégias para processos pedagógicos inclusivos. Nesse sentido, a produção do dispositivo eletrônico para aplicação de provas adaptadas não significa transferir para uma máquina os desafios de avaliar; ao contrário, estes continuavam a repousar nas ações docentes.

Considerou-se que a construção do TAI para o teste em leitura da PB, em diálogo com professores e gestores de escolas, seria um importante espaço de enfrentamento do desafio, simultaneamente, teórico e prático de que a competência leitora constitui traço

latente, cujas habilidades a serem desenvolvidas são expressas em sua Matriz de Referência para a Língua Portuguesa.

Os documentos que fundamentam a matriz para a competência leitora e escritora explicitam os seguintes eixos:

Eixo 1 – Apropriação do sistema de escrita.

Eixo 2 – Leitura.

Eixo 3 – Compreensão e valorização da cultura escrita. (BRASIL, 2015)

Não obstante, nesses documentos é apontado que a escrita não tem sido objeto da PB por questões técnico-metodológicas, pois, seria difícil garantir padronização em uma coleta de informações que, envolvem cerca de dois milhões de respondentes. Outras importantes considerações apresentadas para a matriz se referem ao desenvolvimento da oralidade, que, apesar de sua importância no trabalho pedagógico, não é avaliada devido às limitações impostas à versão impressa da PB e, quanto às habilidades vinculadas ao Eixo 3, explicita-se no documento oferecido aos professores junto com os instrumentos que as mesmas não se constituem em habilidades separadas dos eixos 1 e 2, permeando a concepção do teste, mesmo que os itens façam referência direta às habilidades dos Eixos 1 e 2 e o teste concebido para o diagnóstico da leitura.

Pode-se ponderar que essas considerações são específicas para a versão em lápis e papel, uma vez que a construção de itens para ambiente informatizado possibilitaria superar, em parte, esses limites, ao permitir a realização e correção automatizada de ditados de frases e palavras e de reconhecimento de voz/som dos respondentes em manifestações da oralidade, desdobramentos possíveis para o desenvolvimento de futuras versões para o TAI da PB.

Com apoio em Nevo (1998), ponderou-se que avaliação externa e interna não são necessariamente excludentes e que deve ocorrer um diálogo entre elas, sendo que a PB poderia ser tomada como uma avaliação externa pelos professores, pelo fato de que sua formulação, elaboração e distribuição são definidas por órgão externo às escolas. Não obstante, no projeto incentivou-se esse diálogo entre a PB e as avaliações desencadeadas pelas

professoras das escolas participantes. É na escola que a avaliação encontra seus grandes desafios e as avaliações externas não podem pretender suplantá-los, como, aliás, destacam Adelson, Dickinson e Cunningham (2016), ao evidenciarem que os grandes fatores para efeito de resultados e para a incorporação de novas metodologias, concentram-se nas salas de aula.

Nesse panorama, objetiva-se com o TAI, especificamente com o uso da tecnologia, eliminar algumas deficiências dos testes lineares em lápis-e-papel ou eletrônicos, superar algumas limitações metodológicas, criando condições mais favoráveis para processos avaliativos em pequenas ou grandes escalas, sustentando-se na definição de avaliação de Lukas Mujika e Santiago Etxebarria (2009, p. 91-92):

A avaliação é o processo de identificação, levantamento e análise de informação relevante de um objeto educacional – que poderá ser quantitativa ou qualitativa –, de forma sistemática, rigorosa, planejada, dirigida, objetiva, crível, fidedigna e válida para emitir juízo de valor baseado em critérios e referências preestabelecidos para determinar o valor e o mérito desse objeto a fim de tomar decisões que ajudem a otimizar esse objeto.

Nessa definição, a avaliação deve ser desencadeada para permitir melhorias no processo de alfabetização das crianças e dada a importância da sistematização do ato avaliativo, salienta-se que devem ser levantadas informações relevantes acerca de um objeto educacional. Isso lança luzes nos elementos que constituem os traços efetivamente vinculados à alfabetização, reforçando as preocupações com atividades de pesquisa que atravessam a avaliação, sem com ela se confundir (Cf. Mathison, 2008). Isso reforçou junto às professoras a necessidade de se estudar e refletir aspectos sobre a alfabetização que concentravam os verdadeiros desafios do TAI.

## **2.1 A Formação dos Professores e Gestores na pesquisa**

Paralelamente ao desenvolvimento do TAI e da prova eletrônica, cujas diferenças são definidas em 3,1, as atividades de formação dos professores focaram, em um primeiro momento, em aspectos

conceituais e técnicos da medida educacional e seu lugar na avaliação da aprendizagem.

Na sequência, estudou-se a temática da avaliação na alfabetização, buscando-se alternativas para superar lacunas nos processos avaliativos da competência leitora, levando-se em conta aspectos que fundamentam as decisões da PB para esse aspecto, como uma das etapas da construção do TAI. Nesse momento da formação, ficaram evidentes dois desafios a serem enfrentados: a compreensão por parte dos professores sobre os pressupostos que embasam as habilidades descritas na matriz de Língua Portuguesa da PB bem como sua relevância para o processo de alfabetização e a necessidade de ampla discussão sobre o currículo praticado no final da educação infantil e nos anos iniciais do ensino fundamental.

Percebeu-se, também, que os manuais e documentos que acompanham a PB, ainda que sejam nacionalmente disseminados, necessitam de apropriação pelos professores.

### **3. O DESENVOLVIMENTO DO TAI PARA A PROVINCIA BRASIL**

Nessa seção, pretende-se argumentar que o TAI da PB pode suplantiar parte dos problemas encontrados em sua versão impressa, também denominada lápis e papel.

O TAI (Cf. Piton-Gonçalves, 2013) é uma maneira diferenciada de avaliar, pois o teste seleciona e apresenta as questões conforme o desempenho (acertos e erros) do respondente. Para Olea, Ponsoda e Prieto (1999), o TAI depende essencialmente, mas não unicamente, de dois elementos: existência de banco abrangente de itens, parametrizados, preferencialmente, por modelos da Teoria da Resposta ao Item (TRI), e dispositivo eletrônico para seleção e apresentação do item. A exigência de um Banco de Itens parametrizados pela TRI, associado ao seu lugar na alfabetização, foi decisivo na escolha do teste de leitura da PB, tendo em vista a disponibilização de itens parametrizados pelo Inep, que nos apoiou no fornecimento de informações essenciais sobre o banco de itens. Com base nas potencialidades e limitações para o emprego de plataformas baseadas em TAI, apresentadas em Alavarse e Melo (2013) e Alavarse e Catalani (2016), procurou-se organizar as etapas de aplicação dos testes, utilizando parte dos itens

existentes na versão impressa. Primeiramente, foi elaborada uma plataforma para apresentação da versão eletrônica dos itens de uma prova em lápis e papel, denominado Teste Baseado em Computador (TBC) da Provinha Brasil, no qual houve a transposição do conteúdo dos itens existentes no papel para uma versão idêntica no *tablet*. Os 20 itens que integraram o TBC da Provinha Brasil correspondiam ao teste 1 de leitura de 2016 e ele foi aplicado para uma parte dos alunos, enquanto a outra realizava, paralelamente, o mesmo teste, mas na versão em lápis e papel. A aplicação paralela tinha por objetivo a verificação da permanência ou não das características psicométricas dos itens. Posteriormente, para a versão TAI do teste, foi incorporada à plataforma o dispositivo/ algoritmo de aferição de proficiência, seleção de itens e critério de parada do teste. Os itens que compuseram o TAI eram diferentes do TBC, para evitar acertos que não estivessem relacionados com a proficiência, ou seja, resultantes da memorização das respostas. Em ambas aplicações houve o diálogo com professores e gestores, na perspectiva de evitar que esse recurso tecnológico fosse visto como algo que substituiria a ação docente.

#### **3.1 O Teste Baseado em Computador (TBC) da PB**

Para atender aos pressupostos psicométricos subjacentes à constituição do TAI da PB, construiu-se primeiramente um TBC do teste impresso para diagnóstico da leitura, que foi aplicado para os alunos do 2º ano do ensino fundamental das 15 escolas participantes, cujo objetivo foi observar se as características psicométricas como índice de dificuldade/facilidade, índice de discriminação e 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 TBC é um teste em meio eletrônico que continua tendo as características lineares do teste em lápis e papel, com uma sequência de itens pré-definida e idêntica para todos os respondentes. Uma das vantagens do TBC é a imediata obtenção de resultados do teste, sendo que duas etapas existentes na aplicação em lápis e papel são eliminadas com o teste informatizado: a) o transporte das respostas para uma folha de respostas, aspecto que adquire especial importância para respondentes de determinadas faixas etárias, por exemplo, crianças e

idosos, ou inexperientes na realização de testes; e b) a digitação ou digitalização das respostas para constituição do banco de dados para análise estatística, pois a interação dos respondentes com o dispositivo eletrônico permite que as respostas passem para o banco de dados de forma automatizada, conferindo rapidez na obtenção dos resultados, especialmente em testes realizados em larga escala.

A segunda vantagem do TBC consiste em permitir a incorporação de ferramentas tecnológicas na elaboração dos itens, diversificando e ampliando as tarefas ou problemas propostos aos respondentes. Os itens construídos para o teste computadorizado podem lançar mão das ferramentas tecnológicas de duas diferentes maneiras, sempre possibilitando a ampliação: a) nos modos de apresentar os contextos e/ou objetos auxiliares/suportes na reflexão proposta pelo item e que mobilizam uma resposta do respondente, podendo integrar multimídias e agregar movimento e som às figuras, gráficos, textos e ilustrações já utilizados nos itens de testes de lápis e papel; e b) nas operações cognitivas solicitadas, suplantando as possibilidades de expressar escolhas, descrições, identificações, comparações, relacionamentos, análise e avaliações em torno dos fatos, fenômenos, ou linguagens, objetos da aferição pretendida. Neste estudo nem todas essas vantagens foram plenamente exploradas, tendo em vista a necessidade de partir do banco de itens já existente da PB em leitura, portanto, de itens construídos para a prova impressa.

Assim, para a etapa do TBC da PB em leitura foram formados, aleatoriamente, dois subgrupos de alunos, um para o teste em lápis e papel e outro para o TBC, tendo sido constatado que não existem diferenças significativas nos parâmetros dos itens quanto a esses dois subgrupos, exceto em determinados descritores que requerem maiores cuidados na apresentação em meio digital.

O software para aplicação do TBC da PB foi criado em plataforma WEB, garantindo que qualquer dispositivo com navegador de Internet pudesse ser utilizado para aplicação do teste. Para garantia da comparabilidade das aplicações no lápis e papel e no TBC, uma das necessidades do sistema era a exibição da totalidade do item na tela do dispositivo. Essa necessidade ocasionou testes com itens em diferentes formatos, para que o examinando sempre fosse capaz de visualizar todos os elementos do

item: o número de ordem, o enunciado, todas as alternativas e o botão para avançar para o próximo item.

A transposição dos itens para a tela do *tablet* foi feita utilizando o material disponibilizado pelo Inep em seu portal *online*. O material era composto de arquivos em formato PDF – *Caderno do aluno* e *Caderno do aplicador* –, dos quais foram retirados os textos para os itens.

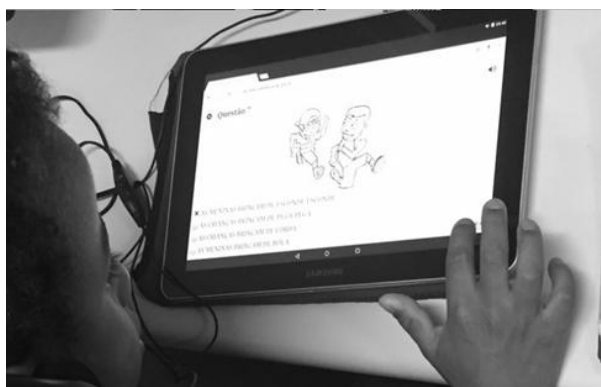
Um dos desafios na construção da versão TBC assentou-se na existência de três tipos diferentes de itens em relação à autonomia de leitura exigida dos (as) alunos (as). O teste em lápis e papel é composto por itens cujos textos são lidos: totalmente pelo professor, parcialmente pelo professor e totalmente pelo aluno.

A leitura feita pelo professor foi substituída pela locução automatizada do texto e foi introduzido um botão de áudio para que o aluno pudesse reproduzir novamente a locução, na mesma quantidade de vezes que o professor poderia repetir a leitura (duas vezes) para a versão impressa da prova. A locução dos textos no TBC da PB usou a Interface de Programação de Aplicativos em inglês *Application Programming Interface* (API) de síntese de voz do Google. Como a leitura feita pelo professor era alternada com a exigência de leitura autônoma do aluno em alguns itens, foram identificadas duas formas de locução informatizada: locução única e dupla locução, sendo que a primeira refere-se aos itens que exigiam uma única locução automatizada, realizada logo que o item era apresentado na tela (p.e. “marque um ‘xis’ no quadradinho que ...”); e a segunda, consistia em uma locução inicial apenas para solicitar que os alunos prestassem atenção (p.e. “leia o texto silenciosamente”) e uma locução secundária, que ocorria após o aluno realizar a leitura do texto-base do item, contendo a enunciação da tarefa cognitiva a ser realizada (p.e. “qual o assunto do texto?”).

Na aplicação do TBC, bem como do TAI da PB, os/as alunos/as não poderiam retornar aos itens já respondidos. Essa decisão foi tomada para aproximação do TBC da PB à dinâmica de aplicação da PB em lápis e papel, na qual todos os alunos são conduzidos pelo professor, que realiza a leitura necessária a cada tipo de item, determinando um tempo médio para as respostas e definindo o avanço de todos para o próximo item.

Na aplicação do TAI, essa decisão se justificou pelo argumento de que o algoritmo do TAI utiliza o histórico de respostas do aluno para escolher os itens futuros e, portanto, voltar no item já respondido e alterar a resposta dificultaria a constituição do algoritmo de seleção do próximo item.

Para o item que requeria uma única locução automatizada (Figura 1), ela era realizada no momento que o texto-base e as alternativas do item eram carregados na tela. O aluno poderia selecionar a repetição dessa locução ou selecionar a alternativa desejada. Somente depois de selecionada a resposta, o botão de próximo item era habilitado.



**Figura 1.** Tela com item que requer uma única locução, sendo respondido por criança do segundo ano de uma escola.

Para o item que requeria dupla locução, a aplicação carregava o texto-base do item na tela e o aluno poderia selecionar a repetição dessa locução inicial ou pressionar o botão para a segunda locução, momento em que as alternativas eram carregadas na tela. Somente depois de selecionada a resposta do item, o botão de próximo item era habilitado. Conforme apontado, o TBC contou com 20 itens, dispostos numa sequência pré-definida para coincidir com a versão lápis e papel da PB em leitura.

As aplicações foram realizadas com a utilização de *tablets* da marca *Samsung Galaxy Tab P7510*, com tela de 10", já existentes nas escolas, e o navegador *Google Chrome*, versão 49. Para garantir que cada aluno/a ouvisse apenas o áudio de sua prova, foram utilizados fones de ouvido durante as aplicações (Figura 2).

A configuração do volume dos dispositivos e dos fones de ouvido foi uma dificuldade adicional e exigiu o auxílio de estagiários no momento da aplicação.



**Figura 2.** Crianças do segundo ano realizando o TBC da PB em uma das escolas.

Para acessar o sistema de aplicação da Provinha, os *tablets* eram conectados na rede sem fio das escolas. A heterogeneidade na quantidade, disposição e limite de acesso dos pontos de acesso de rede sem fio das escolas configurou outro obstáculo, requerendo a distribuição dos alunos de uma turma em mais de uma sala.

O dispositivo previa o encerramento automático do teste após transcorridas 2 horas, caso o teste não fosse encerrado pelo aluno por motivos adversos. No TBC da PB, foram aplicadas 524 provas eletrônicas, com um tempo médio de aplicação de aproximadamente 15 minutos e com 43 segundos para cada item. O tempo total de aplicação por turma costumou ser de 30 a 45 minutos, devido à necessidade de se esperar que todos os alunos de uma turma terminassem a prova para se chamar a turma seguinte, sendo que os primeiros 5 a 10 minutos eram destinados às instruções básicas para inserção do código do aluno e ajuste do volume dos fones de ouvido. Em média, o tempo para aplicação do TBC por turma foi de 25 minutos, enquanto o tempo médio para a aplicação da prova na versão em lápis e papel foi de uma hora.

Observa-se que a construção do TBC, e posteriormente do TAI, apresentou desafios

tecnológicos: por utilizar os dispositivos já existentes nas escolas, por acessar um banco de itens remoto, e pela própria migração de conteúdo de um formato para outro, em especial, as instruções sonoras.

### **3.2 A incorporação do algoritmo para efetivação do TAI da PB**

Um TAI agrega todas as vantagens apontadas para um TBC, mas representa um aprimoramento, pois supera, entre outras características, a linearidade do teste, caracterizada por um conjunto de itens dispostos numa sequência fixa, mesmo quando o respondente pode decidir qual item responder.

O TAI necessita, além de uma plataforma tecnológica, como o TBC, um programa ou algoritmo para administrar os itens.

Para o TAI da PB utilizou-se a mesma plataforma construída para o TBC, agregando um banco com 46 itens e aumentando o número de vezes que o aluno poderia repetir a locução automatizada do texto para 10 vezes.

O algoritmo foi programado para usar as respostas dadas aos itens respondidos a fim de estimar a proficiência e escolher o próximo item com um nível de dificuldade cada vez mais próximo da aferição obtida para o respondente.

O TAI admite que o teste seja diferenciado na quantidade e na complexidade dos itens para cada respondente. Isso é possível porque a seleção dos itens procura aproximar a complexidade do item ao conhecimento do respondente. Mais do que proporcionar desafios possíveis ao respondente, esse procedimento confere maior fidedignidade às estimativas da proficiência. O algoritmo desenvolvido tem como fundamento para aferição da proficiência a Teoria da Resposta ao Item (TRI) (Baker, 2001) e seu objetivo é proporcionar uma dinâmica adaptativa ao teste. O formato concreto do algoritmo é um pacote desenvolvido na linguagem R, em um ambiente (*Linux*, *Windows* ou *Mac*) que tenha instalado o software estatístico R, versão 3.0 ou superior. Foram utilizadas algumas funções provenientes dos pacotes '*catR*' (Magis & Raïche, 2012) e '*irtoys*' (Partchev, 2016), devidamente adaptadas aos objetivos do projeto. O algoritmo deve apresentar três componentes: 1) método de estimação de proficiência; 2) critério para seleção de itens; e 3) critério para finalizar o teste.

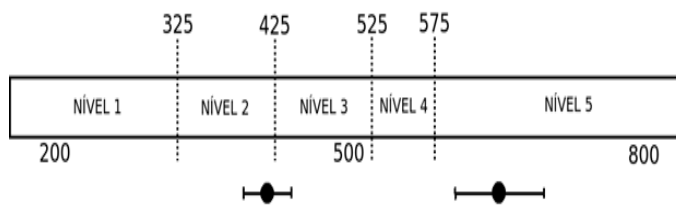
No TAI da PB em leitura, o algoritmo que gerenciou a seleção dos itens foi programado para apresentar um primeiro item, alocado no centro da escala de proficiência. Vale lembrar que foi utilizada a escala já existente para a prova na versão em lápis e papel, uma vez que as características psicométricas dos itens não se alteraram na transposição para o teste eletrônico. Após obter a resposta ao primeiro item, o algoritmo, além de estimar a proficiência, calcular sua precisão e, caso o critério de finalização do teste não seja atingido, seleciona o próximo item. A seleção é feita entre os itens que mais se aproximam da proficiência aferida. Além disso, também foi previsto que o algoritmo garanta a cobertura da matriz de referência, aspecto essencial para o processo de validação na medida educacional e para a consistência pedagógica da estimativa.

A estimação da proficiência foi realizada com base na distribuição esperada a posteriori (EAP) de Bock e Mislevy (1982) com 21 pontos de quadratura. O critério escolhido para seleção de itens foi o da Máxima Informação de Fisher (Barrada, 2010). Conforme já dito, adicionou-se ao critério de Máxima Informação de Fisher, o critério relativo à articulação entre itens e matriz de habilidade, garantindo que os itens respondidos tivessem como condição atender aos eixos 1 e 2 da Matriz de Referência da PB em leitura. Com isso, o algoritmo escolhe os itens mais informativos, equilibradamente, entre os dois eixos de habilidades da matriz, com objetivo de atribuir maior validade de conteúdo ao teste.

Para a finalização do teste, foi utilizado um critério adicional aos aspectos mais utilizados, que são: a) número de itens do teste (mínimo de 8 e máximo de 20 itens); b) limite permitido de incerteza (Erro Padrão máximo, sendo considerado o valor máximo de 35 pontos na estimação da proficiência em uma escala que vai de 0 a 1000. O critério adicional trata do intervalo de confiabilidade na estimação da proficiência e da verificação se ele está contido em um único nível, dentre os cinco níveis de proficiência existentes da PB em leitura.

A Figura 3 ilustra o funcionamento do critério de finalização do teste ao aferir se o intervalo de confiabilidade para a proficiência estimada estava contido em um dos cinco níveis de proficiência da escala da PB em leitura. Esse critério foi usado de forma complementar ao de Erro Padrão máximo.





**Figura 3.** Representação dos intervalos de confiança das proficiências estimadas para os respondentes A e B.

Embora a proficiência do respondente B apresente um intervalo de confiança maior do que a proficiência do respondente A, nota-se que o intervalo de B está inteiramente contido no nível 5. O encerramento do teste para o respondente B ocorreu após responder e acertar apenas cinco itens no teste adaptativo, pois o intervalo já se encontra alocado inteiramente no nível 5, permitindo a interpretação pedagógica e consequente intervenção para esse nível de proficiência em leitura. Para o respondente A, a estimativa ainda não alcançou o objetivo de alocar o intervalo em um único nível de proficiência, ou seja, ainda não está definido se a proficiência estimada pertence ao nível 2 ou 3, mesmo que o número de itens respondidos seja maior (11 itens).

Os dois primeiros critérios são amplamente utilizados em TAI (Barrada, 2012). O terceiro critério, usado para a finalização do teste, por sua vez foi desenvolvido tendo em vista que a PB em leitura visa principalmente informar o professor sobre o nível de proficiência dos alunos, tendo cada nível da escala uma interpretação pedagógica. Nessa perspectiva, o critério de parada que, nos demais TAI, leva em consideração minimizar o erro de medida na estimação da proficiência, estabelecendo um valor máximo para esse erro, no TAI da PB em leitura, diferentemente, priorizou-se a minimização do erro para circunscrever tanto a proficiência como o intervalo de confiança da mesma a um único nível interpretado da escala.

Para aumentar a quantidade de estudantes que participariam da aplicação do TAI da PB em leitura, além dos alunos dos segundos anos, uma parte dos alunos do primeiro ano das 15 escolas também foi chamado para realizar o teste. As aplicações não contemplaram todas as turmas de primeiro ano devido às restrições de datas para aplicação. Era necessário realizar essa aplicação com apenas uma

visita por escola e considerando que a permanência do aluno não pudesse ultrapassar 5 horas, conseguia-se atender um limite máximo de 6 turmas por escola, com tempo médio para aplicação do TAI de 30 minutos para cada turma. Nesse período de tempo, privilegiou-se a aplicação para turmas de segundo ano, tendo em vista constituírem público-alvo desse teste, avaliando as turmas de primeiro no período restante para encerramento do atendimento diário.

O TAI da PB em leitura foi aplicado para 1983 alunos, distribuídos em 78 turmas pertencentes às 15 escolas. A proficiência média para o segundo ano que participou da pesquisa, estimada em 495,28 pontos (numa escala de 0 a 1000 pontos) é próxima ao esperado, dado que foi previamente determinado para os (as) alunos (as) do Brasil a proficiência média de 500 pontos na PB em leitura, conforme relatórios do órgão federal. Por outro lado, o desvio padrão da média para estudantes do segundo ano que participaram da pesquisa foi estimado em 79,04, sendo um valor menor do que o esperado para os estudantes do Brasil, estimado em 100 pontos. Isso pode refletir uma variância menor da população analisada em relação à população original de todo Brasil, para a qual a PB foi delineada. Nota-se, além disso, que a média do primeiro ano foi menor que 500, média da escala de proficiência, validando as metodologias subjacentes à construção do algoritmo, à adequação da plataforma e à elaboração dos itens.

No TAI, verificou-se que os estudantes do segundo ano, fizeram provas pouco maiores (média de 17,75 itens) que os estudantes do primeiro ano (16,74 itens), fato que explica também o tempo maior de prova (11,98 minutos para o segundo ano e 10,88 para o primeiro ano). A redução no tempo médio pode ser devido ao rápido alcance do critério de parada, tendo em vista a existência de nível inferior da escala com intervalos de amplitude maiores, acarretando rápida alocação da proficiência e seu intervalo de confiança nesses níveis.

Comparando o tempo médio do TAI com o do TBC, notou-se ganho médio de 3 minutos, proporcionado pela dinâmica adaptativa do algoritmo. Esta ocorrência pode ser atribuída a diversos fatores, entre eles, que os alunos são expostos a uma quantidade menor de itens no TAI, sendo em média 17,75 itens, uma diminuição de 11% na extensão do

teste em relação ao TBC, para alunos (as) do segundo ano. A diminuição na extensão do teste é fator positivo, uma vez que impede que os alunos errem os itens ao final do teste em virtude do efeito cansaço. As implicações pedagógicas dessa diminuição no tempo total do teste permitem entender que o teste via TAI abrange menor tempo do período de aula, deixando espaço maior para atividades destinadas à aprendizagem.

#### 4. DISCUSSÕES FINAIS

Os resultados preliminares do estudo sugerem que as estimativas das proficiências geradas pelo TAI da Provinha Brasil em leitura se mostraram consistentes relativamente ao arcabouço conceitual da TRI e às metodologias adotadas pelo órgão federal responsável pela elaboração das provas e escalas.

Foi possível destinar momentos de participação de professores alfabetizadores e gestores das 15 escolas na reflexão sobre os aspectos técnicos e metodológicos da medida em avaliação e suas relações com a avaliação diagnóstica proporcionada pela PB em leitura.

O processo de constituição do TBC e do TAI da PB em leitura exigiu encontros periódicos que articulassem a atuação dos pesquisadores das diferentes áreas. Primeiro, na transposição do teste em lápis e papel para o TBC, permitindo atender aos aspectos psicométricos e pedagógicos necessários na elaboração da prova eletrônica (TBC). Segundo, na elaboração do algoritmo, acrescentando etapa ao critério de seleção do item para atender ao processo de validação da matriz de avaliação e alterando etapa no critério de parada para atender ao fato de ser mais relevante pedagogicamente no contexto da intervenção pedagógica para promoção da alfabetização, a maior exatidão na determinação do nível na escala de proficiência no qual se insere tanto a proficiência como seu o intervalo de confiança. Terceiro, o processo desenvolvido para envolver os gestores e professores no estudo, vislumbrando que o incremento de uma plataforma de avaliação não prescinde da atuação dos educadores, posto que sem o debate e envolvimento desses atores, não há incorporação das TIC no processo avaliativo.

Com relação aos resultados dos alunos do segundo ano do ensino fundamental, o TAI reduziu tempo de realização do teste e permitiu que alunos respondessem número menor de itens comparado ao teste em lápis e papel. Também possibilitou teste mais adequado ao desempenho do aluno, visto que o algoritmo selecionou itens, cuja dificuldade se assemelhava à proficiência do respondente, característica que provas frequentemente não conseguem atender e que, por sua vez, garantem menor erro de medida.

Quanto ao erro de medida na estimação da proficiência, cabe destacar 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.

O TAI PB em leitura mostrou seu potencial em suprimir algumas limitações na avaliação diagnóstica da alfabetização e letramento iniciais e possibilitou evidências de que é uma ferramenta de avaliação que pode diagnosticar com maior precisão e abrangência o desempenho em leitura dos alunos da rede. Nesse sentido, de um lado, enseja patamares adequados para implementação de políticas de avaliação para redes de ensino como um todo, no caso de avaliação de sistema, ou para disseminação de processos avaliativos, no caso do uso no interior de escolas e salas de aula, associadas, inclusive, à formação de professores em avaliação da aprendizagem. Por outro lado, evidencia a necessidade de um trabalho articulado e interdisciplinar entre pesquisadores e educadores de diferentes áreas do conhecimento para que as TIC sejam incorporadas também no âmbito da avaliação educacional.

#### 5. REFERÊNCIAS

- [1] J. L. Adelson; E.R. Dickinson; B.C. Cunningham, A multigrade, multiyear statewide examination of reading achievement: examining variability between districts, schools and students. **Educational Researcher**, v. 45, n. 4, p. 258-262, May 2016.
- [2] O. M. Alavarse; E. M. T. Catalani. Alfabetização e TIC: os testes adaptativos informatizados (TAI) como recurso, In: COMITÊ GESTOR DA INTERNET – CGI.br. **Pesquisa sobre o uso das**

- tecnologias de informação e comunicação no Brasil – TIC Educação 2015.** Coord. Alexandre F. Barbosa, São Paulo: Comitê Gestor da Internet no Brasil, 2015, pp. 35-44.
- [3] O. M. Alavarse; W. C. de Melo, Avaliação educacional e testes adaptativos informatizados (TAD): desafios presentes e futuros, In: COMITÊ GESTOR DA INTERNET – CGI.br. **Pesquisa sobre o uso das tecnologias de informação e comunicação no Brasil – TIC Educação 2012**, Coord. Alexandre F. Barbosa, São Paulo: Comitê Gestor da Internet no Brasil, 2013, pp. 103-112.
- [4] F. B. Baker, *The Basics of Item Response Theory* (2nd Ed.), Washington: ERIC Clearinghouse on Assessment and Evaluation, 2001.
- [5] J. R. Barrada, A Method for the Comparison of Item Selection Rules in Computerized Adaptive Testing, 2010.
- [6] J. R. Barrada, Testes adaptativos informatizados: uma perspectiva general, 28, 289–302, 2012.
- [7] R. D. Bock & R. J. Mislevy, Adaptive EAP Estimation of Ability in a Microcomputer Environment, *Applied Psychological Measurement*, 6(4), pp. 431-444, 1982.
- [8] Brasil, Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Nacionais Anísio Teixeira (Inep), **Guia de interpretação de resultados: Provinha Brasil – Leitura, teste 1**, 2015.
- [9] A. Kleiman, Introdução: o que é letramento? In: \_\_\_\_\_ (Org.), *Os significados do letramento: uma nova perspectiva sobre a prática social da escrita*. Campinas, SP: Mercado de Letras, 1995, pp. 13-61.
- [10] J. F. Lukas Mujika; K. Santiago Etxeberria, **Evaluación educativa**. 2. ed. Madrid: Alianza, 2009. pp. 91-92.
- [11] D. Magis; G. Raîche. Random Generation of Response Patterns under Computerized Adaptive Testing with the {R} Package {catR}, *Journal of Statistical Software*, 48(8), 1–31, 2012. Retrieved from <http://www.jstatsoft.org/v48/i08/>
- [12] D. Nevo, 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, pp. 89-97.
- [13] J. Olea; V. Ponsoda; G. Prieto (Ed.). **Tests informatizados: fundamentos y aplicaciones**, Madrid: Pirámide, 1999.
- [14] I. Partchev, (2016). Irtoys: A Collection of Functions Related to Item Response Theory (IRT). Retrieved from <https://cran.r-project.org/package=irtoys>
- [15] J. Piton-Gonçalves, Desafios e perspectivas da implementação computacional de testes adaptativos multidimensionais para avaliações educacionais, 2013, 153 f. Tese (Doutorado)- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2013.
- [16] M. Soares, *Alfabetização: a questão dos métodos*, São Paulo: Contexto, 2016, 384 p.
- [17] B. V. Street, Alfabetización y cultura. *Boletín Proyecto Principal de Educación en América Latina y el Caribe*, Santiago, n. 32, pp. 39-46, dic. 1993.
- [18] S. Mathison. What is the difference between evaluation and research – and why do we care?. In: N. L. Smith; Paul R. Brandon (Ed.), *Fundamentals issues in evaluation*, New York: The Guilford, 2008, pp. 183-196.
- [19] R. Williams, **Cultura**. 2. ed. Tradução de Lólio Lourenço de Oliveira, Rio de Janeiro: Paz e Terra, 2000.
- [20] C. González Mariño; M. de L. Cantú Gallegos; H. E. Camacho Cruz; J. A. Maldonado Mancillas, Prácticas innovadoras de aprendizaje emergentes en el siglo XXI. In: *Revista Ibero-Americana de Sistemas, Cibernética e Informática: RISCO*, 2018.
- [21] J. R. Vargas, Estrategias para la gestión del conocimiento em ambientes mediados, caso de aplicación em la industria petrolera. In: *Revista Ibero-Americana de Sistemas, Cibernética e Informática: RISCO*, 2018.
- [22] O. Munhoz Alavarse et al. Teste Adaptativo Informatizado como recurso Tecnológico para Alfabetização Inicial, In: *Séptima Conferencia Iberoamericana de Complejidad, Informática y Cibernética*, 2017, Orlando. **Memorias** [S.l.: s.n.], 2017. p. 165-169.