

# PESQUISA APLICADA A INSTRUMENTOS DE AVALIAÇÃO EDUCACIONAL:

DESENVOLVIMENTOS E PERSPECTIVAS COM  
BASE NA EXPERIÊNCIA DO CAEd/UFJF





**Organizadoras**

Hilda Aparecida Linhares da Silva  
Liamara Scortegagna

**PESQUISA APLICADA A INSTRUMENTOS  
DE AVALIAÇÃO EDUCACIONAL:**

DESENVOLVIMENTOS E PERSPECTIVAS COM  
BASE NA EXPERIÊNCIA DO CAEd/UFJF





### **ORGANIZADORAS**

Hilda Aparecida Linhares da Silva

Liamara Scortegagna

### **REVISÃO**

Mariângela Innocêncio

### **PROJETO**

Editora da Universidade Federal de Juiz de Fora

Selo Editorial Gestão e Avaliação Educacional

### **CONSELHO EDITORIAL**

Carolina Alves Magaldi (Letras)

Edna Rezende Silveira de Alcântara (Instituto de Artes e Design)

Eduardo Magrone (Educação)

Eliane Medeiros Borges (Educação)

Frederico Braida Rodrigues de Paula (Arquitetura e Urbanismo)

Liamara Scortegagna (Ciência da Computação)

Marcos Tanure Sanabio (Administração e Ciências Contábeis)

Rogéria Campos de Almeida Dutra (Ciências Sociais)

Victor Cláudio Paradela Ferreira (Administração e Ciências Contábeis)

Wagner Silveira Rezende (Direito Privado)

Convidados externos à UFJF: Adolfo Ignacio Calderón (PUC- Campinas) e Fernando M. Reimers (Universidade de Harvard)

Juiz de Fora, 2023

### **Ficha catalográfica elaborada pela Biblioteca Central da UFJF.**

Pesquisa aplicada a instrumento de avaliação educacional: desenvolvimentos e perspectivas com base na experiência do CAEd/UFJF / Organizadores Hilda Aparecida Linhares da Silva, Liamara Scortegagna. – Juiz de Fora, MG: Editora UFJF/ Gestão e Avaliação educacional, 2023.

Dados eletrônicos (1 arquivo: 6,82 mb)

171 p.: il. col.

ISBN: 978-85-93128-64-6

1. Avaliação da educação – Brasil. 2. Tecnologias. I. Silva, Hilda Aparecida Linhares.

II. Scortegagna, Liamara. III. Título.

CDU: 371.26(81)





75

e-VAL: um protótipo para análise e avaliação de níveis de interatividade em itens digitais

*Liamara Scortegagna  
Adriana Rocha Bruno*

105

A WebQuest como proposta de Avaliação Digital no contexto da Aprendizagem Significativa Crítica em Ciências para o Ensino Médio

*Marcello Ferreira  
Danielle Xabregas Pamplona Nogueira  
Olavo Leopoldino da Silva Filho  
Marcos Rogério Martins Costa  
Joaquim José Soares Neto*

143

Uso de imagens nos suportes de Itens de Ciências da Natureza utilizados em Avaliações de Larga Escala

*Ana Carolina Araújo da Silva  
Andréia Francisco Afonso  
Clarice Parreira Senra  
Rita de Cássia Reis  
Wallace Alves Cabral  
Juliana Vicini Florentino Rodrigues*

163

Sobre os Autores

# Apresentação

O Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF) é referência em avaliação educacional em larga escala, formação de gestores da educação pública e desenvolvimento de tecnologias de gestão escolar. Com foco em pesquisa e produção de medidas de desempenho e fatores relacionados à aprendizagem dos estudantes, o CAEd atua há mais de vinte anos em parcerias com o governo federal, as redes municipais e estaduais de ensino e instituições e fundações da área educacional. São muitas as ações e os projetos, inseridos em diferentes contextos, cujo objetivo é garantir a todas as crianças e a todos os jovens o seu direito de aprender.

Neste e-book são apresentadas algumas contribuições das ações de pesquisa do CAEd para o debate sobre a avaliação educacional e suas tecnologias. Com o intuito de fomentar tais ações o CAEd, desde o ano de 2016, tem apoiado e financiado projetos de pesquisa cujo foco são as tecnologias relacionadas à avaliação em larga escala e constituído, ao longo dos anos, um grupo de pesquisadores de diferentes universidades brasileiras dedicados à produção de conhecimentos sobre o tema. Esses conhecimentos têm contribuído para consolidar as parcerias do Centro com os diferentes entes federados e melhor qualificar as atividades de apoio às redes de ensino.

No ano de 2019, o edital de apoio a projetos de pesquisa em instrumentos de avaliação, publicado pelo CAEd, reuniu pesquisas relacionadas às áreas de Ciências Humanas, Ciências da Natureza, Matemática e Língua Portuguesa, além de projetos relacionados à produção de itens digitais interativos e a novas tecnologias para a avaliação da oralidade, mais especificamente, para a avaliação da fluência em leitura. Os materiais que integram este e-book são fruto desses projetos.

Numa perspectiva interinstitucional, os textos que integram o e-book são de autoria de pesquisadores da Universidade de Brasília (UNB), Universidade Federal de São João Del Rei (UFSJ), Universidade Federal de Juiz de Fora (UFJF) e Universidade do Rio de Janeiro (UNIRIO), Universidade do Estado de Minas Gerais (UEMG). Em seu conjunto, apresentam análises ancoradas em amplo material empírico produzido no âmbito das avaliações que o CAEd realiza em nível nacional, estadual e municipal e,

ainda, revisões de literatura sobre a avaliação educacional em larga escala. São artigos que tratam da avaliação educacional e que trazem contribuições para se pensar novas possibilidades para o avanço do conhecimento num campo relativamente pouco explorado pela literatura educacional: as avaliações em larga escala e os instrumentos de que se utilizam. Desse modo, a publicação que ora se apresenta tem o intuito de socializar os conhecimentos produzidos no âmbito das ações de pesquisa do CAEd e de fomentar novas pesquisas no campo da avaliação em larga escala e suas tecnologias.

Uma das frentes das pesquisas desenvolvidas pelo CAEd é aquela que se volta às diferentes áreas de conhecimento com o intuito de estabelecer um diálogo entre os resultados das avaliações em larga escala e os currículos para a educação básica. Nessa frente, o artigo *“A definição dos níveis de complexidade das tarefas de leitura: relações entre currículo e avaliação em larga escala”* discute possíveis contribuições que a definição dos níveis de complexidade de itens que avaliam a leitura pode trazer à reflexão sobre a progressão do currículo de Língua Portuguesa ao longo da educação básica. É também sobre as avaliações de Língua Portuguesa o artigo *“Possibilidades de criação de itens digitais para avaliação da oralidade em Língua Portuguesa usando a plataforma TAO”*, que trata de um eixo das práticas de linguagem ainda pouco abordado pelas avaliações de larga escala: a oralidade. Com foco nas tecnologias que tornam possível a avaliação da fluência em leitura, o artigo *“Tecnologias como suporte para avaliação de fluência em leitura: experiências no âmbito do CAEd/UFJF”* apresenta os avanços que o CAEd vem obtendo na criação de tecnologias que permitam a captura e análise de sons da fala, no âmbito das avaliações de fluência em leitura. Na área de Ciências da Natureza, o artigo *“Uso de imagens nos suportes de itens de Ciências da Natureza utilizados em Avaliações de Larga Escala”* tece considerações sobre os suportes imagéticos dos itens que avaliam os objetos de conhecimento da área e a relação desses suportes com a resolução dos itens pelos estudantes.

Outra frente importante das pesquisas do CAEd se refere à aplicabilidade de tecnologias digitais ao campo da avaliação educacional, que permitam maior interatividade dos estudantes como os testes e assim como a avaliação de habilidades não passíveis de serem abordadas pelas avaliações tradicionais, em papel. Nessa frente, o artigo *“e-Val: um protótipo para análise e avaliação de níveis de interatividade em itens digitais”* apresenta o protótipo de um instrumento cujo objetivo é avaliar o nível de interatividade de itens que integram avaliações, tanto aquelas aplicadas em larga escala quanto as utilizadas pelos docentes em suas salas de aula. Finalmente, na área de Ciências da



Natureza, o artigo “A WebQuest como proposta de avaliação digital no contexto da aprendizagem significativa crítica em ciências para o ensino médio” apresenta uma proposta de avaliação para os componentes curriculares da área a partir de WebQuest, acenando para um modelo de avaliação que dialoga mais proximamente com os interesses dos estudantes do ensino médio e que permite uma maior interação desses estudantes com os objetos de conhecimento.

Como se pode ver, os textos aqui reunidos podem contribuir para fomentar o debate sobre os desafios que a avaliação educacional deve enfrentar em tempos nos quais a escola e seus atores vêm sendo cada vez mais convocados a se conectarem, de forma mais efetiva, aos estudantes, seus interesses e modos de interação com o conhecimento e a informação.

Hilda Aparecida Linhares da Silva (Coordenadora da Pesquisa Aplicada  
a Instrumentos de Avaliação CAEd/UFJF)

Manuel Palácios da Cunha Melo (Coordenador geral do CAEd/UFJF)



# A definição dos níveis de complexidade das tarefas de leitura: relações entre currículo e avaliação em larga escala<sup>1</sup>

*Begma Tavares Barbosa<sup>2</sup>*

*Hilda Aparecida Linhares da Silva<sup>3</sup>*

*Rosângela Veiga Júlio Ferreira<sup>4</sup>*

---

<sup>1</sup> Projeto financiado pela Fundação CAEd/UFJF.

<sup>2</sup> Professora titular aposentada da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: [begmatb@gmail.com](mailto:begmatb@gmail.com)

<sup>3</sup> Professora associada da Faculdade de Educação da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: [hilda.micarello@uab.ufjf.br](mailto:hilda.micarello@uab.ufjf.br)

<sup>4</sup> Professora do Colégio de Aplicação João XXIII, da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: [rosangelaveiga.ferreira@ufjf.edu.br](mailto:rosangelaveiga.ferreira@ufjf.edu.br)



## 1. Introdução

Este texto tem por objetivo apresentar resultados parciais da Pesquisa Aplicada a Instrumentos de Avaliação, área de Língua Portuguesa, desenvolvida no âmbito das atividades do Centro de Políticas Públicas e Avaliação (CAED), da Universidade Federal de Juiz de Fora, Minas Gerais, Brasil. Focaliza, em especial, a proposição de classes de itens que integraram avaliações em larga escala de Língua Portuguesa, de âmbito nacional, estadual e municipal, com base nas teorias do texto e do processamento na leitura e nos parâmetros estatísticos desses itens, gerados a partir da Teoria da Resposta ao Item (TRI), modelo de três parâmetros.

A referida pesquisa insere-se no escopo mais amplo das atividades do CAED as quais englobam ações de apoio às redes públicas de educação básica nos campos da gestão dos sistemas de ensino e formação de seus profissionais. Dentre as atividades fins da instituição, destacam-se os programas de avaliação em larga escala da educação básica nas diferentes áreas de conhecimento por meio da produção e aplicação de instrumentos de avaliação e da análise e divulgação dos resultados dessas avaliações. Os projetos envolvidos na Pesquisa Aplicada a Instrumentos de Avaliação visam à produção de conhecimentos que, a um só tempo, subsidiem as ações do CAED, em especial aquelas relacionadas à composição e à manutenção do Banco do Itens, e permitam o desenvolvimento de novas tecnologias de avaliação, divulgação e apropriação de resultados pelas redes de ensino. A pesquisa se desenvolve, portanto, numa interseção entre o currículo para a educação básica, sua aplicação pelos sistemas de ensino e apropriação pelos estudantes, e a avaliação, como um instrumento de gestão do currículo.

Face às características e objetivos que movem as ações da Pesquisa Aplicada a Instrumentos de Avaliação, os resultados por ela produzidos estão ancorados, por um lado, nas teorias do conhecimento próprias às diferentes áreas e seus componentes curriculares e, por outro, nas teorias em que se fundamentam os testes cognitivos, com base em modelos psicométricos.

No caso específico da pesquisa na área de Língua Portuguesa, no recorte apresentado neste artigo, buscamos responder às seguintes questões: o que torna uma tarefa de leitura, cuja realização demanda a mobilização de determinadas habilidades de leitura, mais simples ou mais complexa? O que os itens que avaliam uma mesma habilidade de leitura podem revelar sobre o modo como



essa habilidade progride, do ponto de vista de seu desenvolvimento pelos estudantes, ao longo de diferentes etapas da educação básica? Nosso entendimento é o de que a resposta a essas duas questões pode, a um só tempo, permitir a construção de testes mais ajustados à etapa de escolarização à qual são aplicados e orientar gestores e professores em relação à progressão do ensino da leitura ao longo da educação básica. Com essas duas frentes de análise, complementares e simultâneas, a pesquisa, pautada em dados da avaliação em larga escala, parte do currículo e a ele retorna, num círculo virtuoso.

Com o intuito de apresentar o percurso empreendido na direção de construir respostas às duas questões anteriormente enunciadas, discutimos, na primeira parte deste artigo, questões relativas à avaliação em larga escala, com foco na noção de “classes de itens”, em que se ancora a pesquisa. Na segunda parte, apresentamos a concepção de leitura que orienta a pesquisa, abordando noções relevantes como as de texto, leitor, leitura e ensino de leitura. Na terceira parte, expomos os principais resultados da pesquisa.

## 2. A noção de classe de itens e sua relação com a validade de conteúdo em testes cognitivos

A noção de classe de itens vem sendo utilizada pelo CAED como um conceito operacional do banco de itens. Tal noção passou a integrar o projeto de construção do repositório de itens do Centro, no ano de 2011, quando teve início um primeiro esforço de elaboração de matrizes de referência para avaliação das diferentes áreas de conhecimento que pudessem subsidiar a organização do referido repositório.

A classe constitui um meio de operacionalizar a avaliação da manifestação de uma habilidade nos diferentes níveis de complexidade em que ela possa se apresentar. Parte-se do princípio de que uma mesma habilidade, descrita nas matrizes de referência para a avaliação, pode apresentar diferentes níveis de complexidade, que não são contemplados na descrição que se faz dessa habilidade. Tome-se como exemplo a habilidade de “localizar informações explícitas em textos”, presente nas matrizes de avaliação do SAEB do 2º, 5º e 9º anos do ensino fundamental e 3º ano do ensino médio. É razoável pensar que essa habilidade se manifeste em níveis de complexidade distintos, a depender



da etapa de escolarização avaliada, pois há uma progressão em seu desenvolvimento no processo de formação do leitor, o que justifica avaliá-la em diferentes etapas de escolarização. Essa progressão é que permite que o ensino promova a formação de um leitor menos experiente para que se torne mais experiente ou mais proficiente.

A definição de classes de itens é, portanto, um meio para explicitar esses diferentes níveis de complexidade, de modo que a avaliação da habilidade seja capaz de apreender diferenças de sua manifestação, o que justifica que haja diferentes pontos de ancoragem dos itens que avaliam tal habilidade na escala de proficiência por meio da qual os resultados da avaliação são apresentados<sup>5</sup>. Nesse sentido, a definição de uma classe de itens se faz na interseção entre: (i) critérios oriundos de uma teoria do conhecimento que contemple a natureza do construto que se está avaliando; (ii) evidências sobre como os estudantes se apropriam desse conhecimento ao longo de seu processo de aprendizagem, obtidas por meio da análise do desempenho dos estudantes nos testes; (iii) uma teoria da avaliação que oriente o modo como se pode aferir essa aprendizagem em diferentes momentos do processo de escolarização. Uma vez definidas as classes de itens que avaliam uma mesma habilidade, é possível acessar aqueles mais adequados às diferentes etapas de escolarização avaliadas. A noção de classe de itens é, portanto, um recurso para assegurar a validade de um teste e está fortemente ancorada nesse conceito, sobre o qual discorreremos brevemente a seguir.

O conceito de validade vem sendo amplamente discutido no campo da psicometria, com diferentes definições. Pasquali (2007), buscando posicionar o campo frente a essas diferentes definições e inspirado nas pesquisas dos autores holandeses Borsboom, Mellenbergh e van Heerden (BORSBOOM; COLS., 2003; BORSBOOYM; COLS, 2004), afirma que a validade

[...] constitui uma propriedade do instrumento de medida e que não tem nada a ver com a idéia de que ela consistiria num julgamento sobre o significado dos escores de um teste, retomando, assim, o sentido original que Kelly (1927) quis dar a esse conceito, ao afirmar que um teste é válido quando mede aquilo que supostamente deve medir (PASQUALI, 2007, p. 105).

---

<sup>5</sup> De acordo com Klein (2014, s. p., grifos da autora), “*Escala de proficiência* é um conjunto de números ordenados, obtido pela Teoria de Resposta ao Item (TRI) que mede a proficiência (habilidade) em uma determinada área de conhecimento. A probabilidade de se acertar um item aumenta à medida que a proficiência (habilidade) aumenta”.



Em psicometria, o que um teste se propõe a medir é chamado construto, sendo, portanto, um referencial para apoiar os resultados de uma avaliação.

No caso das avaliações em larga escala de Língua Portuguesa, impõe-se um desafio para a definição do construto leitura, de forma a permitir o estabelecimento de critérios e meios para medi-lo para que se possa ter suficiente clareza do que está sendo avaliado. Nesse sentido, faz-se necessário eleger uma teoria explicativa da natureza desse construto, uma vez que não existe uma unanimidade em sua definição. No que se refere à pesquisa aplicada a instrumentos de avaliação do CAED, em curso desde o ano de 2016, as teorias do texto, sobretudo as teorias de base cognitivista (SMITH, 1991; KLEIMAN, 1989; KOCH, 1998) e a linguística do texto (KOCH, 1989; KOCH; TRAVAGLIA, 1989), vêm dando sustentação à definição e abordagem do construto leitura, subsidiando os encaminhamentos metodológicos e analíticos para a abordagem dos dados empíricos que consistem em um amplo conjunto de itens já aplicados em avaliações sistêmicas de Língua Portuguesa, do 1º ao 12º anos da educação básica.

Para discorrer sobre validade dos testes, apoiamo-nos, também, nos *Standards for Educational and Psychological Testing* (JESUS; REGO; SOUZA, 2018), que apresentam cinco fontes de evidências da validade de um teste, a saber: “1) evidências baseadas no conteúdo do teste; 2) evidências baseadas no processo de resposta; 3) evidências baseadas na estrutura interna; 4) evidências baseadas na relação com outras variáveis; e 5) evidências baseadas nas consequências da testagem” (AERA; APA; NCME, 2014 apud JESUS; REGO; SOUZA, 2018, p. 8).

O processo de definição de classes de itens atende ao objetivo de assegurar a validade dos testes com base em evidências relacionadas ao seu conteúdo e à sua estrutura interna. Ainda segundo os autores anteriormente referidos, os procedimentos para a produção de evidências de validade baseadas no conteúdo são aqueles que dizem respeito ao levantamento de dados “sobre a representatividade da matriz e dos itens do teste, investigando se esses consistem em amostras abrangentes do domínio que se pretende avaliar” (AERA; APA; NCME, 2014 apud JESUS; REGO; SOUZA, 2018, p. 8). Já os procedimentos para produção de evidências baseadas na estrutura interna consistem no levantamento de “dados sobre a representação do construto, com base nas dimensões avaliadas, na qualidade dos itens e na confirmação de hipóteses derivadas da teoria” (HUTZ, 2009, p. 251, apud JESUS; REGO; SOUZA, 2018, p. 8).



Na pesquisa aqui apresentada, o processo de definição de classes de itens tem início numa descrição, apoiada nas teorias do texto e da leitura de base cognitivista e interacional, na linguística textual, nos estudos de semântica, da natureza das habilidades avaliadas. Tal descrição fundamenta a análise dos elementos dos textos que dão suporte aos itens e o modo como se articulam na manifestação da habilidade. Essa é, portanto, uma estratégia que visa produzir evidências sobre a validade do teste, no sentido de que os itens efetivamente avaliem o que se propõem a avaliar.

A segunda etapa do processo consiste numa incursão à empiria, neste caso, a um conjunto significativo de itens já aplicados em avaliações anteriores e que avaliam uma habilidade específica da matriz de referência. Tal incursão envolve uma análise estatística dos itens com base em seus parâmetros, gerados a partir da TRI. São observados o parâmetro A (que indica se o item discrimina os estudantes com maior proficiência daqueles com menor proficiência em leitura); o parâmetro B (que expressa o nível de dificuldade do item) e o parâmetro C (que expressa a probabilidade de acerto ao acaso)<sup>6</sup>. Essa análise preliminar objetiva excluir da amostra itens que não apresentem qualidade, do ponto de vista estatístico. São, portanto, procedimentos que visam à produção de evidências sobre a validade dos instrumentos baseadas na estrutura interna do teste.

Uma vez selecionados os itens com bons parâmetros, procede-se a uma análise pedagógica desses itens. Nessa etapa, o ponto de ancoragem dos itens na escala de proficiência<sup>7</sup> é considerado para que se observem elementos comuns a um conjunto de itens que se encontram em um mesmo intervalo da escala. Esses elementos – nível de complexidade do texto, gênero a que pertence, forma de construção do comando do item, sinalizações que o texto oferece para a resolução da tarefa de leitura proposta - são a base para a definição de uma mesma classe de itens. No caso dos itens de Língua Portuguesa, a explicitação de uma classe se faz com base numa articulação entre: (i) a habilidade de leitura avaliada (descriptor da matriz de referência); (ii) o nível de complexidade do texto que dá suporte ao item e por meio do qual a habilidade é avaliada; (iii) a natureza das

---

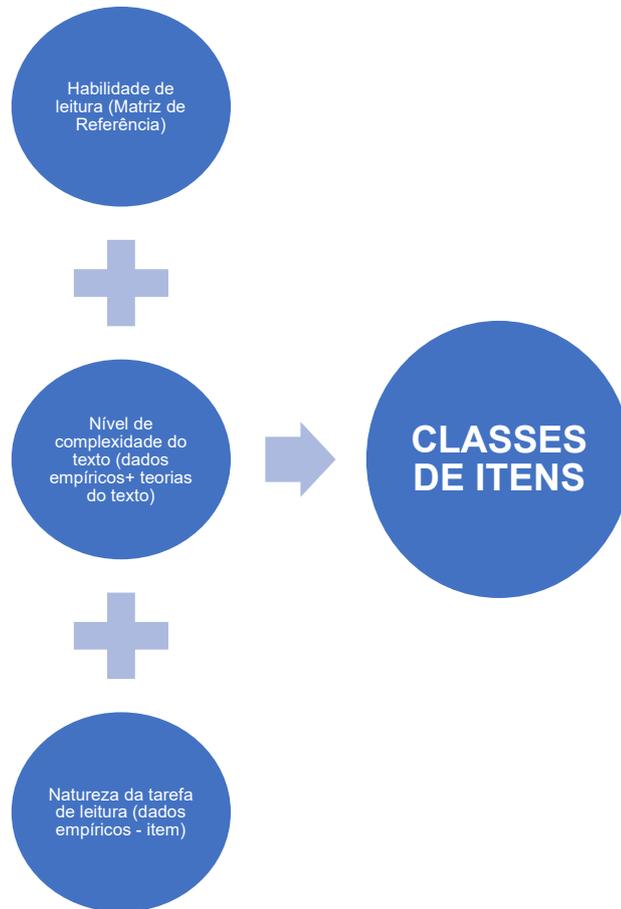
<sup>6</sup> Para o processo de validação dos itens para análise são eliminados: itens com parâmetro A menor que 0,01, visto que esses itens apresentam fraco poder de discriminação entre estudantes que desenvolveram as habilidades avaliadas e aqueles que não as desenvolveram e, ainda, itens com parâmetro C superior a 0,025, por serem itens que apresentam forte probabilidade de acerto ao acaso.

<sup>7</sup> A escala de proficiência coloca numa mesma métrica os itens de um teste e a população que realizou o teste.



sinalizações que o texto oferece para a realização da tarefa de leitura proposta pelo item. O esquema explicativo da definição de classes de itens no âmbito da pesquisa em avaliação pode ser melhor visualizado na figura 1.

**Figura 1. Esquema originário de uma classe de itens**



Fonte: Elaborada pelas autoras.

Como se pode observar pelo exposto, a noção de classe de itens tem uma estreita relação com o conceito de validade de conteúdo em testes cognitivos. Por essa razão, torna-se uma noção relevante para apoiar a constituição de um Banco de Itens para o CAED, assim como para apoiar uma reflexão mais fundamentada por dados empíricos sobre a progressão de habilidades de leitura.



### 3. Currículo, ensino e avaliação em larga escala de Língua Portuguesa: diálogos e articulações

Nesta seção, apresentamos a concepção de leitura que fundamenta a pesquisa em avaliação na área de Língua Portuguesa e que é também aquela que vem sendo expressa nos documentos curriculares nacionais desde fins da década de 1990, com a publicação dos Parâmetros Curriculares Nacionais de Língua Portuguesa (BRASIL, 1998).

Para tanto, recuperamos, brevemente, algumas contribuições teóricas que acreditamos já serem conhecidas por muitos professores de Língua Portuguesa, oriundas, sobretudo, de estudos acadêmicos sobre a linguagem e a leitura, os quais embasam vários documentos oficiais que servem como parâmetros para a organização de nossos programas de ensino de Língua Portuguesa, como os Parâmetros Curriculares Nacionais para o Ensino Fundamental e Médio (BRASIL, 1998) e, mais recentemente, a própria Base Nacional Comum Curricular (BRASIL, 2017).

A concepção que aqui apresentamos assenta-se na premissa de que a leitura é uma construção subjetiva de significados (SMITH, 1991; KLEIMAN, 1989; KOCH, 1998), ou seja, os leitores atuam sobre o texto a partir de um vasto conjunto de conhecimentos acumulados e estruturados em sua mente. Diante de um texto, acionam seus conhecimentos de mundo e seus conhecimentos linguísticos, que podem ser formalizados – como os conhecimentos sobre os textos e sobre a língua (sua gramática, seu léxico) aprendidos na escola – ou não formalizados, como aqueles adquiridos pela vivência em determinada cultura.

Dessa concepção de leitura, entendida como construção subjetiva de sentidos, decorre uma compreensão do texto não como objeto que porta um sentido, mas como objeto que guia o sentido. Assim, o “significado” não estaria no texto, mas o texto, como um mapa (FAUCONNIER, 1997), oferece um conjunto de sinalizações, de pistas que guiam o leitor na tarefa de construção de sentidos que é a leitura. A utilização da palavra “sentidos”, no plural, implica assumir que um mesmo texto pode suscitar diferentes leituras. Um mesmo leitor, por exemplo, pode, em diferentes momentos de sua vida, construir leituras distintas de um mesmo texto. Diferentes leitores, em função de seu arcabouço de conhecimento ou mesmo de seus objetivos de leitura, podem compreender um mesmo texto a partir de diferentes perspectivas, construindo, para ele, diferentes leituras.



Por outro lado, a concepção da leitura como “atividade subjetiva de construção de sentido” não implica assumir que qualquer leitura produzida a partir da interação com um determinado texto seja “autorizável” (POSSENTI, 1999). E aqui temos um dado fundamental para a discussão do ensino e da avaliação da leitura: há certamente leituras que os textos não permitem fazer, ou seja, aquelas que não encontram fundamentação suficiente nas pistas textuais. Um rico exercício escolar de leitura, a propósito, é o de solicitar que os estudantes fundamentem as leituras feitas. Em alguns casos, o resultado desse exercício será o descarte de hipóteses iniciais de leitura. Em outros casos, o professor poderá perceber uma possibilidade de leitura ainda não cogitada ou mesmo uma contribuição coerente com o processo de formação do estudante.

Compreender o que fazem os leitores proficientes quando leem é de fundamental importância para pensar o ensino e a avaliação da leitura. Segundo Kleiman (1993, p. 49), a tentativa de ensinar a ler não seria incoerente com a natureza subjetiva da leitura, “se o ensino da leitura for entendido como o ensino de estratégias de leitura, por uma parte, e como o desenvolvimento das habilidades linguísticas que são características do bom leitor, por outra”.

Leitores proficientes atuam de forma consciente sobre os textos, ou seja, utilizam estratégias “metacognitivas” para atribuir sentido ao que leem. As estratégias metacognitivas de abordagem do texto são utilizadas pelos leitores como forma de automonitorar sua compreensão, buscando eficiência na leitura. É o que acontece, por exemplo, quando alguém, antes de iniciar a leitura de um texto, investiga seu contexto de produção (quem escreve?, quando?, em que suporte?, com que objetivo? etc.); quando avalia um título e constrói hipóteses de leitura; quando considera marcadores textuais relevantes para antecipar informações, dentre outras possibilidades. Boas práticas de formação de leitores devem promover o exercício dessas estratégias, que, no âmbito da pesquisa sobre a qual discorreremos neste texto, são também consideradas no processo de definição das classes de itens que integram as avaliações em larga escala.

Outro tipo de estratégia de leitura são as estratégias cognitivas, mais diretamente relacionadas ao que tratamos como habilidades leitoras. Segundo Kleiman (1993), as estratégias cognitivas de leitura seriam aquelas “operações inconscientes” do leitor. A autora argumenta que o conhecimento utilizado para realizar essas operações é também um conhecimento implícito e envolve o domínio que os falantes de uma língua têm de suas regras, sua sintaxe, seu léxico. Para Kleiman (1993, p. 50):



Dentro dessa visão do processo de leitura, isto é, como um conjunto de estratégias cognitivas e metacognitivas de abordagem do texto, o ensino estratégico de leitura consistiria, por um lado, na modelagem de estratégias metacognitivas, e, por outro, no desenvolvimento de habilidades verbais subjacentes aos automatismos das estratégias cognitivas. Este último tipo de instrução seria realizado através de análise textual característica da desautomatização do processo.

É nesse sentido que acreditamos ser possível o desenvolvimento de habilidades de leitura na escola, já que estamos aqui tratando de uma competência cognitiva. Um modelo de ensino comprometido com esse objetivo envolve a ampliação das capacidades linguísticas dos estudantes, de modo que desautomatizem os processos subjacentes às operações cognitivas. Da mesma maneira, a avaliação da leitura deve considerar esses processos.

Tomemos, para exemplificar essa proposta, duas das habilidades presentes em todas as Matrizes de Referência de Avaliação de Língua Portuguesa: a habilidade de inferir o sentido de uma palavra e a de reconhecer o assunto de um texto.

No primeiro caso, para apreender o sentido de uma palavra desconhecida, os leitores costumam acionar seus conhecimentos lexicais, muitas vezes não formalizados, oriundos da vivência na sua língua; acionam, de sua gramática interna, regras morfológicas; reconhecem relações entre termos cognatos, por exemplo. Esse conjunto de conhecimentos implícitos contribui no processo de produção de inferência lexical.

Um ensino de linguagem comprometido com a ampliação desses conhecimentos e das capacidades linguísticas dos estudantes, dentro das práticas de leitura ou com vistas a elas, contribui para o desenvolvimento da habilidade de inferência lexical. Outra forma de contribuir com o desenvolvimento dessa habilidade é chamar a atenção do aprendiz para as sinalizações textuais relevantes que oferecem apoio à produção de inferência lexical, como sinônimos, antônimos, palavras do mesmo campo semântico, exemplos, definições etc. Quando o professor “ensina” a explorar o contexto linguístico em que uma palavra desconhecida aparece, ele está “ensinando a ler”, ou seja, desenvolvendo uma habilidade de leitura das mais fundamentais: a de inferir o sentido de palavras. Na construção das situações-problema que estruturam os itens que avaliam habilidades de leitura, tais sinalizações são parte dessas situações.



A habilidade de “reconhecer o assunto de um texto”, por exemplo, também aciona conhecimentos implícitos sobre a língua, particularmente conhecimentos sobre os textos, sobre como se estruturam. Os leitores costumam considerar, em primeiro lugar, o título do texto, bem como outras sinalizações mais evidentes: subtítulos, ilustrações, lides etc. Quanto mais proficientes forem, mais elementos textuais os leitores vão considerar, como, por exemplo, a organização tópica do texto, as recorrências lexicais etc. O desenvolvimento da habilidade de reconhecer o assunto de um texto envolveria, portanto, o domínio de um conjunto particular de conhecimento ou de capacidades linguísticas. Do mesmo modo, itens que avaliam essa habilidade o fazem, tendo como suporte textos nos quais o assunto está indicado no título, apresentado como tópico do primeiro parágrafo, ou que deve ser inferido a partir de pistas textuais oferecidas ao longo de todo o texto. Tais diferenciações implicam diferentes níveis de complexidade da tarefa de reconhecer o assunto de um texto e, portanto, diferentes classes de itens que podem avaliar essa habilidade.

Como se pode concluir pelo exposto até aqui, a compreensão da natureza da leitura e de como agem leitores proficientes quando leem é fundamental para as práticas de ensino e, conseqüentemente, de avaliação da aprendizagem. A construção de boas práticas de ensino envolve exercitar, com os leitores em formação, os procedimentos e as estratégias utilizados pelos leitores proficientes, de modo que aqueles passem a agir de forma mais consciente na interação com os textos. Nesse percurso, avaliar a leitura envolve considerar se as leituras produzidas pelos estudantes são leituras fundamentadas pelas sinalizações textuais e em que medida eles se apropriam de estratégias eficientes para abordar o texto. Com base na premissa das classes de itens apoiadas nesses elementos, considerando diferentes níveis de complexidade que caracterizam a progressão em leitura, a divulgação de resultados pode passar a ter uma dimensão de retorno ao currículo cada vez mais precisa. O que os estudantes leem? Que tarefas de leitura realizam? Em textos de que nível? O que determina a gradação da complexidade da leitura de um texto? A procura por possíveis respostas a questões dessa natureza contribui para a discussão em torno do lugar da avaliação em larga escala no processo de implementação de políticas públicas de formação do leitor proficiente estabelecendo um círculo virtuoso entre currículo – avaliação – currículo.

Partindo desse círculo, discutimos de que modo as concepções de leitura, texto e leitor, apresentadas anteriormente, aplicam-se à construção de instrumentos de avaliação em larga escala. Nessa discussão, primeiramente, é importante assinalar o quão difícil é mensurar uma competência



cognitiva. Essa medida é, no entanto, necessária e fundamental no processo de formação de leitores proficientes. As interações de sala de aula são um espaço privilegiado para a avaliação da leitura. Além do acesso às leituras produzidas pelos estudantes, o professor pode conhecer muito sobre como os leitores agem sobre os textos, sobre que estratégias utilizam na leitura de textos de diferentes gêneros, sobre que sinalizações textuais consideram na atividade interpretativa. Essas são informações relevantes para orientar as práticas de ensino da leitura.

No caso das avaliações em larga escala, a aferição da competência leitora ocorre, em geral, por meio da aplicação de testes de múltipla escolha, através dos quais podem ser reunidas informações sobre que tipo de tarefa de leitura o estudante é capaz de resolver, mas não – pelo menos até o presente momento – sobre como atuam sobre os textos para chegarem a uma resposta. O principal instrumento da avaliação em larga escala é, portanto, o conjunto das tarefas de leitura relativas a diferentes habilidades previstas em uma Matriz de Referência, materializadas nos itens que compõem os testes. A informação sobre o domínio de determinada habilidade de leitura pelo estudante advém da sua condição de resolver, com êxito, tarefas relativas a essa habilidade, o que é indicativo de que a habilidade foi desenvolvida. Da obtenção de informações dessa natureza é que se afere sua proficiência em leitura.

Portanto, para que cumpra seus objetivos, à avaliação em larga escala cumpre descrever, de forma fundamentada, cada uma das habilidades de leitura previstas na Matriz de Referência, visando à clareza quanto ao construto avaliado, e construir, para cada uma dessas habilidades, um conjunto de tarefas de leitura com diferentes níveis de complexidade, de modo que se possam produzir medidas representativas da progressão do desempenho dos estudantes em leitura, baseadas no tipo de tarefa que eles conseguem resolver. A definição dos diferentes níveis de complexidade das tarefas de leitura deriva, portanto, da proposição das classes de itens para cada uma das habilidades. Desse modo, uma pergunta importante é: que aspectos determinam o grau de complexidade de uma tarefa de leitura?

A concepção de leitura discutida anteriormente, que concebe o leitor como sujeito atuante na interação com o texto, ajuda a responder à questão. De um lado, tem-se o texto, “mapa” que guia a construção de sentidos, “prevendo” certos procedimentos interpretativos; de outro lado, tem-se o leitor, que seleciona pistas textuais relevantes para construir sentido na leitura. Assim, tarefas



de leitura menos complexas seriam aquelas para cuja resolução o leitor pudesse contar com um conjunto de sinalizações textuais mais evidentes e pontualmente localizadas. Tarefas mais complexas seriam aquelas cuja resolução demandasse o estabelecimento de relações entre várias sinalizações textuais, menos evidentes e mais dispersas pelo texto, exigindo conexões várias e, por vezes, uma apreensão global do sentido do texto. Outro aspecto determinante do nível de complexidade de uma tarefa de leitura é, sem dúvida, o nível de complexidade do próprio texto<sup>8</sup>.

Diante do exposto, com base nos desdobramentos da pesquisa aqui apresentada, apontamos o desafio de que as avaliações em larga escala em Língua Portuguesa considerem três tarefas fundamentais com relação à produção de instrumentos que avaliam o construto leitura: i) a descrição fundamentada da natureza das habilidades de leitura que integram as matrizes de referência; ii) a identificação dos aspectos definidores do nível de complexidade das tarefas de leitura relacionadas a cada uma dessas habilidades; iii) a identificação de aspectos definidores do nível de complexidade dos textos que dão suporte à avaliação dessas tarefas de leitura.

## 4. A pesquisa em avaliação: principais resultados

À guisa de conclusão do presente artigo, apresentamos, nesta seção, uma síntese dos principais resultados da pesquisa em avaliação em Língua Portuguesa, no período entre os anos de 2016 e 2020. São eles: a construção de critérios para a definição do nível de complexidade de tarefas relacionadas a diferentes habilidades de leitura; ii) o estabelecimento de critérios definidores do nível de complexidade dos textos.

Para a construção dos critérios que definem o nível de complexidade de tarefas relacionadas a diferentes habilidades de leitura, foram analisados cerca de 1300 itens, representativos de diferentes habilidades que integram a matriz de referência para a avaliação da leitura. Esses itens foram aplicados em avaliações em larga escala de sistemas de ensino nacional, estaduais e/ou municipais, portanto, todos os itens eram parametrizados e estavam distribuídos, com base em seus

---

<sup>8</sup> Sobre esse tema, desenvolvemos uma discussão no artigo “Avaliações em larga escala de Língua Portuguesa: uma pesquisa sobre a complexidade dos textos que dão suporte a itens que avaliam a leitura” (MICARELLO; BARBOSA; FERREIRA, 2020).



parâmetros, pela escala de proficiência do SAEB. A distribuição dos itens pela escala de proficiência permitiu considerar seu ponto de ancoragem, numa ordem crescente para compreender quais eram as tarefas de leitura mais complexas ou menos complexas, considerando-se cada uma das habilidades analisadas. Essa análise, como já dito em seção anterior deste artigo, foi precedida por uma descrição fundamentada de cada habilidade avaliada.

Os resultados de uma análise preliminar, com base nos aspectos comuns a itens que ancoram num mesmo intervalo da escala de proficiência, aponta para alguns aspectos definidores do nível de complexidade das tarefas de leitura: i) a presença, no texto que dá suporte ao item, de sinalizações de apoio à resolução da tarefa proposta; ii) o tipo de sinalização, face à natureza da habilidade avaliada; iii) a posição dessas sinalizações no texto.

A título de ilustração, tomemos como exemplo a definição de classes de itens para o descritor “inferir o sentido de uma palavra ou expressão”, uma das habilidades mais relevantes à construção de sentidos na leitura. Tarefas menos complexas de leitura caracterizam-se pela presença de sinalizações de apoio “co-textuais” (ORLANDI, 1983, p. 193), ou seja, localizadas no texto e próximas à palavra cujo sentido deve ser inferido. Além desses dois critérios, também a natureza da sinalização é aspecto relevante: a sinonímia é um tipo de sinalização que se destaca como facilitadora da tarefa em relação à presença de outro tipo de sinalização, como a antonímia ou um conjunto de palavras do mesmo campo semântico (DELL’ISOLA, 2005). Considera-se, ainda, a presença de marcação tópica que facilita ao leitor a tarefa de reconhecer o assunto de um texto; um enunciado parafrástico que permite acessar uma informação, dentre outras possibilidades. Para a definição de classes de itens para o descritor em tela, observa-se também a ausência dessas sinalizações, que demanda, na maioria das vezes, a apreensão global do sentido do texto<sup>9</sup>. Assim, com base nos critérios enunciados e na análise do tipo de item que avalia o descritor e sua distribuição pela escala de proficiência, foi possível chegar às classes de itens apresentadas no quadro 1.

---

<sup>9</sup> No caso do descritor “inferir o sentido de uma palavra ou expressão”, também o tipo de palavra cujo sentido deve ser inferido (se própria do uso coloquial da língua ou de seu uso formal) também interfere no grau de complexidade da tarefa. Esse aspecto, no entanto, está associado ao nível de complexidade dos textos, critério de complexidade a ser tratado adiante.



**Quadro 1. Classes de itens que avaliam a habilidade de inferir o sentido de uma palavra ou expressão**

Intervalo da escala de Proficiência	Descrição da classe de itens
Abaixo de 140 pontos	Não foram observados itens para este intervalo.
Entre 140 e 165 pontos	C01D08 – Inferir o sentido de léxico próprio da <b>linguagem coloquial</b> (palavra, expressão idiomática e outras expressões metafóricas) em textos de nível 1, com apoio em pistas co-textuais (sinonímia; palavras do mesmo campo semântico; dedução a partir de enunciado).
Entre 165 e 180 pontos	C02D08- Inferir o sentido de léxico próprio da <b>linguagem coloquial</b> (palavra, expressão idiomática e outras expressões metafóricas) em textos de nível 2, com apoio em pistas co-textuais (sinonímia; palavras do mesmo campo semântico; dedução a partir de enunciado).
Entre 180 e 200 pontos	C03D08- Inferir o sentido de léxico próprio da linguagem coloquial (palavra, expressão idiomática e outras expressões metafóricas) em textos de nível 2, com apoio em pistas co-textuais (sinonímia; palavras do mesmo campo semântico; dedução a partir de enunciado).
Entre 200 e 220 pontos	C04D08 – Inferir o sentido de léxico próprio da linguagem coloquial (palavra, expressão idiomática e outras expressões metafóricas) em textos de nível 2, sem apoio em pistas co-textuais.
Entre 220 e 240 pontos	C05D08 – Inferir o sentido de palavras ou expressões não coloquiais e de metáforas em textos de nível 1 e 2, com apoio de pistas co-textuais.
Entre 240 e 260 pontos	Não foram observados itens para este intervalo.
Entre 260 e 280 pontos	C06D08 – Inferir o sentido de palavras ou expressões não coloquiais e de metáforas em textos de nível 3 e 4, com apoio de pistas co-textuais.
Entre 280 e 300 pontos	Não foram observados itens para este intervalo.
Entre 300 e 320 pontos	C07D08 – Inferir o sentido de palavras ou expressões não coloquiais e de metáforas em textos de nível 3 e 4, sem apoio de pistas co-textuais.
Acima de 320 pontos	Não foram observados itens para este intervalo.

Fonte: Pesquisa aplicada a instrumentos de avaliação, área de Língua Portuguesa, CAED/UFJF.

Observa-se, com base no quadro 1, que a ausência de itens em intervalos da escala anteriores a 140 pontos constitui um indicativo de que a habilidade de inferir o sentido de palavra ou expressões demonstra ter sido desenvolvida apenas por estudantes com proficiência em leitura a partir de 140 pontos. Esse dado aponta para duas possibilidades: ou os estudantes com proficiência menor que 140 pontos da escala não demonstram ter desenvolvido essa habilidade ou não são observados itens que avaliem a habilidade em níveis de complexidade coerentes com o desempenho em leitura desses estudantes. A primeira hipótese parece mais plausível, uma vez que estudantes com proficiência inferior a 140 pontos provavelmente ainda não venceram a etapa da alfabetização inicial.



A ausência de classes representativas de pontos intermediários da escala – entre 200 e 240 pontos e entre 280 e 300 pontos – é indicativa de um limite do corpus analisado, uma vez que não foram encontrados itens válidos (parâmetro A igual ou superior a 0,01) nesses intervalos da escala. Entretanto, a classe imediatamente anterior e a imediatamente superior aos referidos intervalos permitem conclusões sobre o nível de complexidade da tarefa de inferir informações em textos que estudantes com proficiência compatível com os intervalos para os quais não foram observados itens representativos são capazes de realizar.

Finalmente, a ausência de itens que avaliam a habilidade de inferir o sentido de palavra ou expressão em intervalos acima de 320 pontos da escala pode ser indicativa de que o maior nível de complexidade da tarefa é observado a partir desse ponto da escala, o que parece coerente com a natureza da habilidade avaliada: leitores proficientes são aqueles capazes de inferir o sentido de uma palavra ou expressão, mesmo em textos que tratam de temas pouco familiares e que apresentam léxico especializado, com base em pistas co-textuais.

A consistência desses critérios confirmou-se pela análise de outros descritores da matriz de referência. Tomemos, para um segundo exemplo, o descritor “reconhecer o assunto de um texto”. O reconhecimento do tema ou assunto de um texto é tarefa pouco complexa mediante a presença de sinalizações de apoio no título ou no início do primeiro parágrafo. É tarefa mais complexa quando demanda a análise da estrutura tópica do texto ou de palavras e expressões nele recorrentes. Também a natureza das sinalizações é determinante aqui: repetições (a correspondência total entre a informação do texto e a do gabarito do item) ou presença de palavras-chave, no título ou no início do primeiro parágrafo, são sinalizações mais claras e, portanto, mais facilitadoras; paráfrases são sinalizações menos claras. A ausência de sinalizações dificulta a tarefa, demandando a apreensão de sentido global do texto, o estabelecimento de relação entre suas partes.

Se os critérios descritos – a presença de sinalização de apoio, o lugar onde essa sinalização se apresenta e o tipo de sinalização – contribuem para explicar o nível de complexidade das tarefas de leitura que avaliam uma mesma habilidade, eles não foram, no processo de análise, critérios suficientes. Segundo esses critérios, itens com tarefas de leitura identificadas como pouco complexas ancoravam, em alguns casos, em pontos altos da escala de proficiência, deixando clara a necessidade de considerar um outro critério: o da complexidade dos textos-suporte dos itens.



A definição de critérios para estabelecer o nível de complexidade dos textos que servem como suporte aos itens da avaliação em larga escala foi objeto de atenção em etapas anteriores da pesquisa em avaliação na área de Língua Portuguesa<sup>10</sup>. De um modo geral, as investigações preliminares já apontavam para diferenças óbvias, sobretudo entre os suportes de itens que ancoravam em pontos extremos da escala de proficiência: em pontos mais baixos, gêneros narrativos (como os contos infantis e algumas fábulas), bilhetes, poemas rimados, trovas, convites; em pontos mais altos, textos com tratamento literário mais elaborado (como fragmentos de romances e contos da nossa tradição literária, poemas mais elaborados metaforicamente) e textos do campo da ciência e da sociedade com temática especializada<sup>11</sup>.

A tarefa da pesquisa consistiu em analisar diferentes conjuntos de textos, segundo seu intervalo de ancoragem na escala de proficiência, para identificar os aspectos definidores de sua complexidade. Com base nessa análise, chegou-se à classificação dos textos em quatro níveis de complexidade, definidos pelos seguintes critérios: sintaxe, léxico, tema, estrutura textual, extensão.

Do ponto de vista sintático, a classificação de um texto como mais simples ou mais complexo leva em conta: o tamanho do período, a ordem dos componentes (estruturas em ordem direta ou indireta), o tipo de estrutura sintática (se coordenada ou subordinada), a presença ou não de intercalações.

Desse modo, textos sintaticamente mais simples (ou seja, com menor grau de complexidade) caracterizam-se pela presença de períodos curtos (1 linha em média), sem muitas intercalações, com predomínio de estruturas em ordem direta, de estruturas coordenadas e paralelismo sintático. Textos mais complexos se caracterizam pela presença de períodos mais longos (2 linhas ou 3 linhas em média), com intercalações (de estruturas nominais e verbais), predomínio de subordinação, alterações na estrutura canônica das frases (orações reduzidas antecipadas, por exemplo)<sup>12</sup>.

---

<sup>10</sup> Sobre as contribuições da pesquisa em relação à definição de níveis de complexidade dos textos, ver Barbosa; Micarello e Ferreira (2020).

<sup>11</sup> A categoria “gênero” apresentou-se, já no início da investigação, como um dos aspectos definidores do nível de complexidade de uma tarefa de leitura. Trata-se de categoria importante, mas não suficiente para se pensar a progressão no ensino e na avaliação.

<sup>12</sup> A sintaxe dos textos vai se tornando cada vez mais complexa à medida que o léxico empregado se aproxima mais do discurso formal. A extensão dos períodos também é determinante do nível de complexidade do texto, uma vez que períodos longos, com 3 ou mais linhas, por exemplo, exigem um fôlego de leitura que leitores iniciantes não possuem.



Outro aspecto determinante no campo da sintaxe diz respeito aos recursos que articulam orações e períodos, ou seja, a coesão sequencial e referencial. Textos sintaticamente mais simples caracterizam-se por procedimentos de referenciação por pronome pessoal do caso reto, reiteração por repetição ou sinonímia, com referentes próximos. Nos textos mais complexos sintaticamente, há referenciação por pronome pessoal do caso oblíquo (sobretudo registros menos usuais desses pronomes), por elipse, uso do pronome relativo, com referente mais distante.

Do ponto de vista lexical, a classificação de um texto como mais simples leva em conta o predomínio de léxico próprio do uso coloquial da linguagem ou de registro escrito mais informal. Um texto com maior complexidade lexical se define pelo predomínio de palavras e expressões próprias do discurso formal, como também pela utilização de léxico especializado (de campos específicos do conhecimento) e pelo uso metafórico da palavra.

Outro fator que define a complexidade de um texto é o seu tema. Textos que abordam temas de áreas especializadas do conhecimento exigirão que os leitores mobilizem conhecimentos mais formalizados para atuarem sobre eles. Esse é um critério que se relaciona bastante diretamente com o léxico. Textos que classificamos como de alta complexidade exigem certo domínio de temas e vocabulário especializados. Uma estratégia corrente desses textos é exatamente a explicitação do sentido de termos técnicos, como forma de oferecer apoio à construção de sentido, facilitando a leitura. Também o tratamento literário do tema ou a indiretividade no tratamento do tema torna mais complexo o texto.

A estrutura textual é outro fator de complexidade dos textos. Textos de estrutura canônica tendem a ser menos complexos, principalmente se houver marcas que apontem para os componentes dessa estrutura (como marcadores que introduzem o conflito narrativo ou que demarcam a passagem do tempo). Por exemplo: narrativas organizadas segundo a ordem orientação-complicação-desfecho caracterizam textos menos complexos. Narrativas que intercalam trechos de orientação com trechos de ação ou que envolvem *flashbacks* caracterizam textos mais complexos.



Finalmente, a extensão do texto também é critério relevante de complexidade. No entanto, este parece ser um critério menos determinante que os demais. Um bom exemplo disso são alguns poemas curtos que, ainda assim, apresentam-se com alto nível de complexidade, devido à sua elaboração literária, a construções metafóricas, por exemplo. Um resultado importante da pesquisa em avaliação, relativo a esse critério, indica que o tamanho do período é mais relevante que o tamanho do texto, quando se discute a complexidade de textos voltados para leitores iniciantes.

Considerando os critérios apresentados, o que distingue os textos de nível 1 dos textos de nível 2 é, fundamentalmente, a sintaxe. Pequenas alterações da estrutura sintática canônica (como intercalações e antecipações), bem como a extensão dos períodos, podem tornar um texto mais complexo para leitores iniciantes, que não possuem fôlego de leitura. Por outro lado, textos de nível 1 e 2 caracterizam-se por exibir léxico mais familiar, próximo do uso coloquial da linguagem.

A utilização de um léxico menos familiar apresentou-se como elemento definidor da complexidade de textos de nível 3 e 4. Textos de nível 3 exibem, normalmente, léxico próprio do discurso formal, implicando atividade inferencial mais constante. Textos de nível 4 exibem léxico especializado, presente em textos cuja temática também se especializa, textos que abordam questões de diferentes áreas de conhecimento: Ciência e tecnologia, Sociedade etc.

No entanto, para fins de definição do nível de complexidade do texto, a consideração desses traços do léxico (se mais, ou menos, próximos do discurso formal) não deve ser rígida. Em alguns textos, por exemplo, usos formais da língua são seguidos de recursos de “facilitação do discurso” que permitem lidar com um léxico desconhecido. Textos que tratam de temas especializados e que são dirigidos ao leitor criança ou adolescente costumam exibir várias estratégias para facilitar sua recepção (como expressões explicativas, sinonímia, exemplos, dentre outras possibilidades). Léxico e tema são, portanto, critérios importantes para a definição de textos de nível 3 e 4.



## 5. Considerações finais

A questão central deste artigo diz respeito à tarefa fundamental da avaliação em larga escala em Língua Portuguesa, cujo objetivo é aferir a competência leitora dos estudantes. Buscamos apresentar os fundamentos teóricos que têm embasado a pesquisa, para, a partir deles, apontar ações que possam contribuir para a construção de instrumentos mais ajustados ao público a que se destinam as avaliações.

Argumentamos que a possibilidade de construir medidas que permitam concluir sobre a competência leitora dos estudantes advém da verificação de como esses estudantes resolvem um conjunto de tarefas de leitura relativas a cada uma das habilidades descritas na matriz de referência para avaliação, ponto de partida para a elaboração dos testes. Esse conjunto de tarefas precisa abarcar os diferentes níveis de complexidade envolvidos nas tarefas de leitura relativas a uma mesma habilidade. Para tanto, propusemos a seguinte questão que nos guiou na proposição das classes de itens: que aspectos determinam o grau de complexidade de uma tarefa de leitura?

Os resultados da pesquisa em avaliação indicam que o nível de dificuldade de diferentes tarefas de leitura é determinado, de um lado, pela complexidade do texto proposto para leitura e, de outro, pela natureza das sinalizações que esse texto oferece como apoio à construção de sentidos. Buscamos elencar, ao longo deste artigo, os aspectos a serem considerados para a definição dessa complexidade. Do cruzamento desses aspectos, pode-se chegar à compreensão do que faz determinado item ser mais difícil que outro. A construção de tarefas com diferentes níveis de complexidade, com base em critérios linguísticos claros para avaliar uma mesma habilidade, permite que se chegue a uma compreensão do modo como progride o desenvolvimento dessa habilidade do ponto de vista da aprendizagem dos estudantes. O trabalho de analisar, com base nos resultados das avaliações em larga escala, como essa progressão vem ocorrendo, do ponto de vista da aprendizagem dos estudantes, pode contribuir para orientar escolhas pedagógicas mais ajustadas às suas reais necessidades.

Nesse sentido, a investigação dos critérios de complexidade textual leva-nos a propor que a competência de leitura é fortemente determinada, inicialmente, pelo domínio da sintaxe da língua. A análise dos itens já aplicados em avaliações em larga escala permite chegar a algumas conclusões que apontam para desdobramentos possíveis no que concerne ao tema da progressão do ensino



e da aprendizagem da leitura e, conseqüentemente, ao currículo de Língua Portuguesa para a educação básica. Os leitores iniciantes vão, aos poucos, apropriando-se da sintaxe da língua, de seu uso formal, a partir do contato com textos cada vez mais complexos do ponto de vista sintático. O amadurecimento desse leitor se consolida na medida em que ele vai ampliando seu repertório de leitura e, conseqüentemente, seu repertório lexical, familiarizando-se com textos de vários campos do conhecimento, dominando sua estrutura, abordando temas diversos cada vez mais especializados. Essa competência se amplia ainda mais quando se compreende o potencial expressivo da linguagem recriada literariamente.

Com base no exposto neste artigo, finalizamos apontando a importância de compreender currículo e avaliação como dimensões indissociáveis dos processos de ensino e de aprendizagem que, cada vez mais, precisam ser assim concebidos, no âmbito da pesquisa acadêmica, assim como das práticas pedagógicas e de gestão. Se, por um lado, o currículo é o ponto de partida da avaliação, que serve ao seu monitoramento, por outro, os resultados da avaliação também podem e devem oferecer elementos para repensar o currículo e sua progressão, à medida que indicam como os estudantes o vivenciam do ponto de vista de suas aprendizagens.



## Referências

- BARBOSA, B. T.; MICARELLO, H. L.; FERREIRA, R. V. J. Avaliações em larga escala de Língua Portuguesa: uma pesquisa sobre a complexidade dos textos que dão suporte a itens que avaliam a leitura. In: **Pesquisa e debate em educação**, v. 10, p. 1064-1081, 2020.
- BRASIL. **Base Nacional Comum Curricular (BNCC)**. Educação é a Base. Brasília, MEC/CONSED/UNDIME, 2017. Disponível em:
- < <http://basenacionalcomum.mec.gov.br/>>. Acesso em: 10 set. 2021.
- BRASIL. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais (PCNs):** terceiro e quarto ciclos do ensino fundamental: Língua Portuguesa. Brasília: MEC/SEF, 1998.
- DELL'ISOLA, R. L. P. **O sentido das palavras na interação leitor-texto**. Belo Horizonte: Faculdade de Educação da UFMG, 2005.
- FAUCONNIER, G. **Mappings in language and thought**. Cambridge University Press, 1997.
- JESUS, G. R.; RÊGO, R. M. L.; SOUZA, V. V. Evidências de validade de conteúdo da prova de psicologia do ENADE. In: **Estudos em Avaliação Educacional**, v. 29, n. 72, p. 858-884, set./dez. 2018. Disponível em: <<http://dx.doi.org/10.18222/eae.v29i72.4897> >. Acesso em: 20 fev. 2020.
- KLEIMAN, A. B. **Texto e leitor: aspectos cognitivos da leitura**. Campinas, SP: Pontes, 1989.
- KLEIMAN, A. B. **Oficina de leitura: teoria e prática**. Campinas, SP: Pontes, 1993.
- KLEIN, R. Escala de proficiência. In: FRADE, I. C. A. S.; COSTA VAL, M. G.; BREGUNCI, M. G. C. **Glossário Ceale: termos de alfabetização, leitura e escrita para educadores**. Belo Horizonte: FAE/UFMG, 2014.
- KOCH, I. G. V. **A coesão textual**. São Paulo: Contexto, 1989.
- KOCH, I. G. V. **O texto e a construção dos sentidos**. São Paulo: Contexto, 1998.
- KOCH, I. G. V.; TRAVAGLIA, C. **Texto e Coerência**. São Paulo: Cortez, 1989.
- ORLANDI, E. P. **A Linguagem e seu Funcionamento**. São Paulo: Editora Brasiliense, 1983.
- PASQUALI, L. Validade dos Testes Psicológicos: Será Possível Reencontrar o Caminho? In: **Psicologia: teoria e pesquisa**, v. 23, n. especial, p. 99-107, 2007. Disponível em: < <http://www.scielo.br/pdf/ptp/v23nspe/18>>. Acesso em: 20 fev. 2020.
- POSSENTI, S. A leitura errada existe. In: BARZOTTO, V. H. (org.). **Estado de leitura**. Mercado de Letras: Associação de Leitura do Brasil, 1999.
- SMITH, F. **Compreendendo a leitura: uma análise psicolinguística da leitura e do aprender a ler**. Porto Alegre: Artes Médicas, 1991.



# Possibilidades de criação de itens digitais para avaliação da oralidade em língua portuguesa usando a plataforma Testing Assisté par Ordinateur (TAO)<sup>1</sup>

Fabiano Santos Saito<sup>2</sup>

## 1. Introdução

O principal objetivo desta pesquisa aplicada é apresentar considerações teóricas e práticas sobre a possibilidade de produção de itens digitais voltados para a avaliação da oralidade em Língua Portuguesa utilizando a plataforma *Testing Assisté par Ordinateur* (TAO). Esta pesquisa faz parte do projeto “Do papel para a tela: proposição de novas formas de avaliar a proficiência em Língua Portuguesa com itens produzidos para plataformas digitais”, financiado pela Fundação Centro de Políticas Públicas e Avaliação da Educação/Universidade Federal de Juiz de Fora (doravante Fundação CAEd/UFJF) e desenvolvido ao longo do ano de 2020.

Como fundamentação teórica, em linhas gerais, parte-se dos pressupostos da oralidade em Língua Portuguesa, que serão discutidos em relação à avaliação em larga escala. Na parte teórica, também serão apresentados brevemente os princípios de *design* de itens de avaliação e de itens digitais, como

---

<sup>1</sup> Projeto financiado pela Fundação CAEd/UFJF.

<sup>2</sup> Departamento de Educação, Linguística e Letras (DELL), Unidade Carangola, Universidade do Estado de Minas Gerais (UEMG), e-mail: prof.fabiano.saito@gmail.com



subsídio teórico voltado para a prática de construção e estruturação dos itens digitais de avaliação. Ainda no âmbito teórico, será apresentada de forma sumária a Teoria de Resposta ao Item (TRI), que fornece parâmetros para construção de itens de avaliação tecnicamente consistentes.

Em relação à metodologia, este estudo pode ser considerado uma pesquisa aplicada, de abordagem exploratória e qualitativa. Para a produção experimental dos protótipos de itens digitais, foi utilizada a plataforma TAO (*Testing Assisté par Ordinateur*), que é uma ferramenta de produção de itens de avaliação que possui possibilidades e certas limitações, quando se consideram aspectos de *design*. A partir da produção dos protótipos de itens digitais, foi realizada uma análise qualitativa sobre o *design* de itens digitais e sua possibilidade de aplicação em avaliações de larga escala.

Como resultados parciais desta pesquisa, pode-se concluir que a produção de itens digitais na plataforma TAO permite a exploração de suportes multimídia e hipermídia, que, por sua vez, permitem a avaliação de habilidades de compreensão da escuta, o que não são consideradas ou pouco utilizadas em instrumentos de avaliação impressos.

## 2. Considerações sobre avaliação e o processo de digitalização

O processo de avaliação de conhecimentos engloba a avaliação em nível micro ou avaliação interna – como, por exemplo, as provas aplicadas pelos professores nas escolas –, bem como a avaliação em nível macro ou avaliação externa – como as avaliações aplicadas por instituições nacionais e internacionais, que são exteriores ao contexto escolar (cf. NASCENTES et al., 2017).

As avaliações externas geralmente são produzidas por instituições que aplicam testes em larga escala, aplicados em formato impresso, constituídos de vários itens de avaliação, que têm por objetivo mensurar conhecimentos, competências e habilidades aprendidos pelos estudantes no processo de escolarização. Entretanto, pesquisas recentes na área de avaliação têm apontado para desafios a serem enfrentados pelas instituições que buscam inovar os procedimentos utilizados para avaliar



conhecimentos, competências e habilidades de forma massiva, de forma confiável e seguindo os mesmos padrões de qualidade alcançados pelos instrumentos de avaliação já utilizados nos testes impressos. Nesse sentido, Braun (2013, p. 154) aponta que

há um número de objetivos que podem ser vislumbrados para as avaliações baseadas em tecnologias. Destes [objetivos] incluem-se melhorar o alinhamento e acurácia para as medidas dos construtos alvos correntes, facilitar a medição de novos construtos, tais como resolução de problemas e letramento computacional/informacional. Estes dois, bem como outros construtos que se prestam mais às avaliações baseadas em tecnologias, podem contribuir para aumentar a credibilidade e a relevância das ILSA [Avaliações Internacionais em Larga Escala (doravante AILE)], não apenas por fortalecer as ligações com o mundo fora das escolas. Entretanto, a avaliação de novos construtos certamente levantará questões metodológicas desafiadoras. A introdução de estímulos mais complexos, bem como o desejo de se avaliar tanto processos quanto resultados, exigirá modelos psicométricos e estratégias de análise de dados mais sofisticados<sup>3</sup>. (tradução nossa)

Em vista desse cenário complexo e contemporâneo, que exige o reposicionamento das instituições aplicadoras de avaliações em larga escala, há claramente uma demanda para a implementação de testes e avaliações em larga escala do formato impresso para o formato digital. Assim, as instituições que desenvolvem e produzem as avaliações se deparam com algumas questões metodológicas desafiadoras, dentre as quais: como transpor a avaliação do formato impresso para o digital? Como produzir itens de avaliação que serão utilizados em plataformas digitais?

Seguindo os rumos para digitalização e tecnologização das avaliações em larga escala impressas, Braun (2013), Beller (2013), Gierl e Lai (2017) concordam que as avaliações em larga escala enfrentam o desafio de mudar os procedimentos de avaliação baseadas em papel para procedimentos de avaliação baseadas nas tecnologias digitais de informação e comunicação. Os autores apontam para as vantagens de as tecnologias digitais propiciarem o uso de gráficos dinâmicos, imagens animadas, sons e vídeos, que o suporte em papel não possibilita ou possibilita de forma limitada (como o caso dos gráficos e imagens estáticas). Em razão disso, Beller (2013, p. 25) afirma: “[é] óbvio que as

---

<sup>3</sup> There are a number of goals that can be envisioned for technology-based assessments. These include improving alignment and accuracy for measures of current target constructs, and facilitating the measurement of new constructs, such as problem solving and computer/information literacy. These two, as well as other constructs that lend themselves more to technology-based assessments, could contribute to increased credibility and relevance of technology-based ILSAs, not least by strengthening links to the world outside schools. However, the assessment of new constructs will certainly raise challenging methodological issues. The introduction of more complex stimuli, as well as the desire to evaluate both processes and outcomes, will call for more sophisticated psychometric models and data-analytic strategies.



tecnologias inovadoras existentes, tais como *smartphone* e *tablets*, bem como tecnologias futuras, possuem o potencial de mudarem dramaticamente a forma como a avaliação será implementada no futuro”<sup>4</sup>. (tradução nossa)

Beller (2013, p. 26) complementa que uma das potencialidades que as tecnologias digitais têm de mudar a forma de avaliar reside no fato de que tais tecnologias incorporam novos estímulos e podem desenvolver outras competências, uma vez que,

ao enriquecer as situações de avaliação através do uso de multimídia, interatividade e controle sobre a apresentação de estímulo, é possível avaliar uma gama bem mais ampla de construtos do que era possível previamente. Por exemplo, SimScientists usam simulações científicas para dar apoio à poderosas avaliações formativas na aprendizagem de conteúdos de ciência complexos (QUELLMALZ et al., 2009). As tarefas de avaliação baseadas em simulação são dinâmicas, envolventes e interativas. São capazes de avaliar conhecimentos científicos complexos e competências investigativas, que vão bem além das capacidades de testes impressos<sup>5</sup>. (tradução nossa)

Nessa mesma linha de pensamento de que o ambiente digital proporciona uma experiência diferenciada e potencialmente “enriquecida” pelas multimídias e trajetórias de navegação possibilitadas pelos *links*, já dizia Lemke (2002, p. 300) que o hipertexto é mais do que o texto impresso, na medida em que a hipermídia traz mais modos de informação que a mídia impressa: “hipermodalidade é mais do que multimodalidade da mesma maneira que o hipertexto é mais que o texto comum. Não é simplesmente porque justapomos imagem, texto e som; planejamos múltiplas interconexões entre eles, tanto potenciais quanto explícitas”<sup>6</sup> (tradução nossa).

Portanto, para a elaboração de itens de avaliação digitais/interativos, há que se pensar em um modelo de elaboração de itens de avaliação em larga escala que incorporem as multimídias, hipermídias, a navegação e a interatividade como princípios teóricos norteadores para a elaboração de protótipos e a produção de itens a serem testados e implementados.

---

4. It is clear that existing innovative technologies, such as smartphones and tablets, as well as future technologies, hold the potential to dramatically change the way assessment will be implemented in the future.

5. By enriching assessment situations through the use of multimedia, interactivity and control over the stimulus display, it is possible to assess a much wider array of constructs than was previously possible. For example, SimScientists uses science simulations to support powerful formative assessments of complex science learning (Quellmalz et al. 2009). The simulation-based assessment tasks are dynamic, engaging, and interactive. They are capable of assessing complex science knowledge and inquiry skills, which go well beyond the capabilities of printed tests.

6. Hypermodality is more than multimodality in just the way that hypertext is more than plain text. It is not simply that we juxtapose image, text, and sound; we design multiple interconnections among them, both potential and explicit.



Para tanto, é necessário pensar que, para a produção de um item digital, também é preciso considerar as questões de *design*, pois o ambiente digital possui certas especificidades que o material impresso não possui e vice-versa. Para lançar as bases teóricas voltadas para o *design* de itens digitais, serão apresentadas na sequência brevemente duas teorias: a Teoria Cognitiva do Aprendizado Multimídia (MAYER, 2001; MAYER; CLARK, 2003); e o *design* instrucional para conteúdo *eLearning* (ISTRATE, 2009).

### 3. Teoria Cognitiva do Aprendizado Multimídia – revisitada

Em seu livro de 2001, intitulado *Multimedia Learning*, Richard E. Mayer apresenta suas pesquisas na área de Psicologia Educacional, Psicologia Cognitiva e Psicologia aplicada à Educação. Mayer (2001) elabora a Teoria Cognitiva do Aprendizado Multimídia, a partir de experimentos que buscavam medir os impactos do uso de várias mídias em um ambiente computacional controlado, testando as várias combinações possíveis entre as mídias para inferir quais combinações tinham maior impacto na aprendizagem. A partir desses experimentos, de base cognitivista, Mayer (2001) elaborou 7 princípios para o *design* de ambientes multimídia. Após algum tempo, revisando e atualizando a Teoria Cognitiva do Aprendizado Multimídia, formulada em 2001, Mayer e Clark (2003) expandiram os princípios para doze, que são os seguintes:

- 1) Princípio da Coerência – as pessoas aprendem melhor quando palavras, figuras e sons são incluídos ao invés de serem excluídos.
- 2) Princípio da Sinalização – as pessoas aprendem melhor quando pistas que destacam a organização do material essencial são adicionadas.
- 3) Princípio da Redundância – as pessoas aprendem melhor com animação e narração do que animação, narração e texto na tela e quando palavras são apresentadas como narração ao invés de narração com texto na tela.
- 4) Princípio da Contiguidade Espacial – as pessoas aprendem melhor quando palavras e figuras relacionadas são apresentadas próximas do que afastadas umas das outras na página ou tela.



- 5) Princípio da Contiguidade Temporal – as pessoas aprendem melhor quando palavras e figuras relacionadas são apresentadas simultaneamente do que sucessivamente.
- 6) Princípio da Segmentação – as pessoas aprendem melhor quando uma lição multimídia é apresentada em segmentos controlados pelo usuário do que uma unidade contínua.
- 7) Princípio do Pré-treinamento – as pessoas aprendem melhor em uma lição multimídia quando sabem os nomes e os comportamentos dos componentes do sistema.
- 8) Princípio da Modalidade – as pessoas aprendem melhor quando palavras são apresentadas como narração ao invés de texto na tela.
- 9) Princípio Multimídia – as pessoas aprendem melhor com palavras e figuras do que só com palavras.
- 10) Princípio da Personalização – as pessoas aprendem melhor quando as palavras estão em um estilo conversacional do que em um estilo formal.
- 11) Princípio da Voz – as pessoas aprendem melhor quando a narração em lições multimídia são faladas em uma voz humana amigável do que em voz de máquina.
- 12) Princípio da Imagem – as pessoas não aprendem necessariamente melhor em uma lição multimídia quando a imagem do narrador/falante é adicionada à tela.

Dadas as contribuições trazidas pela Teoria Cognitiva da Aprendizagem Multimídia, há que se considerar suas limitações, posto que indica, de modo sintético, as diferenças individuais dos usuários e aprendizes expostos aos ambientes e plataformas educacionais que fazem uso de várias mídias, pois tais diferenças vão além das questões puramente cognitivas, mas também às diferenças de patrimônio cultural ou de vivências experienciais e fatores sociais. Estudantes que têm uma vivência e contato maior com as tecnologias digitais tenderão a ter uma performance diferente dos estudantes que não têm acesso às tecnologias.

O próprio Mayer (2001) reconhece que, cognitivamente, o ser humano possui limitações de memória para processar os vários modos de representação da informação de forma simultânea, além do fato de que certos *designs* podem apresentar mais informação (*enriched environments*) ou menos informação (*poor environments*), sendo que o excesso de informações pode limitar o uso da memória e a capacidade de processamento de informações.



Mesmo assim, tais princípios da Teoria Cognitiva podem auxiliar na produção de itens digitais, indicando quais combinações de mídias podem ser mais efetivas, podem funcionar melhor, pensando-se na limitação física e cognitiva de uso da memória e da capacidade de processamento de informações pelo cérebro humano. Ou seja, é necessário balancear e combinar as mídias presentes em um objeto digital, para que o item digital seja eficiente e cognitivamente processável.

Como pode ser analisado, os princípios da aprendizagem multimídia de Mayer (2003) apontam para orientações gerais que podem ser utilizadas na elaboração de vídeos, animações, apresentações multimídia, materiais para aprendizagem *on-line*, em que a experiência de aprendizagem pode acontecer de modo mais efetivo. Os mesmos princípios podem ser utilizados na elaboração e produção de itens digitais/interativos para serem usados em avaliações de larga escala para plataformas e ambientes digitais.

## 4. O *design* visual e pedagógico para conteúdo de *eLearning*

Em um artigo seminal sobre *design* instrucional para conteúdo *eLearning*, Istrate (2009) elaborou algumas orientações de caráter prático para educadores e profissionais que se engajam na produção de conteúdo para ensino e aprendizagem *on-line*. O autor apresenta questões de *design* que indicam desde a formatação e posição do texto em uma *webpage* até o uso comedido das cores e de elementos animados. Para sintetizar as recomendações de Istrate (2009), seguem as mais importantes:

- 1) Orientação aos usuários por meio de organizadores visuais;
- 2) Uso de um caminho de navegação que direcione os usuários a determinado material (um mapa de localização pode ser útil);
- 3) Utilização de terminologia padrão;
- 4) Destaque de palavras-chave;
- 5) Uso de termos significativos para etiquetar o conteúdo;
- 6) Marcação de listas com símbolos (bolinhas) para cada item;
- 7) Apresentação de uma única ideia por parágrafo;



- 8) Uso de estilo de “pirâmide invertida”, no qual o primeiro parágrafo (chamado de “*lead*” na literatura específica) apresenta sucintamente a essência da informação: os parágrafos seguintes elaboram o que já foi apresentado, trazendo informações complementares;
- 9) Utilização de metade (ou até menos) do número de palavras usadas em material impresso ou convencional.
- 10) Teste de funcionamento do *design* em vários equipamentos (como computador, celular, *tablet* etc.).
- 11) Por outro lado, Istrate (2009) indica o que deve ser evitado em materiais voltados para o *eLearning*:
- 12) Texto piscante e imagens animadas irrelevantes;
- 13) Rolagem horizontal, pois o conteúdo essencial deve ser apresentado em uma tela;
- 14) Vídeos longos e imagens de baixa qualidade;
- 15) Apagamento das fontes que originaram o conteúdo;
- 16) Erros gramaticais e de digitação;
- 17) Negligência em relação ao público com sistemas de baixo desempenho e com acesso lento à Internet.

Tais princípios que norteiam a organização do *design* de material instrucional para a *web* podem ser aproveitados quando da produção de itens digitais a serem utilizados nas avaliações em larga escala. As orientações de Istrate (2009) indicam que o *design* de materiais instrucionais deve ser simultaneamente simples, objetivo, claro, sintético, informativo e eficiente.



## 5. Avaliação da oralidade em Língua Portuguesa

Como se observa, tanto os teóricos da área de avaliação em larga escala quanto os teóricos do *design* de material instrucional apontam para as vantagens da utilização das diferentes mídias (multimídias, hipermídias) que possibilitam a exploração de conhecimentos e habilidades inviáveis em instrumentos impressos.

Em função dessa potencialidade de se explorar as multimídias e hipermídias, o projeto “Do papel para a tela: proposição de novas formas de avaliar a proficiência em Língua Portuguesa com itens produzidos para plataformas digitais” se propôs a investigar a possibilidade de se avaliar a proficiência da oralidade em Língua Portuguesa, produzindo protótipos de itens digitais voltados para a compreensão da escuta.

Neste trabalho, em síntese, considera-se a linguagem humana como um grande multissistema, em que os sujeitos sociais utilizam vários sistemas simbólicos (ou linguagens) para interagir em práticas comunicacionais situadas, como indica a Base Nacional Curricular Comum (doravante BNCC):

As atividades humanas realizam-se nas práticas sociais, mediadas por diferentes linguagens: verbal (oral ou visual-motora, como Libras, e escrita), corporal, visual, sonora e, contemporaneamente, digital. Por meio dessas práticas, as pessoas interagem consigo mesmas e com os outros, constituindo-se como sujeitos sociais. Nessas interações, estão imbricados conhecimentos, atitudes e valores culturais, morais e éticos. (BNCC, 2018, p. 63)

Do ponto de vista biológico, histórico, cognitivista e cultural, no que concerne à linguagem verbal humana, há certo consenso entre os linguistas de se afirmar que a fala se desenvolve antes da escrita nas sociedades letradas, que as habilidades da fala e da escuta se desenvolvem quase de forma natural nos seres humanos expostos à linguagem, enquanto que as habilidades da leitura e da escrita são habilidades aprendidas em contextos de instrução formal (STREET, 1984, 2003; ONG, 2012 [1982]).

A língua pode ser segmentada em categorias discretas (as quatro habilidades linguísticas gerais: fala, escuta, escrita, leitura), que podem ocorrer através das práticas sociais de letramento, no caso da leitura e da escrita; ao passo que as habilidades de fala e escuta podem ser situadas nas práticas sociais relacionadas ao conceito de oralidade (ONG, 2012[1982]); MARCUSCHI, 2001; MARCUSCHI, 2008).



Magalhães (2001) juntamente com Magalhães e Carvalho (2018) reconhecem a oralidade como prática linguística importante para a “socialização humana e participação na vida cidadã”, apontando que são poucos os estudos que trabalham as práticas sociais de oralidade como formas de ensino, aprendizagem e avaliação das capacidades linguísticas na modalidade oral.

Ainda é necessário considerar que os gêneros discursivos se materializam nos textos que circulam em sociedade (BAKHTIN, 1979). Dessa feita, só se torna possível avaliar a proficiência da compreensão da escuta em Língua Portuguesa através do uso de gêneros textuais orais, que são unidades de discurso empíricas, que podem ser testadas e observadas. Portanto, a construção de itens de avaliação para aferição da compreensão da escuta em Língua Portuguesa se torna factível através da seleção de gêneros textuais orais.

Como ainda inexistia uma matriz de referência com as competências e as habilidades linguísticas referentes à oralidade em Língua Portuguesa que devam ser avaliadas em um item de avaliação, ao longo da pesquisa, foi realizado um levantamento dos descritores, competências e habilidades linguísticas relacionados à oralidade presentes na BNCC, que indicam quais conhecimentos, habilidades e competências podem ser avaliados em um item de avaliação prototípico.

## 6. O *design* de itens e a Teoria de Resposta ao Item (TRI)

Para a construção do *design* de um item de avaliação, além dos conhecimentos, competências e habilidades a serem avaliados, além da escolha dos textos-suporte que ensejarão a construção dos itens de avaliação prototípicos, outra questão importante a se considerar é a Teoria de Resposta ao Item (TRI).

Na elaboração de um item de avaliação exemplar são elementos constituintes do *design*: 1) o enunciado do item; 2) o texto-suporte; 3) o comando do item; 4) a resposta padrão (ou gabarito); 5) os distratores (cf. CAED/UFJF, s.d., p. 18-19).

O enunciado, o texto-suporte e o comando do item devem estar afinados com a habilidade ou conhecimento que se deseja avaliar. Em relação às alternativas de resposta, tanto o gabarito quanto os distratores devem ser consistentes com os demais elementos do item de avaliação. A formulação



de alternativas de resposta contendo um item correto, sendo os demais distratores plausíveis, permite que a habilidade focal seja avaliada, possibilitando a produção de métricas e comparação das respostas dadas pelos respondentes.

Elaborar o item para que a chance de acerto ao acaso seja diminuída, que a habilidade avaliada esteja sendo discriminada, que o nível de dificuldade seja adequado ao público-alvo da avaliação garantem a satisfação dos parâmetros indicados na Teoria de Resposta ao Item (TRI). Basicamente, a Teoria de Resposta ao Item (TRI) permite a construção de modelos estatísticos que fornecem formas de representar a probabilidade de acerto do item ao acaso, comparada com a habilidade que está sendo avaliada, juntamente com o nível de dificuldade e as características (parâmetros) dos itens (cf. ARAUJO et al., 2009; GOMES, 2014). Um item de avaliação com respostas autoexcludentes ou distratores óbvios não satisfaz aos critérios de validade e consistência de acordo com o modelo proposto pela TRI, por exemplo.

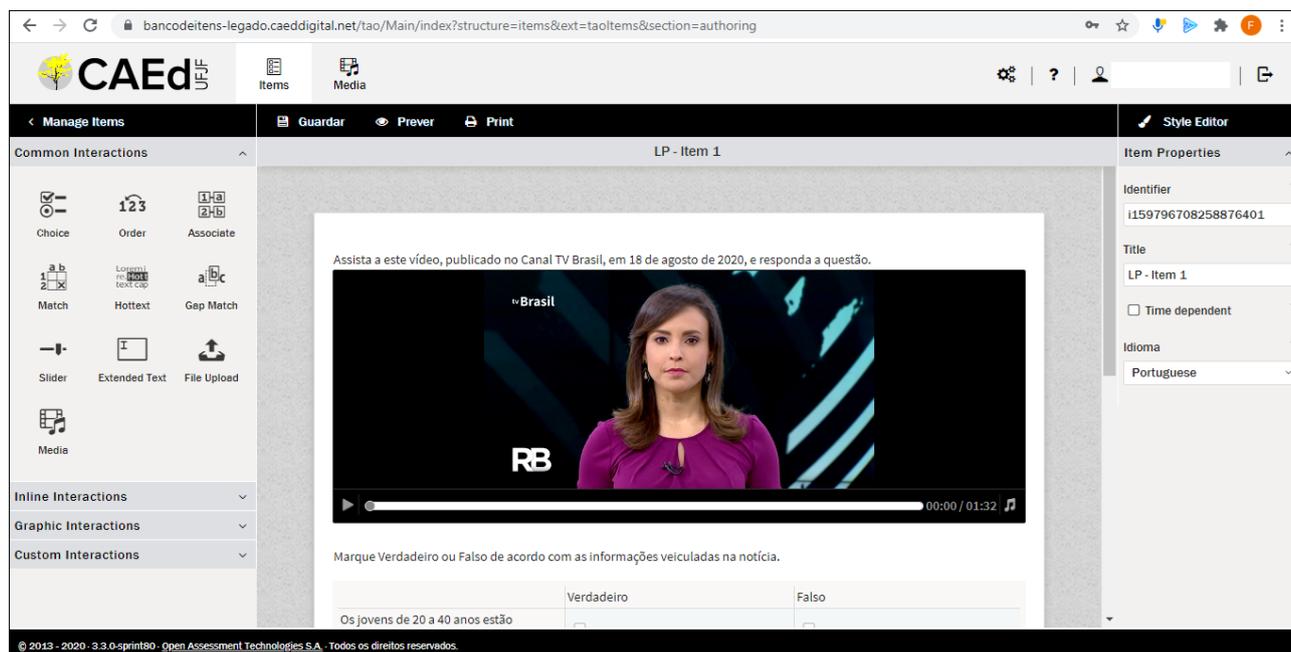
## 7. A plataforma TAO (*Testing Assisté par Ordinateur*)

Realizadas todas as considerações pertinentes à construção e *design* de um item de avaliação prototípico para ser usado em ambiente mediado pelas tecnologias digitais, por fim, há que se considerar o uso de uma plataforma para elaboração dos itens.

Nesta pesquisa foi utilizada a plataforma TAO (*Testing Assisté par Ordinateur*, ou Testagem Assistida por Computador, em português). Essa plataforma possui uma arquitetura aberta, desenvolvida pela Universidade de Luxemburgo, em parceria com o Centro de Pesquisas Públicas Henri Tudor (*Centre de recherche public Henri Tudor*), com licença de uso do tipo *groupware*, que permite a colaboração de diversos usuários, de qualquer lugar e a qualquer tempo. Essa plataforma permite a criação de itens de avaliação digitais, bem como o seu gerenciamento, funcionando como biblioteca ou repositório de itens. A plataforma TAO possui um ambiente de criação bem intuitivo, permitindo o uso de suporte multimídia e a integração de elementos de interatividade no item de avaliação digital. A figura 1 representa a tela da plataforma TAO com os elementos de criação/edição de itens.



Figura 1. Tela de criação/edição na plataforma TAO



Fonte: elaborada pelo autor.

Na tela acima percebe-se que a função de criador/edição de itens na plataforma TAO permite a criação de diversos tipos de item (os ícones à esquerda: múltipla escolha, ordenação, associação, combinação, texto longo), assim como o carregamento de mídia. Na figura acima, observe que o item foi estruturado com uma mídia audiovisual (centro da imagem). No menu da direita, é possível configurar diversas funcionalidades do item elaborado (limitar a quantidade de respostas, a quantidade de vezes de reprodução do vídeo, o tamanho do quadro do vídeo, o tamanho da mídia, configurar embaralhamento de respostas, definir a pontuação, *feedback* automático, dentre outras funções).

A plataforma TAO possui muitas potencialidades para a elaboração de itens digitais no que concerne aos formatos básicos de um item. Entretanto, para construir itens com elementos de interação sofisticados, é necessário o conhecimento aprofundado de códigos de programação, o que é um fator limitador da plataforma.



## 8. Metodologia de pesquisa

Esta pesquisa empregou abordagens e escolhas metodológicas diversas. Durante a fase de revisão de literatura, foi adotada a pesquisa bibliográfica e documental, em que artigos, dissertações, teses, capítulos de livro e livros foram lidos e resenhados para a construção do aporte teórico do estudo.

Em um segundo momento, foi adotada a pesquisa aplicada, de abordagem exploratória, durante a fase de elaboração dos protótipos de itens de avaliação. A plataforma TAO foi utilizada para a testagem de vários *designs* de itens e, a partir da avaliação das potencialidades da plataforma TAO, foram elaborados itens de avaliação prototípicos a partir das bases teóricas. Essa fase foi trabalhosa, uma vez que, para produzir um item, mesmo que experimental ou prototípico, várias etapas são seguidas: 1) definir um público-alvo para o item; 2) definir uma habilidade a ser avaliada, focalizando a compreensão da escuta; 3) selecionar um texto da oralidade em Língua Portuguesa; 4) construir um enunciado e um comando para o item; 5) definir o gabarito e as alternativas distratoras; 6) implementar o item na plataforma digital e testá-lo.

No projeto de pesquisa, foi construído um modelo de teste com itens digitais para ser aplicado a um grupo amostral de respondentes. Entretanto, essa etapa da pesquisa foi suspensa, devido às contingências impostas pela pandemia de Covid-19.

Na impossibilidade de se proceder à testagem com um grupo amostral e realizar um teste piloto, optou-se pela análise qualitativa de alguns itens de avaliação prototípicos produzidos para este estudo.



## 9. Apresentação e análise de dados

Nesta seção serão apresentados alguns itens de avaliação prototípicos ou protótipos de itens digitais produzidos na plataforma TAO, juntamente com uma análise qualitativa sobre o item criado e suas potencialidades de uso em testes de proficiência de larga escala.

### 9.1 Protótipo de item digital 01

Para o item Protótipo 01 de Língua Portuguesa, foi selecionado um vídeo curto do Canal da TV Brasil no *YouTube* (<https://www.youtube.com/user/tvbrasil/videos>), sobre a propagação do vírus Covid-19 entre pessoas jovens (<https://www.youtube.com/watch?v=O0ZYJkXXgT8>). O tema do vídeo foi selecionado devido à sua importância e também pelo fato de o vídeo trazer uma notícia curta, com linguagem jornalística simples e direta. Esse vídeo, considerando-se o conteúdo, pode ser trabalhado nos anos finais do Ensino Fundamental (8º e 9º anos preferencialmente). A duração do vídeo - de 1min e 32s - também é adequada para a plataforma TAO, que não suporta vídeos longos e grandes, o que também se adéqua às postulações de Istrate (2009) e de Mayer e Clark (2003).

Em relação à habilidade ou conhecimento que pode ser avaliado utilizando-se esse vídeo, aponta-se para o seguinte descritor da BNCC:

#### Quadro 1. Descritor de habilidade – Protótipo 01

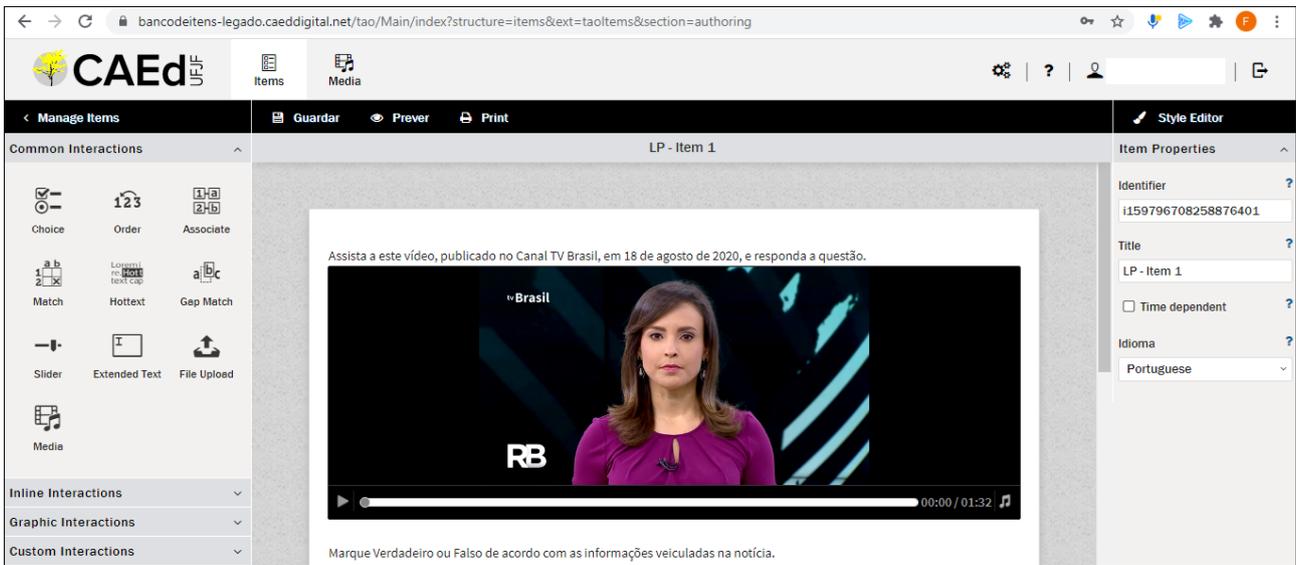
Identificação do descritor	Habilidades especificadas no descritor
EF69LP03	Identificar, em notícias, o fato central, suas principais circunstâncias e eventuais decorrências; em reportagens e fotorreportagens o fato ou a temática retratada e a perspectiva de abordagem, em entrevistas os principais temas/subtemas abordados, explicações dadas ou teses defendidas em relação a esses subtemas; em tirinhas, memes, charge, a crítica, ironia ou humor presente.

Fonte: BNCC, 2018.

Em relação aos procedimentos técnicos, o vídeo foi baixado da Internet utilizando-se telefone celular da marca *Samsung*. O aplicativo do *YouTube* no celular permite o *download* do arquivo de vídeo, possibilitando, ainda, escolher a qualidade e o tamanho do arquivo. Após o *download*, o arquivo foi carregado na plataforma TAO, através da funcionalidade “*media*”.



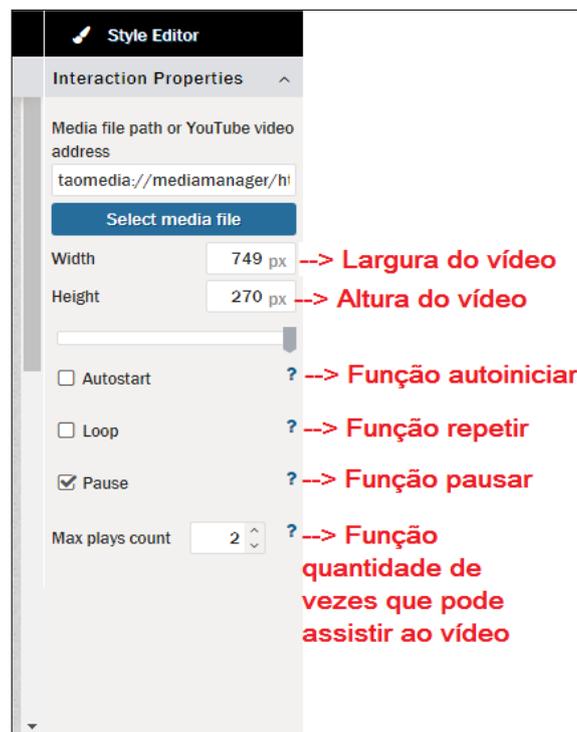
**Figura 2. O vídeo inserido na plataforma**



Fonte: elaborada pelo autor.

Dentro da plataforma TAO, foi possível inserir um comando para a parte do item que possui a mídia audiovisual (“Assista a este vídeo, publicado no Canal TV Brasil, em 18 de agosto de 2020, e responda a questão”) e configurar os controles da mídia:

**Figura 3. Controles de mídia na Plataforma TAO**



Fonte: elaborada pelo autor.

Depois de configuradas as funções de controle da mídia, foi inserido um campo de criação de questão do tipo “*match*” (combinação), que permite criar um quadro de opções do tipo verdadeiro ou falso.

Foram elaboradas 10 proposições (que eram verdadeiras ou falsas de acordo com o conteúdo do vídeo). Na configuração de resposta, foi definido que o respondente deve responder às 10 proposições. Como o número de proposições pode ser definido pelo elaborador de itens, foram selecionadas 10 proposições para aumentar o nível de complexidade da resposta. De acordo com a Teoria de Resposta ao Item (TRI), quanto mais opções a serem consideradas, maior o nível de dificuldade. No caso desse item, se o respondente errar de 1 a 10 proposições, o *score* do item é zero; neste caso, o respondente precisa acertar todas as 10 proposições para ter *score* de 1 ponto. A definição de *score* e a pontuação também podem ser definidos pelo criador do item.

**Figura 4. As 10 proposições no item do tipo “*match*” (combinação) na Plataforma TAO do CAEd**

Marque Verdadeiro ou Falso de acordo com as informações veiculadas na notícia.		
	Verdadeiro	Falso
Os jovens de 20 a 40 anos estão ajudando a espalhar a Covid-19.	<input type="checkbox"/>	<input type="checkbox"/>
A Rússia começou a produzir uma vacina para combater a Covid-19.	<input type="checkbox"/>	<input type="checkbox"/>
A vacina russa passou por todas as fases de teste antes de ser produzida.	<input type="checkbox"/>	<input type="checkbox"/>
O nome escolhido para a vacina russa é Sputnik 4.	<input type="checkbox"/>	<input type="checkbox"/>
Para a OMS, todas as nações devem ser incluídas no projeto de imunização contra o corona vírus.	<input type="checkbox"/>	<input type="checkbox"/>
O Diretor da OMS, Tedros Adhanom, é favorável à nacionalização da vacina.	<input type="checkbox"/>	<input type="checkbox"/>
De acordo com a OMS, os jovens têm consciência de que são transmissores do corona vírus.	<input type="checkbox"/>	<input type="checkbox"/>
De acordo com dados da OMS, a proporção de jovens contaminados cresceu mundialmente.	<input type="checkbox"/>	<input type="checkbox"/>
Pessoas assintomáticas podem transmitir o vírus.	<input type="checkbox"/>	<input type="checkbox"/>
Os idosos são citados como grupo vulnerável à doença.	<input type="checkbox"/>	<input type="checkbox"/>

Fonte: elaborado pelo autor.

A questão do tipo “*match*” com proposições do tipo “verdadeiro” ou “falso” foi criada para testagem dos *designs* possíveis de serem criados na plataforma TAO. A plataforma TAO permite a autocorreção do item, o que facilita a automação dos *scores* atribuídos a cada respondente.



A mídia foi inserida e logo abaixo foi colocado o enunciado com o objetivo do item, as alternativas foram apresentadas em seguida, o que satisfaz aos princípios da contiguidade espacial e da modalidade (MAYER; CLARK, 2003). O vídeo inserido no item é curto, conforme indica Istrate (2009), sua fonte foi mantida e seu conteúdo linguístico e audiovisual possuem relevância e convergência com a habilidade indicada no descritor da BNCC. O respondente consegue resolver o item de forma adequada se assistir e compreender o discurso oral presente no vídeo, portanto, o vídeo avalia a compreensão da escuta e trabalha com a oralidade (MAGALHÃES, 2001).

## 9.2 Protótipos de item digital 02 e 03

No item digital protótipo 02 e 03, foram seguidos procedimentos técnicos similares aos empregados na elaboração do item digital protótipo 01. Para esse item, foi selecionado um vídeo curto do *YouTube* ([https://www.youtube.com/watch?v=CqE76KUAK\\_M](https://www.youtube.com/watch?v=CqE76KUAK_M)), que trabalha com narrativas tradicionais do Nordeste do Brasil, os “causos”. O vídeo apresenta extensão de 5min e 53s. Devido ao tema, o público-alvo a que se destina esse item pode ser de estudantes do 4º e do 5º ano do ensino fundamental, preferencialmente.

Foram executados os mesmos procedimentos para carregar o vídeo na plataforma TAO.

**Figura 5. Frame do vídeo “O caso do lobisomem”, carregado na plataforma TAO**



Fonte: elaborado pelo autor.

A partir da mídia escolhida e carregada, foi possível elaborar quatro (04) itens do tipo “múltipla escolha” e um (01) item do tipo “ordenação das alternativas” utilizando-se o mesmo suporte audiovisual. Devido a questões de espaço, serão apresentados 01 item do tipo “múltipla escolha” (Protótipo 02) e 01 item do tipo “ordenação das alternativas” (Protótipo 03).

Para os itens do tipo “múltipla escolha”, foi utilizada a funcionalidade “choice”, que permite a criação de itens de múltipla escolha. Para o *design* do item, foi escolhido o modelo clássico de questão com cinco alternativas, sendo apenas uma a opção correta (gabarito) e as demais distratoras (CAED/UFJF, s.d.). Esse modelo foi escolhido por ser consistente com a TRI.

Para o Protótipo 02, foi definida a seguinte habilidade específica discriminada no descritor da BNCC:

## Quadro 2. Descritor de habilidade – Protótipo 02

Identificação do descritor	Habilidades especificadas no descritor
EF35LP11	Ouvir gravações, canções, textos falados em diferentes variedades linguísticas, identificando características regionais, urbanas e rurais da fala e respeitando as diversas variedades linguísticas como características do uso da língua por diferentes grupos regionais ou diferentes culturas locais, rejeitando preconceitos linguísticos.

Fonte: BNCC (2018).

Conquanto o *design* do item do tipo múltipla escolha siga um padrão muito similar aos dos itens impressos, há que se destacar que, para se chegar- à resolução do item, é necessário ao respondente assistir ao vídeo e exercitar a compreensão da escuta, focalizando a oralidade (MAGALHÃES, 2001).

## Figura 6. Design do item do tipo “múltipla escolha”

De acordo com o contexto do "causo", qual é o significado da expressão "meter sebo nas canelas"?

- ir embora com pressa.
- despedir-se de quem hospeda.
- ficar extremamente apavorado.
- esconder-se na floresta.
- ficar paralisado de medo.

Fonte: elaborado pelo autor.



Para o item do tipo “ordenação das alternativas” (Protótipo 03), que acompanha a mídia audiovisual “O caso do lobisomem”, foi escolhido o *design* para questões de ordenação das alternativas, através do comando “*order*” da plataforma TAO. Como o texto oral se trata de uma narrativa, que traz uma sequência de fatos, foi possível elaborar esse tipo de questão. Foram selecionados 10 fatos ou acontecimentos da narrativa para que o estudante coloque na ordem ou sequência correta, de acordo com o texto narrado. Foram elaboradas 10 alternativas para serem ordenadas porque isso aumenta o nível de complexidade da tarefa, o que condiz com os princípios da Teoria de Resposta ao Item (TRI), diminuindo as chances de acerto ao acaso.

De modo específico, esse item trabalha com a seguinte habilidade da BNCC:

### Quadro 3. Descritor de habilidade – Protótipo 03

Identificação do descritor	Habilidades especificadas no descritor
EF35LP26	Ler e compreender, com certa autonomia, narrativas ficcionais que apresentem cenários e personagens, observando os elementos da estrutura narrativa: enredo, tempo, espaço, personagens, narrador e a construção do discurso indireto e discurso direto.

Fonte: BNCC (2018).

Há que se ressaltar que, como esse item avalia a compreensão da escuta, a habilidade acima foi adaptada, entendendo-se que, quando se ouve uma narrativa oral ou se assiste a uma narrativa audiovisual, o ouvinte precisa compreender os elementos básicos da narrativa, como personagens, cenário, conflito, enredo, função do narrador etc.

O *design* proposto para esse item permite uma maior interação entre o respondente e o conteúdo apresentado em tela, sendo talvez exigido do respondente o princípio do pré-treinamento (MAYER; CLARK, 2003), pois, para responder à questão, é necessário arrastar as opções do quadro da esquerda e ordená-las no quadro vazio da direita:



**Figura 7. Design do item do tipo “ordenar”**

Coloque os acontecimentos na ordem correta em que eles aparecem no "causo":

- Seu Francisco aparece com a espingarda.
- O lobisomem chega perto da casa e da rede.
- Agenor e Juventina estão namorando no centro da cidade.
- Juventina pede ao pai para Agenor dormir ali por ser tarde da noite.
- Agenor e Juventina caminham pela floresta até a casa da moça.
- Agenor foge da casa de Juventina.
- Agenor fica se remexendo na rede sem sono.
- Seu Francisco entrega a rede para Agenor dormir nela.
- O lobisomem se transforma em Seu Francisco.
- Agenor e Juventina chegam na casa dela.

>

Fonte: elaborado pelo autor.

Para fins de avaliação e ao mesmo tempo para satisfazer aos parâmetros da TRI, no Protótipo de Item Digital 03, o respondente só alcança a pontuação, caso acerte a ordem correta em que os fatos ocorrem na narrativa, o que diminui as chances de acerto ao acaso. Para resolver esse item, ao respondente é demandada a compreensão da escuta, focalizando-se os gêneros da oralidade (MARCUSCHI, 2008).

Os *designs* dos protótipos de item digital 01 e 03 indicam formatos de item inovadores em relação ao protótipo 02, que segue o formato clássico de itens de múltipla escolha.



## 10. Considerações finais

A título de considerações finais, o objetivo deste trabalho foi o de apresentar as bases teóricas para elaboração de itens de avaliação digitais e demonstrar a aplicação prática de tais teorias na elaboração de três protótipos de itens de avaliação digitais, criados através da Plataforma TAO.

As vantagens da utilização de itens digitais em uma avaliação em larga escala digital são as seguintes: 1) os itens digitais permitem o uso de multimídia e hipermídia, sendo, portanto, possível o uso de mídias audiovisuais que circulam nas redes informacionais; 2) os itens digitais permitem a autocorreção; 3) os itens digitais na plataforma TAO permitem formatos de resposta e interação diferentes do modelo clássico de item de múltipla escolha; 4) uma avaliação em larga escala digital diminui os custos com impressão de materiais e a complexa logística de entrega desses materiais nos locais de prova.

Em relação às limitações, há que se considerar a aplicação experimental e a testagem dos itens de avaliação digital para adequar a dificuldade dos itens ao público-alvo, para diminuir a necessidade de pré-treinamento, principalmente em itens digitais com *design* complexo ou sofisticado em demasia. Outro ponto importante a se considerar seria o nível de segurança de dados quando da aplicação das avaliações digitais, pois o resultado da testagem dos itens pode ser diferente, se aplicado em ambiente controlado ou se aplicado para ser feito no celular do avaliando, por exemplo. Além disso, a testagem em ambiente digital depende de uma infraestrutura de tecnologias digitais da informação e comunicação eficientes que garantam uma avaliação digital em larga escala segura, confiável e eficiente.

De forma específica, os protótipos de itens digitais aqui apresentados apontam para a possibilidade de se trabalharem novos conhecimentos nas avaliações em larga escala, neste caso específico, a compreensão da escuta em Língua Portuguesa através do uso de gêneros da oralidade e mídia audiovisual como suporte nestes itens de avaliação. Como limitação deste estudo, aponta-se para a necessidade de validação desses itens digitais de avaliação em testes experimentais controlados aplicados a um grupo amostral para afinação do item e produção de métricas que são importantes na área de avaliação em larga escala.



## Referências

- ARAUJO, E. A. C. de; ANDRADE, D. F. de; BORTOLOTTI, S. L. V. Teoria da Resposta ao Item. **Rev. Esc. Enferm. USP**, 2009; 43(Esp):1000-8. Disponível em: <http://www.scielo.br/pdf/reeusp/v43nspe/a03v43ns.pdf> Acesso em: 03 jan. 2021.
- BAKHTIN, M. Os gêneros do discurso In: BAKHTIN, M. **Estética da criação verbal**. São Paulo: Martins Fontes, 1997 [1979]. p. 278-326.
- BELLER, M. Technologies in Large-Scale Assessments: New Directions, Challenges, and Opportunities. In: DAVIER, M. V.; GONZALEZ, E.; KIRSCH, I.; YAMAMOTO, K. (orgs.). **The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research**. London, New York: Springer Science, 2013.
- BRASIL, Ministério da Educação. **Base Nacional Comum Curricular** (versão final). 2018. Disponível em: [http://basenacionalcomum.mec.gov.br/images/BNCC\\_EI\\_EF\\_110518\\_versaofinal\\_site.pdf](http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf) Acesso em: 17 de jan. de 2021.
- BRAUN, H. Prospects for the Future: A Framework and Discussion of Directions for the Next Generation of International Large-Scale Assessments. In: DAVIER, M. V.; GONZALEZ, E.; KIRSCH, I.; YAMAMOTO, K. (orgs.). **The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research**. London, New York: Springer Science, 2013.
- CAED/UFJF. **Guia de elaboração de itens**. Língua Portuguesa. Disponível em: [http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia\\_De\\_-Elabora%C3%A7%C3%A3o\\_De\\_Itens\\_LP.pdf](http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia_De_-Elabora%C3%A7%C3%A3o_De_Itens_LP.pdf) Acesso em: 18 nov. 2020.
- GIERL, M. J.; LAI, H. The Role of Cognitive Models in Automatic Item Generation. In: RUPP, André A.; LEIGHTON, Jacqueline P. **The Handbook of Cognition and Assessment**. Oxford: Wiley Blackwell, 2017.
- GOMES, L. da S. **A Teoria de Resposta ao Item na avaliação em larga escala**. 2014. 78 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Mestrado Profissional em Matemática em Rede Nacional, Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, RJ.
- ISTRATE, O. Visual and pedagogical design of eLearning content. **eLearning Papers**, n. 17, p. 1-12, 2009.
- LEMKE, J.L. Travels in Hypermodality. **Visual Communication**, v. 1, n. 3, p. 299-325, 2002.
- MAGALHÃES, T. G. **Concepção de oralidade**: a teoria dos PCN e no PNLD X a prática nos livros didáticos. 2007. 208 f. Tese (Doutorado em Letras – Estudos da Linguagem). Programa de Pós-Graduação em Letras, Universidade Federal Fluminense (UFF), Niterói – RJ.
- MAGALHÃES, T. G.; CARVALHO, T. de A. B. Análise do eixo da oralidade na Proposta Curricular de Língua Portuguesa da rede municipal de ensino de Juiz de Fora (MG). **Rev. Bras. Estud. Pedagog.**, Brasília, v. 99, n. 251, p. 111-131, jan./abr. 2018.



MARCUSCHI, L. A. **Produção textual, análise de gêneros e compreensão**. São Paulo: Parábola Editorial, 2008.

MARCUSCHI, L. A. **Da fala para a escrita**. Atividades de retextualização. São Paulo: Cortez, 2001.

MAYER, R. E. **Multimedia Learning**. Cambridge: Cambridge University Press, 2001.

MAYER, R. E.; CLARK, R. C. **E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning**. San Francisco, CA: Jossey-Bass/Pfeiffer, 2003.

NASCENTES, R. M. M.; CONTI, C. L. A.; LIMA, E. F. de. Macro e micro regulações da/na escola: avaliações externas em foco, Universidade Federal de Santa Maria, RS. **Educação - Revista do Centro de Educação**, v. 42, n. 1, p. 99-112, 2017. Disponível em: <https://www.redalyc.org/jatsRepo/1171/117150748008/html/index.html> Acesso em: 15 mar. 2021.

ONG, W. **Orality and Literacy: The Technologizing of the Word**. London and New York: Routledge, 2012 [30th Anniversary of First Edition in 1982, with additional chapters by John Hartley].

PLATAFORMA TAO (versão 3.3.0 usada pelo CAEd). Disponível em: <https://bancodeitens-legado.caeddigital.net/> Acesso em: 15 a 24 ago. 2020.

STREET, B. V. **Literacy in theory and practice**. Cambridge; New York and Melbourne: Cambridge University Press, 1984.

STREET, B. V. What's new in New Literacy Studies? Critical approaches to literacy in theory and practice. **Current Issues in Comparative Education**, Teachers College, Columbia University, v. 5, n. 2, p. 77-91, 2003.



# Tecnologias como suporte para avaliação de fluência em leitura: experiências no âmbito do CAEd/UFJF<sup>1</sup>

*João Victor de Souza<sup>2</sup>*  
*Cristiano Nascimento da Silva<sup>3</sup>*  
*Sarah Matos Rocha Mesquita<sup>4</sup>*  
*Leila Márcia Mafra Martins<sup>5</sup>*  
*Jairo Francisco de Souza<sup>6</sup>*

---

<sup>1</sup> Projeto financiado pela Fundação CAEd/UFJF.

<sup>2</sup> LApIC Research Group, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. joao.souza@ice.ufjf.br

<sup>3</sup> LApIC Research Group, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. cnsilva@ice.ufjf.br

<sup>4</sup> Centro de Políticas Públicas e Avaliação da Educação, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. sarah.mesquita@caed.ufjf.br

<sup>5</sup> Centro de Políticas Públicas e Avaliação da Educação, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. leila@caed.ufjf.br

<sup>6</sup> Departamento de Ciência da Computação, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. jairo.souza@ice.ufjf.br



# 1. Introdução

A avaliação em larga escala é um instrumento que está em constante expansão e tem sido desenvolvido em vários países do mundo. Esse instrumento avaliativo visa proporcionar às redes de ensino um diagnóstico da aprendizagem dos estudantes, para que, a partir desse mapeamento, sejam articuladas políticas públicas e intervenções pedagógicas eficazes na melhoria da educação.

Essas avaliações, que podem ser aplicadas em todo país ou somente nos estados e municípios, são montadas de acordo com uma matriz, cuja elaboração se baseia no currículo estabelecido para cada uma das etapas que o estado ou município planeja avaliar. Nessa matriz constam as habilidades básicas em forma de descritores, que orientam a criação dos itens a serem testados e que irão medir a proficiência, ou seja, o domínio de certos conhecimentos por parte dos estudantes, nas competências mais elementares.

O Centro de Políticas Públicas e Avaliação da Educação (CAEd/UFJF) é uma instituição que promove testes com essas características, além de muitas pesquisas acerca das políticas de educação e da formação de gestores e professores através de seus cursos de pós-graduação. Também tem desenvolvido novas ferramentas para as avaliações de leitura, escrita e, nos últimos três anos, fluência em leitura oral.

A relevância de se avaliar a fluência, habilidade de fundamental importância para a compreensão dos textos lidos pelos estudantes, reside em tomar conhecimento de quais e quantos alunos estão fluentes em leitura.

Diversos estudos, em destaque os desenvolvidos por Cagliari (1992), comprovam que a decodificação precisa e automática das palavras em um texto, desenvolvida já na fase de alfabetização, promove uma interpretação expressiva do material lido, que resulta na compreensão do seu conteúdo. O *National Reading Panel* (NRP), criado pelo Congresso Americano com a finalidade de avaliar a eficácia das variadas estratégias e métodos utilizados para ensinar as crianças a ler, destaca a fluência como um dos pilares para o desenvolvimento da alfabetização e da leitura com compreensão. Para que o estudante leia com fluência, de acordo com Batista (2018), é fundamental que ele automatize o



processo de identificação de palavras, sendo capaz de realizar uma leitura expressiva, que envolva uma adequada atenção aos elementos prosódicos, como entonação, ênfase, ritmo, apreensão de unidades sintáticas.

Segundo Rasinsky (2004), leitores fluentes decodificam as palavras com precisão e automaticamente, sem (ou com o mínimo) uso de seu esforço ou consciência dos recursos que utilizou para tanto. A teoria que apoia esse aspecto da fluência começa com a noção de que os leitores têm recursos atencionais limitados e que se eles gastarem grande parte desses recursos para decodificação de palavras, eles não estarão disponíveis para uso na compreensão.

Desde 2017, o CAEd/UFJF vem promovendo pesquisas para o desenvolvimento de ferramentas que facilitem a elaboração, a montagem, a aplicação e a correção dos testes de fluência em leitura. A automatização dos processos que envolvem cada avaliação é um dos projetos desenvolvidos juntamente com a UFJF. O auxílio computacional proporciona a preparação de testes diversificados, em menos tempo, possibilitando um melhor ajuste na relação entre os itens dentro da própria avaliação, assim como de uma etapa para a outra.

As ferramentas de automatização auxiliam na qualidade dos itens que compõem o teste, pois se gasta menos tempo na elaboração e montagem, desprendendo, assim, mais atenção ao aspecto pedagógico.

Na literatura em Informática na Educação, o uso de Inteligência Artificial, para automatização da avaliação, é verificado em inúmeros domínios, como a avaliação de itens de resposta aberta sobre matemática (LAN et al., 2015), redações (KAKKONEN; MYLLER; SUTINEN, 2006) ou a criação de *feedbacks* individualizados (CAVALCANTI et al., 2020). Em especial no campo da avaliação da fluência em leitura, destacam-se trabalhos que utilizam métodos computacionais para processar o áudio de uma voz gravada digitalmente. Esse método tem como finalidade extrair características da voz que permitam reconhecer as palavras lidas, pausas de sentido utilizadas, bem como entonação, por exemplo. Esse conjunto de técnicas fazem parte de uma subárea da Inteligência Artificial chamada de Reconhecimento Automático de Fala (*Automatic Speech Recognition - ASR*) e seu uso mais frequente no campo da avaliação de fluência está relacionado à avaliação de fluência para aprendizes de segundo idioma (TEJEDOR-GARCÍA et al., 2020; DETEY et al., 2020).



Geralmente, nesse tipo de contexto, o público alvo é o adulto, já fluente em sua língua materna. Infelizmente, existem poucos trabalhos tratando desse tipo de aplicação voltado às crianças e à aplicação em sua língua materna.

A aplicação de reconhecimento de fala para o público infantil possui algumas dificuldades, devido ao fato, principalmente, de ainda estarem em processo de desenvolvimento da fala e da leitura, requerendo adaptações específicas (SHIVAKUMAR; GEORGIU, 2020). Também em avaliação de língua materna, Proença (2018) mostra que existem, na literatura, apenas alguns trabalhos cuja temática é a avaliação e a identificação de problemas na leitura de crianças, de maneira automatizada.

O objetivo deste trabalho é apresentar os avanços na área de avaliação de fluência realizadas no contexto do Centro de Políticas Públicas e Avaliação da Educação (CAEd/UFJF) no que concerne à utilização de abordagens computacionais para avaliação e auxílio na construção de itens. Para isso, o artigo apresenta as tecnologias desenvolvidas pelo CAEd/UFJF em três eixos de trabalho: (1) auxílio computacional para a construção de itens avaliativos; (2) automatização do processo de avaliação em fluência de leitura; (3) democratização do uso de tecnologias de avaliação.

## 2. Fundamentação Teórica

O processo de alfabetização visa proporcionar ao estudante a apropriação do sistema ortográfico e do funcionamento fonológico de sua língua materna. A longa construção desse conjunto de conhecimentos sobre as relações fono-ortográficas da língua amplia as possibilidades de construção desses conhecimentos, em diferentes esferas, permitindo que esse estudante atue com maior autonomia e protagonismo em sua vida social (BNCC, 2018).

Nesse desenvolvimento inicial da escolarização, o estudante é preparado para fazer uma leitura oral de textos de forma fluida. A leitura é uma atividade muito complexa que requer, além da identificação das letras, a decodificação das palavras, o conhecimento dos significados e sua integração sintática. Uma leitura fluente é aquela em que o leitor lê reconhecendo mais rápido as palavras e automatizando algumas estruturas, não cometendo atropelos no ato.



Diversos autores conceituam a fluência em leitura como sendo a associação entre a decodificação e a compreensão leitora. No *Literacy Dictionary* (HARRIS; HODGES, 1995, p. 85), a fluência é “a capacidade de reconhecimento de palavras que leva à compreensão”. Assim sendo, para que o estudante leia com fluência é fundamental que ele automatize o processo de identificação de palavras e seja capaz de realizar uma leitura expressiva, que envolva uma adequada atenção aos elementos prosódicos.

O Centro de Políticas Públicas e Avaliação da Educação (CAEd/UFJF) compreende que a habilidade de ler um texto de forma rápida, suave, despendendo pouco esforço na mecânica necessária à decodificação, é uma meta a ser perseguida nos anos iniciais da escolarização, destacando, assim, a relevância da avaliação da fluência em leitura.

Assim sendo, considera-se que as competências básicas necessárias para uma leitura fluente são: a consciência fonológica, a compreensão do princípio alfabético e a velocidade com que um texto é traduzido em linguagem falada, ou seja, a capacidade de ler um texto adequado a cada faixa etária, com precisão, prosódia e ritmo (PULIEZI, 2015).

A partir dessa concepção e baseados em Maluf e Santos (2017), o teste de fluência do CAEd/UFJF delibera três aspectos essenciais envolvidos em uma leitura fluida:

- (i) a precisão na decodificação da palavra – habilidade de reconhecer ou decodificar as palavras corretamente;
- (ii) a rapidez no reconhecimento das palavras – facilidade com que uma tarefa é realizada;
- (iii) o uso apropriado da prosódia – leitura com expressão apropriada, ritmo e entonação.

Para mensurar esses aspectos, a avaliação em larga escala da fluência em leitura é realizada por meio de um aplicativo, desenvolvido pelo CAEd/UFJF. Essa avaliação é estruturada em três itens: um quadro de palavras, um quadro de pseudopalavras e uma pequena narrativa.

Esses itens possuem critérios pedagógicos, com respaldo em Coscarelli (2002), que os adequam à etapa avaliada, sendo eles: (i) a complexidade silábica das palavras; (ii) o tamanho da palavra; (iii) o número de ocorrências da palavra na língua; (iv) a familiaridade do leitor com as palavras; (v) a probabilidade de uma palavra aparecer em determinado contexto sintático, semântico e pragmático e (vi) a ambiguidade lexical. Além disso, ao pensarmos nos fatores de processamento



lexical, é preciso destacar as correspondências múltiplas na relação grafema e fonema, existentes na Língua Portuguesa. E, por isso, com base nos estudos de Lemle (1980), nesse tipo de avaliação, são consideradas: (i) as variações do /s/, /z/, /r/, /x/; (ii) a concorrência sonora do /g/ e /j/; (iii) a concorrência do /u/ e /l/, em final de sílaba; (iv) /h/ em início de palavra; (v) as variações dos dígrafos /lh/ /nh/ /ch/; (vi) a concorrência sonora do /k/ e /q/ e (vii) a ocorrência do /ç/. Assim sendo, com base em Coscarelli (2002), o presente teste propõe-se a avaliar as palavras a partir de sua complexidade e extensão silábica, levando sempre em consideração as regularidades e as irregularidades presentes no sistema de escrita ortográfico da Língua Portuguesa, como destacadas por Lemle (1980).

A narrativa, terceiro item do teste, é selecionada em corroboração com relatório da Pesquisa Aplicada em Avaliação de Língua Portuguesa do CAEd/UFJF (2016-2019) cuja complexidade textual é determinada por diversas características do texto, tais como: extensão, sintaxe, coesão, estrutura do gênero, familiaridade da temática e tipologia. Essas características permitem a classificação da complexidade dos textos em quatro diferentes níveis, apresentados em Barbosa; Micarello; Ferreira (2018). Conforme as autoras, textos com nível 1 de complexidade são aqueles pouco extensos (até 15 linhas), de sintaxe simples, em que predominam períodos curtos e estruturas coordenadas ou léxico mais próximo do coloquial e adequado à criança, além de serem compostos por temas cotidianos. Por sua vez, textos com nível 2 de complexidade possuem sintaxe mais complexa, em relação ao nível 1, com pequenas alterações na estrutura sintática canônica, léxico mais usual ou próprio do discurso formal.

Durante o processo de aplicação do teste, o avaliador grava a leitura dos itens feita pelo estudante. Cada leitura de item é equivalente a um áudio de até um minuto. Desse modo, a leitura de cada estudante é avaliada a partir de um total de três áudios – um por item. Os áudios são sincronizados do celular do aplicador para uma nuvem do banco de dados do CAEd/UFJF, sendo a partir do *upload* desses áudios que o processo de correção se inicia.

A equipe do Centro de Políticas Públicas e Avaliação da Educação seleciona, criteriosamente, os corretores e treina-os, com o respaldo em uma chave de correção, para que as correções dos áudios sejam iniciadas.



Ao final do processo avaliativo, o teste, pautado nos aspectos estatísticos, psicométricos e pedagógicos fornece uma alocação dos estudantes em perfis (pré-leitor, iniciante e fluente), cujas características descrevem as habilidades desenvolvidas por eles até o momento da avaliação.

### 3. Avaliação automática

A avaliação de fluência em leitura gera um grande volume de áudios que são avaliados por especialistas. A automatização do processo de avaliação desses áudios visa permitir consistência, rapidez e economia financeira para a geração dos relatórios da avaliação. Contudo, a automatização dessa avaliação não é uma tarefa trivial e envolve avanços na pesquisa em Inteligência Artificial e Processamento de Sinais. Nesse sentido, foram criados modelos de Inteligência Artificial para a utilização de algoritmos de reconhecimento automático de fala nos áudios registrados por aparelhos celulares (SOARES et al., 2018). Esses algoritmos são responsáveis por analisar os arquivos de áudios de leitura e extrair as características utilizadas para a correção.

O funcionamento desses algoritmos é baseado em modelos probabilísticos que representam os elementos acústicos, fonéticos e gramaticais do padrão de leitura do cenário em que o sistema deverá ser utilizado. Tais modelos podem ser divididos em três tipos: acústico, léxico e linguagem. O modelo acústico, de cunho probabilístico, é responsável por mapear as ondas sonoras e identificar quais são os fonemas, que são as unidades sonoras que compõem a língua, presentes em cada instante do áudio. O mapeamento do vocabulário com sua transcrição fonética é representado pelo modelo léxico. Por sua vez, o modelo de linguagem é representativo da gramática de entrada, estabelecendo as relações possíveis entre as palavras do vocabulário.

Esses modelos requerem o processamento de conjuntos específicos de dados de observação para a utilização no problema, com a finalidade de produzir as probabilidades adequadas para os modelos acústicos, léxico e de linguagem com base nas observações recebidas. Na área de Computação, esse processo é denominado fase de treinamento dos modelos. Após a fase de treinamento, os modelos são capazes de reconhecer os padrões que devem ser identificados quando o sistema for lidar com



novos arquivos de áudio. No caso em tela, diversas gravações de leituras das crianças, o conteúdo dos itens avaliativos e os resultados dos corretores de avaliações anteriores foram utilizados como base para a criação dos modelos acústico e léxico.

O modelo de linguagem, que representa a gramática, é criado durante o momento em que o sistema avalia o áudio, baseado no item que consta na avaliação. Ao invés de criar uma gramática que descreve por completo a Língua Portuguesa, é criada uma gramática que representa apenas a possibilidade correta de leitura do item. Essa estratégia é possível devido ao fato de o item já ser conhecido previamente, fazendo com que o sistema tente comparar o conteúdo do item com o conteúdo do áudio, impedindo, assim, que léxicos não contidos no item sejam identificados. Para essa técnica, dá-se a nomenclatura de alinhamento forçado (GOMES JR et al. 2019).

Inicialmente, o processamento automático foi implementado para as narrativas e preparado para extrair os dados da avaliação cuja finalidade é determinar o perfil leitor do estudante avaliado. Nesse tipo de item, são consideradas a última palavra lida pelo estudante e a precisão de sua leitura ao longo do texto. Esse processo é dividido em duas etapas, a primeira delas destinada ao alinhamento do item para descobrir a última palavra lida ou a quantidade de palavras lidas (QPL) e, em seguida, um segundo alinhamento para descobrir quais palavras foram lidas corretamente na leitura (quantidade de palavras corretas - QPC). Em suma, são considerados: a quantidade de palavras lidas, determinada pela numeração constante na última palavra lida, e a quantidade de palavras lidas corretamente, dentre as lidas. Após esse processo de identificação, é realizada a razão QPC/QPL para identificar a precisão da leitura. A figura 1 a seguir representa uma visão geral das etapas realizadas no processo automático.

**Figura 1. Fluxograma do processamento automático de áudios**



Fonte: Elaboração dos autores (2021).

Com a extração desses dados, foi possível fazer uma avaliação inicial em dois perfis, os estudantes fluentes e os não-fluentes, utilizando o mesmo limiar de classificação que os corretores humanos. Ao compararmos a classificação dos perfis gerados pelos dados do algoritmo com a realizada pelos corretores, foi possível identificar uma alta taxa de concordância.

Aqui, apresentamos o resultado de experimentos realizados com dados de quatro avaliações em larga escala realizadas pelo CAEd/UFJF: Prefeitura de Teresina (SAETHE); Programa de Avaliação da Educação Básica do Espírito Santo (PAEBES); Programa Mais Alfabetização (PMALFA) e Programa de Alfabetização em Regime de Colaboração (PARC) cuja abrangência foi nacional.

Todas essas avaliações tiveram seus áudios de leitura das narrativas processados pelo sistema automático, proporcionando a extração dos resultados. Em seguida, foram aplicados filtros para remover os áudios: (i) que possuem problemas com os dados dos corretores; (ii) com problemas nos arquivos ou que não puderam ser processados; (iii) muito próximos ao limiar de classificação, no qual pequenos erros podem causar uma mudança de perfil; e (iv) *outliers*, resultados cuja diferença entre o QPL e o QPC se distanciam muito do padrão dos dados, determinado por uma fórmula estatística para indicar o ponto de corte.

Por fim, foram feitos testes para avaliar a performance do algoritmo com os dados restantes da base. Tendo como respaldo os dados dos corretores humanos, foram feitos comparativos para verificar se o perfil atribuído ao estudante pelo algoritmo obtinha o mesmo resultado. Caso isso ocorresse, ele era considerado como um acerto do algoritmo (verdadeiro positivo). A tabela 1 mostra os resultados obtidos pelo processamento automático, em comparação com o total restante após a aplicação dos filtros.

**Tabela 1. Acurácia do processamento automático pelo total da base avaliada**

Avaliação	Taxa de acerto	Total avaliado
SAETHE	96%	80%
PAEBES	98%	83%
PMALFA	98%	84%
PARC	94%	89%

Fonte: Elaboração dos autores (2021).



Os experimentos realizados mostram que a classificação automática corrobora, na maior parte das vezes, com o mesmo perfil determinado pelos corretores manuais. Para completar o processo de avaliação automática do teste de fluência, cuja composição é de três itens: quadro de palavras, quadro de pseudopalavras e uma narrativa, estão sendo criadas soluções para avaliar automaticamente, também, os itens de quadro de palavras e de pseudopalavras.

## 4. Suporte à criação de itens

No contexto da avaliação de fluência, promovida pelo CAEd/UFJF, para cada caderno de teste, são elaborados três tipos de itens: um quadro de palavras, um quadro de pseudopalavras e uma pequena narrativa. Cada item varia sua dificuldade de acordo com o ano letivo que receberá a avaliação da leitura. Os itens de palavras e de pseudopalavras variam em termos de quantidade e de complexidades linguísticas que determinam a seleção dos vocábulos. Já o item de texto varia de acordo com o tema, a extensão, a sintaxe e o léxico que determinam seu nível de complexidade (BARBOSA; MICARELLO; FERREIRA, 2018).

A criação de cada item requer muito tempo e criatividade, uma vez que, para montar o quadro de palavras, é preciso verificar as diversas palavras da Língua Portuguesa e selecionar as que se adéquam à etapa de escolarização avaliada. No quadro de pseudopalavras, há uma maior complexidade no processo de elaboração, visto que é necessário garantir que tais palavras inventadas sejam, de fato, inexistentes. No item de texto, sua seleção requer que haja correspondência entre o nível de complexidade textual e a etapa avaliada. Durante o processo de seleção do texto, além da complexidade, o CAEd/UFJF preocupa-se com a manutenção da originalidade e com a escolha de uma temática que não seja corriqueira ao trabalho em sala-de-aula. Assim sendo, a única intervenção que o Centro de Políticas faz nessas narrativas é a fragmentação de alguns trechos.

Visando facultar esse processo de montagem dos cadernos, foram criadas ferramentas computacionais para assistir o elaborador na proposição do quadro de palavras e de pseudopalavras. Além disso, criou-se uma ferramenta que auxilia o elaborador a determinar qual o nível de complexidade de determinado texto.



A figura 2 representa a tela da ferramenta utilizada para auxiliar na criação de um quadro de palavras. Essa ferramenta permite que o usuário selecione a extensão das palavras que serão buscadas em um *corpus* da Língua Portuguesa, suficientemente grande, pré-processado pela solução, e outras características linguísticas que determinam a complexidade desses léxicos, tais como: canonicidade, classe gramatical, tonicidade e presença de grafemas e fonemas.

Um dificultador da criação dos quadros de palavras é a identificação e não repetição de palavras com diferentes níveis de dificuldade para o estudante ler, sendo definida pela frequência do uso da palavra na língua. Na solução desenvolvida, a frequência da palavra é dada pela sua recorrência no *corpus*. As palavras identificadas são contabilizadas e divididas em sextis de acordo com sua frequência, de forma que quanto menor o valor do sextil, maior a frequência na língua portuguesa.

Para cada característica de correspondência de grafema e fonema, assim como a de canonicidade, há três opções de seleção: “Presente”, “Não Presente” e “Indiferente”. A princípio, todas as correspondências estão indicadas pela opção “Indiferente”, pois essa opção não afeta as palavras que aparecerão, podendo ou não possuírem a característica em questão. A opção “Presente” indica que todas as palavras buscadas possuirão essa característica; já a “Não Presente” determina que nenhuma das palavras buscadas terá a característica selecionada.

Uma vez selecionadas as características - sua presença, ausência ou indiferença -, o sistema gera uma lista randômica de palavras existentes no *corpus* para compor o quadro. O quadro em que as palavras aparecem, conforme mostrado na figura 2, é temporário, pois, ao solicitar nova busca, as palavras que estão nele serão substituídas por outras aleatórias. Assim sendo, para a permanência de palavras selecionadas na composição do quadro final de palavras, deve-se selecionar a seta no canto superior direito de cada palavra desejada, de forma a transferi-la para o quadro ao lado no qual as palavras poderão ser reordenadas.

O ícone no canto superior esquerdo do quadro de permanência das palavras tem o objetivo de informar a quantidade de palavras escolhidas em cada característica, como: quantidade de palavras canônicas e não canônicas; quantidade de palavras oxítonas e de outras tonicidades; e quantidade de palavras em cada correspondência de grafema e fonema. Após selecionadas todas as palavras desejadas para compor o quadro, o ícone no canto superior direito do quadro permite o *download* do quadro como um arquivo em formato CSV.



**Figura 2. Tela de geração do quadro de palavras**

**Quadro de Palavras**

Tamanho do Quadro: 10 GERAR

Classes Gramaticais  
 Verbo  Adjetivo  Substantivo  Outros

Tonicidade  
 Oxítona  Paroxítona  Proparoxítona

Extensão  
 Monossílaba  Dissílaba  Trissílaba  Polissílaba

Frequência  
 Indiferente

Canonicidade  
 Indiferente

Grafema Ç  
 Indiferente

Grafema H no começo  
 Indiferente

Fonema /s/  
 Indiferente

Fonema /z/  
 Indiferente

Fonema /r/  
 Indiferente

Fonema /rr/  
 Indiferente

Quando menor o sexto, maior a frequência

Fonema /x/  
 Indiferente

Fonema /g/  
 Indiferente

Fonema /j/  
 Indiferente

Fonema /k/  
 Indiferente

/u/ no final da sílaba  
 Indiferente

/l/ no final da sílaba  
 Indiferente

Grafema NH  
 Indiferente

Grafema LH  
 Indiferente

Grafema CH  
 Indiferente

sai → brigão → covil → cão

pedrao → pitêu → gestão → ganho

pistom → vial

Fonte: Elaboração dos autores (2020).

Uma tela similar a essa é utilizada para criação dos quadros de pseudopalavras. Para confecção dos itens com essa característica, a solução computacional desenvolvida no CAEd/UFJF é capaz de gerar palavras inexistentes na Língua Portuguesa de maneira aleatória cujas características são definidas pelo operador do sistema: extensão, canonicidade, tonicidade e presença de certos fonemas ou grafemas.

A geração das pseudopalavras utiliza uma palavra-semente, que é uma palavra existente na Língua Portuguesa, escolhida aleatoriamente de acordo com as características selecionadas pelo usuário. A partir dessa palavra-semente, uma série de operações é aplicada para desconstruir a palavra-semente original e gerar opções de pseudopalavras para a composição do quadro de pseudopalavras. É importante destacar que a frequência da palavra-semente é importante nesse processo para gerar pseudopalavras mais fáceis de serem pronunciadas, uma vez que serão geradas a partir de palavras mais fáceis de serem pronunciadas.

As pseudopalavras são geradas através de dois métodos. O método 1 é a junção de sílabas de palavras-semente, selecionadas aleatoriamente, de forma a manter a sílaba na mesma posição em que se encontrava na palavra-semente da qual ela veio. Isso faz com que sílabas do começo de palavras sejam alocadas no começo da nova pseudopalavra. Após a geração dessas pseudopalavras com aspectos morfológicos da palavra de origem, é aplicado um filtro de acordo com as características

selecionadas para apresentação de apenas pseudopalavras válidas, o que faz com que a quantidade de pseudopalavras geradas pelo método 1 tenda a ser menor. O método 2 elege pseudopalavras existentes, que cumpram as características selecionadas, conforme feito no quadro de palavras, mas a partir da alteração de letras ou de sílabas da palavra-semente.

No quadro 1, é possível observar as pseudopalavras geradas a partir de cada método, de acordo com as características escolhidas, sendo as características selecionadas como “Indiferente”, não mostradas na tabela. Dado que uma característica selecionada é com relação a extensão ser trissilábica, cada pseudopalavra gerada pelo método 1 provém de 3 (três) palavras selecionadas aleatoriamente, uma para cada sílaba. A tabela mostra também, no método 2, por exemplo, que a palavra-semente “Barato” originou duas pseudopalavras “Barado” e “Bireto”, confirmando que há, nessas palavras inventadas, um padrão morfológico com relação à palavra de origem. Em “Barado”, temos a manutenção da posição de todas as vogais e substituição do fone /t/ pelo /d/. Na pseudopalavra “Bireto”, todas as consoantes são mantidas, conforme a palavra-semente, além da vogal final, muito embora a vogal inicial e a medial tenham sido substituídas.

### Quadro 1. Resultados da geração de pseudopalavras

Características			
Extensão	Trissílaba	Grafema H no começo	Não presente
Frequência	Sextil 1	Fonema /r/	Presente
Canonicidade	Presente	Fonema /x/	Não presente
Método 1			
Pseudopalavras:	Parami	Nemaro	Fetura
Método 2			
Palavras-semente:	Barato	Perito	Buraco
Pseudopalavras:	Barado	Poroto	Muraco
	Bireto	Cerito	Burano

Fonte: Elaboração dos autores (2021).

Para apoiar a criação de narrativas dentro dos parâmetros de complexidade exigidos para a avaliação de fluência, um conjunto de estatísticas é gerado a partir de uma proposta de texto para que possa ser feita uma análise da dificuldade linguística de sua composição, de acordo com Barbosa;



Micarello; Ferreira (2018). Ainda de acordo com as autoras, seriam definidores da complexidade textual: frequência de classes gramaticais, palavras distintas, canonicidade, tonicidade, extensão, número de sentenças, entre outros. Dessa forma, foi confeccionado um sistema web para analisar algumas das variáveis definidoras da complexidade da narrativa durante a sua confecção, o que é ilustrado na figura 3.

**Figura 3. Tela de análise da complexidade do texto com o texto “A cigarra e a formiga”**

### Análise da Complexidade do Texto

<p><b>A CIGARRA E A FORMIGA</b>                  Era uma vez uma cigarra que vivia saltitando e cantando pelo bosque, sem se preocupar com o futuro. Esbarrando numa formiguinha, que carregava uma folha pesada, perguntou:                  - Ei, formiguinha, para que todo esse trabalho? O verão é para gente aproveitar! O verão é para gente se divertir!                  - Não, não, não! Nós, formigas, não temos tempo para diversão. É preciso trabalhar agora para guardar comida para o inverno.                  Durante o verão, a cigarra continuou se divertindo e passeando por todo o bosque. Quando tinha fome, era só pegar uma folha e comer. Um belo dia, passou de novo perto da formiguinha carregando outra pesada folha.                  A cigarra então aconselhou:                  - Deixa esse trabalho para as outras! Vamos nos divertir. Vamos, formiguinha, vamos cantar! Vamos dançar!                  A formiguinha gostou da sugestão. Ela resolveu ver a vida que a cigarra levava e ficou encantada. Resolveu viver também como sua amiga. Mas, no dia seguinte, apareceu a rainha do formigueiro e, ao vê-la se divertindo, olhou feio para ela e ordenou que voltasse ao trabalho. Tinha terminado a vidinha boa.                  A rainha das formigas falou então para a cigarra:                  - Se não mudar de vida, no inverno você há de se arrepender, cigarra! Vai passar fome e frio.                  A cigarra nem ligou, fez uma reverência para rainha e comentou:                  - Hum!! O inverno ainda está longe, querida! Para cigarra, o que importava era aproveitar a vida, e aproveitar o hoje, sem pensar no amanhã. Para que construir um abrigo? Para que armazenar alimento? Pura perda de tempo.                  Certo dia o inverno chegou, e a cigarra começou a tritar de frio. Sentia seu corpo gelado e não tinha o que comer. Desesperada, foi bater na casa da formiga. Abrindo a porta, a formiga viu na sua frente a cigarra quase morta de frio.                  Puxou-a para dentro, agasalhou-a e deu-lhe uma sopa bem quente e deliciosa.                  Naquela hora, apareceu a rainha das formigas que disse à cigarra:                  - No mundo das formigas, todos trabalham e se você quiser ficar conosco, cumpra o seu dever: toque e cante para nós. Para cigarra e paras formigas, aquele foi o inverno mais feliz das suas vidas.</p>	<p><b>Estatísticas Gerais:</b>                  Total de Palavras: 358                  Frases: 34                  Parágrafos: 16                  Sinais de Pontuação: 74</p> <p><b>Palavras Canônicas: 67</b>                  Dissílaba: 36                  Monossílaba: 25                  Trissílaba: 6</p> <p><b>Palavras Não Canônicas: 291</b>                  Monossílaba: 105                  Trissílaba: 57                  Dissílaba: 99                  Polissílaba: 30</p> <p><b>Tonicidade:</b>                  Oxítona: 187                  Paroxítona: 171</p>	<p><b>Correspondências:</b>                  Grafema Ç: 2                  Grafema H no início: 4                  Grafema SS: 7                  Grafema NH: 14                  Grafema LH: 12                  Grafema CH: 1                  Fonema /s/: 82                  Fonema /z/: 18                  Fonema /f/: 56                  Fonema /r/: 82                  Fonema /x/: 2                  Fonema /g/: 0                  Fonema /j/: 46                  Fonema /k/: 43                  /ul/ no final da sílaba: 40                  /l/ no final da sílaba: 4</p>
---	--	---

GERAR

Obs.: No caso das reticências, elas são computadas nos sinais de pontuação, mas não no aspecto da frase. Só são consideradas como um final de frase se forem final de parágrafo também.

Fonte: Elaboração dos autores (2020).

## 5. Suporte computacional à sala de aula

Uma das frentes de trabalho na avaliação de fluência é a democratização da avaliação por meio da oferta de ferramentas para utilização dos docentes. Nesse sentido, entende-se como relevante o acesso à metodologia e às tecnologias usadas para avaliação de fluência em leitura através de pequenos pacotes de *software* que possam ser facilmente utilizados em sala de aula. Assim sendo, destaca-se a importância de discutir a primeira solução desenvolvida nesta frente de trabalho: um jogo computacional para celulares que permite a avaliação de leitura, a partir de quadros de palavras.

Para projetá-lo, foi necessário apropriar-se do conceito de gamificação, na expectativa de que a avaliação em fluência passasse de um processo formal e exaustivo para uma atividade lúdica, que prenda a atenção dos estudantes e os motive (SILVA et al., 2019). Essa gamificação como intervenção pedagógica recebeu o nome de *Space Gems*.

O jogo possui um cenário contextualizado em um futuro longínquo no qual megacorporações precisam movimentar itens entre planetas de uma galáxia distante. Os jogadores são colocados como pilotos das naves que precisam coletar e entregar pedras preciosas entre os planetas, definindo as melhores rotas de percurso. A cada entrega, os jogadores garantem pontos para suas equipes, ganhando ao final a equipe que obtiver mais pontos. O jogo é baseado no estilo *pickup and delivery*, em que cada jogador, controlando uma nave, pode escolher entre planetas com algumas pedras preciosas e planetas que precisam delas, coletando os recursos, quando disponíveis, ou trocando por pontos, quando houver planetas com demanda por uma pedra preciosa contida na nave. Para fazer uma entrega, os jogadores podem enfrentar obstáculos e estabelecer estratégias para encontrar a rota mais lucrativa, o que estimula a cooperação entre membros do mesmo grupo em busca do objetivo final.

A abordagem utilizada opera em três estágios distintos: captação dos áudios, processamento e geração dos resultados. Em um primeiro momento, o docente precisa definir o contexto de configuração do jogo, escolhendo as palavras desejadas, cadastrando os alunos participantes e definindo a duração da partida em função do número de rodadas. Com a configuração finalizada, o sistema distribui os alunos em grupos aleatórios e gera os artefatos necessários para o início da partida. Para interagir com o jogo, os jogadores podem utilizar um *smartphone* compartilhado, fornecido pela instituição ou eleito por um dos estudantes.

Caso haja disponibilidade, outros *smartphones* podem ser utilizados para a interação dos jogadores, acompanhando o andamento das equipes. A cada rodada, um aluno é nomeado pelo próprio jogo para interagir ativamente, fornecendo os comandos para o direcionamento de sua nave. A cada movimento para outro planeta, o estudante deverá ler uma palavra escolhida pelo sistema, dentre aquelas que foram configuradas pelo aplicador, para realizar o movimento.

Nesse momento, o áudio é captado e enviado ao sistema para ser armazenado e, posteriormente, avaliado. Esse processo se repete três vezes dentro de uma mesma rodada, promovendo a gravação de três leituras de palavras. Ao finalizar as jogadas, ocorre a mudança de estudante ativo, aquele escolhido pelo sistema para dar prosseguimento ao jogo. Esse processo ocorre até que todos os estudantes da equipe tenham participado.

Após as jogadas serem concluídas, o sistema calcula o resultado final e estabelece a pontuação de cada uma das equipes, baseando-se somente na entrega de recursos entre os planetas.



É importante ressaltar que todos os jogadores devem ler o mesmo número de palavras e nenhum deles é penalizado ou eliminado, caso a leitura não possua uma pronúncia adequada. Os áudios com as leituras dos estudantes são processados no serviço de reconhecimento de fala e as estatísticas com as leituras da partida são geradas. Todas as informações sobre a avaliação das leituras, como taxas de acerto e indicação sobre fonemas pronunciados incorretamente, são disponibilizadas somente após o final da partida para o professor.

Com isso, é possível manter todos os alunos engajados durante o processo de avaliação, seja enquanto controlam ativamente a nave ou enquanto se reúnem com seus colegas de equipe para traçar estratégias de percurso. A abordagem empregada não só evita desconfortos entre os alunos, não os penalizando por leituras erradas - não eliminando jogadores ou exibindo a pontuação antes do final da partida - mas também promove a cooperação entre os jogadores que possuem um objetivo em comum. Além disso, a solução garante que o processo de avaliação não seja exaustivo, nem para o professor, nem para os estudantes e, ainda, fornece um meio para os avaliadores verificarem o desempenho de seus alunos na leitura de palavras.

## 6. Conclusões e trabalhos futuros

O presente trabalho apresentou o entendimento do processo de avaliação de fluência em leitura pelo CAEd/UFJF e os avanços computacionais para melhoria do processo de confecção de itens e processamento dos áudios para geração de resultados automatizados das avaliações em larga escala.

A solução para processamento automático dos áudios de leitura da narrativa apresenta resultados satisfatórios, correspondendo, em sua maioria, à classificação realizada por corretores manuais. Diferentes desafios encontrados por esse tipo de avaliação foram tratados ao longo dos anos de pesquisa, como a garantia de qualidade do resultado gerado automaticamente para áudios com presença de variedades linguísticas, leituras com presença excessivas de pausas (leituras silabadas), identificação de hesitações na leitura de palavras (falsos começos), entre outros, o que torna efetiva a sua utilização.



Atualmente, o processamento da leitura do texto fornece informações sobre quantidade de palavras lidas e quantidade de palavras lidas corretamente, mas uma melhora no processamento do texto permitirá que outras informações úteis também sejam extraídas, como as relacionadas à prosódia, de forma a explicitar o ritmo, as pausas de sentido, as acentuações e a entonação de cada palavra lida.

Tendo em vista o bom resultado no caso de leitura de textos, atualmente, as frentes de trabalho estão focadas no processo de automatização da avaliação de leitura dos quadros de palavras e pseudopalavras, cujo objetivo é alcançar um bom resultado também nesses itens. A leitura dos itens de palavras e pseudopalavras possui peculiaridades que fazem com que o modelo usado para avaliação automática da leitura da narrativa não seja muito efetivo, visto que elas possuem uma ocorrência muito maior de falsos começos e de pausas intrapalavras, sendo essas hesitações e pausas mais duradouras do que as presentes na leitura da narrativa.

Para os próximos passos, será importante conseguir identificar e avaliar a prosódia durante a leitura dos itens, que inclui a observação de pausas de sentido, acentuações e entonações. Para a avaliação da qualidade da leitura de uma determinada palavra, faz-se necessária uma forma de identificar se os fonemas esperados de serem pronunciados estão presentes e corretos. Com uma avaliação adequada da prosódia, será possível fornecer um resultado mais adequado do QPC, de forma a produzir uma classificação completa do nível de fluência de leitura de um estudante.

Como mencionado, a gamificação foi uma estratégia pensada para permitir aos próprios docentes aplicarem uma avaliação de forma lúdica e receberem um relatório sobre os resultados de cada participante. Isso permite um acompanhamento de perto da evolução dos estudantes ao longo do ano letivo, assim como uma avaliação da própria didática utilizada, dado que o desenvolvimento e resultados desses estudantes refletirão isso, de forma que as intervenções pedagógicas sobre os métodos educacionais possam ser alteradas com mais rapidez e obtenham maior eficácia.



## Referências

- BARBOSA, B. T.; MICARELLO, H. A. L. da S.; FERREIRA, R. V. J. **Relatório semestral da pesquisa CAED 2016-2019 Língua Portuguesa 5º ao 12º ano 1º e 2º semestres de 2018**. Juiz de Fora/MG: CAEd/UFJF, 2019.
- BATISTA, A. A. G. Alfabetização, leitura e ensino de português: desafios e perspectivas curriculares. **Revista Contemporânea da Educação**. Rio de Janeiro, v. 6, n. 12, agosto/dezembro 2011. Disponível em: <<https://revistas.ufrj.br/index.php/rce/article/view/1638/1486>>. Acesso em: 11 fev. 2021.
- BRASIL. **Base Nacional Comum Curricular: Ensino Médio**. Brasília: MEC/Secretaria de Educação Básica, 2018.
- CAGLIARI, L. C. Da importância da prosódia na descrição de fatos gramaticais. In: ILARI, I. **Gramática do Português Falado: níveis de análise linguística**. Campinas: Editora da UNICAMP, 1992.
- CAVALCANTI, A. P. et al. **Análise Automática de Feedback em Ambientes de Aprendizagem Online**. Anais do XXXI Simpósio Brasileiro de Informática na Educação. SBC, 2020. p. 892-901.
- COSCARELLI, C. V. Entendendo a leitura. **Revista de Estudos da Linguagem**. Belo Horizonte: UFMG. v. 10. n. 1, p.7-27, jan/jun. 2002.
- DETEY, S. et al. Computer-assisted assessment of phonetic fluency in a second language: a longitudinal study of Japanese learners of French. **Speech Communication**, v. 125, p. 69-79, 2020.
- GOMES JR., J. et al. Uso de Alinhadores Forçados para Avaliação Automática em Larga Escala da Fluência em Leitura. **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2019. p. 61.
- HARRIS, T. L.; HODGES, R, A. **The literacy dictionary**. Newark, DE: International reading Association, 1995.
- KAKKONEN, T.; MYLLER, N.; SUTINEN, E. Applying latent Dirichlet allocation to automatic essay grading. **International Conference on Natural Language Processing**. Springer, Berlin, Heidelberg, p. 110-120, 2006.
- LAN, A. S. et al. Mathematical language processing: Automatic grading and feedback for open response mathematical questions. **Proceedings of the Second ACM Conference on Learning @ Scale**, p. 167-176, 2015.
- LEMLE, M. **Guia teórico do alfabetizador**. São Paulo: Ática, 1990.
- MALUF, M. R.; SANTOS, M. J. dos. (Org.) **Ensinar a ler: das primeiras letras à leitura fluente**. Curitiba: CRV, 2017.



PIKULSKI, J. J.; CHARD, D. J. **Fluency: Bridge Between Decoding and Reading Comprehension**. First published: 09 nov. 2011. Disponível em: <<https://ila.onlinelibrary.wiley.com/doi/abs/10.1598/RT.58.6.2>>. Acesso em: 11 fev. 2021.

PULIEZI, S. **Fluência e compreensão na leitura de textos: um estudo com crianças do 4º ano do ensino fundamental**. 157f. 2015. Tese (Doutorado em Educação: Psicologia da Educação) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2015.

PROENÇA, J. D. L. **Automatic assessment of reading ability of children**. Tese de Doutorado. Universidade de Coimbra, 2018.

RASINSKY, T. V. **Assessing Read Fluency**. Ed. PREL - Pacific Resources for Education and Learning, 2004.

SHIVAKUMAR, P. G.; GEORGIU, P. Transfer learning from adult to children for speech recognition: Evaluation, analysis and recommendations. **Computer speech & language**, v. 63, p. 101077, 2020.

SILVA, W. A. et al. Talk2Me: Uma abordagem computacional para auxiliar na identificação de falhas no processo de alfabetização. **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2019. p. 723.

SOARES, E. et al. Avaliação automática da fluência em leitura para crianças em fase de alfabetização. **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2018. p. 11.

TEJEDOR-GARCÍA, C. et al. Assessing pronunciation improvement in students of English using a controlled computer-assisted pronunciation tool. **IEEE Transactions on Learning Technologies**, v. 13, n. 2, p. 269-282, 2020.



# e-VAL: um protótipo para análise e avaliação de níveis de interatividade em itens digitais<sup>1 2</sup>

*Liamara Scortegagna<sup>3</sup>*

*Adriana Rocha Bruno<sup>4</sup>*

## 1. Introdução

Um processo de avaliação da aprendizagem envolve complexa teia de elementos e ações cujo foco principal deveria ser produzir diagnósticos e, a partir destes, produzir e oferecer possibilidades para a aprendizagem. Entretanto, normalmente, está relacionado à produção de informações sobre determinada realidade, e docentes, por meio de instrumentos de avaliação, devem aferir o aprendizado de seus estudantes e indicar, a partir daí, o que precisa ser feito para que eles tenham condições de avançar no processo de aprendizagem.

Além da avaliação da aprendizagem, escolas, secretarias estaduais e municipais de educação desenvolvem o que é denominado de avaliações externas, que possuem objetivos e procedimentos diferentes das avaliações realizadas pelos professores nas salas de aula. Essas avaliações são desenvolvidas e aplicadas por outras instituições e em larga escala, com o objetivo de apresentar, normalmente, um diagnóstico que é utilizado de modo a classificar estudantes quanto ao seu

---

<sup>1</sup> Projeto financiado pela Fundação CAEd/UFJF.

<sup>2</sup> Texto original publicado na Revista Pesquisa e Debate em Educação, v.12, n.1, (2022), página 1-26, e36983, disponível em: <https://periodicos.ufjf.br/index.php/RPDE/article/view/36983>.

<sup>3</sup> Departamento de Ciência da Computação, Universidade Federal de Juiz de Fora, e-mail: [liamara@ice.ufjf.br](mailto:liamara@ice.ufjf.br).

<sup>4</sup> Escola de Educação/Departamento de Didática, Universidade Federal do Estado do Rio de Janeiro, e-mail: [arbruno2208@gmail.com](mailto:arbruno2208@gmail.com).



conhecimento em determinadas áreas e por meio de instrumentos específicos. Assim, são aplicadas de forma padronizada e para um grande número de pessoas. Esses tipos de avaliações, além de mapearem conhecimentos “basilares” desejados e apresentados em documentos produzidos por e para políticas públicas, muitas vezes ranqueiam escolas, estudantes, docentes e condicionam, em grande parte, o recebimento de verbas e benefícios. Porém, compreendemos que tais instrumentos são importantes *feedbacks* para que professores e escolas possam agir em prol de mudanças necessárias que possam se reverter em melhorias para as aprendizagens.

Com as inovações tecnológicas, especialmente com os dispositivos digitais e em rede da Cultura Digital, as possibilidades de automatização no desenvolvimento, aplicação e correção de itens passaram a habitar o campo das avaliações, ampliando recursos para avaliar a aprendizagem dos alunos e/ou, naquelas em larga escala, para apresentar diagnósticos. Esse novo cenário tem transformado os contextos avaliativos, fomentando mudanças que implicam a digitalização do formato impresso e, gradativamente, a produção de itens digitais e interativos em meio digital.

Diante disso, os itens, ou seja, as questões que compõem essas avaliações, também estão passando por um processo de intensa transformação, deixando de ser estáticos e lineares, vindo a apresentar algum tipo de interatividade.

Pautados num paradigma em que a presencialidade é fundamental para a aprendizagem e para a avaliação, vivemos o conflito provocado por essa crença, fundamentado em referências passadas. Diversos estudos (SILVA; SANTOS, 2006; FILATRO, 2004; PETERS, 2003; SILVA, 2003; PALLOFF; PRATT, 2002; MORAES, 1997; BELLONI, 2001, entre outros) apontam para a necessidade de modelos híbridos de educação, em que coexistam possibilidades diversas para o ensino e a aprendizagem e que rompam com o paradigma da presencialidade como possibilidade única para a aprendizagem.

Nessa direção, propomos que as perspectivas de avaliação, aqui incluída a avaliação em larga escala, possam avançar para modelos híbridos, ainda que essa seja uma mudança em médio-longo prazos.

Para que possamos conhecer e entender a diferença entre os tipos de itens encontrados nas avaliações, apresentamos a classificação descrita por Scortegagna (2020) e Bruno (2019b), em que constam: i) Item impresso: desenvolvido para ser respondido em papel impresso, apresenta questões com alternativas em múltiplas escolhas, sendo apenas uma delas correta; traz questionamento(s) em seu enunciado, podendo também conter imagens e/ou informações adicionais para o respondente;



a correção é realizada por pessoas/especialistas, manualmente; ii) Item digital: itens em formato digital, sem que tenham sido concebidos nesse formato; comumente se dá pela transferência de itens desenvolvidos do papel impresso para o formato digital, diferenciando-se pela aplicação e pela correção que ocorrem de forma automatizadas. Qualquer item ou avaliação impresso pode ser transposto para o meio digital e é disponibilizado como um espelho do que já fora programado para o meio impresso; e iii) Item digital interativo: planejados, criados e produzidos no formato digital interativo, obedecem às regras e possibilidades do digital (interatividade, ubiquidade, conectividade, inter-relação, animação etc.), bem como permitem e promovem interferência do usuário no objeto. São itens tecnologicamente enriquecidos que possibilitam, num ambiente digital, interações entre o respondente e o tipo de item por meio do qual ele responde, tais como simulação, associação, seleção, preenchimento, modo de arrastar, entre outras, e permitem, com correções automáticas, uma avaliação ampla dos processos cognitivos e de habilidades.

Bruno (2019a) ressalta, ainda, que há uma diferença significativa entre item digital e item digital interativo. O primeiro constitui a transposição do impresso para o formato digital, enquanto o segundo se realiza por meio da interatividade, com as possibilidades e as potencialidades que o digital e em rede oferecem. Portanto, não falamos apenas de recursos tecnológicos, mas também – e especialmente – de uma mudança significativa na concepção de educação e de avaliação.

A escolha pelo tema sobre itens digitais e digitais interativos e a aplicação em avaliação em larga escala fundamenta-se nas pesquisas realizadas por meio de participação em Editais de pesquisa do Centro de Políticas Públicas e Avaliação da Educação (CAEd) da Universidade Federal de Juiz de Fora (UFJF), desenvolvidas pelas autoras deste manuscrito nos períodos de agosto de 2018 a dezembro de 2019 e fevereiro de 2020 a janeiro de 2021.

No primeiro período do estudo (2018-2019), a pesquisa intitulada “Tecnologias, metodologias digitais e processos cognitivos na elaboração e na aplicação de itens de avaliação para a Educação Básica” foi desenvolvida pela pesquisadora Adriana Rocha Bruno, tendo sido mapeadas: tecnologias disponíveis para produção, veiculação e aplicação de itens para avaliações *on-line*; possibilidades para curadorias digitais na educação, especificamente para banco de itens avaliativos; metodologias e recursos digitais inovadores, com base nos de *startup*, para o desenvolvimento de itens de avaliação para a educação básica; e descobertas das neurociências e das ciências envolvendo os



processos cognitivos, bem como as relações entre o acesso a informações, leitura e aprendizagem em papel e no digital. No período posterior (2020-2021), foi desenvolvida pelas autoras deste texto a pesquisa “Itens digitais e interativos em E-avaliação de larga escala: processos cognitivos, estratégias e produtos didáticos na relação entre meio impresso e meio digital”. Tinha-se como objetivo realizar uma revisão teórico-bibliográfica sobre o tema e, com isso, propor uma arquitetura tecnológica e pedagógica com parâmetros para a utilização na avaliação de itens digitais e interativos em diferentes áreas do conhecimento, além de pesquisar e selecionar plataformas e itens digitais e interativos para análises e validação da arquitetura do instrumento proposto.

O presente artigo apresenta um recorte da pesquisa realizada entre 2020 e 2021 com o instrumento “e-Val: Avaliação da Interatividade em Itens Digitais”, no qual são analisados os níveis de interatividade, possibilitando uma reflexão sobre seu processo de desenvolvimento e sobre a aplicação em itens digitais disponíveis nas plataformas MathemaTIC 1 e TAO2.

Para cumprir com nosso objetivo, este artigo está estruturado em quatro seções. Além da introdução, a seção 2 traz um referencial teórico-bibliográfico sobre o que entendemos acerca do tema interatividade; a organização dos elementos tecnológicos e pedagógicos que embasam o desenvolvimento da arquitetura para o *design* e a análise de itens digitais e interativos; a integração dos elementos e a definição de indicadores, critérios, relevância e escala de interatividade; o cálculo e prototipagem da arquitetura que resultou na versão final do protótipo do instrumento “e-Val: Avaliação da Interatividade em Itens Digitais”. Já a seção 3 apresenta a testagem e a validação do e-Val em itens digitais selecionados das plataformas MathemaTIC e TAO. Por fim, as considerações finais ilustram as contribuições da pesquisa, bem como os trabalhos futuros vislumbrados.



## 2. O que entendemos por interatividade: bases pedagógicas e tecnológicas para o desenvolvimento do e-Val

Para o desenvolvimento de um instrumento como o e-Val, que será apresentado na sequência, foi necessário intenso estudo acerca da interatividade. Sabemos que a Cultura Digital promoveu mudanças muito significativas para o mundo. Dentre as mais importantes, estão: i) a relação com a informação, que passa a ser produzida e acessível a um número exponencialmente maior de pessoas e ii) a relação entre as pessoas, que passam a interagir entre si, via redes sociais digitais, ampliando, pluralizando e potencializando as formas de relacionamento humano. Tais mudanças implicam formas muito diferentes de interação, processo que chamamos de interatividade.

Silva (2014, p. 102) associa a interatividade ao “[...] diálogo, comunicação, troca entre interlocutores humanos, entre humanos e máquinas, e também versa sobre a possibilidade de agir e intervir sobre o programa e o conteúdo”. Segundo o mesmo autor (SILVA, 1998), a interatividade integra sujeito e objetivo, de modo a levar o sujeito a interferir, modificar, cocriar com o objeto, propondo a fusão de sujeito com objetivo, promovendo a articulação entre redes e conexões. É, portanto, diferente de interação, que cinde o emissor do receptor.

Por outro lado, Primo (2000) aborda a relação entre interatividade e reatividade. Para o referido pesquisador, o interativo promove e implica autonomia; já o reativo atua num campo de escolhas apresentadas por outrem. Em seus estudos, traz o que denomina interação mútua e interação reativa. A interação mútua se dá por meio da relação entre e pela interconexão dos sistemas envolvidos. Já a interação reativa, por ser um sistema fechado, implica relações lineares, do tipo estímulo-resposta.

Tais ideias são integradas aos estudos de Bruno (2010; 2007) sobre plasticidade social e tecnológica, em que são apresentados cinco princípios/elementos. A autora compreende que é possível associar a plasticidade orgânica a outros fenômenos da vida humana, estando tal flexibilidade, portanto, presente também nos processos de aprendizagem, de relações sociais, culturais, entre outros. Nessa direção, para a referida pesquisadora, a aprendizagem deve ser entendida como caracteristicamente plástica, dada a sua possibilidade de transformação.



Por fim, Bruno (2007) alerta que a possibilidade de novas conexões celulares ao longo de existência humana é extraordinária, sendo que, quanto mais rico for o ambiente, de modo a estimular atividades mentais, maior o impacto sobre as capacidades cognitivas, emocionais, perceptuais, de memória, entre outras.

As plasticidades social e tecnológica, desenvolvidas por Bruno (2010. p. 175-176), foram fulcrais para o desenvolvimento do e-Val e, com isso, para a constituição da base pedagógica. Nesse sentido, a autora traz os seguintes princípios:

- flexibilidade: rompe com as barreiras que impediriam desdobramentos e integrações de ideias, de pessoas, de informações, de conhecimentos, de funções, pois cria trilhas de possibilidades e emergências;
- conectividade: possibilidade de interligação de temas, experiências, de modo a se desdobrar em outras conexões / ligações sem forma definida;
- integração: diferente de agrupamento, essa característica indica os processos ocorrentes entre eventos plurais, não fragmentando ou excluindo grupos e contextos, mas criando elos de ligação;
- abertura: com sistemas de entrada e saídas múltiplas, em fluxo e em constante emergência, a plasticidade não possui uma organização ou estrutura pré-estabelecida e não obedece a padrões rígidos e fixos;
- dinamicidade: como organismos vivos em constante devir, os elos se interconectam e se integram (desintegram / reintegram), mas se re/des-constituem por meio de conflitos, assumindo funções até aquele momento inimagináveis.

Por fim, integrados aos estudos sobre interatividade, destacamos o que Silva (2014) classifica como os três fundamentos da interatividade:

- Participação-intervenção: “participar é interferir em sua mensagem, é construir coletivamente a comunicação e a aprendizagem” (SILVA, 2014, p. 256). O foco desse fundamento, para os estudos ora propostos, integra-se à ideia de que a ação implica intervenção e interferência. Portanto, quanto mais interferência/intervenção sobre o objeto, mais interativo ele será.



- Bidirecionalidade-hibridação: “comunicar pressupõe bidirecionalidade [...] é produção conjunta dos interlocutores” (SILVA, 2014, p. 256). Por um lado, a bidirecionalidade aponta para duas (ou várias) direções, mas em movimentos de A para B e de B para A, podendo mesmo se direcionar para C, D, entre outros. Por outro, quando integramos a bidirecionalidade à hibridação, promovemos – e ainda podemos provocar – a integração, e mesmo a fusão, segundo Silva (2014), de sentidos, ideias, ações.
- Permutabilidade-potencialidade: envolve múltiplas redes articulatórias de comunicação e de conhecimento. Implica, ainda, informações em redes de conexões e liberdade para permutar, virtualizar, simular, associar e significar.

Os itens digitais apresentam muitas vantagens, como custo inferior, interatividade (por meio da integração de diversas mídias), motivação, reaproveitamento e integração de itens, dentre outros, que são muito ricos tanto para as aprendizagens quanto para a avaliação.

Compreendemos que a elaboração dos itens digitais pode trabalhar com níveis de interações, os quais: i) se adequem aos usuários, considerando seus contextos; ii) extrapolem os recursos disponíveis na Plataforma, integrando dispositivos externos; iii) promovam níveis diversos de interatividade, em acordo com as concepções de avaliação e os contextos; iv) permitam mapeamento dos processos, trajetos, performances, tempos na construção de respostas ao item; e v) propiciem múltiplos percursos para desenvolvimento de respostas por meio da integração de mídias.

Compreendemos que não há zero interatividade, pois todo contato, por menor que seja, com um objeto – sendo esse objeto fruto de criação e/ou transformação humana –, implica baixa interatividade, mas não supomos interatividade nula. Porém, o escopo de baixa interatividade no campo da avaliação assume grau de contato com os elementos, de modo que, quanto maior o contato e interferência, maior a interatividade. Integrando tais aspectos, retomamos os estudos de Primo (2000). Desse modo, os itens poderão ser classificados em reativos ou interativos, trazendo uma relação entre reatividade e interatividade. O referido pesquisador explica que o interativo promove e implica autonomia, enquanto o reativo atua num campo de escolhas apresentadas por outrem. Por fim, articulam tais esquemas os estudos de Bruno (2010) sobre a plasticidade social e tecnológica. Lembramos que, para Silva (1998), a interatividade integra sujeito e objetivo, de modo a levar o sujeito a interferir, modificar, cocriar com o objeto.



As avaliações em larga escala impressas, até então desenvolvidas em massa no Brasil e em diversos países, podem ser classificadas no campo da reatividade, enquanto diversas das avaliações de aprendizagem desenvolvidas nas escolas, especialmente aquelas enriquecidas com tecnologias, se aproximam mais da interatividade. Há que se considerar, para que haja mediação com tecnologias digitais e em rede: os acessos aos dispositivos tecnológicos; redes de conexão à Internet; nível de afinidade do usuário com as tecnologias; e, ainda, as concepções de avaliação e os contextos da(s) instituição(ões) envolvida(s).

**a.** Plasticidade social e tecnológica, reatividade, sujeito e objetivo: bases para a definição dos elementos pedagógicos

Quanto maior o contato e interferência/manipulação, maior a interatividade. Os estudos para desenvolvimento dos elementos pedagógicos que compõem o instrumento contemplaram: Primo (2000) cujo estudo inspirou a nossa classificação de itens digitais em reativos ou interativos, sendo a reatividade integrada ao instrumento, uma vez que a interação mútua está contemplada nos demais itens; os estudos de Bruno (2010) sobre a plasticidade social e tecnológica, cujos elementos – flexibilidade; conectividade, integração, abertura e dinamicidade – ofertaram parâmetros para articulação com os estudos empreendidos; e, por fim, Silva (1998), para quem a interatividade integra sujeito e objetivo, sendo que, quanto mais o sujeito interferir, modificar, cocriar com o objeto, maior será a interatividade.

O item como um todo, inclusas as alternativas de respostas, apresentam maior (3), menor (1) ou intermediária (2) possibilidades de interatividade a partir dos seguintes aspectos pedagógicos:

- » Flexibilidade/abertura: trilhas de possibilidades com sistemas de entrada e saídas diversas para resolução do item. Por exemplo: ainda que a resposta seja única, compreende que existem múltiplos caminhos para sua resolução/resposta;
- » Integração: cria elos e articulação entre áreas do conhecimento ou entre conhecimentos da mesma área e/ou possibilita ações coletivas e colaborativas para sua resolução e/ou arquitetura percursos hipertextuais. Engloba também a permutabilidade, ou seja, possibilidade de associações, simulações e a conectividade, ou seja, a interligação de temas, de experiências, de modo a se desdobrar em outras conexões / ligações;



- » **Intervenção:** promove/favorece a interferência/manipulação do usuário no objeto/item. A intervenção incorpora também a bidirecionalidade na medida em que promove ações do usuário no item em múltiplas direções, em múltiplos formatos e com diversas mídias e recursos.

Observação: quanto mais interferência/intervenção sobre o objeto, mais interativo ele será;

- » **Reatividade:** esse aspecto foi inserido na tabela durante os estudos para que o usuário tenha dimensão de quão reativo é o seu item. Ou seja, quanto menos interativo, mais reativo é o item. Isso significa que o item, em caso de maior reatividade, implica maior passividade do usuário, relações e articulações mais limitadas – especialmente para efeitos de análises/avaliações. Assim, naquele momento, pensou-se que, se o item aponta para reatividade 3, esse fato indicaria que o usuário apresentou a resposta sem que interagisse com o item e com a máquina, não oferecendo a ele (estudante) oportunidades de intervenção, de testagem de alternativas, trilhas/percursos para resposta. E, ao avaliador, apenas o resultado, sem oportunizar identificar/conhecer o percurso de resolução/resposta do/a estudante. Entretanto, por ser um aspecto contrário aos demais, numa “contradimensão”, compreendemos que não seria bom estar no protótipo final de modo que ele foi retirado.

A partir desses aspectos/elementos pedagógicos, foram criados os indicadores avaliativos para o protótipo e-Val.

**b. *Design* de Interface e *Design* de Interação:** bases para a definição dos elementos tecnológicos

Ao projetar e desenvolver itens digitais e interativos, a ênfase deve ser colocada também nas maneiras pelas quais os usuários possam acessar, manipular e navegar através do conteúdo apresentado na tela a partir de uma interface. Sims (1997) declara que a ideia de interação está intrínseca à prática efetiva de ensino e à descoberta individual, não sendo adequado que a atividade educacional, independentemente de ser avaliativa ou não, esteja limitada a conteúdos ou ações em que a interatividade é trivializada apenas por menus, objetos clicáveis e sequências



lineares de navegação. Ademais, apresenta alguns elementos interativos, os quais denomina de níveis de interatividade e são de fundamental importância ao planejar um item digital, bem como para a sua implementação em meios tecnológicos.

Os níveis de interatividades definidos por Sims (1997) constituem-se em: Objeto, Linear; De suporte; De atualização; Construtiva; Reflexiva; De simulação; Hiperlinkada; Contextual não-imersiva e Virtual imersiva.

A justificativa para a utilização dos níveis apresentados tem como base as explicações de Passos e Behar (2011). Os autores ressaltam que os níveis “Objeto”, “Linear” e “De suporte” servem como recursos ao funcionamento do sistema, ajudando o aluno na navegação. Ademais, destacam como as alternativas que oferecem níveis de interatividade que mais podem produzir situações propícias aos aprendizados a “Contextual não-imersiva” e a “Virtual imersiva”.

Ainda sobre o nível “De atualização”, Passos e Behar (2011) afirmam que pode apresentar-se em sistemas bastante complexos, dependendo do tamanho do banco de dados disponível. É possível projetar desde um simples perguntas e respostas até um robô de conversação de inteligência artificial. Já no caso da interatividade “Reflexiva”, as respostas apresentadas pelo sistema têm origem na escrita de outros alunos, o que proporciona certo grau de interação com os colegas ou, ainda, com o sistema.

Já sobre a interatividade “Hiperlinkada”, os autores observam que, mesmo que os *links* ofereçam ao aluno uma grande quantidade de informação, esse tipo de interatividade pode restringi-lo às opções disponibilizadas. Assim, pode desmotivá-lo a buscar por outros conteúdos disponíveis na rede, limitando sua autonomia.

Sims (1997) afirma que, em todas as interações, quanto mais a “Atualização” estiver relacionada à resposta individual do aluno, mais esforço será preciso em termos de *design* e desenvolvimento. Para Passos e Behar (2011), isso se aplica, especialmente, às interatividades “Construtivas”, “De simulação”, “Contextual não-imersiva” e “Virtual imersiva”, nas quais a construção gráfica é necessariamente mais elaborada. Para os autores, nesses casos, os níveis de interatividade e as possibilidades de atuação do aluno frente ao artefato digital aumentam proporcionalmente ao aprimoramento de *design* e tecnologia de cada interface.



De acordo com Sims (1997), quanto maior o nível de interatividade, maior requinte é necessário em termos de *design*. O autor ainda afirma que, “[...] concentrando-se no projeto do conteúdo didático, do *design* gráfico e da comunicação para implementar interações que motivem e envolvam o aluno, o contínuo sucesso funcional e a eficácia das aplicações interativas de ensino estará assegurado” (SIMS, 1997, p. 169).

Assim, entende-se que não é suficiente ter à disposição apenas um bom recurso tecnológico. Precisa-se, na realidade, de um planejamento adequado do conteúdo e do *Design* da Interface.

Além dos elementos do *Design* de Interface, com base nos níveis de interatividade já apresentados, vamos integrar os elementos do *Design* de Interação, que têm por objetivo otimizar as interações entre o usuário e o produto, sistemas, artefatos digitais ou ambientes.

O *Design* de Interação está circunstanciado em cinco elementos: palavras, representações visuais, objetos físicos ou espaço, tempo e comportamento. Os quatro primeiros elementos foram definidos por Smith (2007) e o último, por Silver (2007).

Especificamente, os elementos “Palavras”, “Representações visuais” e o “Espaço” *permitem* (ação) a interação, enquanto o “Tempo” e o “Comportamento” *definem* (reação) a interação.

Normalmente, as interações são atos tangíveis, como pressionar ou clicar em um botão ou agarrar/arrastar uma alça, por exemplo. Não importa se as coisas que permitem a ação são virtuais ou físicas.

Dessa forma, para a análise de itens digitais com maior interatividade e orientação no desenvolvimento, devem-se observar os elementos que “permitem a ação” e “definem a reação” do usuário. Sendo assim, quanto mais bem planejados e aplicados numa interface digital, maior será o nível de interatividade. Também devemos observar que alguns elementos do *Design* de Interação, quando utilizados de forma isolada, apresentam menor nível de interatividade.

As palavras incluem textos, como etiquetas de botões, que ajudam a fornecer aos usuários a quantidade certa de informações – são as “interações” em forma de verbos (ação). Quando utilizadas de forma isolada, a interatividade é considerada baixa. Porém, quando combinadas



com outros elementos do *Design* de Interação, como representações visuais e o tempo, por exemplo, utilizando os objetos ou espaços para a manipulação, o nível de interatividade aumenta consideravelmente.

As representações visuais são elementos gráficos, como imagens, tipografia e ícones que auxiliam na interação do usuário (ação). Quando utilizadas de forma isolada, apresentam um *design* mais atrativo para o usuário, porém, ainda com baixa interatividade. Quando combinadas com outros elementos, podem aumentar significativamente o nível de interatividade.

Os objetos físicos ou espaço referem-se ao meio através do qual os usuários interagem com o artefato digital (ação) – com o qual ou dentro do qual os usuários interagem –, como um *laptop* via *mouse* ou um telefone celular através dos dedos, por exemplo. Os elementos desse grupo apresentam ações e comandos tecnológicos que podem ser inseridos juntamente com os elementos “palavras” e “representações visuais”. Dessa forma, transformam um simples texto ou imagem estáticos em ações de nível alto de interatividade.

O elemento tempo refere-se às mídias que mudam com o tempo, como animações, vídeos e sons (ação e reação), isto é, o conteúdo que muda com o tempo, com som, vídeo ou animação em forma de duração. Esse elemento, quando tecnologicamente aplicado em interfaces digitais utilizando os elementos anteriores descritos, corresponde à alta interatividade, principalmente pela possibilidade de inclusão de animações, simulações, bem como pela interação direta do usuário com a interface do artefato digital.

O comportamento é o elemento que se preocupa com a forma como os quatro elementos anteriores definem as interações que uma interface digital oferece, como aquelas que os usuários podem executar ações em um *site*, por exemplo. O comportamento também se refere a como a interface digital reage às entradas dos usuários e fornece *feedback*. Inclui ação ou operação e apresentação ou reação. Em relação ao nível de interatividade, é possível afirmar que, quanto mais ações e reações dos usuários com a interface digital, maior será a interatividade. Assim, ele pode variar do nível mais baixo para o nível mais alto de interatividade.



- c. Integração dos elementos tecnológicos e pedagógicos e a definição de indicadores, critérios, relevância e escala de interatividade

A integração dos elementos objetivou unir, numa única arquitetura, os elementos tecnológicos e pedagógicos para que, ao se analisar ou planejar um item digital, este seja realizado de forma completa.

Para utilizarmos tais elementos, foi necessário, além da integração, a definição de indicadores com índices de relevância, critérios, bem como sua adequação numa escala para mensurar o nível de interatividade.

Quanto aos indicadores, a arquitetura apresenta 6 (seis) indicadores pertencendo ao grupo dos elementos tecnológicos e 5 (cinco) ao grupo dos elementos pedagógicos. Tais indicadores possuem maior ou menor relevância na composição do nível de interação de um item digital. Assim, para que possamos mensurar tal relevância, foi necessário atribuir uma escala de 1 a 3, sendo 1 para menos relevante e 3 para mais relevante. Essa escala representa o grau de importância de cada indicador na composição do nível de interatividade do objeto digital.

Ressaltamos que há “níveis de interatividade” que funcionam como recursos do funcionamento do sistema, ou seja, ajudam o aluno na navegação e têm a função de suporte. Dessa forma, consideramos possuir menor relevância na interatividade (indicadores do e-Val sobre Suporte e *feedback* e Bidirecionalidade). Já os níveis que representam situações em que o aluno é imerso virtualmente em um mundo no qual o ambiente responde a ações e movimentos desse discente propiciam mais situações de interatividade e potencializam as aprendizagens. Nesse caso, classificamos como alta relevância (indicadores do e-Val sobre: Textos e elementos gráficos, Espaço/meio de interação, Movimento/liberdade do aluno no item, Flexibilidade e abertura e Intervenção). Todavia, há outros níveis que são muito complexos para mensurar, pois há variações que dependem, por exemplo, do tamanho do banco de dados do sistema utilizado e das respostas textuais apresentadas pelos alunos, podendo apresentar relevância significativa ou não. Para essas situações, definimos como média relevância para neutralizar a mensuração e não interferir no resultado final (indicadores do e-Val sobre: Diálogo entre aluno e conteúdo, Comportamento dos elementos no espaço, Integração e Permutabilidade).



A tabela 1 apresenta os indicadores, seu grau de relevância e os grupos de elementos aos quais pertencem.

**Tabela 1. Indicadores, grau de relevância e grupo**

Indicador	Grau de relevância	Grupo de Indicadores
Textos e elementos gráficos	3	Tecnológico
Espaço/meio de interação	3	Tecnológico
Comportamento dos elementos no espaço	2	Tecnológico
Movimento/liberdade do aluno no item	3	Tecnológico
Suporte e <i>feedback</i>	1	Tecnológico
Diálogo entre aluno e conteúdo	2	Tecnológico
Flexibilidade e abertura	3	Pedagógico
Integração	2	Pedagógico
Permutabilidade	2	Pedagógico
Intervenção	3	Pedagógico
Bidirecionalidade	1	Pedagógico

Fonte: elaborada pelos autores (2021).

Para cada indicador, foram desenvolvidos três critérios, considerados também como opções de escolha pelo avaliador ao analisar um item digital e que objetivam classificar o nível de interatividade. Para esses critérios, utilizamos a escala 1 para “Baixa Interatividade”; 2 para “Média interatividade”; e 3 para “Alta Interatividade”, sendo:

Baixa interatividade: refere-se a itens digitais que apresentam interatividade baixa (já que, em nossa compreensão, não há interatividade nula/zero) e se dá por meio do acesso/contato com o item, com opções limitadas para o usuário (uma alternativa para se chegar à resposta), permitindo que ele faça algumas (poucas) escolhas. Itens desse nível não oferecem abertura e controlam o usuário tanto por meio de enunciados fechados e individualizados quanto por meio de alternativas/respostas apresentadas em formato uma única alternativa ou um *menu* restrito, sem que se estimulem ou intencionem, por parte da formulação do item, relações entre conhecimentos e/ou áreas, estimulando a passividade. O item não tem objetivo e compromisso com o processo, mas tão somente com o resultado final. Quanto aos aspectos tecnológicos,



os elementos gráficos e textos são estáticos. Normalmente, não apresentam orientações para o aluno sobre como percorrer para desenvolver a questão, bem como não apresentam opção de ajuda ou retorno. O usuário somente consegue interagir com a questão por meio da opção “clique” e com possibilidade de se mover somente para frente ou para trás, com sequências lineares e pré-definidas. Ademais, após a ação do aluno em finalizar, o sistema apresenta o próximo item do processo avaliativo também de forma linear e pré-definida. Itens digitais desse nível costumam ser versões digitais de itens produzidos para formatos impressos em papel.

**Média Interatividade:** refere-se a itens digitais que apresentam interatividade relativa – por isso, mediana –, podendo ser permitida ao usuário consulta, com opções menos limitadas para o usuário (duas ou três alternativas para se chegar à resposta), possibilitando que ele faça algumas escolhas, mas ainda de forma controlada. O usuário interage apenas com os enunciados, com leve abertura, mas ainda individualizados, tendo alternativas/respostas apresentadas em formato de “uma única alternativa” ou “um menu restrito”, com escassas relações entre conhecimentos e/ou áreas, estimulando a reatividade e não a proatividade. Os textos e os elementos gráficos se apresentam no formato de mídias, como imagens, vídeos, sons com algum tipo de animação em seu enunciado. Apresentam meios para que o aluno interaja de forma controlada na questão/resposta por meio de recursos limitados, tais como selecionar, arrastar, manipular ou desenhar com auxílio, por exemplo, de *mouse* ou *touch screen*. Ainda que se utilizem recursos de animação e audiovisuais, itens de média interatividade não estimulam a interferência e/ou criatividade do usuário, controlando suas possibilidades. O aluno consegue se mover na questão de forma intuitiva por meio de uma sequência não linear ou pré-definida, com possibilidade de *feedbacks*. Após a ação do aluno em finalizar, o sistema apresenta o próximo item do processo avaliativo numa sequência não linear.

**Alta interatividade:** refere-se a itens digitais que apresentam possibilidades diversas e flexíveis, múltiplos caminhos para resolução do item, ou seja, permite uso de mais do que um recurso para que se chegue à resposta esperada. Criam elos e articulações entre áreas do conhecimento, ações coletivas e colaborativas para sua resolução e/ou arquitetura percursos hipertextuais e promovem interferência/manipulação do usuário no objeto/item. O item só “funciona” por meio da interferência do usuário, ou seja, a interação e as interferências são provocadas tanto por meio do enunciado quanto por meio dos recursos disponíveis para sua resolução. Os textos



e os elementos gráficos se apresentam no formato de vídeos e/ou sons interativos, animações, imagens dinâmicas, simuladores, entre outros, e, em seu enunciado, podem ser oferecidas respostas e informações adicionais como pistas para resolução do item. Apresenta meios para que o aluno interfira na questão/resposta, tais como selecionar, arrastar, manipular ou desenhar com auxílio, por exemplo, de *mouse* ou *touch screen* e a utilização de recursos como calculadora, bloco de anotações, acesso a *links* de informações adicionais, espaço para experimentação, simulação, entre outros. Há orientações para o aluno sobre como percorrer para desenvolver a questão de forma explícita (textual, audiovisual ou elementos gráficos) e *feedbacks*, de tal modo que ele pode se mover na questão de forma livre e autônoma. As escolhas das ações do aluno determinam a sequência e a continuidade do processo avaliativo (uso de *chatbots*/inteligência artificial etc.). Ou seja, evidenciam-se preocupação e interesse no processo de resolução e não somente no resultado final.

Conforme já descrito anteriormente, partimos do princípio de que não há zero interatividade, pois todo contato, por menor que seja, com um objeto – sendo esse objeto fruto de criação e/ou transformação humana – implica baixa interatividade, mas não supomos interatividade nula.

**d.** Cálculo e prototipagem da Arquitetura para o *design* e análise de itens digitais interativos: instrumento “e-Val: Avaliação da Interatividade em Itens Digitais”

Após integrar os elementos tecnológicos e pedagógicos e transformá-los em indicadores com grau de relevância e critérios com níveis de interatividade, realizamos a prototipagem da arquitetura utilizando uma planilha eletrônica e aplicando uma fórmula para o cálculo final do nível de interatividade. Tal ação resultou no instrumento “e-Val: Avaliação da Interatividade em Itens Digitais”.

**l)** Calculando o nível de interatividade

Conforme já descrito anteriormente, os indicadores estão agrupados em dois grupos: “tecnológicos” e “pedagógicos”. Cada indicador possui três critérios, aos quais foi atribuído um conceito para cada, que pode ser 1, 2 ou 3. O nível de interatividade é determinado a partir da soma dos conceitos atribuídos a cada critério do indicador, ponderado pelo grau de relevância



que esse indicador possui na avaliação. O grau de relevância, atribuído numa escala de 1 a 3, sendo 1 para menos relevante e 3 para mais relevante, representa o grau de importância de cada indicador na composição do nível de interatividade do objeto digital. Na nota final, considera-se um peso de 50% para o grupo de indicadores “tecnológicos” e 50% para os indicadores “pedagógicos”. Os graus de relevância para cada indicador foram definidos na tabela 1 apresentada no item anterior.

Para determinar o Nível de Interatividade (NI) do item digital, os indicadores foram compilados em dois grupos: os indicadores tecnológicos, com soma dos graus de relevância 14; e os indicadores pedagógicos, com soma dos graus de relevância 11, conforme os valores indicados na tabela 1. O Conceito final é determinado a partir da Eq. 1 a seguir:

$$CF = \frac{(\sum_{t=1}^6 I_t \times GR_t + \sum_{p=1}^5 I_p \times GR_p)}{2}$$

Em que:

CF = Conceito Final

$I_t$  = t-ésimo Indicador tecnológico

$GR_t$  = grau de relevância do t-ésimo Indicador tecnológico

$I_p$  = p-ésimo Indicador pedagógico

$GR_p$  = grau de relevância do p-ésimo Indicador pedagógico

O resultado da aplicação da equação está expresso numa escala entre 1 e 3, sendo que a gradação da escala é mostrada na tabela 2.

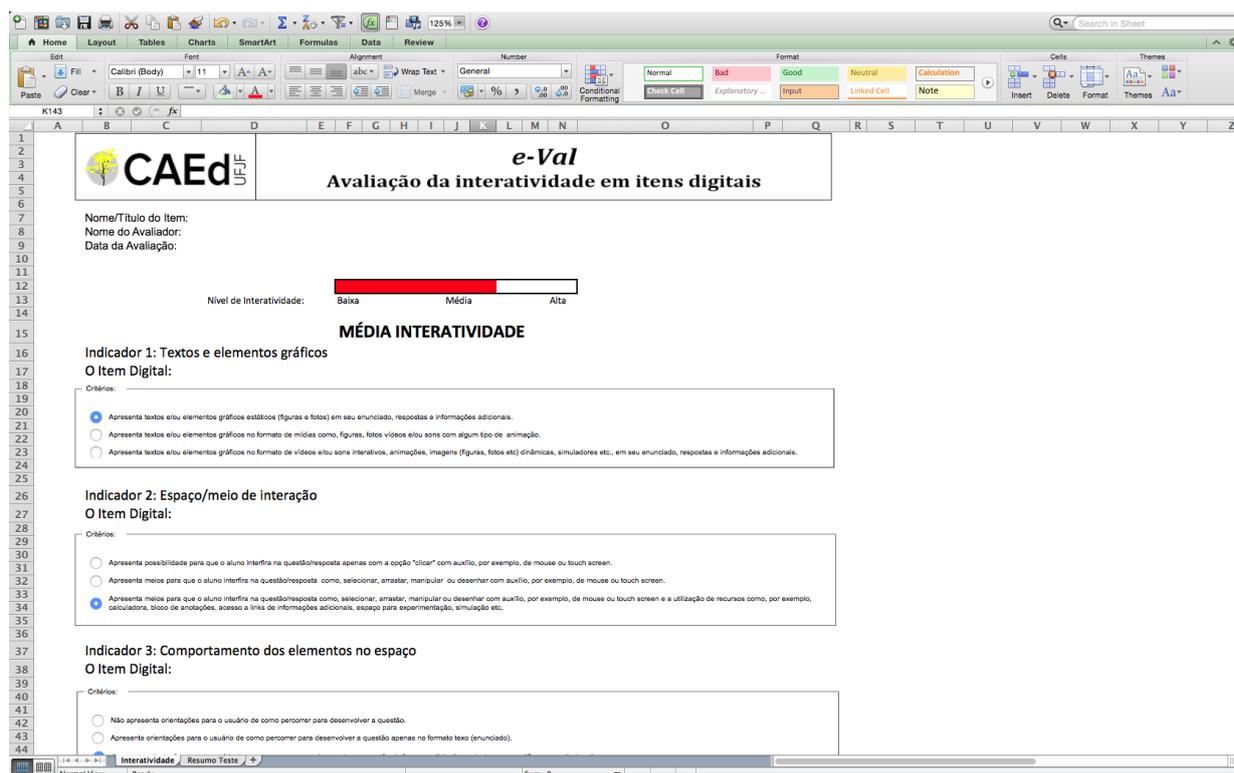
**Tabela 2. Gradação do conceito final do nível de interatividade do objeto digital**

Nível de Interatividade do Objeto Digital	Valor do NI Calculado
Baixa Interatividade	$1 \leq NI < 1,5$
Média Interatividade	$1,5 \leq NI < 2,5$
Alta Interatividade	$NI > 2,5$

Fonte: elaborada pelos autores (2021).

Na figura 1, apresentamos a tela inicial do instrumento e-Val com alguns indicadores e critérios prontos para serem utilizados pelos avaliadores de itens digitais, bom como a barra de interatividade que permite avaliar valores intermediários entre cada conceito, ou seja, quando o resultado da avaliação do objeto resulta num valor entre dois dos conceitos possíveis.

**Figura 1. Tela inicial do “e-Val: Avaliação da Interatividade em Itens Digitais”**



Fonte: elaborada pelos autores (2021).

### 3. Testagem e validação do e-Val em itens digitais

O desenvolvimento e prototipagem do instrumento e-Val perpassou 6 (seis) versões até chegar à atual. Em cada uma das quais se incluíram testes e aplicações reais em itens digitais, objetivando resultados confiáveis e compatíveis com as bases teóricas e os estudos realizados.

A versão 6 do e-Val foi aplicada na avaliação de itens digitais encontrados em plataformas nacionais e internacionais destinadas à avaliação em larga escala, bem como à avaliação da aprendizagem e reforço do processo de ensino e aprendizagem.

Neste recorte da pesquisa, apresentamos a aplicação do instrumento e-Val em alguns itens digitais disponibilizados nas plataformas MathemaTIC e TAO.

A MathemaTIC é uma plataforma digital para avaliação com a qual os alunos realizam atividades de matemática através de itens digitais e apresenta uma estrutura adaptativa, utilizando várias estratégias de aprendizagem. A plataforma, criada pelo Ministério da Educação de Luxemburgo em colaboração com a Vretta<sup>5</sup>, dispõe de exercícios matemáticos interativos, tendo sido adaptada às necessidades dos alunos em coerência com os planos de estudos daquele país.

Já a plataforma TAO permite desenvolver e gerenciar itens digitais em todas as áreas do conhecimento. *Computer-Based Testing*, em inglês; ou *Testing Assisté par Ordinateur*, em francês, foi criada pela Universidade de Luxemburgo e, hoje, é mantida, principalmente, pela *Open Assessment Technologies* (OAT).

No Brasil, o CAEd-UFJF disponibiliza uma versão da plataforma TAO<sup>6</sup> para o desenvolvimento de testes de itens digitais, a qual apresenta potencialidades no conjunto de ferramentas para o desenvolvimento de itens.

A seguir, apresentamos os resultados das análises de alguns itens digitais e digitais interativos a partir da aplicação do instrumento e-Val.

---

<sup>5</sup> A Vretta é uma empresa de tecnologia da educação que pesquisa e desenvolve novas formas com que os alunos se envolvem com a matemática (<https://www.vretta.com/index.html>).

<sup>6</sup> Disponível em <https://bancodeitens-legado.caeddigital.net/>.



**a.** Nome/Título do item: “Viagem de Avião” (Cycle 3)

- » Plataforma MathemaTIC
- » Localização: Endereço MathemaTIC: Módulos/Resolução de problemas/Avião
- » Resultado *e-Val*: “Alta Interatividade”
- » Observações: O item apresenta uma tela inicial (figura 2) com um contexto/história objetivando a solução de um problema. Na tela, é apresentado um “Diário de Bordo” no qual o aluno pode ouvir (som) o problema. Este também está no formato de texto, fora da imagem do diário. O aluno, ao ouvir ou ler o problema, inicia selecionando se a pergunta é sobre “tempo, distância ou quantidade”. Ao responder a opção correta, o aluno recebe um “diamante” como forma de prêmio, sendo solicitado a ele que selecione palavras/texto do problema que identifiquem a resposta. Ou seja, se a resposta foi “tempo”, o aluno deverá selecionar as palavras que a este correspondem no problema. As palavras selecionadas e corretas (em amarelo), automaticamente, aparecem no “Diário de Bordo”, no qual o aluno deverá, na sequência, arrastá-las para representar, neste caso, o tempo que demora o voo, o horário de saída e chegada, entre outros pontos. Na sequência, outras questões são apresentadas e, a cada resposta correta, novas premiações são dadas ao aluno no formato de “diamantes” e de estrelas (gamificação) até o término da verificação do item. O discente pode interferir no item com a opção “clicar”, “ouvir som”, “arrastar” e “digitar”, há permutabilidade (possibilidade de associações, simulações e/ou conectividade), bidirecionalidade (ações do usuário no item em múltiplas direções) e flexibilidade e abertura, não havendo, porém, integração ou elos de conexão explícita entre áreas do conhecimento.



**Figura 2. Item digital “Viagem de Avião” (Cycle 3) – MathemaTIC**



Fonte: MathemaTIC<sup>7</sup>

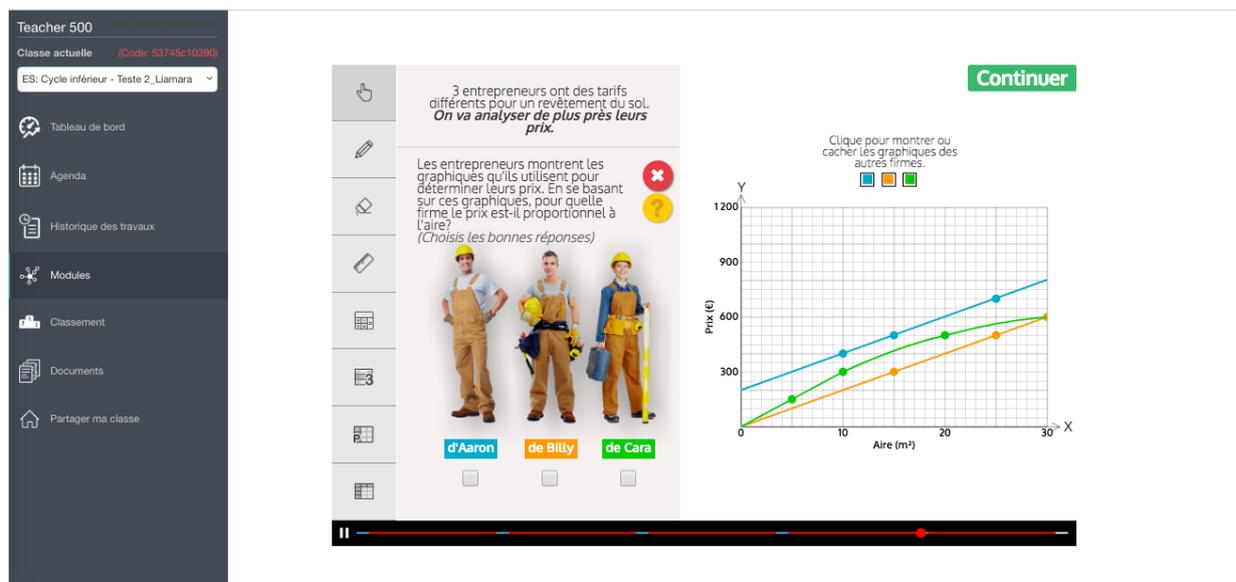
**b. Nome/Título do item: “Les devis” (Cycle inférieur)**

- » Plataforma MathemaTIC
- » Localização: Endereço MathemaTIC: Módulos/Proportionnalité/Les devis.
- » Resultado *e-Val*: “Alta Interatividade”
- » Observações: O item – conforme disposto na figura 3 a seguir – inicia apresentando imagens/fotos de profissionais dentro de um contexto de história em que é possível observar medidas/proporcionalidades. O aluno opta por uma das imagens e, na sequência, é apresentado um espaço quadriculado onde ele pode mover as linhas “x” e “y” para obter a medida solicitada na questão. Ao finalizar, o discente clica em continuar e outras questões são apresentadas. Além da opção de interação de clicar e arrastar no quadriculado, há, também, em algumas questões do item, a opção de resposta de inclusão de números. Não há tempo definido para o desenvolvimento do item. O item apresenta elementos gráficos e textos dinâmicos, meios para que o aluno interfira na resolução do desafio, apresenta

<sup>7</sup> Disponível em <https://www.mathematic.org/>

orientações de como percorrer na questão de forma explícita, livre e autônoma, há diálogo entre o aluno e conteúdo, flexibilidade e abertura, permutabilidade e bidirecionalidade, não havendo, entretanto, integração ou elos de conexão explícita entre áreas do conhecimento.

**Figura 3. Item digital “Les devis” (Cycle inférieur) – MathemaTIC**



Fonte: MathemaTIC<sup>8</sup>

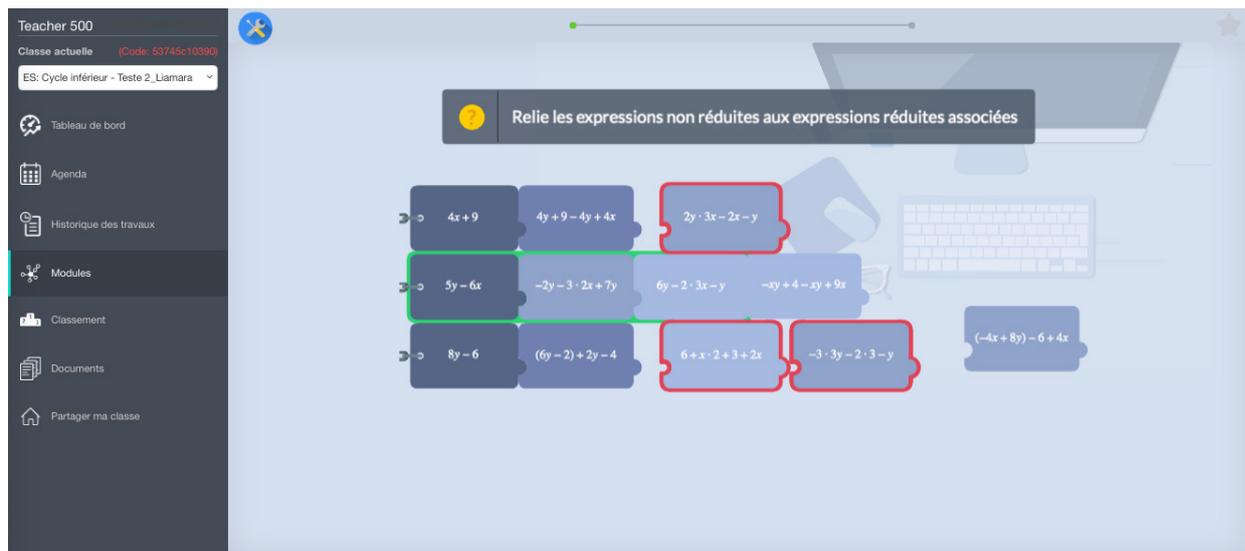
**c. Nome/Título do item: “Puzzie: calcul littéral-développer” (Cycle inférieur)**

- » Plataforma MathemaTIC
- » Localização: Endereço MathemaTIC: Módulos/Calcul littéral/Puzzie: calcul littéral-développer
- » Resultado e-Val: “Média Interatividade”
- » Observações: O item – conforme disposto na figura 4 – inicia solicitando que o aluno opte por um nível de dificuldade e, após a escolha, apresenta uma questão na forma de texto e as propostas de cálculo. O aluno deverá observar cada proposta de cálculo, verificar a sequência de desenvolvimento correta e arrastar a parte do desenvolvimento que consta a sequência correta. Ao finalizar, o aluno clica no botão “submeter”, e o sistema apresenta o que está correto ou não, sendo que há a opção de continuar a tentativa ou, ainda, reiniciar a

<sup>8</sup> Disponível em <https://www.mathematic.org/>

questão. O item é gamificado, o aluno pode interferir no item com a opção “clique e arrastar” e há permutabilidade (possibilidade de associações, simulações e/ou conectividade), porém, não há flexibilidade e abertura. Também não há integração ou elos de conexão entre áreas do conhecimento, bem como não há bidirecionalidade (ações do usuário no item em múltiplas direções).

**Figura 4. Item digital “Puzzie: calcul littéral-développer” (Cycle inférieur) – MathemaTIC**



Fonte: MathemaTIC<sup>9</sup>

**d. Nome/Título do item: G12001D18**

- » Plataforma TAO
- » Localização: <https://bancoeditens/bancodeitens.rdf#i153969516299082751>
- » Resultado e-Val: “Baixa Interatividade”
- » Observações: O item digital – disposto na figura 5 a seguir – apresenta uma questão de história e orienta o aluno a selecionar as opções em que constem as expressões que representam as “características” solicitadas. É um item típico da transposição do papel impresso para um meio digital sem nenhum tratamento ou cuidado com a interação.

<sup>9</sup> Disponível em <https://www.mathematic.org/>

### Figura 5. Item digital "G12001D18" – TAO

Entre 1945 e 1991, a Guerra Fria ampliou as tensões entre países capitalistas e socialistas, que protagonizaram diferentes formas de conflito e manifestação de poder. Selecione abaixo as expressões que retratam as características geográficas envolvidas nesse contexto.

 You can select maximum of 2 choices

- Reconfiguração geopolítica
- Transformação cultural
- Unificação étnica
- Homogeneização econômica
- Fronteiras naturais

Fonte: TAO<sup>10</sup>

**e.** Nome/Título do item: G12003D18

- » Plataforma TAO
- » Localização: [https://bancodeitens/bancodeitens.rdf#i15451295876780\\_6040](https://bancodeitens/bancodeitens.rdf#i15451295876780_6040)
- » Resultado *e-Val*: “Baixa Interatividade”
- » Observações: Apesar de o item – conforme disposto na figura 6 – apresentar orientações para o aluno de como percorrer para desenvolver a questão de forma intuitiva e explícita, apresenta meios para que ele interfira na resposta como arrastar e promover a bidirecionalidade. O item apresenta baixa interatividade, chegando a quase média interatividade.

<sup>10</sup> Disponível em <https://www.taotesting.com/>



**Figura 6. Item digital "G12003D18" – TAO**

Leia o texto a seguir sobre alterações no meio ambiente, e complete cada espaço em branco com a palavra correspondente. [Clique ou arraste a palavra até o espaço em branco no texto. Cada palavra pode ser utilizada uma vez.](#) <sup>11</sup>

dasmatamento    efeito estufa    Dióxido de Carbono (CO<sub>2</sub>)

[...] Além de provocar o \_\_\_\_\_, a queimada é uma das principais responsáveis pelo aumento do \_\_\_\_\_ em nosso planeta e o consequente aquecimento global. Isso ocorre pois é muito grande a liberação de \_\_\_\_\_, um dos principais gases do efeito estufa, na atmosfera durante uma queimada.

Disponível em: <<https://bit.ly/283wQG>>. Acesso em: 10 dez. 2018. Fragmento.

Fonte: TAO<sup>11</sup>

**f. Nome/Título do item: Q12001D18**

- » Plataforma TAO
- » Localização: <https://bancodeitens/bancodeitens.rdf#i153805417317352656>
- » Resultado e-Val: “Baixa Interatividade”
- » Observações: O item (figura 7) apresenta orientações para o aluno de como agir para desenvolver a questão de forma intuitiva e explícita; o enunciado e respostas apresentam pelo menos duas possibilidades flexíveis, com sistemas de entrada e saídas, para resolução do item; favorece a interferência do aluno com o vídeo e a escolha da opção correta e, ainda, promove a bidirecionalidade.

<sup>11</sup> Disponível em <https://www.taotesting.com/>

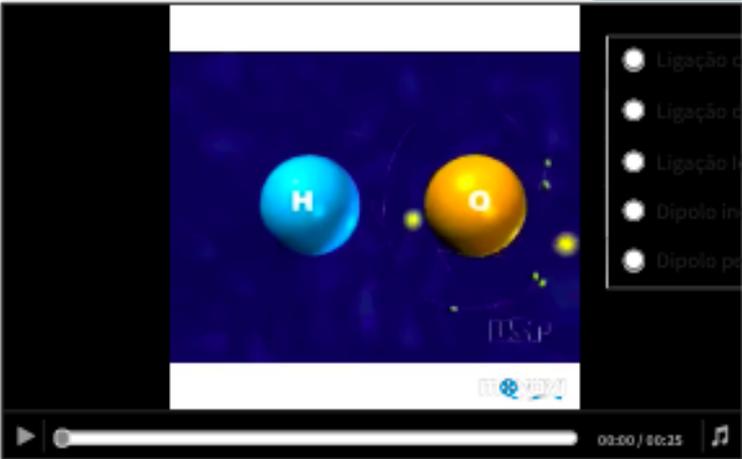
Figura 7. Item digital "Q12001D18" – TAO

A água é uma substância encontrada no planeta Terra e ocupa, aproximadamente, 80% dessa superfície. Assista ao vídeo abaixo e depois responda à questão.

Esse vídeo apresenta a interação química responsável pela formação de uma molécula da água. Qual é o tipo de interação química presente nessa molécula?

You must select at least 1 choice

- Ligação covalente.
- Ligação de hidrogênio
- Ligação iônica
- Dipolo induzido - dipolo induzido.
- Dipolo permanente - dipolo permanente.



Disponível em: <a href="https://www.youtube.com/watch?v=ThoD-SAczw6">https://www.youtube.com/watch?v=ThoD-SAczw6</a>. Acesso em: 20 set. 2018. \*Adaptado para fins didáticos.

Fonte: TAO<sup>12</sup>

Nas análises empreendidas na plataforma TAO, o e-Val auferiu “baixa interatividade” em todos os itens. Em alguns poucos casos, o índice chegou a quase média interatividade, mas permanecendo em baixa. Tais dados não podem ser vistos como exclusivamente relacionados à plataforma, mas também às propostas de itens desenvolvidos. A TAO apresenta instrumentos para a produção de itens. A maioria deles se baseia em recursos que eram desenvolvidos para o meio impresso: relacionar colunas, marcar opção correta, entre outros. Ainda que tenham sido adaptados para o *on-line* – relacionar colunas pode vir em formato de arrastar ou preencher lacunas –, a lógica permanece como a do analógico. Contudo, há abertura para que mudanças ocorram, pois é uma plataforma amigável, aberta e que permite *plugins* e integrações com outros recursos e formatos. Porém, tal feito só se dará quando houver demandas, o que depende de mudanças estruturais nos processos didático-pedagógicos dos desenvolvedores de itens. Dito de outra forma: é preciso mudar a visão de avaliação dos desenvolvedores de itens, de professores, bem como de outros atores envolvidos nesse processo.

<sup>12</sup> Disponível em <https://www.taotesting.com/>



## 4. Considerações finais

No presente artigo, objetivamos apresentar o processo de desenvolvimento e validação do “e-Val: Avaliação da Interatividade em Itens Digitais”, um instrumento que diagnostica os níveis de interatividade de itens digitais.

Para tal, buscamos elucidar o contexto da aplicação de um instrumento como o e-Val, que se constitui na elaboração e análise de itens digitais e interativos utilizados na composição de avaliação em larga escala, utilizados por escolas, secretarias estaduais e municipais de educação. Tais avaliações (compostas por itens, hoje também digitais) têm o objetivo não apenas de mapear os conhecimentos dos alunos, mas, especialmente, pautar a elaboração de políticas públicas, muitas vezes ranqueando as escolas, estudantes, docentes e, por vezes, ainda condicionando o recebimento de verbas e benefícios.

Ressaltamos que, devido às inovações tecnológicas, especialmente com os dispositivos digitais e em rede da Cultura Digital, as possibilidades de automatização no desenvolvimento, aplicação e correção de itens que constituem as avaliações em larga escala passaram a ser uma realidade iminente e estão fomentando mudanças que implicam a digitalização do formato impresso e, gradativamente, a produção de itens digitais e interativos em meio digital. Assim, os itens estão deixando de ser estáticos e lineares, passando a apresentar algum tipo de interatividade.

Para definir interatividade, pautamo-nos nos estudos de Silva (2014), Primo (2000) e Bruno (2010; 2007) e compreendemos um item digital interativo como itens planejados, criados e produzidos no formato digital, com níveis de interatividade. Ademais, obedecem às regras e possibilidades do digital (interatividade, ubiquidade, conectividade, inter-relação, animação, entre outras), bem como promovem a interferência do usuário no objeto. São tecnologicamente enriquecidos e possibilitam, num ambiente digital, interações entre o respondente e o tipo de item por meio do qual ele responde, tais como simulação, associação, seleção, preenchimento, modo de arrastar, entre outras, além de correções automáticas, permitindo uma avaliação ampla dos processos cognitivos e de habilidades.

Bruno (2019a) apresenta uma diferença significativa entre item digital e item digital interativo, sendo o primeiro a transposição do impresso para o formato digital e o segundo realizado por meio da interatividade, com as possibilidades e as potencialidades que o digital e em rede oferecem.



Para o desenvolvimento do instrumento e-Val, produzimos elementos pedagógicos e tecnológicos que contemplem as características necessárias para o planejamento, *design* e análise do nível de interatividade que um item deve apresentar para cumprir com o seu objetivo de avaliação, bem como a definição de indicadores com índices de relevância, critérios e a adequação destes numa escala para mensurar o nível de interatividade, a qual definimos por “baixa, média e alta interatividade”.

A validação do e-Val foi realizada na aplicação da análise do nível de interatividade em itens digitais disponíveis nas plataformas virtuais MathemaTIC e TAO.

O e-Val demonstrou ser um dispositivo eficaz para auxiliar os desenvolvedores de itens digitais no planejamento e *design* da interatividade, bem como para a análise do nível de interatividade em itens já desenvolvidos.

A análise dos itens digitais disponíveis nas plataformas citadas apresenta variados níveis de interatividade, sendo que os itens da plataforma MathemaTIC apresentam *design* mais atrativo, elementos de gamificação, bem como apresentam média e alta interatividade quase na totalidade. Já os itens digitais disponibilizados na plataforma TAO mostram que, em alguns poucos casos, o índice chegou a quase média interatividade, mas permanecendo em baixa interatividade. As análises nessa plataforma mostraram que ainda há prevalência da transposição dos itens desenvolvidos em meio impresso para o meio digital, com foco para os conteúdos específicos e/ou ações em que a interatividade é trivializada por objetos ou opções de respostas clicáveis e sequências lineares de navegação, além de raras associações entre conhecimentos, pares ou estratégias didático-pedagógicas para se chegar às respostas possíveis.

Ressaltamos que o instrumento e-Val integra uma pesquisa com objetivos e ações que contemplam a análise de itens digitais e digitais interativos com dimensões mais complexas e que objetivaram, desde 2018, a busca pelo conhecimento do que existe em termos de avaliações em larga escala utilizando o meio digital e, especificamente, a interatividade no desenvolvimento de itens.

As pesquisas futuras intencionam o desenvolvimento de itens digitais interativos exemplares a partir dos indicadores e índices de relevância apresentados no instrumento e-Val, os quais estão fortemente pautados em pesquisas científicas comprovadas. Ademais, pretendemos dar continuidade às pesquisas bibliográficas com o objetivo de ampliar a construção do conhecimento sobre o tema de avaliações digitais em larga escala e itens digitais interativos, em parceria com pesquisadores brasileiros e estrangeiros.



## Referências

- BELLONI, M. L. **Educação a Distância**. Campinas/SP: Ed. Autores e Associados, 2001.
- BRUNO, A. R. **A aprendizagem do educador: estratégias para a construção de uma didática on-line**. Tese (Doutorado em Educação) – Programa de Pós-Graduação em Educação: Currículo, Pontifícia Universidade Católica de São Paulo, São Paulo, 2007. Disponível em: <https://tede2.pucsp.br/bitstream/handle/9974/1/Adriana%20Rocha%.pdf>. Acesso em: 15 jan. 2021.
- BRUNO, A. R. **Relatório de atividades de pesquisa em avaliação 2016-2019/Octubro**. CAEd, 2019a.
- BRUNO, A. R. **Relatório de atividades de pesquisa em avaliação 2018-2019/Octubro**. CAEd, 2019b.
- BRUNO, A. R. Travessias invisíveis: plasticidade, diferença e aprendizagem em redes rizomáticas de formação de adultos educadores nos ambientes *online*. In: XV ENCONTRO NACIONAL DE DIDÁTICA E PRÁTICA DO ENSINO: convergências e tensões no campo da formação e do trabalho docente, 15, 2010. **Anais eletrônicos ...** Belo Horizonte: Autêntica, 2010. Disponível em: [http://endipe.fae.ufmg.br/livros/Livro\\_3](http://endipe.fae.ufmg.br/livros/Livro_3). Acesso em: 15 jan. 2021.
- FILATRO, A. **Design instrucional contextualizado: educação e tecnologia**. São Paulo: Editora Senac São Paulo, 2004.
- MORAES, M. C. **O paradigma educacional emergente**. Campinas-SP: Papirus, 1997.
- PALLOFF, R. M.; PRATT, K. **Construindo comunidades de aprendizagem no ciberespaço: estratégias eficientes para sala de aula online**. Trad. V. Figueira. Porto Alegre: Artmed, 2002.
- PASSOS, P. C. S. J.; BEHAR, P. A. Interação e Interatividade através das interfaces de materiais educacionais digitais. **Novas Tecnologias na Educação - CINTED-UFRGS**, v. 9, n. 1, jul. 2011.
- PETERS, O. **Didática do Ensino a Distância: experiências e estágio da discussão numa visão internacional**. Trad. I. Kayser. S. Leopoldo/RS: Editora Unisinos, 2003.
- PRIMO, A. F.T. Interação mútua e reativa: uma proposta de estudo. **Revista Farmecos**, n.12, jan. 2000. p. 81-92.
- SCORTEGAGNA, L. **Relatório de atividades de pesquisa em avaliação 2020-2021/Março**. CAEd, 2020.
- SILVA, M. Que é interatividade. **Boletim técnico do Senac**, Rio de Janeiro, v. 24, n. 2, mai/ago. 1998. p. 27-35.
- SILVA, M. **Sala de aula Interativa: educação, comunicação etc**. 7. ed. São Paulo: Ed. Loyola, 2014.



SILVA, M. (org). **Educação** online: teorias, práticas, legislação, formação corporativa. São Paulo, Edições Loyola, 2003.

SILVA, M; SANTOS, E. **Avaliação da aprendizagem em educação** online. São Paulo: Edições Loyola, 2006.

SILVER, K. **What Puts the Design in Interaction Design**. 2007. Disponível em: <https://www.uxmatters.com/mt/archives/2007/07/what-puts-the-design-in-interaction-design.php>. Acesso em: 16 jul. 2020.

SMITH, G. C. Foreword: What is interaction design. In **Designing interactions**. Cambridge, MA: MIT Press, 2007. Disponível em: [http://www.designinginteractions.com/downloads/DesigningInteractions\\_foreword.pdf](http://www.designinginteractions.com/downloads/DesigningInteractions_foreword.pdf). Acesso em: 17 jul. 2020.

SIMS, R. Interactivity: a Forgotten Art? **Computer in Human Behavior**, v. 13, n. 2, p. 157-180, may. 1997.



---

# A WebQuest como proposta de Avaliação Digital no contexto da Aprendizagem Significativa Crítica em Ciências para o Ensino Médio<sup>1 2</sup>

*Marcello Ferreira<sup>3</sup>*

*Danielle Xabregas Pamplona Nogueira<sup>4</sup>*

*Olavo Leopoldino da Silva Filho<sup>5</sup>*

*Marcos Rogério Martins Costa<sup>6</sup>*

*Joaquim José Soares Neto<sup>7</sup>*

---

<sup>1</sup> Financiado pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd) da Universidade Federal de Juiz de Fora (UFJF).

<sup>2</sup> Texto original publicado na Revista Pesquisa e Debate em Educação, v. 12, n. 1, p. 1-32, 2022. DOI: <https://doi.org/10.34019/2237-9444.2022.v12.35023>.

<sup>3</sup> Instituto de Física, Universidade de Brasília, e-mail: marcellof@unb.br.

<sup>4</sup> Faculdade de Educação, Universidade de Brasília, e-mail: daniellen@unb.br

<sup>5</sup> Instituto de Física, Universidade de Brasília, e-mail: olavolsf@unb.br.

<sup>6</sup> Instituto de Letras, Universidade de Brasília, e-mail: marcosrmcosta15@gmail.com.

<sup>7</sup> Instituto de Física, Universidade de Brasília, e-mail: jjsoaresneto@gmail.com.



## 1. Introdução

As Tecnologias Digitais de Informação e Comunicação (TDIC) têm sido aplicadas e qualificadas no contexto educacional. Isso não ocorre, por óbvio, sem que sejam também crescentes os debates acerca da pertinência e da efetividade de tal adoção. Referências como Kenski (2009), por exemplo, indicam que todas as aprendizagens, em todos os tempos, são mediadas pelas tecnologias disponíveis.

É relevante destacar, ainda, que os estudantes desta geração são altamente conectados e imersos em uma cultura não apenas global, mas, sobretudo, digital, marcada por interconexões, pela interação e cooperação em comunidades virtuais, assim como pela inteligência coletiva (KÄMPF, 2011; LÉVY, 2003). Com isso, são considerados impróprios os processos de ensino e de avaliação baseados em uma ou poucas estratégias didático-pedagógicas. Isso não é posto apenas para reiterar o que se tem mostrado evidente, mas para sublinhar que as alternativas passam por assumir que as práticas sociais afeitas a esses estudantes – e, em particular, aquelas de âmbito educacional – devem observar a natividade digital e todas as decorrências cognitivas e culturais adjacentes (COELHO; COSTA; MATTAR NETO, 2018).

Emergiram, pois, outras formas de aprender que pressupõem novas formas de ensinar. Essa é a visão de Costa, Duqueviz e Pedroza (2015) que afirmam a existência da expectativa de que essa configuração de uso das tecnologias influencie os processos de aprendizagem, a partir de mudanças provocadas nas formas de interação e de socialização das informações mediadas por elas.

Assim, é necessário que a práxis pedagógica com o uso das TDIC seja dimensionada a partir de uma outra forma de ensinar, perpassada pela inclusão digital dos professores, traço fundante de suas formações profissionais (KENSKI, 2009; MASETTO, 2000). Nessa perspectiva, a integração da cultura digital na mediação pedagógica depende mais fortemente do trinômio formação-engajamento-motivação que de infraestrutura e equipamentos (KEMPFER, 2012).

Como resultado dessa integração, os processos educacionais passam a se dedicar à mobilização de aprendizagens significativas. Desenvolver ambientes e estratégias educativas centradas nos estudantes e na promoção da sua autonomia é fundamental para induzir e qualificar a permanência e o sucesso escolar; adequar esses ambientes e dispositivos a um padrão não linear de aprendizagem é crucial para que não se perca a sua coordenação às demandas cognitivas próprias da cultura digital.



Salienta-se que o ambiente formal de ensino permite estabelecer um elemento focal, centralizador das informações hipertextuais e capaz de impedir, como afirma Goleman (2013), o caráter dispersivo inerente ao ambiente virtual, em especial a rede mundial de computadores (web).

Nesse sentido, há que se reconhecer a pertinência de uma mudança também nos processos de ensino e aprendizagem, com foco em uma perspectiva formativa da avaliação, capaz de agregar as duas dimensões mencionadas, em vez de obliterar uma em prol da outra. De fato, a maneira de avaliar com frequência remete à intencionalidade do processo educativo como um todo (MOREIRA; MASINI, 2016). Diferentemente de processos tradicionais de ensino e avaliação, aqueles de ensino e aprendizagem atentos aos movimentos socioculturais da contemporaneidade pressupõem o desenvolvimento, por parte dos estudantes, de pensamento crítico, reflexivo, profundo, investigativo, metódico e criativo em diversos contextos. Isso se pode materializar por uma abordagem individualizada, não se restringindo ao ensino mecânico, baseado em memorização e ao mero ateste de conhecimentos, mas à indicação de *feedbacks* que conduzam o estudante ao objetivo a ser alcançado.

Para a superação dessa situação mecanicista de ensino, pode-se lançar mão de uma abordagem que proponha uma aprendizagem que seja significativa (AUSUBEL, 1969; 2000), atenta a conhecimentos-base, ou âncoras, materiais significativos, abordagens multifacetadas e foco na ressignificação produtiva de saberes. Essa abordagem, na perspectiva desta pesquisa, é compatível com a *aprendizagem on-line*, isto é, com a aprendizagem que ocorre no ambiente virtual, eletrônico.

A aproximação entre a *abordagem da aprendizagem significativa*, de Ausubel (2000), e a *aprendizagem on-line* é possível. Isso decorre, porque, como investigamos neste estudo, os elementos precípuos da aprendizagem significativa podem ser articulados via apresentação neste formato por meio de sequências de ensino que podem, por sua vez, serem construídas com o apoio de recursos digitais.

Diante dessas considerações, neste estudo, investiga-se qual modelo de avaliação digital de caráter formativo poderia subsidiar uma aprendizagem significativa crítica, com recorte na área de Ciências e na etapa do Ensino Médio. Esse recorte se justifica, em primeiro lugar, pelo fato de que, de acordo com o Índice de Desenvolvimento da Educação Básica (IDEB), o Ensino Médio brasileiro tem se apresentado como a etapa da Educação Básica de menor rendimento entre as demais – por exemplo, em 2019, obteve a média 4,2, que foi abaixo da meta projetada de 5,2 (INEP, 2019). Isso denota, entre outros aspectos, a necessidade de melhoria dos processos relacionados ao Ensino Médio, a



fim de se garantirem estratégias que impulsionem melhores resultados, com base em ações mais efetivas voltadas ao rendimento escolar e à trajetória dos estudantes. Em segundo lugar, justifica-se o recorte na área de Ciências por ser esta formada por um grupo de disciplinas que, usualmente, são reputadas como de maior grau de dificuldade, envolvendo, não raro, elementos de maior abstração no processo de aprendizagem. A partir de um modelo de avaliação digital, baseado nos pressupostos de avaliação formativa e da aprendizagem significativa crítica, este trabalho apresenta e analisa uma proposta de ensino por meio das WebQuests.

A WebQuest é um formato de ensino orientado no qual os recursos utilizados para resolver uma determinada tarefa são oriundos da Web. Bernie Dodge divulgou essa proposta pela primeira vez, em 1995, partindo, sobretudo, das primeiras experiências de uso da web nas unidades escolares norte-americanas (DODGE, 1995). Para o referido autor, a WebQuest se define como “[...] uma investigação orientada, na qual algumas ou todas as informações com as quais os aprendizes interagem são originadas de recursos da internet, opcionalmente suplementadas com videoconferências”<sup>8</sup>. Dodge (1995) propõe, com base nas dimensões do pensar de Marzano (1992), dois tipos de WebQuest: as curtas e as longas. As *WebQuest curtas* servem para a aquisição e a integração do conhecimento e, por isso, são executadas, em geral, entre uma e três aulas. Já as *WebQuest longas* servem para a ampliação e o refinamento do conhecimento, desenvolvendo-se, comumente, desse modo, no intervalo de uma semana ou até de um mês.

Autores, como March (2004), asseveram que as WebQuest não são *novidade* no espaço pedagógico, uma vez que existem redes de aprendizagem orientadas no âmbito de outras metodologias de ensino. Mesmo assim, a WebQuest é interessante por utilizar, desafiadora e organizadamente, recursos da Web – que, em geral, são tratados de maneira excessivamente genérica ou específica ou são desprezados pelos professores em seu cotidiano nas salas de aula do século XXI. De acordo com Santos e Barin (2014, p. 2):

A WebQuest tem se mostrado uma metodologia efetiva no uso da internet voltada ao processo educacional, visto que parte da definição de um tema e objetivos são propostos pelo professor, o qual realiza uma pesquisa inicial e seleciona uma série de *links* referentes ao assunto, para a consulta de seus alunos, evitando então, uma coleta de dados que não terá importância para a pesquisa.

---

<sup>8</sup> Utilizamos a tradução proposta por Jarbas Novelino Barato, que está disponível em: [https://www.dm.ufscar.br/~jpiton/downloads/artigo\\_webquest\\_original\\_1996\\_ptbr.pdf](https://www.dm.ufscar.br/~jpiton/downloads/artigo_webquest_original_1996_ptbr.pdf) Acesso em: 01 jun. 2021.



Para melhor compreensão das potencialidades da WebQuest, a sua investigação está estruturada, neste artigo, em diferentes momentos. Metodologicamente, a investigação foi segmentada em três partes. Na primeira – inserida no tópico a seguir –, apresentam-se os pressupostos teóricos que ancoram a proposta da WebQuest. Na segunda, é explicitada a metodologia em que foi baseada a proposta supracitada. Na terceira, é descrito o modelo de WebQuest com itens de múltipla escolha e, em seguida, é apresentada aplicação desse modelo em um projeto-piloto que coletou uma amostra de dados com nove alunos do Ensino Médio.

## 2. Delimitação dos pressupostos teóricos

O modelo de ensino com a adoção de WebQuests se desenvolve no interior de marcos teóricos que podem incluir pressupostos avaliativos. É importante, portanto, explicitar tais pressupostos, de modo a se procurar verificar em qual medida o modelo proposto os encampa concretamente.

Sacristán e Gómez (2007) sustentam que avaliar significa estimar, calcular, taxar, valorizar, apreciar ou apontar valor, atribuindo, assim, valoração a alguma coisa. Por isso, avaliar é uma atividade essencialmente humana e está presente em vários momentos do cotidiano do indivíduo, inclusive no contexto educacional. Os autores definem, ainda, que:

Avaliar se refere a qualquer processo por meio do qual alguma ou várias características de um/a aluno/a, de um grupo de estudantes, de um ambiente educativo, de objetivos educativos, de materiais, professores/as, programas, etc., recebem a atenção de quem avalia, analisam-se e valorizam-se suas características e condições em função de alguns critérios ou pontos de referência para emitir um julgamento que seja relevante para a educação (SACRISTÁN; GÓMEZ, 2007, p. 298).

Dessa forma, entendemos que a determinação de valor ou mérito sobre o que é avaliado consiste em um processo sistemático que se dá a partir de critérios estabelecidos previamente, uma vez que julgarão o objeto de análise e o seu resultado. Portanto, a adoção de critério, no processo avaliativo, implica, conseqüentemente, uma tomada de decisão.

Oliveira (2011) define, por sua vez, que a avaliação deve ser um processo contínuo que fornece informações, com procedimentos sistematizados, sobre a complexidade educacional, de modo que os agentes escolares decidam acerca de intervenções e de mudanças necessárias para garantir a



aprendizagem do aluno. A avaliação pode levar as escolas, constantemente, à reflexão, à negociação e ao aperfeiçoamento do processo de ensino-aprendizagem, além de possibilitar o monitoramento da rede de ensino.

Nesses diversos momentos do processo avaliativo, é possível destacar as diferentes funções que a avaliação pode assumir: diagnóstica, formativa e somativa (ARREDONDO; DIAGO, 2009). Dentre essas funções, o modelo que se adota neste trabalho se baseia na avaliação formativa. A pesquisa sugere que a avaliação formativa deva ser preferida, em vez da avaliação somativa e classificatória, por diversas razões. Em primeiro lugar, porque a avaliação formativa incentiva a autorregulação do estudante que, ao utilizar os conhecimentos adquiridos e aqueles recentemente apresentados, pode explorar as próprias habilidades de aprender a aprender. Em segundo lugar, a avaliação formativa pressupõe, segundo Perrenoud (1999), que o ato de avaliar não faz sentido por si apenas, mas pelo próprio processo de ensino e aprendizagem. Finalmente, diferentemente da avaliação somativa que é, predominantemente, pontual – isto é, em determinados períodos de tempo, o teste ocorre –, a avaliação formativa é contínua, porque ocorre em diversos e diferentes períodos de tempo ao longo de todo o plano de ensino proposto pelo professor e/ou unidade escolar. Portanto, a avaliação formativa é a mais adequada quando se prevê uma aprendizagem significativa, nos moldes da proposta de Ausubel (1969; 2000), que exige etapas importantes e sequenciais, como aquelas do levantamento dos conceitos prévios, os chamados subsunçores, de sua organização e de processos analíticos e sintéticos de aprendizagem (FERREIRA et al., 2020).

Ressalta-se que a concepção de avaliação formativa tem o poder de, a depender de sua proposição, contribuir para a reestruturação do fazer pedagógico, independentemente da modalidade (presencial ou digital) em que se insira. Nesse sentido, ela é capaz de articular as dimensões políticas e formais já mencionadas, muito mais que a avaliação de caráter somativo, tais como as avaliações em larga escala. Isso não implica, evidentemente, que ambas não possam coexistir, mas representando objetivos distintos do processo de ensino e aprendizagem.

Por isso, a pesquisa indica que a perspectiva de um aprendizado significativo, incorporada a um método de avaliação – primordialmente, formativo (PERRENOUD, 1999) –, é capaz de promover práticas interativas de busca de informação para o ataque a problemas investigativos voltados à pesquisa. Esse pressuposto teórico tende a ser mais eficiente, porque reconhece a ambiência com



que os alunos já estão familiarizados (conhecimentos prévios e universo digital) e o seu processo de (inter)ação com os conhecimentos (avaliação formativa, na perspectiva da aprendizagem significativa).

Ainda no escopo avaliativo, o modelo de WebQuest aqui proposto apropria-se da técnica de Geração Automática de Itens (AIG) que, segundo Gierl e Haladyna (2012), representa uma área de pesquisa recente, na qual teorias cognitivas e psicométricas específicas são associadas às práticas de construção de testes, com o objetivo de produzir itens de teste digitais. A AIG pretende desenvolver processos avaliativos por meio do uso de tecnologias inovadoras para o desenvolvimento e a validação de itens.

Tendo como referencial os mencionados autores, compreende-se que cada item possui um objeto de conhecimento e uma demanda cognitiva pretendida, considerando-se dois tipos de domínios, quais sejam: (a) usar conhecimentos, habilidades e estratégias de maneiras complexas; e (b) concentrar-se em uma única capacidade cognitiva cujas habilidades representam estruturas mentais complexas, que crescem lentamente.

A demanda cognitiva, por sua vez, refere-se à complexidade mental na execução de uma tarefa. Destarte, a tarefa pode ser um item de teste em que o estudante seleciona entre opções ou cria uma resposta a um item, pergunta ou comando. Essa versatilidade na proposição da interação do item é um fator que contribui para a implementação da WebQuest, fundamentada nessa proposta da AIG.

Acompanhando a crescente popularidade da mídia digital e o uso da internet, a transição do teste impresso para o digital vem ocorrendo na última década, surgindo, assim, novos tipos de testes e recursos de teste. De acordo com Gierl e Haladyna (2012), um Teste Adaptativo Computadorizado (CAT) é de caráter digital e usa um modelo que implementa um processo de seleção e de administração de itens, pontuando as respostas do examinado e atualizando sua estimativa de capacidade após a administração de cada item. Esse processo de seleção de novos itens, com base nas respostas do examinado aos itens administrados anteriormente, é continuado até que uma regra seja cumprida.

Para Gierl e Haladyna (2012), os testes computadorizados digitais oferecem muitas vantagens para os estudantes e os educadores, em comparação com os testes tradicionais baseados em papel. Os computadores permitem o desenvolvimento de tipos de itens inovadores e formatos alternativos de itens. Ademais, os itens em testes digitais podem ser pontuados imediatamente, fornecendo



aos examinados um *feedback* instantâneo – o que, no teste impresso, é impossível no caso de aplicação em massa. Outra possibilidade é que os computadores permitem testes contínuos e sob demanda – o que também é inviável, em larga escala, com o teste impresso. Finalmente, a vantagem mais importante do teste digital é que ele favorece a medição de desempenhos mais complexos, integrando itens de teste e mídia digital para melhorar, substancialmente, a medição do pensamento complexo.

Segundo Pasquali (2013), o CAT representa uma aplicação da Teoria de Resposta ao Item (TRI), porque compreende uma teoria do traço latente relacionada a testes de habilidade ou de desempenho. O termo teoria do traço latente se refere a uma família de modelos matemáticos que relaciona variáveis observáveis (itens de um teste, por exemplo) e traços hipotéticos não-observáveis ou aptidões, estes responsáveis pelo aparecimento das variáveis observáveis, ou melhor, das respostas ou comportamentos emitidos pelo sujeito que são as variáveis observáveis. Assim, temos um estímulo (item) que é apresentado ao sujeito e este responde a ele (PASQUALI; PRIMI, 2003).

Os autores descrevem que, na TRI, a resposta que o sujeito dá ao item está relacionada ao nível que ele possui no traço latente, compreendendo-se, assim, que este é a causa e a resposta do sujeito é o efeito. O nível no traço latente, por sua vez, é construído com base em relações hipotéticas entre as respostas observadas, o sujeito e o seu nível nesse mesmo traço latente.

De acordo com Pasquali (2013), a TRI se fundamenta em dois axiomas fundamentais, a saber:

o desempenho do sujeito em uma tarefa (item) se explica em função de um conjunto de fatores ou traços latentes. Logo, o desempenho é o efeito e os traços latentes são a causa ( $\theta$ );

a relação entre o desempenho na tarefa e o conjunto dos traços latentes pode ser descrita por uma expressão chamada Curva Característica do Item (CCI).

Com base nesses axiomas, o uso da TRI no CAT pressupõe que a sequência dos itens apresentados ao examinando depende do seu desempenho no item anterior. Isso pode se constituir em uma maneira de implementar, de maneira automática, a estrutura de aprendizagem significativa, baseada em subsunçores e organizadores prévios, na perspectiva da avaliação formativa.



Em resumo, o modelo de WebQuest proposto neste estudo utiliza os seguintes pressupostos teóricos: a aprendizagem significativa (AUSUBEL, 1969; 2000), a avaliação formativa (PERRENOUD, 1999; SACRISTÁN; GÓMEZ, 2007), o modelo AIG a partir do CAT, seguindo a Teoria do Traço Latente aplicada à TRI (GIERL; HALADYNA, 2012; PASQUALI, 2013; PASQUALI; PRIMI, 2003).

### 3. Metodologia

De forma geral, este trabalho é de caráter descritivo, porque, conforme Gil (2002, p. 42) aponta, se caracteriza no conjunto das pesquisas que “[...] têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou, então, o estabelecimento de relações entre variáveis”. Em nosso caso, busca-se descrever as potencialidades da WebQuest utilizando-a no modelo avaliativo digital voltado ao ensino de Ciências do Ensino Médio. Além disso, este estudo também se caracteriza como uma pesquisa exploratória, uma vez que pesquisas desse tipo “[...] têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições” (GIL, 2002, p. 41). Como foi dito na introdução, retomamos a proposição de Dodge (1995) à luz das recentes pesquisas da área da educação e das potencialidades advindas com a internet e das TDIC.

De maneira específica, este estudo se define como uma pesquisa tecnológica que, segundo Freitas Júnior e Sousa (2018, p. 9), emprega os mesmos meios da pesquisa científica, mas com um objetivo claro: a construção de um artefato. “Por isso, podemos afirmar que a pesquisa de natureza tecnológica possui características que a tornam extremamente adequada para o desenvolvimento de *softwares* e a criação de artefatos, como aqueles produzidos pela área de TIC”.

A proposição de uma pesquisa tecnológica se justifica em razão do objetivo proposto: a construção de um modelo de WebQuest como instrumento de avaliação digital. Com isso, a proposição corrobora Bunge (1985, p. 231, tradução nossa), ao admitir que a tecnologia pode ser vista como “[...] o campo do conhecimento relativo ao projeto de artefatos e ao planejamento de sua realização, operação, ajuste, manutenção e monitoramento, à luz do conhecimento científico”.



Segundo Freitas Júnior e Sousa (2018), a *Design Science Research Methodology* (DSRM) é uma metodologia de pesquisa amplamente empregada no desenvolvimento de investigações tecnológicas, isto é, o seu foco está na produção de um novo artefato. Por isso, busca gerar conhecimento que seja aplicável e útil para a solução de problemas, melhoria de sistemas existentes e, ainda, criação de novas soluções e/ou artefatos. Para tanto, a DSRM compreende cinco etapas.

A etapa 1 do DSRM é reservada para o processo de identificar o problema e a sua motivação. Após esse momento de sensibilização, a etapa 2 busca definir os objetivos da solução proposta. A etapa 3, de *design* e desenvolvimento, compreende projetar e desenvolver o artefato que se pretende construir, empregando, para isso, todas as tecnologias que se julgar necessárias. A etapa 4, de demonstração, prevê a implantação do artefato em um ambiente controlado, próximo do real, para a realização de testes de desempenho, procurando identificar a sua completude e, principalmente, em que medida os objetivos propostos estão sendo contemplados. Na etapa 5, realiza-se a avaliação da solução proposta, com a participação do usuário a quem ela se destina, de modo a aferir a sua precisão, viabilidade, completude e adequação aos requisitos especificados (FREITAS JÚNIOR; SOUSA, 2018).

Como requisito para validação da medida a ser produzida pela proposta de avaliação, também são levantadas evidências de validade com base no conteúdo e no processo de resposta, tendo como referência os *Standards for Educational and Psychological Testing*, os quais são propostos pelas seguintes instituições internacionais: *American Educational Research Association* (AERA); *American Psychological Association* (APA); *National Council on Measurement in Education* (NCME) (AERA; APA; NMCE, 2014).

Com o uso desses métodos de validação, foi realizada uma pesquisa teórica que buscou incorporar os pressupostos apresentados na seção anterior. A segunda etapa foi uma pesquisa tecnológica na qual se construiu o modelo WebQuest. Na subseção a seguir, apresentam-se os pressupostos metodológicos da constituição da medida avaliativa da WebQuest.



### 3.1 Pressupostos metodológicos para constituição da medida avaliativa de aprendizagem significativa em Física

No que se refere aos pressupostos metodológicos para a constituição da medida avaliativa de aprendizagem significativa em Física, proposta neste estudo, são considerados os relativos à Engenharia da Avaliação (EA), à nova taxonomia Marzano e Kendall (2008) e ao Modelo Rasch de Multifacetas (MFRM).

De acordo com Luecht (2012), a EA é uma maneira de abordar todo o processo de projeto e desenvolvimento de testes, itens e escalas de pontuação. O autor argumenta que a EA não é uma tecnologia específica e nem um modelo psicométrico. É uma estrutura abrangente para reconceitualizar e, em seguida, projetar e implementar, baseando-se em princípios, um sistema de ponta a ponta para construir, de maneira eficiente e consistente, itens e testes que suportam uma ou mais escalas de pontuação válidas.

Com isso, a AE emprega princípios e processos de engenharia e fortes controles de qualidade para direcionar o projeto, o desenvolvimento e a montagem, bem como a análise, a pontuação e os relatórios de exames e os resultados de avaliação. São quatro os processos empregados em EA, a saber: (1) mapear construtos e modelagem de evidências; (2) construir modelos de tarefas e mapas de modelos de tarefas; (3) projetar modelos e escrever itens; (4) calibrar as tarefas de avaliação e o controle de qualidade.

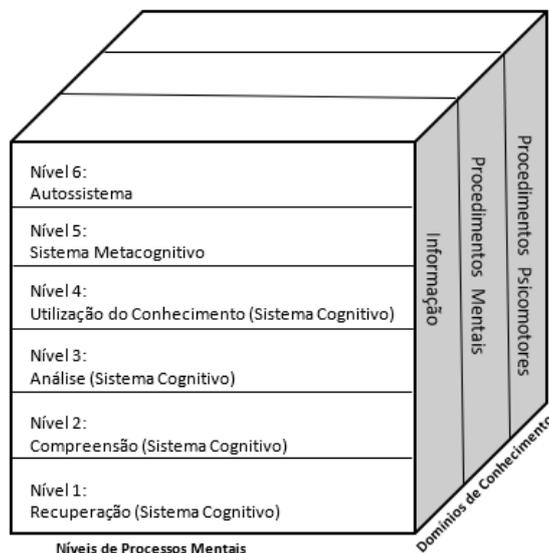
O mapeamento de construtos é uma descrição ordenada de conhecimentos e desempenho de habilidades ao longo de uma escala de pontuação pretendida. O desenvolvimento de um mapa de construtos envolve a elaboração cuidadosa de um conjunto ordenado de proficiência. Então, o mapa de construtos também pode ser ampliado com base nos modelos que descrevem o desempenho esperado, incluindo exemplos de produtos ou outras evidências baseadas no desempenho de um determinado estudante, quanto ao desenvolvimento de habilidades cognitivas e o conhecimento implícita ou explicitamente exigidos por cada nível de proficiência – o que está em consonância com a proposta da aprendizagem significativa, de Ausubel (1969; 2000).



Para orientar a elaboração do mapa de construtos, utiliza-se como referência a Nova Taxonomia, de Marzano e Kendall (2006). Esses autores propuseram outra taxonomia, após compreenderem que a taxonomia de Bloom (BLOOM et al., 1956) não contemplava todos os aspectos e as possibilidades de ensino e aprendizagem – o que, inclusive, já era enfatizado no texto original. De acordo com os autores, a taxonomia de Bloom, ao considerar que as diferenças entre os níveis da taxonomia se baseiam em graus de dificuldade de processamento mental, desconsidera o princípio de que mesmo o mais complexo dos processos pode ser aprendido no nível em que é realizado com pouco ou nenhum esforço consciente. Marzano e Kendall (2006) compreendem, por sua vez, que, embora os processos mentais não possam ser ordenados hierarquicamente em termos de dificuldade, eles podem ser classificados por controle, pois alguns processos exercem controle sobre a operação de outros.

A Nova Taxonomia é composta de uma estrutura bidimensional na qual uma dimensão corresponde aos seis níveis de processos mentais e a outra, aos três domínios de conhecimento. Os processos mentais são organizados em seis níveis, representando o grau de complexidade da demanda para realização da operação.

**Figura 1. Estrutura da Nova Taxonomia**



Fonte: Tradução da proposta de Marzano e Kendall (2006, p. 14).

Ainda para a constituição do modelo de medida, referenciamos o Modelo Rasch de Multifacetado (MFRM). Segundo Linacre (1994), esse modelo surgiu da necessidade de considerar que métodos de medida convencionais não apreendem outros fatores que podem influenciar a avaliação e/ou



os resultados dos testes. Corrobora com esse pensamento a proposta de Eckes (2009) para quem as situações de avaliação são caracterizadas por conjuntos distintos de fatores envolvidos direta ou indiretamente na obtenção dos resultados da medição. No MFRM, esses fatores são reconhecidos como facetas. Assim, uma faceta pode ser definida como qualquer fator, variável ou componente da avaliação que, supostamente, afeta os resultados dos testes de maneira sistemática. Ainda segundo Eckes (2009), essa definição incluiu facetas que são de interesse substantivo (por exemplo, examinandos), bem como facetas que se supõe que contribuam com erros sistemáticos de medição (por exemplo, avaliadores, tarefas, entrevistadores, hora do teste). Também, as facetas podem interagir entre si de várias maneiras. Elementos de uma faceta (por exemplo, avaliadores) podem influenciar os resultados dos testes quando alinhados com subconjuntos de elementos de outra faceta (por exemplo, examinadores do sexo feminino ou masculino). Além das interações bidirecionais, interações de ordem superior entre elementos específicos ou subconjuntos de elementos de três ou mais facetas podem ser consideradas e afetar as pontuações dos testes de maneira sutil.

Voltando à discussão trazida por Linacre (1994), destacam-se os seguintes avanços em relação aos modelos anteriores quando se faz uso do MFRM:

- a) a contribuição de cada faceta para as observações é dominada por um único parâmetro, com um valor independente de todos os outros parâmetros do quadro de referência;
- b) os parâmetros se combinam de maneira aditiva para produzir as observações. A aditividade implica que todos os parâmetros compartilhem uma escala linear. A linearidade não apenas ajuda o analista a entender o significado subjacente aos dados, mas também fornece uma base útil para análises adicionais;
- c) a estimativa de qualquer parâmetro depende da acumulação de todas as classificações nas quais participa, mas é independente dos valores particulares de qualquer uma dessas classificações.

Apreendidas essas vantagens, cabe registrar, ainda, a relevância do uso do MFRM em avaliações educacionais. Para Oliveira (2019, p. 53), “[...] a inovação desse modelo é, justamente, sua capacidade de estimar as facetas e com isso estimar de forma mais precisa a habilidade do examinando”. Com isso, os resultados da medida e os eventuais enviesamentos podem ser apresentados e interpretados objetivamente.



Segundo Toffoli e Simon (2018), o modelo MFRM permite análises para os efeitos individuais decorrentes de cada elemento que faz parte da avaliação (examinando, avaliador, tarefas etc.). Dessa forma, o modelo dispõe da possibilidade de obter informações que possam servir de diagnóstico, em nível individual, a respeito do funcionamento de cada elemento, além do estudo das influências de outros fatores incluídos nas análises.

Silva (2019, p. 137) demonstra que “a substituição da TCT pela MFRM é uma alternativa plenamente viável e com mais benefícios”, sobretudo pela característica do modelo que pressupõe a proficiência por alinhamento ao padrão de acerto e erro no teste e velocidade de disponibilização dos resultados, o que potencializa a interpretação e a intervenção pedagógica.

Considerando esses benefícios da MFRM para o seu uso, o modelo de avaliação conta, neste trabalho, com quatro facetas, as quais se baseiam na seguinte equação, que foi adaptada de Linacre (1994):

$$\log \log \left( \frac{P_{nijk}}{P_{nijk-1}} \right) = B_n - D_i - I_j - S_k \quad (1)$$

Em que:

$B_n$  = Habilidade do examinando, segundo a Teoria da Aprendizagem Significativa

$D_i$  = Dificuldade do item

$I_j$  = Função de interatividade do item

$S_k$  = Subárea da Física

Passamos, agora, a explicar cada uma das quatro facetas ( $B_n$ ;  $D_i$ ;  $I_j$  e  $S_k$ ) que compreende esse modelo de MFRM. No que se refere à  $B_n$  (Habilidade do examinando, segundo a Teoria da Aprendizagem Significativa), os níveis de aprendizagem significativa, para efeitos do modelo avaliativo tal como o estamos entendendo, estão contidos na escala de proficiência. Essa escala foi construída para este estudo, durante a pesquisa. A referida escolha está alinhada ao mapeamento de construção de tarefas. A seguir, apresentamos a escala e suas evidências em quatro níveis constituídos em níveis análogos e retroativos à já referida perspectiva ausubeliana (AUSUBEL, 1969; 2000).



**Quadro 1. Proposta de Taxonomia a partir da Teoria da Aprendizagem Significativa**

Níveis de Performance	Evidências
<b>Nível 4: REELABORAR</b>	NE 4: Avaliações e pesquisas de modo satisfatório e com interação crítica. Nesse nível, as operações são executadas com alto grau de desafio, alta velocidade e precisão perfeita. Observam-se indícios sustentados de aprendizagem significativa por meio de construções e relações, com adequada reelaboração e expressão e formulações críticas.
<b>Nível 3: AVALIAR</b>	NE 3: interpretações, aplicação de princípios, reconhecimento de características e prototipagem de modo satisfatório. Nesse nível, as operações são crescentemente complexas em sistemas simples, com média precisão e velocidade, grau mediano de desafio e com produtos rotineiros realizados rapidamente e com alta precisão. Observam-se indícios sustentados de aprendizagem significativa, por meio de construções e relações, com adequada reelaboração e expressão.
<b>Nível 2: INTERPRETAR</b>	NE 2: Análises, comunicação e representações em caráter intermediário. Nesse nível, as operações são moderadamente complexas em sistemas simples, com média precisão e velocidade, com grau baixo de desafio e com produtos rotineiros realizados rapidamente e com precisão satisfatória. Evidenciam-se indícios de aprendizagem significativa por meio de construções e relações, sem a adequada reelaboração e expressão.
<b>Nível 1: EXECUTAR</b>	NE1: Explicações, previsões, cálculos em caráter introdutório ou precário. Nesse nível, as operações são executadas de forma simples e envolvendo pequenos conjuntos de objetos de conhecimentos, com baixa precisão e velocidade. Não há indícios consistentes de aprendizagem significativa e os resultados, apesar de estarem relacionados ao objeto, são imprecisos, inconsistentes ou apresentam erros importantes de conceitos.

Fonte: Elaboração dos autores, baseada em contribuições de Marzano e Kendall (2006), Ausubel (1969; 2000) e Moreira (1999; 2017a).

Quanto à *Di* (Dificuldade do item), consiste em um parâmetro psicométrico relacionado aos acertos do examinando. Esse parâmetro é medido em um piloto do teste. Em outras palavras, esse item será (re)formulado a partir dos resultados obtidos pelos examinandos. Para tanto, faz-se necessário um piloto de teste – o qual abordaremos no tópico 4 deste trabalho.

Referente à *Ij* (Função de interatividade do item), é um item que foi apropriado de alguns autores para elucidar esse aspecto. Tomaz (2011) propõe a diferenciação dos conceitos de interação e interatividade. Segundo a autora, “[...] “a interatividade surge relacionada com os atributos tecnológicos que promovem trocas entre homem-sistema e a interação como uma relação que se estabelece numa situação, seja com um conteúdo, tutor ou sistema tecnológico associado” (TOMAZ, 2011, p. 2655).



Dentre as dimensões de interatividade, destaca-se a interatividade para uma aprendizagem significativa, a qual realça os aspectos relacionados com o envolvimento cognitivo e não tanto os aspectos físicos presentes na interação. A respeito dessa dimensão, Sims (1999) reporta ao estudo de Hannafin (1989), o qual, por sua vez, ao desenvolver o conceito de engajamento, sugeriu que "[...] estamos preocupados com a maneira pela qual a instrução promove o engajamento cognitivo - o processamento intencional e proposital do conteúdo da lição" (HANNAFIN, 1989, p. 170). Para isso, é necessário que o processo de aprendizagem foque menos nos aspectos físicos e evidentes da interação e mais nas atividades cognitivas e pensantes do recurso de aprendizado incorporado. Para fornecer um contexto para a aprendizagem significativa, Hannafin (1989) identificou um conjunto de cinco funções interativas (navegação, consulta, verificação, elaboração, controle processual) e sugeriu um conjunto de atividades para apoiar essas funções.

A interatividade e a aprendizagem significativa também são consideradas, no que diz respeito ao modelo de esquema da memória humana, em que as informações são percebidas como armazenadas em uma rede de nós interconectados. "A força do conhecimento se baseia não apenas no número de nós existentes, mas mais na qualidade e número de interconexões entre os nós" (PARRISH, 1996, p. 2). A interatividade que incentiva o processamento cognitivo mais profundo (CRAIK; LOCKHART, 1972), potencialmente, leva à construção dessas interconexões.

Hannafin (1989) afirma que a interação pode ser considerada como uma realização de uma ou mais instruções, funcionando desde o simples fornecimento de controle processual até a viabilização de níveis diferenciados de processamento cognitivo. A seguir, são apresentadas, no quadro 2, as funções de interação consideradas neste estudo:



**Quadro 2. Funções de interação e respectivas premissas psicológicas, estratégias comuns de interação e estratégias adicionais de engajamento**

Funções de interação	Premissas psicológicas	Estratégias comuns de Interação	Estratégias adicionais de engajamento
<b>Navegação</b>	Habilidades metacognitivas; Orientação para os componentes da lição.	Menus/botões de opção.	Apresenta opções de estrutura de hipertexto; Questiona por que as seções escolhidas foram selecionadas.
<b>Consulta</b>	Apoio ao conhecimento prévio; Habilidades metacognitivas; Assimilação de respostas ao esquema.	Menu estruturado por consulta; Perguntas de linguagem natural; Opções para obter mais informações; Referências a informações relacionadas.	Pede resposta prevista; Questiona por que a pergunta é importante; Pede para identificar perguntas e conceitos relacionados.
<b>Verificação</b>	Recuperação do conhecimento codificado para Short-Term Memory (STM) <sup>9</sup> ; Reforço da aprendizagem.	Questões embutidas; <i>Feedback</i> apropriado para respostas; Ramificação condicional.	Pede estimativas de confiança; Pede aos usuários que gerem perguntas que avaliem suas habilidades; Pede respostas em tempo real.
<b>Elaboração</b>	Apoio ao conhecimento prévio; Codificação reforçada; Propagação da ativação entre nós relacionados; Maior facilidade de recuperação	Estratégias induzidas "pensar em"; Introduzir relacionamentos com conteúdo familiar; Fornecimento de exemplos.	Pede outros casos em que os conceitos foram aplicados; Pede explicações sobre porque as respostas estão corretas ou não; Emprega diálogo cooperativo para ampliar a contribuição disponível
<b>Controle processual</b>	Habilidades metacognitivas; Short-Term Memory (STM) não sobrevalorizado	Exemplos de operações solicitadas: "Pressione o botão para continuar..." "Toque na tela quando tiver visto o suficiente para responder"	Pede resumos, escritos com as próprias palavras; Pede para gravar notas ou pontos pouco claros do conteúdo; Pede para gerar perguntas sobre o conteúdo abordado.

Fonte: Tradução e adaptação da proposta de Hannafin (1989).

<sup>9</sup> Short-Term Memory (STM) é a capacidade de reter pouca informação por um curto tempo.

Partindo das funções, outro fator examinado é a faceta *Sk* (Subárea da Física no Ensino Médio). Esse item é apoiado no entendimento de que a Física, como conhecimento organizado, é uma herança de mais de dois milênios de estruturação discursiva, que se ocupa dos fenômenos naturais, com campos epistemológico, teórico e metodológico próprios. Para tanto, adotou-se a versão mais contemporânea da divisão dessa disciplina, na perspectiva dos currículos escolares básicos (CARVALHO et al., 2010). Essa versão é a que apresentamos e a que foi considerada para esse modelo:

- *Mecânica*: estuda o movimento. É subdividida em: cinemática, dinâmica, gravitação, estática e hidrostática;
- *Óptica e Ondulatória*: estuda os fenômenos ligados à luz e pode ser subdividida em óptica geométrica e óptica física; estuda as ondas e engloba a acústica, que trata das ondas sonoras;
- *Termologia*: estuda os fenômenos térmicos. É subdividida em termometria, calorimetria e termodinâmica.;
- *Eletromagnetismo*: estuda os fenômenos elétricos e magnéticos. É subdividida em eletrostática, eletrodinâmica e eletromagnetismo;
- *Física moderna e contemporânea*: envolve relatividade geral e especial, mecânica quântica e física nuclear, além de tópicos de Astronomia.

Essas subáreas compuseram as facetas dos itens digitais avaliativos em Física no modelo WebQuest. Dessa maneira, os conteúdos disciplinares correspondentes a essa área do conhecimento foram utilizados como objeto de ensino, os quais, por sua vez, constituíram o teste com relação à sua interveniência (ou não) nos desempenhos dos examinandos. No próximo tópico, utilizados os conhecimentos teóricos e metodológicos apresentados nesta seção, trazemos à luz o modelo de WebQuest relacionado ao ensino de Física dirigido ao Ensino Médio – objetivo deste trabalho.



## 4. Constituição do modelo de WebQuest

### 4.1 Da teoria à prática

Ausubel (2000) define que a aprendizagem de determinado conjunto de conceitos, associados a um campo conceitual ou parte dele, é a internalização desse conjunto de conceitos pela estrutura cognitiva do aprendiz. Desse processo de internalização decorre a necessidade de reorganização da estrutura cognitiva, bem como da própria informação em uma interação bidirecional.

Com efeito, a estrutura cognitiva é compreendida como estando hierarquicamente organizada, segundo conceitos que se vinculam de modo relacional, constituindo uma árvore ou grafo. Tal estrutura não está vazia, mas contém elementos prévios, os subsunçores, também de caráter conceitual, que podem servir de ponto de partida para a aprendizagem de novos conceitos. Segundo Ausubel (1969), a aprendizagem pode ser metaforicamente concebida como um processo de ancoragem dos novos conhecimentos em uma estrutura cognitiva por meio de subsunçores.

Ausubel (1969) define, então, *aprendizagem significativa* como o processo pelo qual uma nova informação estabelece uma relação com aspectos específicos e relevantes da estrutura cognitiva já apresentada pelo indivíduo. Tal modo de aprendizagem diferencia-se, portanto, de uma *aprendizagem mecânica*, que ocorre sem referência substancial à estrutura cognitiva do aprendiz e implica uma assimilação arbitrária pela memorização. “Assim, estabelecem-se dois polos opostos de um processo que, de fato, vive nas regiões intermediárias, entre uma aprendizagem totalmente significativa e uma aprendizagem absolutamente mecânica” (SILVA FILHO; FERREIRA, 2018, p. 107).

A aprendizagem significativa é definida pela assimilação de significados claros, precisos, diferenciados e transferíveis (MOREIRA, 1999). Para que tal aprendizagem ocorra de maneira eficaz, é preciso que se construa uma ponte cognitiva entre os subsunçores do aprendiz e o conjunto de conceitos a serem aprendidos por ele. Essa ponte cognitiva se constrói, segundo Ausubel (1969), por meio dos organizadores prévios. Diante daquilo que já foi dito, fica claro que o levantamento prévio dos subsunçores é indispensável para o estabelecimento da aprendizagem significativa.



O elemento diacrônico da teoria ausubeliana decorre da possibilidade de enriquecimento da estrutura cognitiva, que estabelece conjuntos de subsunçores, cada vez mais amplos, capazes de induzir uma apreciação progressivamente mais profunda das experiências do indivíduo. Isso é algo que não ocorre, de maneira geral, na aprendizagem mecânica.

Os aspectos que sustentam a ideia de aprendizagem significativa são compatíveis com a perspectiva da WebQuest, uma vez que, como sustenta Dodge (1995, p. 11):

As WebQuests podem ser aperfeiçoadas com elementos motivacionais que envolvam a estrutura básica de investigação, dando aos aprendizes um papel a ser desempenhado (cientista, detetive e repórter, por exemplo), criando uma personalidade fictícia com a qual os participantes deverão interagir via e-mail, e apresentado um cenário dentro do qual os participantes irão trabalhar [...].

Com esses aspectos motivacionais considerados, pode-se apreender que a WebQuest é uma proposta que abarca os subsunçores dos aprendizes.

Esses aspectos são considerados na proposta da WebQuest, a ser descrita no próximo subtópico. Isso decorre, porque, segundo Dodge (1995, p. 10), tanto as WebQuests de curta quanto aquelas de longa duração são planejadas “[...] deliberadamente para fazer o melhor uso possível do tempo do aprendiz. É questionável o benefício da ‘surfagem’ pela rede sem uma tarefa clara na cabeça; e muitas escolas devem racionar bastante o tempo de conexão dos alunos”. Em outras palavras, a WebQuest faz o uso reflexivo-crítico dos recursos da web, o que está em consonância com a proposta de Moreira (2017b).

## 4.2 O modelo de WebQuest para o ensino de Física

Segundo Dodge (1995), a WebQuest é uma atividade orientada para a investigação na qual as informações com as quais os discentes interagem advêm de recursos da internet. Essa estratégia representa uma possibilidade de engajamento de estudantes, pois estimula tanto a pesquisa quanto o pensamento crítico e criativo – elementos que servem de índice da ocorrência de aprendizagem significativa crítica (MOREIRA, 2017a; 2017b). Antes de expor o modelo preparado para o ensino de Física, vamos retomar os componentes básicos de uma WebQuest.



Conforme a proposta de Dodge (1995, p. 11), há cinco componentes básicos da WebQuest, a saber:

- 1) Introdução: desperta o interesse e orienta a execução da tarefa;
- 2) Tarefa: explicita o produto esperado ao término da tarefa, observado o nível cognitivo compatível;
- 3) Processo: indica passos procedimentos intermediários para a realização da tarefa e os recursos necessários;
- 4) Avaliação: indica a perspectiva de desempenho, bem como instrumentos e indicadores;
- 5) Conclusão: retoma e resume o aspecto central da experiência, incentiva a reflexão acerca do processo e o reconhecimento do que foi aprendido.

Partindo dessa estrutura básica, o modelo proposto é constituído pelo seguinte processo:

**Figura 2. Estrutura proposta**



Fonte: Elaboração dos autores.

As etapas que constituem a WebQuest são:

- a) Introdução: apresenta-se a temática-alvo, apontando-se as principais características;
- b) Tarefa: descreve o objetivo final da tarefa e indica o processo que o usuário deve realizar para alcançar o objetivo proposto.

De acordo com Simão (2001, p. 60), quando da realização da tarefa, deve-se entender que:

Os passos necessários para o cumprimento da tarefa são descritos no Processo, que pode ainda servir para orientar os estudantes na forma ou estratégia para organizar as informações coletadas para o desenvolvimento do produto na tarefa. Essa organização pode ser na forma de bloco de notas, fluxograma, mapas conceituais, coletânea de endereços de sites na WWW e outros, todos desenvolvidos com auxílio do computador através de software apropriado. No Processo, pode ainda ser sugerida uma escala de tempo para o desenvolvimento de cada etapa ou passo no cumprimento integral da tarefa, buscando otimizar o tempo dos alunos.

- c) **Avaliação:** são apresentadas orientações acerca da avaliação da tarefa. Por isso, são disponibilizados itens digitais interativos e de múltipla escolha. Essa definição difere da versão original de WebQuest, uma vez que, na proposta original, são indicados itens abertos.

Entende-se que, no item da avaliação, a composição de itens de múltipla escolha com feedbacks automáticos atendem melhor às avaliações em larga escala e de cunho formativo. Isso acompanha o referencial de aprendizagem no qual essa proposta está baseada (MOREIRA, 2012; AUSUBEL, 2000).

- d) **Conclusão:** constitui o *feedback* geral da WebQuest. Nessa etapa, há o espaço em que se comenta o desempenho do examinando na WebQuest, situando-o quantos aos seus acertos e erros e habilidades demonstradas.
- e) **Amplie seus conhecimentos:** faz-se, nesse tópico, uma síntese do conteúdo, que, em vez de reduzir as informações, tem como objetivo ampliar os horizontes da aprendizagem. Pode-se fazer isso de diferentes maneiras, a partir dos recursos digitais: desde a indicação de bibliografia complementar e aportes audiovisuais, até videoaulas que aprofundam o conhecimento.

Assim, as WebQuests desta pesquisa não possuem problemas abertos, que podem receber diferentes soluções por diferentes estudantes. Por estarem organizadas segundo itens fechados (múltipla escolha, com alternativas de A até E), as WebQuests são, elas mesmas, problemas fechados. Mesmo assim, procurou-se manter a ideia fundamental das WebQuests, que é a atuação sobre um problema contextualizado, articulando diferentes aspectos de um ou mais objetos de conhecimento.

Outro elemento foi buscado nesse modelo proposto: o caráter ampliativo das WebQuests. Buscou-se relacionar esse aspecto com o fundamento teórico da pesquisa, visto que o critério para aferição de significação da aprendizagem é a capacidade do estudante de articular seus conhecimentos em contextos diversos daqueles em que os aprenderam. Tais características são importantes para que se diferencie a WebQuest (no formato proposto) e um teste contendo questões de múltipla escolha. Destaca-se, ainda, que tais características fazem com que a resolução de uma WebQuest demande tempo significativo na sua execução – o que não ocorreria em um teste de múltipla escolha.



A seguir, apresentamos um exemplo de introdução e de tarefa de WebQuest no formato proposto.

**Figura 3. Exemplo de comando da WebQuest**

**Equação Geral dos Gases - experimento**

Data de Criação: 26/08/2020  
 Cod. do Criador: 38556790100

**Introdução:**

A **equação geral dos gases** estabelece uma relação entre três grandezas fundamentais para a descrição do estado de um gás, são elas: pressão, volume e temperatura. Ela também nos fornece a explicação para uma série de fenômenos que ocorrem no nosso dia a dia. Dentre tais fenômenos está o funcionamento da panela de pressão. Nesse caso, para compreender esses fenômenos, deve-se conhecer a relação entre a pressão, temperatura e o volume dos gases. Nesta WebQuest, você será levado a desenvolver suas ideias sobre a relação entre essas variáveis e, finalmente, compreender o fenômeno do cozimento de alimentos usando-se uma panela de pressão.

**Tarefa:**

Nesta WebQuest você usará um simulador PhET que você pode encontrar [aqui](#). Este simulador permite alterar controladamente as variáveis termodinâmicas que compõem o fenômeno a ser estudado. O objetivo da sua empresa é desenvolver uma panela de pressão inovadora. Entretanto, para tal, será preciso compreender os princípios físicos que subjazem aos fenômenos físicos que relacionam a temperatura, a pressão e o volume de gases em recipientes, situação fisicamente análoga àquela de uma panela de pressão. É aí que você, o físico da empresa, entra...



Com este simulador [aqui](#), você irá desenvolver as principais ideias relacionadas à Lei dos Gases. No processo de investigação da Lei dos Gases, devemos manipular as variáveis intervenientes de maneira controlada, para chegarmos a uma representação correta do fenômeno estudado. Assim, caberá a você desenvolver tais conhecimentos para que sua empresa, de posse deles, possa trabalhar com seus engenheiros ou desenhistas industriais no sentido de produzir uma panela de pressão inovadora.

**Conclusão:**

Obter uma equação em Física que represente um fenômeno natural não é nada fácil. Trata-se de saber que variáveis são importantes para aquele fenômeno. Trata-se, também, de saber como essas variáveis se relacionam entre si. Isso significa, na maioria absoluta das vezes, de se conseguir manter um controle sobre tais variáveis, mantendo-se algumas fixas e variando-se outras, de modo a ver como essas relações se dão em situações mais simples... Até que o fenômeno completo possa ser verificado em toda a sua generalidade.

Fonte: Elaboração dos autores.

Nessa atividade, o objeto de conhecimento corresponde à Lei dos Gases. Para tanto, segue-se uma apresentação contextualizada e, depois, uma tarefa é indicada. Para realizá-la, o examinando deverá abrir um *link*, explorar uma ferramenta digital e, então, simular, virtualmente, as ideias teóricas apresentadas. Há, com esse processo, uma experiência ativa com a faceta apresentada para a medição da habilidade buscada no item, o que a descaracteriza como uma mera apresentação contextualizada. Ela é uma atividade imersiva que trabalha com diferentes recursos e atividades com o alunado em seu processo de aprendizagem.

Na figura a seguir, há um exemplo de WebQuest, com itens de múltipla escolha, desenvolvido com base na faceta da Lei dos Gases. Para a elaboração dos itens de cada WebQuest, foram desenvolvidas instruções gerais. Com efeito, os elaboradores devem considerar: a) conteúdos expressos na classificação e codificação de subárea da Física por conjunto de objetos de conhecimentos na Educação Básica; b) demanda cognitiva: os itens devem considerar um objeto de conhecimento e serem constituídos em gradação crescente de complexidade de demanda cognitiva; e c) níveis segundo a função de interatividade digital no item.

**Figura 4. Exemplo de item constitutivo da WebQuest**

Item Nº: 1	Valor do Item: 2	ID do Item: 10
Este item refere-se à lei de Boyle, como elemento preparatório para a compreensão da lei dos Gases.	Você deve manipular o aplicativo PhET que se encontra <a href="#">aqui</a> (aba: propriedade dos gases), satisfazendo os critérios associados à lei de Boyle. Com isso, você deverá buscar a representação gráfica do fenômeno.	
Criador: 38556790100	Data de Criação: 8/26/2020 12:00:00 AM	Gabarito: C
Dificuldade: 1	Interatividade: I5	Habilidade: EM13CNT301
Sub-área: OC6-08		
Competência: CNM3		

**Descrição Breve:**  
 Neste item trabalhamos com a ideia da lei de Boyle, que descreve o comportamento do gás ideal quando se mantém sua temperatura constante (transformação isotérmica). Pede-se que o estudante use um aplicativo PhET para encontrar o comportamento da pressão em termos do volume, mantendo-se a temperatura constante.

**Texto do Item:**  
 A lei de Boyle, formulada pelo químico irlandês Robert Boyle (1627-1691), descreve o comportamento do gás ideal quando este se mantém a uma temperatura constante (transformação isotérmica). No processo, faz-se necessário aplicar lentamente uma força sobre essa tampa, pois desse modo não alteraremos a temperatura do gás. Para responder a este item, você deverá manipular o simulador PhET que se encontra [aqui](#). Para as variações no volume, considere que o recipiente apresentado no simulador é, inicialmente, um cubo perfeito, com o comprimento da aresta lá apresentado e que, ao mover uma das paredes, apenas esta dimensão é alterada, levando o cubo a um paralelepípedo. A partir da manipulação do simulador PhET, assinale a alternativa que apresenta a relação mais adequada entre a pressão e o volume, quando se mantém a temperatura constante.

**Alternativas:**

A:

B:

C:

Fonte: Elaboração dos autores.

No quadro a seguir, apresentam-se, de forma sucinta, os objetos de conhecimento considerados na WebQuest, bem como se indicam a subárea e o código adotados.

**Quadro 3. Classificação e codificação de subárea da Física por conjunto de objetos de conhecimentos na Educação Básica**

SUBÁREA	CÓD. OC	OBJETO DE CONHECIMENTO
	OC1	Conhecimentos básicos e fundamentais - Noções de ordem de grandeza. Notação Científica. Sistema Internacional de Unidades. Metodologia de investigação: a procura de regularidades e de sinais na interpretação física do mundo. Observações e mensurações: representação de grandezas físicas como grandezas mensuráveis. Ferramentas básicas: gráficos e vetores. Conceituação de grandezas vetoriais e escalares. Operações básicas com vetores.
<b>Mecânica:</b> cinemática, dinâmica, gravitação, estática e hidrostática.	OC2	O movimento, o equilíbrio e a descoberta de leis físicas – Grandezas fundamentais da mecânica: tempo, espaço, velocidade e aceleração. Relação histórica entre força e movimento. Descrições do movimento e sua interpretação: quantificação do movimento e sua descrição matemática e gráfica. Casos especiais de movimentos e suas regularidades observáveis. Conceito de inércia. Noção de sistemas de referência inerciais e não inerciais. Noção dinâmica de massa e quantidade de movimento (momento linear). Força e variação da quantidade de movimento. Leis de Newton. Centro de massa e a ideia de ponto material. Conceito de forças externas e internas. Lei da conservação da quantidade de movimento (momento linear) e teorema do impulso. Momento de uma força (torque). Condições de equilíbrio estático de ponto material e de corpos rígidos. Força de atrito, força peso, força normal de contato e tração. Diagramas de forças. Identificação das forças que atuam nos movimentos circulares. Noção de força centrípeta e sua quantificação. A hidrostática: aspectos históricos e variáveis relevantes. Empuxo. Princípios de Pascal, Arquimedes e Stevin: condições de flutuação, relação entre diferença de nível e pressão hidrostática.
	OC3	Energia, trabalho e potência - Conceituação de trabalho, energia e potência. Conceito de energia potencial e de energia cinética. Conservação de energia mecânica e dissipação de energia. Trabalho da força gravitacional e energia potencial gravitacional. Forças conservativas e dissipativas.
	OC4	A Mecânica e o funcionamento do Universo - Força peso. Aceleração gravitacional. Lei da Gravitação Universal. Leis de Kepler. Movimentos de corpos celestes. Influência na Terra: marés e variações climáticas. Concepções históricas sobre a origem do universo e sua evolução.
<b>Óptica e Ondulatória:</b> óptica geométrica, óptica física, ondas e acústica	OC5	Oscilações, ondas, óptica e radiação - Feixes e frentes de ondas. Reflexão e refração. Óptica geométrica: lentes e espelhos. Formação de imagens. Instrumentos ópticos simples. Fenômenos ondulatórios. Pulsos e ondas. Período, frequência, ciclo. Propagação: relação entre velocidade, frequência e comprimento de onda. Ondas em diferentes meios de propagação.

SUBÁREA	CÓD. OC	OBJETO DE CONHECIMENTO
<b>Termologia:</b> termometria, calorimetria e termodinâmica.	OC6	O calor e os fenômenos térmicos - Conceitos de calor e de temperatura. Escalas termométricas. Transferência de calor e equilíbrio térmico. Capacidade calorífica e calor específico. Condução do calor. Dilatação térmica. Mudanças de estado físico e calor latente de transformação. Comportamento de Gases ideais. Máquinas térmicas. Ciclo de Carnot. Leis da Termodinâmica. Aplicações e fenômenos térmicos de uso cotidiano. Compreensão de fenômenos climáticos relacionados ao ciclo da água.
<b>Eletromagnetismo:</b> eletrostática, eletrodinâmica e eletromagnetismo.	OC7	Fenômenos Elétricos e Magnéticos - Carga elétrica e corrente elétrica. Lei de Coulomb. Campo elétrico e potencial elétrico. Linhas de campo. Superfícies equipotenciais. Poder das pontas. Blindagem. Capacitores. Efeito Joule. Lei de Ohm. Resistência elétrica e resistividade. Relações entre grandezas elétricas: tensão, corrente, potência e energia. Circuitos elétricos simples. Correntes contínua e alternada. Medidores elétricos. Representação gráfica de circuitos. Símbolos convencionais. Potência e consumo de energia em dispositivos elétricos. Campo magnético. Ímãs permanentes. Linhas de campo magnético. Campo magnético terrestre.
<b>Física moderna e contemporânea:</b> relatividade, mecânica quântica e física nuclear	OC8	Física Moderna e Contemporânea – Radiação do corpo negro. A natureza da luz. Efeito fotoelétrico. Fluorescência e fosforescência. O funcionamento do laser. Relatividade Geral e Especial. Mecânica Quântica. Física Nuclear. Tópicos de Astronomia.

Fonte: Elaboração dos autores, de acordo com a BNCC (BRASIL, 2017) e Carvalho et al. (2010).

Há, ainda, a *Demanda cognitiva*. Nesse aspecto, os itens da WebQuest devem acolher um objeto de conhecimento e ser constituído em gradação de complexidade de demanda cognitiva. Esses níveis são entendidos a partir da sequência apresentada no quadro 1, de N1 a N4, representados na matriz a seguir, elaborada a partir da Base Nacional Curricular Comum (BNCC) e da Taxonomia de Níveis de Aprendizagem Significativa.



**Figura 5. Análise da matriz BNCC para a área de Ciências da Natureza e Matemática segundo Taxonomia de Aprendizagem Significativa**

COMPETÊNCIAS	HABILIDADES																									
	EM13CNT204	EM13CNT208	EM13CNT209	EM13CNT104	EM13CNT201	EM13CNT202	EM13CNT206	EM13CNT308	EM13CNT309	EM13CNT101	EM13CNT105	EM13CNT303	EM13CNT103	EM13CNT301	EM13CNT305	EM13CNT106	EM13CNT107	EM13CNT203	EM13CNT205	EM13CNT302	EM13CNT304	EM13CNT306	EM13CNT307	EM13CNT102	EM13CNT207	
	TAXONOMIA DE NÍVEIS DE APRENDIZAGEM SIGNIFICATIVA																									
	N1: EXECUTAR			N2: INTERPRETAR							N3: AVALIAR							N4: REELABORAR								
CNM1	✓	✓	✓	✓							✓	✓	✓	✓		✓	✓								✓	
CNM2					✓	✓	✓											✓	✓							✓
CNM3								✓	✓					✓	✓					✓	✓	✓	✓			

Fonte: Elaboração dos autores<sup>10</sup>.

O terceiro elemento refere-se aos *níveis de interatividade digital* no item. Esse elemento está fundamentado nas propostas de Tomaz (2011), Sims (1999) e Hannafin (1989) – estudos abordados na metodologia deste estudo. Os níveis de interatividade considerados nesta proposta de WebQuest podem ser acompanhados no quadro 4, que retoma as cinco funções de interação expostas no quadro 2.

**Quadro 4. Funções de interação**

Nível de interatividade	Função da interação
I1	Navegação
I2	Consulta
I3	Verificação
I4	Elaboração
I5	Controle processual

Fonte: Elaboração dos autores, com base nas contribuições de Sims (1999) e Hannafin (1989).

<sup>10</sup> As siglas utilizadas para indicar as habilidades e as competências são as mesmas que constam na BNCC (BRASIL, 2017).

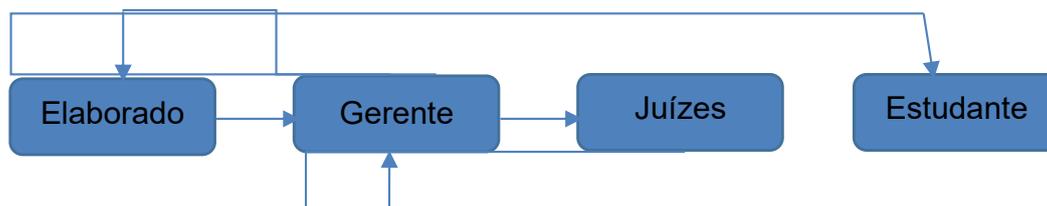
Cada nível de interatividade se relaciona com as premissas psicológicas, as estratégias comuns de interação e as estratégias adicionais de engajamento. O quadro 4 retoma as funções de interatividade associando-as hierarquicamente a níveis que permitem a extrapolação em características operacionais das WebQuests propostas.

### 4.3 Projeto-piloto da WebQuest com itens de múltipla escolha

Para executar esse modelo de WebQuest composto por itens de múltipla escolha, foi realizada a implementação de um projeto-piloto em que foram desenvolvidas dez WebQuests, cada qual com três ou quatro itens, todos com cinco alternativas, seguindo as orientações da classificação e codificação dos objetos de conhecimentos (quadro 3), de demanda cognitiva (figura 6) e de interatividade (quadro 4).

Como forma de concretizar o modelo teórico proposto, foi realizada uma implementação computacional em um *site* totalmente operacional, disponível no seguinte endereço: <https://thewebquest.net/thewebquest>. De modo geral, o sistema implementado no site funciona de acordo com papéis assumidos pelos usuários, a saber: administrador, elaborador, gerente, juiz, estudante. Para cada papel, há funcionalidades que lhe são precípuas. Assim, foi possível estabelecer um fluxo de construção de WebQuests, que tem a seguinte dinâmica:

**Figura 6. Fluxo de produção, gerenciamento, julgamento e realização das WebQuests**



Fonte: Elaboração dos autores.

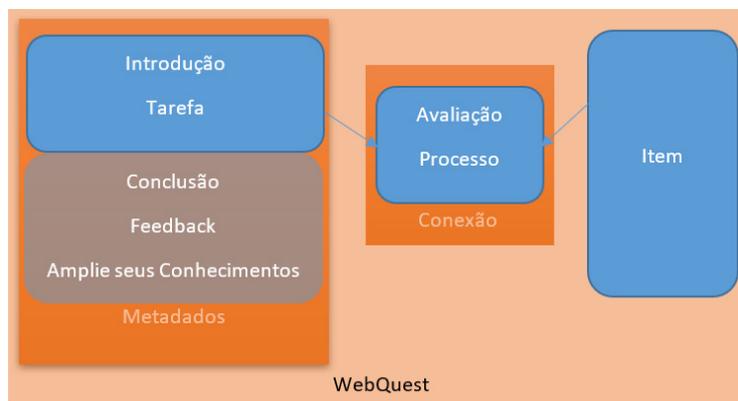
Seguindo esse fluxo, o elaborador é acionado para construir as WebQuests. Uma vez elaborada, a WebQuest é enviada ao gerente, que a analisa e verifica se ela pode ser submetida aos juízes. Caso haja necessidade de ajustes, o gerente deve devolver ao elaborador. O processo se repete até que a WebQuest seja considerada pronta para ser submetida à apreciação dos juízes, para que façam



sua análise. Os juízes analisam as WebQuests como especialistas na área e devolvem o resultado da análise para o gerente. A depender da análise dos juízes, o gerente devolve a WebQuest para o elaborador ou a libera para apresentação aos examinandos.

Ao administrador cabe tarefas de mais baixo nível, como atribuição dos papéis no sistema, cadastramento das facetas, dos índices de interatividade e suas descrições, entre outras ações. Do ponto de vista da estrutura de banco de dados, as WebQuests estão organizadas da seguinte forma:

**Figura 7. Estrutura de dados das WebQuests, na forma de relacionamento entre seus diversos constituintes**



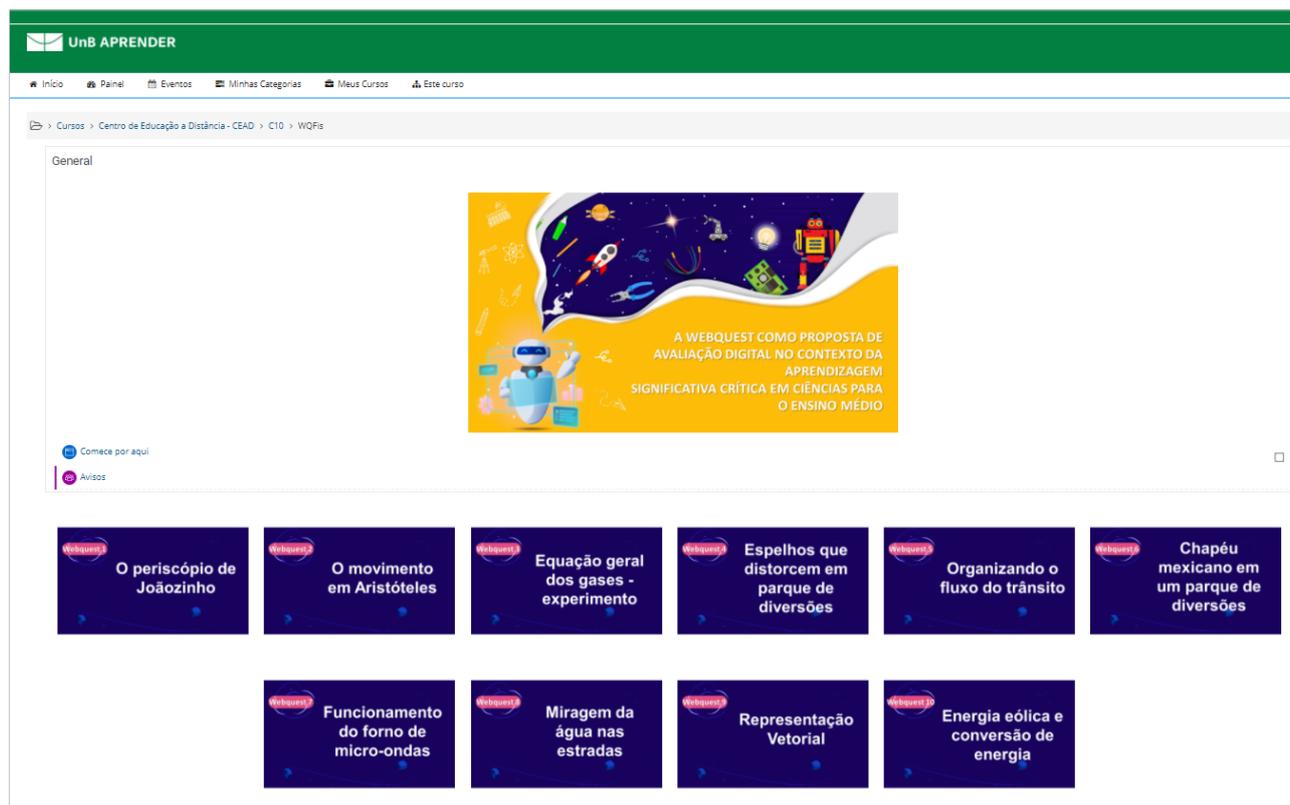
Fonte: Elaboração dos autores.

Na figura 7, as setas indicam que o relacionamento entre os metadados da WebQuest e os elementos de conexão (Avaliação e Processo) são do tipo “um para muitos”, ou seja, os mesmos metadados de uma WebQuest podem ser associados a diferentes estruturas de Avaliação e Processo (conexão). A mesma possibilidade ocorre entre os itens e a estrutura de conexão, de modo que um item pode estar referido a muitas estruturas de conexão. Dessa forma, é possível conectar diferentes conjuntos de itens a diferentes conjuntos de metadados, formando, assim, com uma mesma base existente, diversas WebQuests. Essa organização é particularmente importante para a implementação de uma estrutura CAT, que demanda maior flexibilidade no processo de apresentação de itens diferentes (a depender das características do examinando) a um mesmo conjunto de metadados.

Nessa organização, de forma geral, os elementos de Conclusão, *Feedback* e Amplie Seus Conhecimentos só ficam visíveis para o estudante após a realização da WebQuest. Ou seja, se o estudante não responder aos itens da WebQuest, ele não avança para as outras etapas. Outro fator a se considerar é o relacionamento entre os metadados da WebQuest e os elementos de conexão, como Avaliação e Processo.

Essa estrutura também foi utilizada em um Ambiente Virtual de Aprendizagem (AVA). O ambiente escolhido foi o *Moodle*, na versão 3.8, por meio da Plataforma Aprender 2, da UnB (2021). Essa escolha ocorreu, porque o grupo de pesquisa deste estudo tinha maior afinidade com esse ambiente – o que não impede que outros estudos apliquem o mesmo sistema em outro tipo de AVA, como, por exemplo, o *Teams* da Microsoft ou o *Google Classroom*. A seguir, apresentam-se figuras da implementação do referido sistema de WebQuest dentro da plataforma Aprender 2, da UnB.

**Figura 8. Exemplo da versão de implementação das WebQuest no Moodle**



Fonte: Elaboração dos autores, extraída da plataforma Aprender 2, da UnB (2021).



**Figura 9. Exemplo da versão de implementação das WebQuest no Moodle**

> Cursos > Centro de Educação a Distância - CEAD > C10 > WQFis > Equação geral dos gases - experimento > Equação geral dos gases - experimento

## Equação geral dos gases - experimento

### Introdução

### Equação Geral dos Gases - experimento

**Introdução**

A equação geral dos gases estabelece uma relação entre três grandezas fundamentais para a descrição do estado de um gás, são elas: pressão, volume e temperatura. Ela também nos fornece a explicação para uma série de fenômenos que ocorrem no nosso dia a dia. Dentre tais fenômenos está o funcionamento da panela de pressão. Nesse caso, para compreender esses fenômenos, deve-se conhecer a relação entre a pressão, temperatura e o volume dos gases. Nesta WebQuest, você será levado a desenvolver suas ideias sobre a relação entre essas variáveis e, finalmente, compreender o fenômeno do cozimento de alimentos usando-se uma panela de pressão.

**Tarefa**

Nesta WebQuest você usará um simulador PHET que você pode encontrar aqui. Este simulador permite alterar controladamente as variáveis termodinâmicas que compõem o fenômeno a ser estudado. O objetivo da sua empresa é desenvolver uma panela de pressão inovadora. Entretanto, para tal, será preciso compreender os princípios físicos que subjazem aos fenômenos físicos que relacionam a temperatura, a pressão e o volume de gases em recipientes, situação fisicamente análoga àquela de uma panela de pressão. É aí que você, o físico da empresa, entra...

Com este simulador aqui, você irá desenvolver as principais ideias relacionadas à Lei dos Gases. No processo de investigação da Lei dos Gases, devemos manipular as variáveis intervenientes de maneira controlada, para chegarmos a uma representação correta do fenômeno estudado. Assim, caberá a você desenvolver tais conhecimentos para que sua empresa, de posse deles, possa trabalhar com seus engenheiros ou desenhistas industriais no sentido de produzir uma panela de pressão inovadora.



[Prosseguir](#)

Você completou 0% da lição

Fonte: Elaboração dos autores, a partir do Aprender 2, da UnB (2021).

Na figura 8, há a apresentação de dez WebQuests. Na figura 9, há o exemplo da mesma WebQuest que aborda a Lei de Gases, representada inicialmente nas figuras 3 e 4.

Dando consecução ao processo de análise e validação do modelo de WebQuest, foi realizado, inicialmente, o levantamento de evidências de validade com base no conteúdo, a partir do processo de análise de juízes e do cálculo do Coeficiente de Validade de Conteúdo (CVC). Após as alterações empreendidas a partir dos resultados da análise de juízes, as WebQuests foram inseridas em uma amostra, como dito, por meio de um projeto-piloto. Nessa amostra, as WebQuest foram submetidas às respostas de estudantes da Educação Básica pública do Distrito Federal, selecionados com base nos cadastros fornecidos por professores indicados pelos integrantes da pesquisa. Cartas-convite foram enviadas aos responsáveis pelos estudantes que consentiram, junto com seus tutelados, em participar da pesquisa. Ao total, participaram nove estudantes, das três séries do Ensino Médio, na faixa etária dos 15 aos 17 anos de idade. Na primeira etapa, foi realizado o processo de análise semântica que, segundo Pasquali (1999), serve para verificar se todos os itens do teste são compreensíveis para todos os membros da população à qual se destina o instrumento. Foram, então,

realizadas três sessões *on-line* com os grupos, mediados por duas alunas do curso de Pedagogia da UnB. Em cada sessão, uma mediadora fez a leitura de cada WebQuest e questionou os alunos sobre a sua compreensão.

Na segunda etapa, a de análise de processo de resposta, os estudantes responderam e analisaram as WebQuests. Essa etapa ocorreu entre dezembro de 2020 e janeiro de 2021. Como dito, foi utilizada a versão *Moodle* para apresentar os sistemas desenvolvidos (figuras 8 e 9) aos estudantes.

No formulário disponibilizado, cada estudante respondeu e descreveu o respectivo processo de resposta de três ou quatro itens das quatro WebQuests disponibilizadas, totalizando, assim, 13 itens por estudante – com 65 alternativas no total. Em síntese, foi solicitado, para cada resposta, o registro do processo cognitivo para o respectivo entendimento, na perspectiva da Aprendizagem Significativa (AUSUBEL, 1969; 2000; MOREIRA, 2017a), bem como o relato das eventuais dificuldades.

Considerando essa amostra, foi possível constatar as seguintes observações. Primeiro, alguns examinandos supuseram, em suas descrições acerca do processo cognitivo durante a realização da WebQuest, que a organização desta, por si, deveria levá-los à resposta correta. Essa falsa impressão gera a demanda por avaliar as orientações a respeito do modelo avaliativo como um todo, porque o uso interno de seus recursos pode desencadear falsas associações cognitivas no processo de aprendizagem.

O segundo achado é que alguns usuários não vislumbraram justificativas contundentes para o acerto. Embora o objetivo da WebQuest tenha sido explorar o recurso digital na sistematização do processo de aprendizagem sobre o conteúdo disciplinar – no caso, conteúdos de Física –, uma parcela dos estudantes-respondentes foi incapaz de afirmar isso de forma contundente. Por isso, esse aspecto deve ser considerado e mais bem evidenciado em uma eventual reformulação dos itens que compõem a WebQuest, para que fique mais perceptível ao estudante o objetivo de pesquisa sistemática do tema por meio de recursos digitais.

Notou-se que houve posicionamento que acionou a sequência cognitiva definida por Ausubel (1969; 2000) para a aprendizagem significativa; em particular, a ancoragem (ou subsunção), com base na potencialização desencadeada pelos *links* como materiais ou a especialização/contextualização



da resposta. Por isso, em algumas das situações relatadas – e particularmente nas descrições das respostas aos primeiros itens –, pode-se notar a utilização de recursos de subordinação conceitual, como previsto por Ausubel (2000) em processos dessa natureza.

É digno de nota, também, que alguns dos usuários examinados fizeram constar que as terminologias utilizadas pelos itens estariam fora dos seus alcances prévios e assinalaram que não tiveram acesso curricular à parte dos objetos abordados nos itens. Esse último aspecto, particularmente, parece guardar relação com o período letivo que temos vivenciado no Brasil a partir da pandemia da Covid-19. De forma geral, observou-se, em quatro WebQuests, uma oscilação nos índices de acertos, o que se pode prever haja maior ou menor correlação das justificativas apresentadas pelos examinandos, em relação às habilidades associadas aos itens. Esse aspecto deverá ser objeto de reavaliação pela equipe que desenvolveu essa proposta.

Os ganhos com essa pesquisa amostral foram os seguintes: primeiramente, houve a validação da proposta com base no processo de resposta do aluno-usuário, sendo obtidas, ao total, três fontes de dados, a saber: a) a descrição individualizada dos processos de resposta às WebQuests da amostra analisada; b) o descritivo das respostas de cada estudante às WebQuests; e c) os relatórios de aplicação do teste na versão *Moodle* das WebQuests. Em segundo lugar, os resultados parciais dessa amostra permitiram evidenciar estatísticas relacionadas aos acertos e aos erros das respostas, mas, principalmente, identificar tendências; erros sistemáticos nas respostas; categorias emblemáticas de erros; dados sobre atendimento (ou não) da habilidade proposta pelo item; e indicativos de reformulação. Em terceiro lugar, esse projeto-piloto com a realização de uma proposta de WebQuest abre uma perspectiva pedagógica a ser explorada: WebQuests com itens de múltipla escolha com base em um processo estruturado de ensino e aprendizagem não presencial.



## 5. Avanços necessários: à guisa de uma conclusão

Neste artigo, o objetivo de construir um modelo de WebQuest como instrumento de avaliação digital para o ensino de Ciência, sobretudo para o estágio do Ensino Médio, foi cumprido. Para tanto, o estudo foi segmentado em três partes. Na primeira parte, foi apresentada uma discussão fundamentada sobre os pressupostos educacionais e avaliativos que poderiam sustentar uma proposta de WebQuest com itens de múltipla escolha. Na segunda, foram abordados os pressupostos metodológicos desta pesquisa tecnológica. Na terceira parte, foi descrito o modelo de WebQuest com itens de múltipla escolha, apresentando parte dos resultados alcançados no projeto-piloto que contou com a participação de nove estudantes do Ensino Médio.

Com este estudo, foi possível apreender que a composição de itens de múltipla escolha e com *feedbacks* automáticos dá indicativos satisfatórios – pelo menos, inicialmente – às avaliações em larga escala e de cunho formativo, podendo adequar-se à estratégia da WebQuest. Isso acompanha o referencial de aprendizagem no qual essa proposta está baseada (MOREIRA, 2012; AUSUBEL, 2000), bem como sustenta a abertura para uma nova seara pedagógica a ser explorada na relação entre TIC e Educação.

Uma limitação da pesquisa em seu estágio foi a quantidade diminuta dos dados empíricos coletados na amostra do projeto-piloto. Percebe-se que se faz necessária a realização de mais testes, em maior escala, bem como a análise fatorial, a calibragem dos dados e a análise Rash. Outra limitação foi a consideração de apenas quatro das dez WebQuest, sendo, portanto, preciso explorar as outras WebQuest a partir de outras aplicações com alunos do Ensino Médio. Além disso, foram analisados apenas dados de alunos-usuários do Distrito Federal, o que circunscreve a pesquisa a uma localidade. Para ampliar as possibilidades de aplicação da WebQuest, nota-se que é imprescindível acolher pré-testes de examinandos oriundos de diversas e distintas localidades do Brasil, dada a dimensão continental do País e a desigualdade social de suas regiões.



## Referências

- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION - AERA; AMERICAN PSYCHOLOGICAL ASSOCIATION -APA; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION - NCME. **Standards for educational and psychological testing**. Washington, DC: APA, 2014.
- ARREDONDO, S. C.; DIAGO, J. C. A avaliação na educação. In: ARREDONDO, S. C.; DIAGO, J. C. **Avaliação educacional e promoção escolar**. Curitiba: IBPEX; São Paulo: Unesp, 2009. p.27-78
- AUSUBEL, D. P. **Educational psychology: A cognitive view**. New York: Holt, Rinehart & Winston, 1969.
- AUSUBEL, D. P. **The acquisition and retention of knowledge: a cognitive view**. Boston: Kluwer, 2000.
- BRASIL, Ministério da Educação. **Base Nacional Curricular Comum**. Ensino Infantil e Ensino Fundamental. Brasília: Conselho Nacional da Educação, 2017.
- BLOOM, B. S. et al. **Taxonomy of educational objectives**. New York: David Mckay, 1956.
- BUNGE, M. **Treatise on basic philosophy**. Boston: D. Reidel, v. VII, 1985.
- CARVALHO, A. M. P. et al. **Ensino de Física**. São Paulo: Cengage Learning, 2010.
- COSTA, R. S.; DUQUEVIZ, S. B. C.; PEDROZA, R. L. C. Tecnologias Digitais como instrumentos mediadores da aprendizagem dos nativos digitais. **Revista Quadrimestral da Associação Brasileira de Psicologia Escolar e Educacional**, São Paulo, v. 19, n. 3, set.-dez, p. 603-610, 2015.
- COELHO, P. M. F; COSTA, M. R. M.; MATTAR NETO, J. A. Saber Digital e suas Urgências: reflexões sobre imigrantes e nativos digitais. **Educação & Realidade**, Porto Alegre, v. 43, n. 3, p. 1077-1094, 2018.
- CRAIK, F.; LOCKHART, R. Levels of processing: A framework for memory research. **Journal of Verbal Learning & Verbal Behavior**, v. 11, p. 671-684, 1972.
- DODGE, Bernie. WebQuests: a technique for internet – based learning. **The Distance Educator**, v.1, n 2, p. 10-13, 1995.
- ECKES, T. Many-facet Rasch measurement. In: TAKALA, S. (Ed.). **Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (Section H)**. Strasbourg, France: Council of Europe/ Language Policy Division, 2009. p. 1-53.
- FERREIRA, M. *et al.* Unidade de Ensino Potencialmente Significativa sobre óptica geométrica apoiada por vídeos, aplicativos e jogos para smartphones. **Revista Brasileira de Ensino Física**, v. 42, e20200057, p. 1-3, jun. 2020.



FREITAS JUNIOR, V. **Guia para escrita de artigos científicos: uma perspectiva da pesquisa tecnológica.** Sombrio: Instituto Federal Catarinense, 2018.

GIERL, M.; HALADYNA, T. **Automatic Item Generation: theory and practice.** Routledge: New York and London, 2012.

GIL, A. C. **Como elaborar projetos de pesquisa.** 4.ed. São Paulo: Atlas, 2002.

GOLEMAN, D. **Foco: a atenção e seu papel para o sucesso.** Tradução de Cássia Zanon. Rio de Janeiro: Objetiva, 2013.

HANNAFIN, M. J. Interaction Strategies and Emerging Instructional Technologies: Psychological Perspectives, **Canadian Journal of Educational Communication CJEC**, v.18, n. 3, p. 167 -180, 1989.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA – INEP. **Índice de Desenvolvimento da Educação Básica – IDEB – Resultados e Metas.** 2019. Disponível em: <http://ideb.inep.gov.br/resultado/> Acesso em: 01 jun. 2021.

KÄMPF, C. A geração Z e o papel das tecnologias digitais na construção do pensamento. **ComCiência**, n. 131, p. 1-4, 2011.

KEMPFER, L. M. A. **Impactos da Cibercultura na mediação pedagógica nas escolas públicas municipais do ensino fundamental da cidade de Goiânia (GO).** Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade de Brasília, Brasília, 2012.

KENSKI, V. M. **Educação e tecnologias: o novo ritmo da informação.** Campinas: Papirus, 2009.

LÉVY, P. **O que é virtual?** Tradução de Paulo Neves. Rio de Janeiro: Editora 34, 1996.

LINACRE, J. M. **Many-facet Rasch Measurement.** Chicago: Mesa Press, 1994.

LUECHT, R. M. Automatic Item Generation for Computerized Adaptive Testing. In: GIERL, M. J.; HALADYNA, T.M. **Automatic item generation: theory and practice.** New York: Routledge, 2012, p. 196-216.

MARCH, T. The learning power of WebQuests. **Educational Leadership**, v. 61, n. 4, p. 42-47, 2004.

MASETTO, M. T. Mediação pedagógica e o uso da tecnologia. In: MORAN, J. M.; MASETTO, M. T.; BEHRENS, M. A. **Novas tecnologias e mediação pedagógica.** 12 ed. Campinas: Papirus, 2000, p. 133-173.

MARZANO, R. J. **A different kind of classroom: teaching with dimension with dimensions of learning.** Alexandria: Association for Supervision and Curriculum Development, 1992.

MARZANO, R. J.; KENDALL, J. **The new Taxonomy of Educational Objectives.** 2. ed. Nova Iorque: Corwin Press, 2007.



- MOREIRA, M. A. **Aprendizagem significativa**. Brasília: Editora da UnB, 1999.
- MOREIRA, M. A. **Unidades de Ensino Potencialmente Significativas – UEPS**. Temas de ensino e formação de professores de ciências. Natal: EDUFRN, 2012.
- MOREIRA, M. A. **Teorias de aprendizagem**. São Paulo: EPU, 2017a.
- MOREIRA, M. A. Grandes desafios para o ensino da física na educação contemporânea. **Revista do Professor de Física**, v. 1, n. 1, p. 01-13, 2017b.
- MOREIRA, M. A.; MASINI, E. F. A. **Aprendizagem significativa: a teoria de David Ausubel**. São Paulo: Moraes, 2016.
- OLIVEIRA, A. P. M. **Prova Brasil como política de regulação da rede pública do Distrito Federal**. 2011. 274f. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação, Universidade de Brasília, Brasília, 2011.
- OLIVEIRA, A. R. **O modelo multifacetado de Rasch na construção de subescalas**. 2019. 100f. Dissertação (Mestrado em Desenvolvimento, Sociedade e Cooperação Internacional) - Centro de Estudos Avançados Multidisciplinares, Universidade de Brasília, Brasília, 2019.
- PARRISH, P. **Interactivity in computer-aided learning (CAL)**. 1996. Disponível em: [http://www.comet.ucar.edu/pub\\_html/mpi/interac3.html](http://www.comet.ucar.edu/pub_html/mpi/interac3.html). Acesso em: 01 jun. 2021.
- PASQUALI, L. **Psicometria: teoria dos testes na psicologia e na educação**. Petrópolis: Vozes, 2013.
- PASQUALI, L.; PRIMMI, R. Fundamentos da teoria da resposta ao item: TRI. **Avaliação psicológica**, Porto Alegre, v. 2, n. 2, p. 99-110, dez. 2003.
- PERRENOUD, P. **Avaliação: da excelência à regulação das aprendizagens – entre duas lógicas**. Trad. Patrícia Chittoni Ramos. Porto Alegre: Artes Médicas Sul, 1999.
- SACRISTÁN, J. G; GÓMEZ, A. I. P. **Comprender e transformar o ensino**. Tradução de Ernani da Fonseca. 6. ed. Porto Alegre: ARTMED, 2007.
- SANTOS, T. R. dos; BARIN, C. S. Problematização da Metodologia Webquest na Prática Educativa: potencialidades e desafios. **Revista Tecnologias na Educação**, ano 6, n. 11, p. 1-11, dez. 2014.
- SILVA, W. **Proposta de uma metodologia para a produção e interpretação de medidas educacionais em avaliação em larga escala por meio da utilização da Modelagem Rasch com duas ou mais facetas**. 2019. 150 f. Tese (Doutorado em Educação) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Educação, Rio de Janeiro, 2019.
- SILVA FILHO, O. L.; FERREIRA, M. Teorias da aprendizagem e da educação como referenciais em práticas de ensino: Ausubel e Lipman. **Revista do Professor de Física**, v. 2, n. 2, ago. 2018.



SIMÃO, P. P. M. WebQuest - um uso inteligente da Internet na escola. **Caderno do Professor**, n. 7, p. 55-64, fev. 2001.

SIMS, R. Interactivity on stage: Strategies for learner-designer communication. **Australian Journal of Educational Technology**, 1999, v. 15, n. 3, p. 257-272, Summer 1999.

TOFFOLI, S. F. L.; SIMON, C. V. B. A utilização do modelo multifacetado de Rasch na análise das influências dos avaliadores sobre as avaliações com itens abertos. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, n. 101, p. 1303-1323, 2018.

TOMAZ, J. M. C. Efeitos do nível de interatividade num recurso educativo digital - diferenças na satisfação e na aprendizagem. **Atas do II Congresso Internacional TIC e Educação**. Lisboa, 2011, p. 2654-2675. Disponível em: <http://ticeduca.ie.ul.pt/atas/pdf/318.pdf> Acesso em: 01 jun. 2021.

UNIVERSIDADE DE BRASÍLIA- UNB. A WebQuest como proposta de avaliação digital no contexto da aprendizagem significativa em Ciências para o Ensino Médio. **Aprender 2**. 2021. Disponível em: <https://aprender2.unb.br/course/view.php?id=1459>. Acesso em: 01 jun. 2021.



# Uso de imagens nos suportes de Itens de Ciências da Natureza utilizados em Avaliações de Larga Escala<sup>1</sup>

*Ana Carolina Araújo da Silva<sup>2</sup>*

*Andréia Francisco Afonso<sup>3</sup>*

*Clarice Parreira Senra<sup>4</sup>*

*Rita de Cássia Reis<sup>5</sup>*

*Wallace Alves Cabral<sup>6</sup>*

*Juliana Vicini Florentino Rodrigues<sup>7</sup>*

---

<sup>1</sup> Projeto financiado pela Fundação CAEd/UFJF.

<sup>2</sup> Departamento de Educação, Universidade Federal de Juiz de Fora, e-mail: anacarolina.silva@ufjf.br

<sup>3</sup> Departamento de Química, Universidade Federal de Juiz de Fora, e-mail: andreia.afonso@ufjf.br

<sup>4</sup> Departamento de Educação, Universidade Federal de Juiz de Fora, e-mail: clarice.senra@ufjf.br

<sup>5</sup> Departamento de Educação, Universidade Federal de Juiz de Fora, e-mail: rita.reis@ufjf.br

<sup>6</sup> Departamento de Ciências Naturais, Universidade Federal de São João del-Rei, e-mail: wallacecabral@ufsj.edu.br

<sup>7</sup> Universidade Federal de Juiz de Fora, e-mail: juvicini@hotmail.com



## 1. Introdução

Este artigo apresenta um recorte de uma pesquisa sobre itens de Ciências da Natureza utilizados em avaliações em larga escala de programas estaduais e municipais. Nessa pesquisa ampliada, refletimos sobre a estrutura de um item representativo de uma tarefa e nos questionamos quais as possíveis estruturas de comando, suporte e distratores que poderiam expressar diferentes níveis de aprendizagem de uma habilidade. Ao longo de nossos estudos, uma indagação sempre foi recorrente - como o uso de representações nos suportes de itens de Ciências da Natureza pode influenciar na compreensão e na execução da tarefa solicitada?

Desse modo, neste texto, apresentamos algumas reflexões sobre o uso de imagens nos suportes de itens de Ciências da Natureza utilizados em avaliações de larga escala. Para este estudo, partimos do pressuposto que, na Educação em Ciências, a utilização de imagens em sala de aula auxilia no processo de construção de significados e de processos científicos. As imagens são importantes, pois possibilitam representar não apenas fenômenos, mas também noções ligadas a esses fenômenos, ou conceitos, e a compreensão em um nível teórico (CAVALCANTE et al., 2012). Entre a utilização de imagens, por exemplo, podemos citar as representações simbólicas de partículas submicroscópicas em diferentes modelos explicativos propostos pela Ciência para compreender e explicar diferentes fenômenos por ela estudados.

É comum observarmos, nos itens em Ciências da Natureza, a utilização de suportes para auxiliar o/a estudante a compreender a situação apresentada no enunciado do item. Esses suportes podem envolver: textos, equações químicas, esquemas, diagramas, imagens (figuras, desenhos, fotografias etc), tabelas e gráficos. Independentemente do tipo de suporte, ele deve estar alinhado ao descritor, ao comando e ser relevante para a solução/entendimento da situação problematizada no item. Deve-se evitar a utilização de suportes apenas para ilustrar a situação apresentada no comando, de modo que não ofereçam nenhuma informação relevante para a resolução do problema apresentado.

Nesse sentido, a seleção de um suporte deve estar atrelada à compreensão do descritor, uma vez que, nos casos em que o descritor envolve aspectos macroscópicos, é desejável que o suporte também apresente o fenômeno ou que envolva abordagens fenomenológicas. O mesmo raciocínio vale para aqueles descritores que envolvem aspectos teóricos e representacionais. Por sua vez, essa



adequação aos níveis de conhecimento também deve respeitar o nível de escolarização ao qual o item se destina, sendo coerente com a faixa etária dos/as estudantes e com as representações usualmente utilizadas pelas/os docentes na Educação Básica. Essa adequação também contribui para que o item possibilite uma boa discriminação e não se configure nem como um obstáculo à resolução da tarefa, tampouco como um facilitador.

Em relação ao uso de imagens, elas são importantes recursos para a comunicação de ideias científicas, para a visualização de fenômenos do ponto de vista macroscópico e/ou microscópico, além de contribuírem para a inteligibilidade de diversos textos científicos e na sua conceitualização (PEREIRA; BIANCO, 2016). Para terem um efeito positivo, elas devem fornecer informações sobre as relações entre os elementos concretos a serem descritos no texto melhorando a sua compreensão. Assim, além de as imagens atuarem como facilitadoras na recuperação das informações em longo prazo, elas possuem também efeitos sobre a cognição que parecem ser melhores quando fornecem uma maneira para os estudantes interpretarem o que leram ou ouviram falar, ligando ou organizando a informação textual (SANA et al., 2016).

Em nossos estudos, temos observado que a maneira como um fenômeno (científico ou cotidiano) é apresentado nos itens interfere na sua interpretação e na construção cognitiva que o/a estudante faz para responder ao comando. Para Jaloto e Medeiros (2016), o sentido único de uma imagem é questionável, pois

[...] uma mesma imagem pode significar de diferentes modos, pois integra um sistema simbólico de representação influenciado por princípios que organizam possibilidades de representação e de significação numa dada cultura; tanto a estrutura e a forma da imagem intervêm na constituição dos sentidos, quanto as posições que os sujeitos ocupam numa dada formação discursiva. (JALOTO; MEDEIROS, 2016, p. 2876)

Logo, ao selecionar uma imagem para compor o item, quer seja como um suporte para a execução da tarefa, quer seja como distrator, devemos ter em mente as leituras possíveis que ela despertará na criança e no/a jovem que a lerá. É uma imagem que apresenta um contexto muito específico de um grupo cultural? É uma imagem que, em alguma medida, estaria presente no contexto social ou escolar do sujeito? Ela apresenta ou representa o verbalismo característico do conhecimento científico? Qual o nível de domínio da linguagem científica necessária para interpretá-la? Essas e outras questões devem ser levadas em conta quando utilizamos as imagens em avaliações em larga escala.



Diante dessas considerações, questionamo-nos: como o uso de imagens no suporte pode influenciar na resolução da tarefa pelos/as estudantes e no grau de dificuldade do item? Assim, realizamos uma análise documental dos itens de Ciências da Natureza presentes na Revista do Professor do Programa de Avaliação da Educação Básica do Estado do Espírito Santo - PAEBES, por serem itens já publicados. Adotamos os pressupostos teóricos da análise de conteúdo (BARDIN, 1995) que nos permitiu analisar os itens que contêm representações, levando em consideração como a área de Educação em Ciências tem interpretado o uso de imagens no ensino.

## 2. Referencial Teórico

Ao longo da nossa trajetória enquanto pesquisadores/as do projeto anteriormente citado, percebemos que ainda são incipientes pesquisas no Brasil sobre avaliação em larga escala e suas implicações pedagógicas, principalmente na área de Ciências da Natureza. Assim, visando contribuir com esse campo, nesta pesquisa, investigaremos as imagens utilizadas nos suportes dos itens.

Acreditamos que o uso de imagens, seja em itens ou em materiais pedagógicos, não pode ser visto como mero recurso para “embelezar” ou ajudar a vender materiais e, sim, como uma forma de linguagem que pode produzir diferentes sentidos quando entra no jogo discursivo, podendo potencializar (ou gerar confusão) no processo de ensino e aprendizagem.

Na área de Ciências da Natureza, as imagens podem ter um papel central na construção e comunicação de conhecimentos científicos pelos/as estudantes. Porém, a sua utilização deve ser feita a partir de diferentes critérios. No que tange à avaliação em larga escala, a utilização de imagens em itens apresenta referência em diferentes habilidades, tal como será visto em pesquisas que envolvem o Sistema Nacional de Avaliação da Educação Básica (SAEB).

A pesquisa de Medeiros et al. (2017), apesar de não ter um enfoque nas imagens presentes nos itens, apresenta resultados importantes a partir das respostas dos/as estudantes no SAEB. Por exemplo, por meio da análise de um item, os pesquisadores mostraram que o uso de setas em um fluxograma no suporte pode contribuir para a construção de múltiplos sentidos pelos/as estudantes.



O significado da seta na representação se mostrou um fator limitante para a descrição da cadeia [alimentar]; todos os estudantes entrevistados mostram, pelo menos inicialmente, construir sentidos distintos ao daquele aceito pela ciência (MEDEIROS et al., 2017, p. 19).

As análises das entrevistas com os/as estudantes desta pesquisa (IDEM, 2017) mostraram também as dificuldades no entendimento de certos conceitos e, como caminho, eles/as tendem a produzir sentidos que fazem referência a elementos próximos de suas realidades. Portanto, trata-se de uma estratégia de resposta às demandas cognitivas apresentadas. Essa constatação vai ao encontro da pesquisa de Martins, Gouvêa e Piccinini (2005). As pesquisadoras, durante entrevistas com estudantes do 3o e 4o ciclos do Ensino Fundamental, mostraram as estratégias de leitura das imagens realizadas por eles/as. Em suas considerações,

a análise destas revelou que, na busca de uma significação para a imagem, eles se engajam em procedimentos elaborados que envolvem análises dos elementos composicionais, buscas na memória por experiências relevantes, estabelecimento de relações com situações do seu cotidiano (incluindo experiências escolares) (MARTINS; GOUVÊA; PICCININI; 2005, p. 39).

Ainda sobre a demanda cognitiva presente em um item, para Medeiros et al. (2017), certos elementos de um item (texto base e enunciado, por exemplo) podem agregar dificuldades que não estão relacionadas ao conhecimento que se quer medir. Isso ocorre, principalmente, quando temos o uso de palavras no sentido conotativo ou porque a imagem está total ou parcialmente desconectada do texto verbal.

Sobre as imagens, objeto de interesse neste artigo, entendemos que elas podem ser importantes recursos para a comunicação científica, tal como discutem Martins, Gouvêa e Piccinini (2005). Para elas, as imagens podem facilitar a inteligibilidade de um texto ou desempenhar um papel na constituição de ideias científicas e na sua conceitualização. Porém, é preciso superar a ideia de que a leitura de uma imagem não dá margem para a polissemia, pelo contrário, a interpretação da imagem pode levar a sentidos diferentes daqueles esperados pelo/a elaborador/a de um item. Nesse contexto, concordamos que

do mesmo modo que não basta saber ler a palavra para dar sentido a um texto, também nas representações visuais os sentidos possíveis ultrapassam a simples identificação visual de seus componentes. Aprender a ler textos verbais e/ou imagéticos é muito mais do que decodificar signos (SOUZA, 2014, p. 128).



Trabalhando nessa perspectiva de que a imagem é um objeto simbólico que produz sentidos na relação com a historicidade e a exterioridade, Jaloto e Medeiros (2016) avaliaram as causas dos erros dos/as estudantes devido a diferentes interpretações da imagem. Tal pesquisa teve como material de análise entrevistas cognitivas realizadas com estudantes do 9o ano do Ensino Fundamental sobre as imagens de questões de Ciência da Natureza do SAEB.

Para analisar o material construído, os pesquisadores (JALOTO; MEDEIROS, 2016) fizeram uso do conceito de iconicidade, referindo-se ao grau com que a imagem se relaciona com a realidade que busca representar. “Quanto maior o grau de iconicidade, menor a abstração. Desenhos figurativos, esquemas, signos normalizados e representações em forma de palavras reduzem a iconicidade da imagem e aumentam sua abstração” (SILVA, G. S.; BRAIBANTE; PAZINATO, 2013; SOUZA, 2014 apud JALOTO; MEDEIROS, 2016, p. 2877). Além desse conceito, os autores também analisaram a relação dos textos verbais com as imagens.

Os resultados desta pesquisa evidenciaram que as “representações imagéticas nas questões analisadas, o uso de setas e símbolos normalizados, aliado às desconexões entre a imagem e os textos verbais associados a elas, contribui para a construção de múltiplos sentidos” (JALOTO; MEDEIROS, 2016, p. 2885). Dentre os elementos que podem ser utilizados nas imagens, o uso de setas tem um papel importante nas interpretações dos fenômenos representados. Vale destacar que as setas podem ser utilizadas em diferentes situações no Ensino de Ciências, por exemplo, podemos utilizá-las para representar direção, intensidade, movimento, causalidade, entre outras. Como destacam Martins e Queiroz (2010)

elas atuam em ambientes semióticos muito distintos, provendo usuários e intérpretes com diversas competências e habilidades cognitivas. Uma classificação e análise da variedade de setas e de suas operações permitem-nos entendê-las funcional e morfológicamente, elucidando operações relacionadas a tarefas cognitivas específicas (percepção, categorização, memória, atenção, etc.). (MARTINS, QUEIROZ, 2010, p. 2)

Na pesquisa realizada por Martins e Queiroz (2010), que envolveu a análise das setas encontradas nas imagens de 144 livros didáticos, as setas foram classificadas a partir de 14 categorias. Esse estudo mostrou que as setas podem assumir diferentes funções, como, por exemplo: suprir limitações típicas de imagens, relacionadas à proporção, profundidade, movimento e dimensão e podem ser



utilizadas também “para direcionar a análise para determinado foco, para nomear coisas e estão presentes em objetos que as utilizam habitualmente como marcadores ou indicadores” (MARTINS; QUEIROZ, 2010, p. 15).

Outro conceito que também emerge no estudo de imagens no Ensino de Ciências é o da iconicidade. Para Souza (2014, p. 128), “o grau de realidade com que uma imagem se relaciona com a realidade que representa passa por diferentes níveis, desde a igualdade total (iconicidade máxima) até a abstração total”. A partir disso, a pesquisadora compreende que imagens com baixa iconicidade podem trazer problemas para o Ensino de Ciências, uma vez que, dentro de sala de aula, convivem estudantes com diferentes níveis de leitura de representações. Isso pode acontecer pela dificuldade de perceber a construção do objeto representado até a sua observação real.

Como conclusão de sua pesquisa, Souza (2014) indica a importância de o/a docente trabalhar com imagens, perpassando os diferentes níveis de iconicidade. Essa movimentação entre os diversos níveis, tendo o/a docente como mediador/a, possibilita a explicitação dos conhecimentos pelos/as alunos/as. Dessa forma, acreditamos que, a partir desse trabalho em sala de aula, os impactos mencionados nas pesquisas anteriores podem ser minimizados. Um dos caminhos para essa mudança pode ser por meio do incentivo ao trabalho com imagens na formação inicial e continuada de professores, tal como é realizado na pesquisa de Silva et al. (2006).

Por outro lado, encontramos, na literatura, o trabalho de Pereira e Bianco (2016) no qual mostram que o uso de imagem em um item facilita sua resolução. Para chegar a essa conclusão, os autores elaboraram duas questões com o mesmo objeto de aprendizagem, porém, uma com imagem e outra sem. Após a aplicação em distintas escolas de Ensino Fundamental, perceberam que “o uso de imagens nas avaliações se mostrou muito mais eficaz do que textos, importante ressaltar que as questões eram exatamente iguais mudando somente o enunciado que diferia em texto ou imagem” (PEREIRA; BIANCO, 2016, p. 796).

Apesar dessa generalização apresentada na pesquisa de Pereira e Bianco (2016), concordamos que o uso de imagens em itens deve ser estudado com mais profundidade. Por isso, neste trabalho, propomos a analisar as imagens presentes em itens de Ciências da Natureza utilizados em avaliações de



larga escala a partir de alguns aspectos já citados anteriormente, tais como: i) iconicidade da imagem, ii) sua relação com texto verbal, iii) limitações presentes na representação, iv) elementos presentes na imagem (setas, legenda, ...) e v) nível cognitivo associado à tarefa apresentada.

Um importante instrumento utilizado para analisar o nível cognitivo é a nova taxonomia de Bloom que pode ser utilizada para apoiar o planejamento didático-pedagógico, a estruturação dos conteúdos, assim como auxiliar os processos de avaliação. A nova Taxonomia de Bloom do Domínio Cognitivo é estruturada em seis categorias: lembrar, entender, aplicar, analisar, avaliar e criar (FERRAZ; BELHOT, 2010). Essas categorias podem ser consideradas graus de dificuldades, começando com habilidades mais simples até as mais complexas. No quadro 1, a seguir, apresentamos os níveis da Taxonomia de Bloom que auxiliaram na análise de dificuldade dos itens apresentados nesta pesquisa.

**Quadro 1. Estrutura do processo cognitivo na Taxonomia de Bloom – revisada**

<b>Lembrar</b>	Significa lembrar informações aprendidas anteriormente, com foco na memorização. Verbos: Reconhecendo e Reproduzindo
<b>Entender</b>	Relacionado a capacidade do estudante reproduzir com suas próprias palavras, a informação que viu, ouviu ou leu.
<b>Aplicar</b>	Relacionado a executar ou usar um procedimento numa situação específica e pode também abordar a aplicação de um conhecimento numa situação nova.
<b>Analisar</b>	Relacionado a dividir a informação em partes relevantes e irrelevantes, importantes e menos importantes e entender a inter-relação existente entre as partes.
<b>Avaliar</b>	Relacionado a realizar julgamentos baseados em critérios e padrões qualitativos e quantitativos ou de eficiência e eficácia.
<b>Criar</b>	Significa colocar elementos junto com o objetivo de criar uma nova visão, uma nova solução, estrutura ou modelo utilizando conhecimentos e habilidades previamente adquiridos.

Fonte: Adaptado de Ferraz; Belhot (2010, p. 429).

O conhecimento e a utilização da taxonomia podem auxiliar professores, assim como elaboradores na tarefa de planejar e sistematizar a avaliação (TREVISAN; AMARAL, 2016). Para analisar as imagens dos itens segundo os aspectos destacados anteriormente, optamos pela análise do conteúdo (BARDIN, 1995). Na seção a seguir, apresentamos o percurso adotado para esta pesquisa.



### 3. O caminho metodológico

Nesta seção, apresentamos os caminhos metodológicos adotados ao longo da pesquisa em avaliação. Para tanto, empregamos a opção metodológica de natureza qualitativa, que possui as seguintes características: (i) coletam-se os dados, em geral, em cenário natural, sendo o/a pesquisador/a o instrumento-chave nesse processo; (ii) é descritiva; (iii) preocupa-se com o processo em vez de simplesmente com resultados ou produtos; (iv) tende a analisar os dados de forma indutiva; (v) enfatiza o “significado” (BOGDAN; BIKLEN, 1994).

Para esta pesquisa qualitativa, utilizamos a análise documental para a coleta de dados. Para Triviños (1987, p. 111), a "análise documental fornece ao investigador a possibilidade de reunir uma grande quantidade de informação sobre leis estaduais de educação, processos e condições escolares, planos de estudo, requisitos de ingresso, livros-texto etc". A fonte de dados, deste estudo, são os itens publicados em *Revistas do Professor*<sup>8</sup> - Ciências da Natureza 9º ano do Ensino Fundamental/ 3º ano do Ensino Médio. Portanto, os dados obtidos a partir da análise dessas revistas podem ser “[...] designados por qualitativos, o que significa ricos em pormenores descritivos relativamente a pessoas, locais e conversas, e de complexo tratamento estatístico” (BOGDAN; BIKLEN, 1994, p. 16).

Dentre os métodos de análise documental descritos na literatura, adotamos os pressupostos teóricos da análise de conteúdo (BARDIN, 1995). A análise de conteúdo busca caracterizar as vivências do sujeito e suas percepções sobre um objeto e seus fenômenos, proporcionando, assim, o levantamento de indicadores (quantitativos ou não) que permitem, a partir de processos sistemáticos, a produção de inferência de conhecimentos (BARDIN, 1995; CAVALCANTE, 2014).

Com base em Bardin (2006), as etapas de uma “análise de conteúdo” pressupõem: primeira etapa - pré-análise; segunda etapa: exploração do material e terceira etapa - tratamento dos resultados, inferência e interpretação que envolve a relação entre os dados obtidos e a fundamentação teórica.

A pré-análise, fase em que se organiza o material a ser investigado, envolve quatro etapas, quais sejam: Leitura flutuante, Escolha dos documentos, Formulação das hipóteses e dos objetivos; Referenciação dos índices e elaboração de indicadores (BARDIN, 2006). Portanto, essa etapa

---

<sup>8</sup> As *Revistas do Professor* são coleções para a divulgação dos resultados do Programa de Avaliação da Educação Básica do Espírito Santo (PAEBES).



compreendeu a escolha do material de estudo que foram as Revistas do Professor publicadas entre os anos 2017 e 2019. Após a seleção dos materiais, realizamos a leitura flutuante das revistas para observar o comportamento dos itens e as possíveis relações com as competências da BNCC do 1° ao 9° ano do Ensino Fundamental e do 1° ano ao 3° do Ensino Médio. No estudo desses itens, chamou-nos a atenção a utilização de diferentes representações. Portanto, a definição do corpus de análise compreendeu a escolha de itens que tivessem a utilização de diferentes representações.

A segunda etapa, a exploração dos materiais, consistiu na definição de categorias, identificação das unidades de registro e das unidades de contexto nos documentos. Para isso, analisamos um conjunto de itens frente aos dados pedagógicos fornecidos pelas revistas de diversos projetos municipais e estaduais<sup>9</sup> de avaliações em larga escala, totalizando 107 itens analisados. Desses itens analisados, encontramos 51 com algum tipo de representação no suporte. Esses itens foram categorizados como itens que possuem imagens, esquemas, modelos, quadro/tabelas, e equações químicas. Para esta pesquisa, apresentamos a análise de dois itens do PAEBES da edição de 2019 que possuem imagens. A escolha desses dois itens se deu pelos níveis de desempenho estudantil em Ciências da Natureza sendo um item para o padrão mais alto de proficiência (avançado) e mais baixo (abaixo do básico). Os itens como um todo (enunciado, suporte, comando, alternativas e distratores) são as unidades de contexto desta pesquisa. Já as unidades de registro são os suportes que possuem imagens. A análise desses suportes poderá nos fornecer as características e as informações sobre o item e a sua tarefa e, ainda, se a imagem possui setas, legendas e iconicidade da imagem. Essas características, além de serem os indicadores desta pesquisa, são as unidades de registro. Neste artigo, procuramos compreender os dados e as informações obtidas a partir da análise das unidades de contexto que são os itens como um todo.

Por fim, a terceira etapa consiste no tratamento dos resultados, inferência e interpretação dos itens e das suas imagens presentes. Nessa etapa, destinada ao tratamento dos resultados, ocorrem a condensação e o destaque das informações para análise, culminando nas interpretações inferenciais, constituindo-se como o momento da intuição, da análise reflexiva e crítica (BARDIN, 2006).

---

<sup>9</sup> Prova Floripa, Avaliando SC, Santa Catarina em Somativa, Santa Catarina, Avalie BA, SADEAM



## 4. Análises e Resultados

Ao olharmos para o conjunto de itens analisados ao longo da pesquisa, percebemos que, de modo geral, a imagem utilizada como suporte cumpre o papel de mediadora entre a tarefa solicitada e a avaliação do conhecimento adquirido pelo/a estudante, além de ser um estímulo visual para motivá-lo/a à resolução do item. Porém, a presença de imagens não pode ser relacionada à facilidade do item, como mostraremos na análise de duas imagens presentes em dois itens de Biologia, aplicados no 3º ano do Ensino Médio, pelo PAEBES, em 2019, e publicados na Revista do Professor - Ciências da Natureza (CAED, 2019).

O primeiro item (figura 1) tem como objetivo avaliar a interpretação dos/as estudantes a respeito do ciclo da água, mais especificamente, “reconhecer a etapa de evaporação da água em seu ciclo” (CAED, 2019, p. 66). Diante disso, a interpretação, de acordo com a tarefa proposta, tem sentido de reconhecimento, exigindo, portanto, o desenvolvimento do primeiro nível do cognitivo (quadro 1), segundo a Taxonomia de Bloom Revisada (FERRAZ; BELHOT, 2010).

Para Trevisan e Amaral (2016), o reconhecimento do conceito avaliado está relacionado ao nível cognitivo “Lembrar”, o que possibilita ao/à estudante reproduzir ideias e conteúdos. Além disso, as tarefas desse nível também podem se aproximar de ações como distinguir, selecionar, reproduzir e recordar uma informação relevante, que foi memorizada (TREVISAN; AMARAL, 2016).



**Figura 1. Item aplicado pelo PAEBES**

(B12002105) A imagem abaixo representa o ciclo da água.

Disponível em: <[http://www.escolovar.org/agua\\_ciclo.htm](http://www.escolovar.org/agua_ciclo.htm)>. Acesso em: 24 nov. 2014. \*Adaptada para fins didáticos.

Nessa imagem, qual fenômeno é indicado pelo número 3?

- A) Condensação.
- B) Decomposição.
- C) Evaporação.
- D) Precipitação.
- E) Transpiração.

Fonte: CAEd (2019, p. 65).

Diante dos sentidos que podem ser atribuídos a “Lembrar”, voltamos nosso olhar para a imagem, no sentido de analisá-la quanto às possibilidades de o/a estudante ter utilizado como um meio de resgatar de sua memória conceitos de definições pertinentes ao ciclo da água, uma vez que esse assunto é abordado nas aulas de Ciências desde os anos iniciais do Ensino Fundamental.

A abordagem do ciclo da água nos anos iniciais do Ensino Fundamental vai ao encontro das orientações da Base Nacional Comum Curricular (BNCC), que indica o desenvolvimento das seguintes habilidades no 5º ano do Ensino Fundamental:

(EF05CI02) Aplicar os conhecimentos sobre as mudanças de estado físico da água para explicar o ciclo hidrológico e analisar suas implicações na agricultura, no clima, na geração de energia elétrica, no provimento de água potável e no equilíbrio dos ecossistemas regionais (ou locais).

(EF05CI03) Selecionar argumentos que justifiquem a importância da manutenção da cobertura vegetal para a manutenção do ciclo da água, a preservação dos solos, dos cursos de água e da qualidade do ar atmosférico (BRASIL, 2017, p. 341).



O mesmo documento justifica a indicação dessas habilidades à medida que,

[...] ao iniciar o Ensino Fundamental, os alunos possuem vivências, saberes, interesses e curiosidades sobre o mundo natural e tecnológico que devem ser valorizados e mobilizados. Esse deve ser o ponto de partida de atividades que assegurem a eles construir conhecimentos sistematizados de Ciências, oferecendo-lhes elementos para que compreendam desde fenômenos de seu ambiente imediato até temáticas mais amplas (BRASIL, 2017, p. 331).

No caso do item apresentado na figura 1, a representação utilizada para o ciclo da água é semelhante àquelas retratadas nos livros didáticos. Por exemplo, para o 5º ano do Ensino Fundamental, um dos livros didáticos de Ciências (LEMBO; COSTA; VESPASIANO, 2017), aprovado pelo Programa Nacional do Livro Didático (PNLD) em 2019, mostra a representação do ciclo da água, semelhante à imagem apresentada nesse item, com os nomes das etapas retratadas pelas setas.

Portanto, essa representação da figura 1 pode ter sido reconhecida pelos/as avaliados/as, que estavam matriculados/as no 3º ano do Ensino Médio, uma vez que é comum ser encontrada nos livros didáticos, quando a temática é tratada. Logo, a memorização dos conceitos pode ter sido acionada para a resolução do item.

Diante dessa hipótese, buscamos identificar a categoria em que o item foi classificado na escala de proficiência (padrão de desempenho), divulgada na Revista do Professor (anteriormente citada). Na revista, ele foi classificado na categoria “Abaixo do básico” por ter uma ancoragem (na escala de proficiência que vai até 500) de até 250 pontos.

Essa ancoragem também pode ser justificada pelo fato de a imagem, apesar de não ser colorida, trazer elementos, como nuvens, montanhas e árvores, que são conhecidos pelos/as estudantes. Diante deles, podemos afirmar que o grau de iconicidade tende a ser elevado, levando a menor abstração (JALOTO; MEDEIROS, 2016). Vale ressaltar que “a **iconicidade** se refere ao grau com que a imagem se relaciona com a realidade que busca representar” (JALOTO; MEDEIROS, 2016, p. 2877, grifo dos autores).

Outros elementos, em um primeiro momento, podem não ser compreensíveis - lençóis de água, oceanos, lagos e rios -, pois apresentam uma tonalidade escura, o que dificultaria a interpretação. Logo, acreditamos que tiveram seus nomes escritos (legendas) para facilitar o entendimento da trajetória da água, no ciclo, em seus diferentes estados físicos. Isso corrobora com o trabalho de Jaloto e Medeiros (2016).



Esse auxílio na interpretação da imagem também pode ter sofrido influência das setas indicativas e seus respectivos números, que indicaram a ordem de leitura do ciclo da água. Segundo Medeiros et al. (2017), o conceito de seta pelo/a estudante é importante para a compreensão de que ela indica uma sequência, um fluxo.

As setas foram importantes para distinguir o gabarito (letra c – Evaporação) do distrator (letra a – Condensação). Este último poderia ser considerado como a resolução, se a ponta da seta não estivesse indicando o sentido da mudança de estado físico. Como destacam Martins e Queiroz (2010), as setas exibem muitas e variadas funções, podem ser utilizadas para suprir limitações apresentadas pela imagem, indicar sentido, representar mudança ou transformação, entre outras funções.

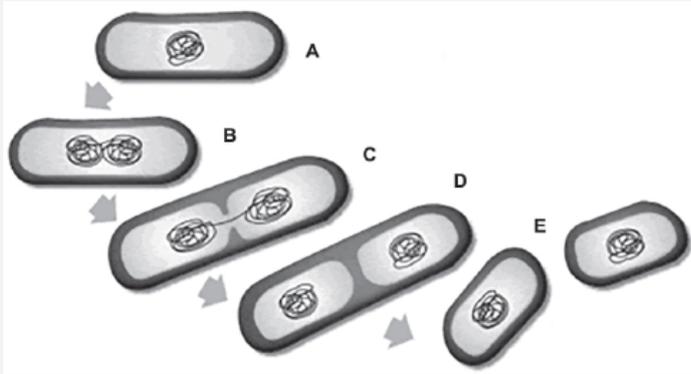
Diante da análise da primeira imagem, avaliamos que a clareza, assim como as setas, nomes e números nela presentes, foram importantes não só para a compreensão do ciclo da água representado, mas também para a resolução do item, uma vez que este não apresenta um enunciado com informações que podem auxiliar o/a estudante ou até mesmo algo contextualizando a situação do item.

Trazemos agora para a análise um segundo item (figura 2) que também se utiliza de recursos como setas e letras que representam etapas de um fenômeno biológico. Tal item tem como objetivo identificar o desenvolvimento da seguinte habilidade: “Reconhecer a rapidez da reprodução assexuada como uma vantagem desse tipo de reprodução em relação a sexuada” (CAED, 2019, p. 75). Portanto, este, assim como o anterior (figura 1), se restringe ao primeiro nível do desenvolvimento cognitivo da Taxonomia de Bloom Revisada - reconhecer (FERRAZ; BELHOT, 2010), cujo significado está no quadro 1.



**Figura 2. Item aplicado pelo PAEBES**

(B120770G5) A imagem abaixo apresenta as etapas de um tipo de reprodução.



Disponível em: <<https://goo.gl/dJV5C>>. Acesso em: 24 out. 2017. \*Adaptado para fins didáticos.

Constitui uma característica desse tipo de reprodução

- A) a alta taxa de crescimento.
- B) a geração de poucos descendentes.
- C) a troca de material genético.
- D) o alto gasto energético no processo.
- E) o aumento da variabilidade genética.

Fonte: CAEd (2019, p. 74).

A reprodução assexuada é um dos conteúdos programáticos previstos para o 2º ano do Ensino Médio, de acordo com o Currículo Básico Escola Estadual do Espírito Santo (SEDU, 2009). Dessa forma, espera-se que, no 3º ano (ano escolar avaliado pelo PAEBES), o/a estudante tenha construído conhecimentos sobre esse assunto.

Na BNCC do Ensino Médio, a área de Ciências da Natureza e suas Tecnologias não apresenta uma habilidade específica que oriente a abordagem da reprodução e seus tipos. Mas, na Competência 2, cuja escrita é “Construir e utilizar interpretações sobre a dinâmica da Vida, da Terra e do Cosmos para elaborar argumentos, realizar previsões sobre o funcionamento e a evolução dos seres vivos e do Universo, e fundamentar decisões éticas e responsáveis” (BRASIL, 2018, p. 542), podem ser mobilizados os conceitos relacionados à reprodução e hereditariedade (BRASIL, 2018).

O mesmo documento indica essa mesma mobilização no 8º ano, sendo a habilidade respectiva a ela: “(EF08CI07) Comparar diferentes processos reprodutivos em plantas e animais em relação aos mecanismos adaptativos e evolutivos” (BRASIL, 2017, p. 349). Sendo assim, o/a estudante pode ter contato com definições sobre reprodução sexuada e assexuada, diferenciando-as, desde os anos finais do Ensino Fundamental.

Esse fato é confirmado por livros didáticos de Ciências aprovados pelo PNLD de 2018 que, assim como na representação para o ciclo hidrológico, também apresentam representações semelhantes (USBERCO et al., 2018) àquela utilizada no item retratado na figura 2, desde o 8º ano do Ensino Fundamental.

Em face às semelhanças entre as imagens utilizadas como suporte nos itens e as representações presentes em livros didáticos aprovados no PNLD (LEMBO; COSTA; VESPASIANO, 2017, USBERCO et al., 2018), podemos inferir que a memorização possa ser acionada para a resolução do item, pelo reconhecimento da imagem. Entretanto, o item foi categorizado como “avançado” por ter ancorado acima de 375 pontos, em uma escala de proficiência que possui, como pontuação máxima, 500 pontos (CAED, 2019).

Diante dessa categorização, buscamos indícios na imagem que pudessem justificá-la. Ela mostra uma célula, cujo material genético se duplica para, então, haver a formação de descendentes. Como não há troca de material genético, a reprodução assexuada, em questão, proporciona o crescimento da população e é feita de forma mais rápida, gerando mais descendentes, quando comparada com a reprodução sexuada.

Contudo, ao contrário do que vemos nas representações utilizadas nos livros didáticos que apresentam legendas e indicações dos processos para fins de ensino, a imagem retratada na figura 2 não apresenta qualquer legenda. Por isso, acreditamos que possa haver dificuldades na compreensão de que não houve troca de material genético na célula representada na letra B, uma vez que parece haver dois deles em uma mesma célula, além de uma indicação de uma provável interação na célula da letra C. Dessa forma, reforçamos a necessidade da clareza das representações para a inteligibilidade do conceito (MARTINS; GOUVÊA; PICCININI, 2005; PEREIRA; BIANCO, 2016).



As legendas são um aspecto a ser considerado quando analisamos a ancoragem do item. Sua ausência pode elevar o grau de dificuldade, o que justifica a ancoragem do item, próximo ao limite máximo (500 pontos). Elas poderiam ser inseridas ao invés das letras (de A a E), uma vez que as setas já cumprem a finalidade de apontar o direcionamento da leitura da imagem, sendo essa uma de suas finalidades (MARTINS; QUEIROZ, 2010).

Essa indicação por setas e legendas pode ser importante pelo fato de que esse item, assim como o anterior analisado (figura 1), não apresenta um enunciado que possa auxiliar o/a estudante a compreender a imagem.

Além disso, ao nos voltarmos para os distratores, aquele da letra B pode induzir o/a estudante a relacioná-lo ao gabarito com base na imagem, já que uma célula (letra A) gerou apenas outras duas (letra E), o que pode ser interpretado como a geração de poucos descendentes.

Um outro fator que pode ter contribuído para a ancoragem do item pode ser o baixo grau de iconicidade da imagem, o que leva o/a estudante a construir o conhecimento de forma abstrata. Como apontado anteriormente, o estudo de Souza (2014) mostra a relação entre imagens com baixa iconicidade e problemas enfrentados no Ensino de Ciências. Por isso, assim como Cavalcante et al. (2012), acreditamos na importância das imagens para a representação dos conceitos que se quer avaliar.



## 5. Considerações finais

A partir da presente pesquisa, entendemos que nossos resultados contribuem para a consolidação de uma área de investigação ainda recente no campo do Ensino de Ciências. A sua importância é justificada pela possibilidade de um olhar mais crítico em torno da utilização de imagens em itens.

Vimos que a imagem aparece de duas maneiras no item, seja como ilustração de uma situação apresentada no texto ou sem uma apresentação prévia, direcionando o leitor para o comando logo após a imagem. Inclusive, nos casos em que não há uma legenda da imagem no item, é possível perceber que isso pode ser um dificultador na interpretação da imagem. A forma como a imagem é colocada no item pode influenciar nos sentidos a serem produzidos. Recursos complementares apresentados na imagem, por exemplo, setas, nomes e números, podem favorecer sua compreensão e resolução.

Por fim, em razão da linguagem específica utilizada na área e da própria natureza polissêmica da imagem, é preciso uma atenção especial para que as leituras realizadas pelos/as alunos/as coincidam com as desejadas pelos/as elaboradores/as. Assim, entendemos que a utilização de imagens em itens da área de Ciências da Natureza necessita de um cuidado especial por parte dos/as elaboradores/as ou pelos/as professores/as.



## Referências

BARDIN, L. **Análise de conteúdo (L. de A. Rego & A. Pinheiro, Trads.)**. Lisboa: Edições 70. (Obra original publicada em 1977), 2006.

BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 1995.

BOGDAN, R. C.; BIKLEN, S. K. **Investigação qualitativa em educação**: uma introdução à teoria e aos métodos. Portugal: Porto Editora, 1994.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular – Ensino Fundamental**. Brasília: MEC, 2017.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular – Ensino Médio**. Brasília: MEC, 2018.

CAEd. Programa de Avaliação da Educação Básica do Espírito Santo. **Revista do Professor – Ciências da Natureza**, 2019. Disponível em: <https://avaliacaoemontoramentoespíritosanto.caeddigital.net/#!/colecoes>. Acesso em: 29 jul. 2021.

CAVALCANTE, A. L. B. L.; BARROS, V. T. de O.; ROCHA, P. R. Z.; PEREIRA, F. A. F.; PERASSI, R. L. de S.; REMOR, C. A. M. **Revista Científica de Design**, v.3, n.1, p.183-192, jul. 2012.

CAVALCANTE, R. B.; CALIXTO, P.; PINHEIRO, M. M. K. Análise de conteúdo: considerações gerais, relações com a pergunta de pesquisa, possibilidades e limitações do método. **Informação & Sociedade: Estudos**, João Pessoa, v.24, n.1, p. 13-18, jan./abr. 2014.

ESPÍRITO SANTO. Secretaria de Educação. **Currículo Básico Escola Estadual**. Ensino Médio: área de Ciências da Natureza, v. 02. Vitória: SEDU, 2009.

FERRAZ, A. P. do C. M.; BELOT, R. V. Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. **Gestão & Produção**, v.17, n.2, p.421-431, 2010.

JALOTO, A.; MEDEIROS, L. Leitura de imagens no Saeb Ciências: um estudo a partir do erro dos estudantes. **Revista de Ensino de Biologia (SBEnBio)**, n. 9, p. 2875-2886, 2016.

LEMBRO, A.; COSTA, I.; VESPASIANO, S. S. **Odisséia: Ciências 5º ano**, v. 05. São Paulo: SEI, 2017.

MARTINS, I.; GOUVÊA, G.; PICCININI, C. Aprendendo com Imagens. **Ciência e Cultura**, v.. 57, n. 57, p. 38-40, 2005.

MARTINS, L.; QUEIROZ, J. Morfologia para setas em livros didáticos: Uma abordagem semiótica. **Arcos Design**, v.5, n.2, p. 2-16, dez. 2010.

MEDEIROS, L.; COSTA, D. O.; GUEDES, A. T.; REGO, A. M. J. Aprendendo com as respostas dos estudantes a itens dos testes das avaliações de larga escala: lições do Sistema Nacional de Avaliação da Educação Básica. **Revista Com Censo**. v. 4, n. 4, nov. 2017.

PEREIRA, S. M.; BIANCO, A. A. G. Recursos imagéticos para o ensino das ciências a crianças disléxicas. **CCNEXT - Revista de Extensão**, v.3 - n. Ed. Especial, p. 793-797, 2016.

SANA, T. C. V.; SOUZA, D. D. D.; ARROIO, A. O papel das imagens fotográficas no processo de construção do conhecimento em sala de aula de Química. **Revista Debates em Ensino de Química**, v.2, n.1, p. 5-17, abr. 2016.

SILVA, H. C. da; ZIMMERMANN, E.; CARNEIRO, M. H. da S.; GASTAL, M. L.; CASSIANO, W. S. Cautela ao usar imagens em aulas de ciências. **Ciência e Educação**, v. 12, n. 2, p. 219-233, abr. 2006.

SILVA, G. S.; BRAIBANTE, M. E. F.; PAZINATO, M. S. Os recursos visuais utilizados na abordagem dos modelos atômicos: uma análise nos livros didáticos de Química. **Revista Brasileira de Pesquisa em Educação em Ciências**, v. 13, n. 2, p. 159-182, out. 2013.

SOUZA, L. H. P. Imagens científicas e ensino de ciências: uma experiência docente de construção de representação simbólica a partir do referente real. **Cadernos Cedes**, v. 34, n. 92, p. 127-131, out. 2014. Disponível em: <https://www.cedes.unicamp.br/>. Acesso em: 30 jul. 2021.

TREVISAN, A. L.; AMARAL, R. G. A Taxonomia revisada de Bloom aplicada à avaliação: um estudo de provas escritas de Matemática. **Ciência & Educação**, Bauru, v. 22, n. 2, p. 451-464, out. 2016.

TRIVIÑOS, A. N. S. **Introdução à pesquisa em ciências sociais**: a pesquisa qualitativa em educação. São Paulo: Atlas, 1987.

USBERCO, J.; MARTINS, J. M.; SCHETCHMANN, E.; FERRER, L. C.; VELLOSO, H. M. **Companhia das Ciências – 8º ano**: Ensino Fundamental, anos finais. 5 ed. São Paulo: Saraiva, 2018.



---

# Sobre os Autores

**Adriana Rocha Bruno** é Professora do Departamento de Didática e do PPGEduc, da Escola de Educação, da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). É vice-coordenadora do Curso de Pedagogia a distância (LIPEAD) e é membro do NDE (Núcleo Docente estruturante) da UNIRIO. Pós-Doutora em Educação pelo Instituto de Educação da Universidade de Lisboa-Portugal, Doutora e mestre em Educação: Currículo - PUC-SP e Graduada em Pedagogia (UNIP). Foi professora da Educação Básica por 17 anos. Pesquisadora de produtividade PQ-2. É líder do Grupo de Pesquisa Aprendizagem em Rede - GRUPAR, que conta com pesquisadores e estudiosos envolvidos em investigações sobre as Docências contemporâneas, a Aprendizagem do adulto em ambientes em rede e a formação docente em meio à Cultura Digital. É autora do Livro: Formação de professores na cultura digital: aprendizagens do adulto, educação aberta, emoções e docências, publicado pela Editora da UFBA em 2021 e a versão em E-Book está aberta no site da editora. Contato: arbruno2208@gmail.com

**Ana Carolina Araújo da Silva** realizou seu doutorado em Educação pela Universidade Federal de Minas Gerais (2015) e seu mestrado em Educação pela Universidade Federal de Mato Grosso (2010). É professora da Universidade Federal de Juiz de Fora e do Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública - CAEd/FACED/UFJF e do Programa de Pós-Graduação em Educação PPGE/UFJF. Atuou na Universidade Federal de Santa Catarina, no Centro de Formação e Atualização dos Profissionais da Educação Básica (CEFAPRO) e na Escola Estadual Porfíria Paula de Campos em Mato Grosso na área de Ensino de Química. Desenvolve pesquisa na área de avaliação em Ciências da Natureza no CAEd/UFJF. Contato: anacarolina.silva@ufjf.br



**Andréia Francisco Afonso** realizou Mestrado em Zoologia no Museu Nacional do Rio de Janeiro e Doutorado em Ciência na Universidade Federal de São Carlos. É professora adjunta da UFJF, onde atua no curso de Licenciatura em Química e no Programa de Pós-Graduação em Química. Atua também no Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública - CAEd/FACEd/UFJF e na Especialização em Ensino de Ciências e Matemática nos anos iniciais da FACEd/UFJF. É líder do grupo QuInFormação e pesquisadora do CAEd-UFJF, no projeto de pesquisa em avaliação, na área de Ciências da Natureza. Contato: andreaia.afonso@ufjf.br

**Begma Tavares Barbosa** graduou-se em Letras, com Mestrado em Linguística pela Universidade Federal de Juiz de Fora e doutorado em Linguística pela Pontifícia Universidade Católica do Rio de Janeiro. É professora titular aposentada do Colégio de Aplicação João XXIII da Universidade Federal de Juiz de Fora, onde atuou no Ensino Médio e na formação de professores através da orientação de estágio curricular. No campo da pesquisa dedicou-se, principalmente, ao tema do letramento literário. Atualmente é pesquisadora do Centro de Avaliação da Educação Básica (CAEd), onde desenvolve investigação sobre a produção de sentido na leitura, com vistas ao aperfeiçoamento das matrizes de avaliação. Contato: begmatb@gmail.com

**Clarice Parreira Senra** realizou doutorado em Ciência, Tecnologia e Educação e Mestrado profissionalizante em Ensino de Ciências e Matemática no CEFET/RJ. É professora adjunta do Departamento de Educação da Universidade Federal de Juiz de Fora, onde atua nos cursos de Licenciatura em Física e Pedagogia. Desenvolve trabalhos na área de robótica educacional e ensino por projetos. É pesquisadora do CAEd-UFJF, no projeto de pesquisa em avaliação, na área de Ciências da Natureza. Contato: clarice.senra@ufjf.br

**Cristiano Nascimento da Silva** é pós-graduado em Banco de Dados e graduado em Análise e Desenvolvimento de Sistemas, ambos pelo Centro Universitário Claretiano. É graduando do bacharelado em Ciência da Computação pela Universidade Federal de Juiz de Fora. Participou de um estágio por 2 anos, no qual trabalhou utilizando Inteligência Artificial, desenvolvimento de aplicações e tratamento de dados, com o objetivo de reconhecer e avaliar automaticamente a leitura de crianças. Contato: cnsilva@ice.ufjf.br



**Danielle Xabregas Pamplona Nogueira** é Doutora e Mestre em Educação pela Universidade de Brasília (UnB). Especialista em Administração Escolar e Metodologia da Educação Superior. Licenciada em Educação Artística (habilitação em Música) pela Universidade do Estado do Pará (UEPA). É professora do Departamento de Políticas Públicas e Gestão da Educação da Faculdade de Educação da Universidade de Brasília (UnB), atuando principalmente nos seguintes temas: políticas públicas, gestão e tecnologias educacionais. Contato: daniellen@unb.br

**Fabiano Santos Saito** realizou Doutorado em Linguística (2015) e Mestrado em Linguística (2011), ambos pelo Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora (UFJF). Possui especialização *latu sensu* em Televisão, Cinema e Mídias Digitais (2012) pela UFJF e especialização em Docência com ênfase na Educação Básica (2022) pelo Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG), Campus Arcos. Possui experiência de 10 anos como professor da educação básica, atualmente leciona na Universidade do Estado de Minas Gerais (UEMG), no Departamento de Educação, Linguística e Letras (DELL) da Unidade Carangola, atuando no Curso de Letras. Atua desde 2020 nas pesquisas sobre elaboração de itens digitais para avaliação da compreensão da escuta em Língua Portuguesa na Fundação Centro de Políticas Públicas e Avaliação da Educação (Fundação CAEd). Contato: prof.fabiano.saito@gmail.com

**Hilda Aparecida Linhares da Silva** realizou Pós-doutorado e Doutorado em Educação na PUC-Rio e Mestrado em Educação na UFJF. É professora associada da Universidade Federal de Juiz de Fora, onde atua no curso de Licenciatura em Pedagogia, no Programa de Pós-Graduação em Educação da Faculdade de Educação e no Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública do CAEd/UFJF. É Líder do grupo LINFE, que realiza pesquisas sobre infâncias, linguagem e educação. É Especialista em Linguagem do CAEd/UFJF, onde coordena projetos de pesquisa em avaliação e atua em projetos de avaliação e desenvolvimento profissional em diversos estados da federação. Atua como consultora do INEP em avaliações nacionais de Língua Portuguesa para os anos iniciais do ensino fundamental. Contato: hilda.micarello@uab.ufjf.br



**Jairo Francisco de Souza** é Doutor em Informática pela PUC-Rio e Mestre em Engenharia de Sistemas e Computação pela COPPE/UFRJ. É professor associado da Universidade Federal de Juiz de Fora, onde atua no Departamento de Ciência da Computação e leciona no Programa de Pós-Graduação em Ciência da Computação da UFJF. É pesquisador do Laboratório de Aplicações e Inovação em Computação (LApIC), onde realiza pesquisas em Inteligência Artificial aplicada à Educação envolvendo aprendizagem adaptativa, avaliação de fluência em leitura, sistemas educacionais de recomendação e knowledge tracing. Contato: [jairo.souza@ice.ufjf.br](mailto:jairo.souza@ice.ufjf.br).

**João Victor de Souza** é bacharel em Ciência da Computação pela Universidade Federal de Juiz de Fora. Participou de projetos com temas relacionados à Inteligência Artificial e Processamento de Linguagem Natural, em contextos como detecção e prevenção de notícias falsas, sistemas de recomendação e reconhecimento automático de fala. Associado ao Laboratório de Aplicações e Inovação em Computação (LApIC), e atualmente fazendo parte da Fundação do Centro de Políticas Públicas e Avaliação da Educação CAEd/UFJF, participa do projeto de avaliação automática de fluência em leitura. Contato: [joao.souza@caed.ufjf.br](mailto:joao.souza@caed.ufjf.br)

**Joaquim José Soares Neto** é professor titular do Instituto de Física da Universidade de Brasília (UnB). Realizou a graduação e o mestrado na UnB, doutorado na Aarhus University (Dinamarca) e pós-doutorado no California Institute of Technology (Caltech), nos Estados Unidos. Trabalha em pesquisa na área de Avaliação da Educação Superior e da Educação Básica por mais de 18 anos. Foi Presidente do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep); é Conselheiro da Câmara de Educação Superior do Conselho Nacional de Educação (CNE); e é Presidente da Associação Brasileira de Avaliação Educacional (ABAVE). Foi Relator no Conselho Nacional de Educação da Base Nacional Comum Curricular (BNCC). Atualmente, atua no Programa de Pós-Graduação em Desenvolvimento, Sociedade e Cooperação Internacional do Centro de Estudos Avançados Multidisciplinares (Ceam/UnB) e no Programa de Pós-Graduação em Ensino de Física, ambos da UnB. Contato: [jjsoaresneto@gmail.com](mailto:jjsoaresneto@gmail.com)



**Juliana Vicini Florentino Rodrigues** possui graduação em Ciências Exatas, em Bacharelado em Química e em licenciatura em Química pela Universidade Federal de Juiz de Fora. Possui mestrado em Química pela mesma universidade. Atualmente compõe a equipe de Ciências da Natureza como analista em instrumentos de avaliação na área de Química na Fundação CAEd. Contato: juvicini@hotmail.com

**Leila Márcia Mafra Martins** é pós-graduada em Gestão educacional, Faculdade Metodista Granbery (onde exerceu o cargo de professora por 23 anos em todos os segmentos até o curso de Pedagogia). É graduada em Letras, assim como em Artes, pelo CES-JF Centro de Ensino Superior de Juiz de Fora. Trabalhou (2011/2022) no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF), como Analista de Instrumentos de Avaliação, na equipe de Escrita em elaboração de itens de leitura, produção de texto e fluência em leitura oral, com atuação desde a implementação e elaboração dos primeiros testes em larga escala de fluência em leitura oral, com participação na criação de ferramentas e aplicação de pilotos. Contato: mafra\_leila@yahoo.com.br

**Liamara Scortegagna** é professora do Departamento de Ciência da Computação (DCC) da Universidade Federal de Juiz de Fora (UFJF). Doutora em Engenharia de Produção - Mídia e Conhecimento pela UFSC, Mestre em Ciência da Computação pela UFSC e Graduada em Informática pela UnC. Atua nos Programas de Pós-Graduação em Educação Matemática (PPGEM) e em Gestão e Avaliação da Educação Pública (PPGP) da UFJF. É pesquisadora da Fundação CAEd/UFJF, coordenadora do curso de Graduação em Licenciatura em Computação (LiCOMP/UFJF) e membro dos grupos de pesquisa: GRUPAR – Grupo de Pesquisa Aprendizagem em Rede, NIDEEM - Núcleo de investigação, Divulgação e Estudos em Educação Matemática - Tecnologias de Informação e Comunicação na Educação Matemática e LApIC - Laboratório de Aplicações e Inovação em Computação. É autora do livro: Objetos de Aprendizagem pela editora CEAD/UFJF. Contato: liamara@ice.ufjf.br.



**Marcello Ferreira** é Professor e Vice-Diretor no Instituto de Física da Universidade de Brasília (UnB). Doutor em Educação em Ciências pela UFRGS, mestre em Ensino de Física e licenciado em Física pela UnB. Realizou Estágio Pós-Doutoral na UFSCar, na área de Ensino de Ciências e Tecnologias. Pesquisa e orienta nos temas: fundamentos teórico-metodológicos para o ensino de ciências e física; formação de professores de física; estudos foucaultianos; educação a distância; e políticas públicas da educação. Docente permanente no Programa de Pós-graduação em Ensino de Física da UnB, em Educação em Ciências da UFRGS e em Gestão e Avaliação da Educação Pública da UFJF. Pesquisador CAEd/UFJF na área de avaliação digital em física. Possui experiência como gestor de políticas educacionais, tendo exercido os cargos de Coordenador de Articulação Acadêmica e Coordenador-Geral de Programas e Cursos em EaD na Capes. Foi também professor, Coordenador de Planejamento, Desenvolvimento e Avaliação e Pró-Reitor de Assuntos Estudantis e Comunitários da Universidade Federal do Pampa. Atualmente, é Pró-Reitor de Pós-Graduação da Sociedade Brasileira de Física (SBF). Contato: marcellof@unb.br

**Marcos Rogério Martins Costa** é Doutor e Mestre em Letras pela Universidade de São Paulo (USP); graduado em Letras (Português/Linguística) pela mesma instituição e em Pedagogia pela Universidade de Santo Amaro (UNISA). Atua, atualmente, como membro da Equipe Multidisciplinar do Centro de Educação a Distância da Universidade de Brasília. Entre 2020 e 2021, foi professor substituto no Instituto de Letras da Universidade de Brasília (UnB). É CEO da startup Escrita com Ciência. Pertence à Academia Contemporânea de Letras (ACL), ocupando a cadeira 21, cujo patrono é Érico Veríssimo, e à Academia Nacional de Ciências, Letras e Artes (ANACLA), ocupando a cadeira 7, cujo patrono é José Saramago. Possui experiência em Língua Portuguesa, Linguística Geral, Semiótica, Análise do Discurso (AD-Francesa) e Estudos do Círculo de Bakhtin. Suas outras áreas de interesse são Ciências da Linguagem, Modelos de Ensino e Aprendizagem e Novas Mídias. Contato: marcosrmcosta15@gmail.com



**Olavo Leopoldino da Silva Filho** possui graduação em Filosofia pela Universidade de Brasília (1999), graduação em Física pela Universidade de Brasília (1986), graduação em Ciências da Computação pela Universidade de Brasília (1994), graduação em Matemática pela Universidade de Brasília (1988), Mestrado em Filosofia pela Universidade de Brasília (2001), mestrado em Física pela Universidade de Brasília (1989) e Doutorado em Física pela Universidade de Brasília (1995). Atualmente é professor associado da Universidade de Brasília. Tem experiência na área de Física, com ênfase em Interpretação da Mecânica Quântica, em particular na interface quântico-clássico. Em filosofia, trabalha com lógica e linguagem, em particular com a aplicação da lógica na análise de línguas naturais. Contato: olavolsf@unb.br

**Rita de Cássia Reis** realizou Doutorado em Educação, área de concentração Educação em Ciências na UFMG e Mestrado em Química, na área de concentração Educação em Química na UFJF. É professora adjunta da Universidade Federal de Juiz de Fora, onde atua nos cursos de Licenciatura em Pedagogia e Licenciaturas em Química, também atua no Programa de Pós-Graduação em Educação da Faculdade de Educação/UFJF. É Líder do grupo NARRATIVAS, que realiza pesquisas e ações extensionistas sobre formação de professores, desenvolvimento profissional, pesquisa narrativa e ensino de ciências. É pesquisadora do CAEd/UFJF, no projeto de pesquisa em avaliação, na área de Ciências da Natureza. Contato: ritaeduquim@hotmail.com

**Rosângela Veiga Júlio Ferreira** realizou Doutorado e Mestrado em Educação na UFJF. É professora do Ensino Básico Técnico e Tecnológico (EBTT) do Colégio de Aplicação João XXIII/UFJF. Atualmente, é vice-líder do grupo de pesquisa Linguagens, Infâncias e Educação (FACEd/UFJF). É Especialista em Linguagem do CAEd/UFJF, onde pesquisa sobre avaliação em leitura e atua como professora do mestrado profissional do Programa de Pós-Graduação em Gestão e Políticas Públicas (PPGP-CAEd/UFJF). Contato: rosangelaveiga.ferreira@ufjf.br



**Sarah Matos Rocha Mesquita** é doutoranda em Linguística pela Universidade Federal de Juiz de Fora (UFJF) e Mestre em Linguística pela UFJF. É Analista de Instrumentos de Avaliação no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF), na equipe de Itens e Testes dos Anos Iniciais, cuja atribuição é a elaboração dos itens e dos testes de leitura (Língua Portuguesa e Matemática), Escrita e Fluência em Leitura. Contato: sarah.mesquita@caed.ufjf.br

**Wallace Alves Cabral** realizou Mestrado e Doutorado em Educação no PPGE-UFJF. É professor adjunto da UFSJ, onde atua no curso de Licenciatura em Química e no Programa de Pós-Graduação em Educação do Departamento de Ciências da Educação. Atuou também no Programa de Pós-Graduação em Ensino de Ciências da UFOP. É líder do grupo GEPLEC, que visa refletir sobre o funcionamento da linguagem em aulas de Ciências, buscando aporte teórico no referencial teórico da Análise do Discurso de Linha Francesa. Sou pesquisador do CAEd-UFJF, no projeto de pesquisa em avaliação, na área de Ciências da Natureza. Contato: wallacecabral@ufsj.edu.br



ufjf



SELO EDITORIAL  
GESTÃO E AVALIAÇÃO  
EDUCACIONAL

CAEd UFJF