

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE
MINAS GERAIS - *CAMPUS* OURO BRANCO
SISTEMAS DE INFORMAÇÃO

João Augusto Colem da Costa Val

**SISTEMA ADAPTATIVO DE RECOMENDAÇÃO DE QUESTÕES DO
ENEM BASEADO EM TEORIA DE RESPOSTA AO ITEM**

JOÃO AUGUSTO COLEM DA COSTA VAL

**SISTEMA ADAPTATIVO DE RECOMENDAÇÃO DE QUESTÕES DO
ENEM BASEADO EM TEORIA DE RESPOSTA AO ITEM**

Trabalho de conclusão de curso apresentado ao Curso de Sistemas de Informação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais - *Campus Ouro Branco* para a obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Prof. Luciana Lourdes Silva,

Coorientador: Prof. Márcio Assis Miranda

Ouro Branco - MG
2026

C838s Costa Val, João Augusto Colem da.

Sistema adaptativo de recomendação de questões do Enem: baseado em teoria de resposta ao item. /João Augusto Colem da Costa Val. – 2025.

14f.il.col.

Orientadora: Luciana Lourdes Silva.

Trabalho de Conclusão de Curso (Sistemas de Informação) – Instituto Federal de Minas Gerais. *Campus* Ouro Branco, 2025.

1. Teoria de resposta ao item. 2. Sistemas de recomendação. 3. Aprendizagem adaptativa. 4. Educação. 5. Enem. I. Silva, Luciana Lourdes. II. Instituto Federal de Minas Gerais. *Campus* Ouro Branco. III. Título.

CDU: 004.8



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
CAMPUS OURO BRANCO

Av. Afonso Sardinha, nº 90, Bairro Pioneiros, CEP: 36.420-000, Ouro Branco - Minas Gerais

(31) 3742-2149 – gabinete.ourobranco@ifmg.edu.br

ANEXO II – ATA DE CONCLUSÃO DE TCC

Aos 19 dias do mês de janeiro de 2026, às 20:00 horas,
João Augusto Colem da Costa Val, aluno(a) regularmente matriculado no Curso
de Sistemas de Informação do Instituto Federal de Minas Gerais, campus Ouro Branco,
matrícula 0066824, concluiu o seu Trabalho de Conclusão de Curso por meio de:

() Publicação do artigo intitulado _____
na revista/conferência _____,
cujo comprovante de aceitação será anexado a esta ata, recebendo a nota _____ pelo
trabalho. Eu, na qualidade de orientador do aluno, lavrei a presente ata atestando a
conclusão do trabalho, a qual será assinada por mim e pelo aluno.

Professor Orientador

Aluno

(X) Defesa em sessão pública realizada às 20:00 horas, na sala auditório do
Instituto Federal de Minas Gerais, campus Ouro Branco, na presença da banca
examinadora composta pelos docentes:

1 - Daniela Costa Terra

2 - Ederson Naves Fernandes Gonçalves Júnior

3 - Marcio Assis Miranda

do artigo intitulado Sistema Adaptativo de Recomendação de Questões do ENEM
Baseado em Teoria de Resposta ao Item



**MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
CAMPUS OURO BRANCO**

Av. Afonso Sardinha, nº 90, Bairro Pioneiros, CEP: 36.420-000, Ouro Branco - Minas Gerais

(31) 3742-2149 – gabinete.ourobranco@ifmg.edu.br

A banca examinadora, após reunião em sessão reservada, deliberou pela aprovação do referido trabalho, atribuindo a nota 97,6. Eu, na qualidade de presidente da banca examinadora, lavrei a presente ata que será assinada por mim, pelos demais examinadores e pelo aluno.

Observações pertinentes à defesa:

NOME E ASSINATURA DOS COMPONENTES DA BANCA E DO ORIENTADO



Documento assinado digitalmente
LUCIANA LOURDES SILVA
Data: 24/01/2026 22:13:17-0300
Verifique em <https://validar.iti.gov.br>

Professor Orientador: *Luciana Lourdes Silva*



Documento assinado digitalmente
DANIELA COSTA TERRA
Data: 26/01/2026 17:26:49-0300
Verifique em <https://validar.iti.gov.br>

Examinador 1: *Daniela Costa Terra*



Documento assinado digitalmente
EDERSON NAVES FERNANDES GONCALVES JUN
Data: 26/01/2026 16:35:55-0300
Verifique em <https://validar.iti.gov.br>

Examinador 2: *Ederson Naves Fernandes Gonçalves Júnior*



Documento assinado digitalmente
MARCIO ASSIS MIRANDA
Data: 26/01/2026 12:02:41-0300
Verifique em <https://validar.iti.gov.br>

Examinador 3: *Marcio Assis Miranda*



Documento assinado digitalmente
JOAO AUGUSTO COLEM DA COSTA VAL
Data: 23/01/2026 13:09:13-0300
Verifique em <https://validar.iti.gov.br>

Aluno(a): *João Augusto Colem da Costa Val*



MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
CAMPUS OURO BRANCO

Av. Afonso Sardinha, nº 90, Bairro Pioneiros, CEP: 36.420-000, Ouro Branco - Minas Gerais

(31) 3742-2149 – gabinete.ourobranco@ifmg.edu.br

DECLARAÇÃO ANTI-PLÁGIO

Eu, João Augusto Colem da Costa Val, estudante do curso Bacharelado em Sistemas de Informação do IFMG – Campus Ouro Branco, declaro, para os devidos fins e efeitos, e para fazer prova junto ao IFMG – Campus Ouro Branco, que, **sob as penalidades previstas no art. 299 do Código Penal Brasileiro**, que é de minha criação o Trabalho de Conclusão de Curso que ora apresento.

Art. 299 do Código Penal Brasileiro, que dispõe sobre o crime de *Falsidade Ideológica*:

“Omitir, em documento público ou particular, declaração que dele devia constar, ou nele inserir ou fazer inserir declaração falsa ou diversa da que devia estar escrita, com o fim de prejudicar direito, criar obrigação ou alterar verdade sobre fato juridicamente relevante:

Pena – reclusão, de 1 (um) a 5 (cinco) anos, e multa, se o documento é público, e reclusão de 1 (um) a 3 (três) anos, e multa, se o documento é particular.

Parágrafo único. Se o agente é funcionário público, e comete o crime prevalecendo-se do cargo, ou se a falsificação ou alteração é de assentamento de registro civil, aumenta-se a pena de sexta parte”.

Este crime engloba plágio e compra fraudulenta de documentos científicos. Por ser verdade, e por ter ciência do referido artigo, firmo a presente declaração.

Ouro Branco, 23 de janeiro de 2026



Documento assinado digitalmente
JOAO AUGUSTO COLEM DA COSTA VAL
Data: 23/01/2026 13:07:42-0300
Verifique em <https://validar.it.gov.br>

Assinatura do aluno: _____

RESUMO

Este artigo aborda o desafio do estudo personalizado para o ENEM, descrevendo o desenvolvimento de um sistema adaptativo de recomendação de questões baseado na Teoria de Resposta ao Item (TRI). Muitos estudantes têm dificuldade em selecionar questões de estudo adequadas, o que leva à desmotivação. O trabalho envolveu o desenvolvimento de um protótipo de aplicação web que recomenda questões progressivas do ENEM, utilizando a TRI para personalizar o percurso de aprendizagem. O protótipo inclui funcionalidades essenciais como autenticação, avaliação inicial de proficiência, recomendação adaptativa, dashboard de desempenho e revisão de questões. O estudo incluiu a construção de um banco de dados de questões, a implementação de um algoritmo de recomendação baseado em TRI e a realização de uma avaliação com 27 estudantes do primeiro ano do Ensino Médio para avaliar a percepção de utilidade e progresso. Os resultados validaram a hipótese central, revelando percepções positivas em todas as dimensões avaliadas: alta usabilidade (média 4,51), forte engajamento (média 4,35) e taxa de recomendação de 92,6% (25 de 27 estudantes), indicando forte aceitação da abordagem baseada em TRI.

Palavras-chave: ENEM. Teoria de Resposta ao Item. Sistemas de Recomendação. Aprendizagem Adaptativa. Educação.

ABSTRACT

This research addresses the challenge of personalized study for the ENEM (Brazilian National High School Exam) by developing an adaptive question recommendation system based on Item Response Theory (IRT). Students may struggle to select appropriate study questions, leading to frustration. This work involved the development of a prototype web application that recommends progressive ENEM questions, leveraging IRT to personalize the learning path. The prototype includes essential features such as authentication, initial proficiency assessment, adaptive recommendation, performance dashboard, and question review. The study included the building of a question database, implementing an IRT-based recommendation algorithm, and conducting a pilot study with 27 first-year high school students to assess user perception of utility and progress. The results validated the central hypothesis, revealing positive perceptions across all evaluated dimensions: high usability (mean 4.51), strong engagement (mean 4.35), and a recommendation rate of 92.6% (25 out of 27 students), indicating strong acceptance of the IRT-based approach.

Keywords: ENEM. Item Response Theory. Recommendation Systems. Adaptive Learning. Education.

SUMÁRIO

1	Introdução	9
2	Contribuições	9
3	Referencial Teórico	10
3.1	Teoria de Resposta ao Item (TRI)	10
3.1.1	Modelo Logístico de Três Parâmetros	10
3.1.2	Escala de Proficiência do ENEM	10
3.2	Sistemas de Recomendação e Aprendizagem Adaptativa	11
3.3	Trabalhos Relacionados e Lacunas na Literatura	11
4	Metodologia	11
4.1	Desenho do Estudo	11
4.2	O Sistema Desenvolvido	12
4.3	Funcionamento do Sistema de Recomendação	13
4.3.1	Estimativa Inicial de Proficiência	13
4.3.2	Algoritmo de Seleção de Questões Adaptativa	13
4.3.3	Atualização da Proficiência	14
4.4	Configuração dos Parâmetros do Algoritmo	17
5	Resultados e Discussão	18
5.1	Resultados	18
5.1.1	Análise Qualitativa	18
5.2	Discussão	18
6	Limitações e Ameaças à Validade	19
7	Conclusão	20

1. Introdução

O Exame Nacional do Ensino Médio (ENEM) é uma das principais avaliações educacionais em larga escala do Brasil, criado em 1998 e consolidado a partir de 2009 como mecanismo central de seleção para ingresso nas instituições de ensino superior [INEP 2021b]. Anualmente, mais de 4 milhões de estudantes se inscrevem no exame, que avalia competências e habilidades desenvolvidas ao longo do ensino médio em quatro áreas do conhecimento [INEP 2024a]. Desde 2009, a correção utiliza a Teoria de Resposta ao Item (TRI), que permite a estimação de proficiências considerando os parâmetros psicométricos de cada item e o padrão de respostas do participante [INEP 2021b, INEP 2021a].

No entanto, essa complexidade metodológica cria desafios práticos significativos: identificar questões adequadas ao nível de proficiência torna-se uma tarefa complexa, pois as questões possuem características probabilísticas que variam conforme o perfil do respondente. Questões excessivamente simples não impulsionam o avanço, enquanto itens demasiadamente complexos podem gerar frustração e desmotivação, especialmente para estudantes que estão iniciando sua preparação.

Diante desse cenário, foi desenvolvido um protótipo de aplicação web que oferece recomendação progressiva e adaptativa de questões do ENEM, fundamentada na TRI. A hipótese central que guiou esta investigação foi que um sistema de recomendação adaptativo de questões, embasado na TRI, tem o potencial de amplificar a percepção de avanço e a motivação dos estudantes em sua preparação para o ENEM. Este trabalho explora como a integração rigorosa dos princípios psicométricos da TRI com algoritmos de recomendação pode criar uma ferramenta educacional que personalize efetivamente o estudo, apresentando questões em níveis apropriados de dificuldade que desafiam o estudante sem gerar frustração, promovendo assim um aprendizado mais eficaz e motivador desde os estágios iniciais do Ensino Médio.

2. Contribuições

Este trabalho apresenta quatro contribuições principais para o campo de sistemas de recomendação educacionais baseados em Teoria de Resposta ao Item. A primeira contribuição consiste na construção de um banco de dados estruturado que integra microdados oficiais do INEP com textos e alternativas de questões do ENEM, categorizadas por competências e habilidades, com parâmetros TRI (dificuldade b_j , discriminação a_j e acerto casual c_j) devidamente integrados. A segunda contribuição compreende o desenvolvimento de um algoritmo de recomendação adaptativo híbrido que combina busca em intervalos crescentes ao redor da proficiência estimada, seleção por discriminação e atualização de proficiência que incorpora a percepção de competência do usuário através de estimativa ponderada, demonstrando uma aplicação rigorosa dos princípios psicométricos da TRI em sistemas de recomendação. A terceira contribuição reside no desenvolvimento de uma plataforma web funcional que implementa o sistema completo de recomendação adaptativa, oferecendo funcionalidades essenciais como autenticação, classificação inicial de proficiência, recomendação adaptativa, dashboard de desempenho e revisão de questões. A quarta contribuição consiste em evidências empíricas obtidas em uma avaliação com 27 estudantes do primeiro ano do Ensino Médio, que revelaram alta usabilidade (média 4,51), forte engajamento (média 4,35) e taxa de recomendação de 92,6%.

3. Referencial Teórico

Este estudo integra três conceitos fundamentais: a Teoria de Resposta ao Item (TRI), os Sistemas de Recomendação e a Aprendizagem Adaptativa.

3.1. Teoria de Resposta ao Item (TRI)

A Teoria de Resposta ao Item (TRI) supera as limitações da Teoria Clássica dos Testes (TCT) ao elevar o item à condição de unidade primária de análise [Andrade et al. 2000]. A probabilidade de um indivíduo responder corretamente a um item é modelada como uma função que considera os parâmetros intrínsecos do item (dificuldade, discriminação e acerto casual) e a proficiência do indivíduo [Andrade et al. 2010]. Como definido por Andrade [Andrade et al. 2010], “a TRI é um conjunto de modelos matemáticos que considera o item como unidade básica de análise e procura representar a probabilidade de um examinando dar uma certa resposta a um item como função dos parâmetros do item e do(s) traço(s) latente(s) do indivíduo”. No contexto brasileiro, a TRI é utilizada em avaliações de larga escala como o SAEB e o ENEM, garantindo a comparabilidade das notas entre diferentes edições do exame. Embora sua utilização em avaliações somativas seja amplamente reconhecida, o potencial da TRI para aprimorar a avaliação formativa e a recomendação de estudos ainda constitui um campo fértil para a inovação [Andrade 2020].

3.1.1. Modelo Logístico de Três Parâmetros

O modelo de TRI utilizado no ENEM é o logístico de três parâmetros (ML3), desenvolvido por Birnbaum [Birnbaum 1968]. Este modelo expressa a probabilidade $P(x_j = 1|\theta, a_j, b_j, c_j)$ de um participante com proficiência θ responder corretamente a um item j através da seguinte função:

$$P(x_j = 1|\theta, a_j, b_j, c_j) = c_j + \frac{(1 - c_j)}{1 + e^{-Da_j(\theta - b_j)}} \quad (1)$$

Neste modelo, o parâmetro a_j ($a_j > 0$) representa a discriminação do item, indicando o poder de diferenciação entre participantes que dominam e não dominam a habilidade avaliada. O parâmetro b_j expressa a dificuldade do item na mesma escala da proficiência, sendo que quanto maior seu valor, mais difícil é a questão. O parâmetro c_j ($0 < c_j < 1$) representa a probabilidade de acerto casual, ou seja, a probabilidade de um examinando de proficiência muito baixa acertar o item por acaso. O traço latente θ representa a proficiência do indivíduo, enquanto D é uma constante de escala (geralmente $D = 1,7$ para aproximar a função logística à distribuição normal).

3.1.2. Escala de Proficiência do ENEM

No ENEM, as proficiências são calculadas em uma escala métrica padronizada. Inicialmente, os valores são calculados em uma escala (0,1) com média 0 e desvio-padrão 1. Para melhor interpretação, essa escala é transformada para a escala (500,100), onde 500 representa o desempenho médio dos concluintes regulares do ensino médio de

2009 e 100 representa o desvio-padrão. A transformação é realizada através da fórmula $\theta_{(500,100)} = 100 \cdot \theta_{(0,1)} + 500$. Essa escala permite que questões e participantes sejam posicionados na mesma “régua”, facilitando a interpretação pedagógica. Uma questão é posicionada no valor da escala em que a probabilidade de acerto está próxima de 0,65, indicando que participantes nesse nível possuem alta probabilidade de dominar o conteúdo avaliado. A TRI considera a coerência das respostas, esperando que participantes que acertam questões difíceis também acertem questões fáceis, pois a aquisição do conhecimento ocorre de forma cumulativa. Padrões incoerentes são interpretados como possíveis acertos ao acaso, contribuindo menos para o aumento da proficiência estimada. A estimação da proficiência no ENEM utiliza o método Expected a Posteriori (EAP), que combina a informação das respostas do participante com uma distribuição a priori das habilidades, considerando não apenas o número de acertos, mas também quais itens foram acertados e errados, bem como os parâmetros desses itens.

3.2. Sistemas de Recomendação e Aprendizagem Adaptativa

Sistemas de recomendação educacionais propõem materiais de estudo e atividades que se alinham ao perfil cognitivo de cada aluno [Silva et al. 2022]. Quando integrados a plataformas de ensino, originam a aprendizagem adaptativa, que ajusta o nível de dificuldade e o tipo de conteúdo em tempo real [Chen 2017]. O princípio da zona de desenvolvimento proximal de Vygotsky [Vygotsky 1978] fundamenta essa abordagem: o conteúdo deve estar entre o nível de desenvolvimento real e o potencial do estudante. A integração da precisão psicométrica da TRI com sistemas de recomendação permite criar ferramentas que personalizam efetivamente o aprendizado.

3.3. Trabalhos Relacionados e Lacunas na Literatura

A literatura aborda a aplicação da TRI em avaliações brasileiras [Andrade 2020] e sistemas de recomendação educacionais [Silva et al. 2022]. Diversos estudos exploram plataformas para simulados do ENEM [Braida and Rezende 2025] e personalização do estudo [Corrêa et al. 2023]. Contudo, a maioria das plataformas comerciais que afirmam empregar TRI recorrem a modelos de regressão simplificados [Estuda.com sd], e a maioria dos sistemas de recomendação educacionais não utiliza métricas psicométricas robustas, limitando-se a classificações subjetivas de dificuldade.

Identifica-se uma lacuna: a escassez de ferramentas que integrem os modelos psicométricos da TRI para oferecer recomendação verdadeiramente adaptativa. Este estudo preenche essa lacuna, investigando como a aplicação rigorosa dos princípios da TRI pode potencializar sistemas de recomendação. A diferenciação deste trabalho reside na integração explícita entre os parâmetros psicométricos da TRI (dificuldade, discriminação e acerto casual) e o algoritmo de recomendação.

4. Metodologia

A metodologia adotada foi de natureza aplicada e exploratória, com uma abordagem mista (qualitativa e quantitativa).

4.1. Desenho do Estudo

O estudo foi estruturado em duas fases. A primeira fase compreendeu o desenvolvimento do protótipo e a construção do banco de dados de questões do ENEM, integrando

microdados do INEP com textos e alternativas das questões. O banco de dados foi constituído por questões de todas as quatro áreas do conhecimento, com parâmetros TRI (dificuldade, discriminação, acerto casual) e categorizadas por competências e habilidades. O algoritmo de recomendação baseado em TRI foi implementado no backend, enquanto os frontends foram construídos para autenticação e interface do estudante. A segunda fase consistiu na avaliação com usuários: 27 estudantes do primeiro ano do Ensino Médio, selecionados por conveniência, realizaram uma única sessão de uso da plataforma, respondendo à avaliação inicial de proficiência e ao conjunto diário de questões recomendadas. Os dados foram coletados por meio de um questionário de percepção estruturado em quatro blocos: (1) dados básicos sobre preparação para o ENEM; (2) usabilidade, utilizando questões adaptadas do System Usability Scale (SUS); (3) percepção de adaptação e dificuldade das questões recomendadas; e (4) motivação e engajamento. Todos os itens foram respondidos em escala Likert de 1 a 5. A análise dos dados quantitativos foi realizada por meio de estatística descritiva e coeficientes de correlação de Pearson.

4.2. O Sistema Desenvolvido

O protótipo desenvolvido consiste em uma plataforma web¹ projetada para proporcionar uma experiência de estudo personalizada e adaptativa para o ENEM. A plataforma foi construída sobre uma arquitetura de monorepo com três aplicações: *authentication*, *dashboard* e *backend*. O backend segue Clean Architecture com camadas: *domain* (lógica de negócio), *infra* (repositórios e serviços), *presentation* (controladores HTTP) e *main* (configuração). A implementação utilizou Node.js com TypeScript e Express para o backend, React com TypeScript para os frontends, PostgreSQL como banco de dados, e API RESTful com autenticação JWT.

O sistema implementa cinco funcionalidades: autenticação e gerenciamento de conta, classificação inicial de proficiência (avaliação diagnóstica com questões pré-selecionadas), recomendação adaptativa de questões (núcleo do sistema), dashboard de desempenho (evolução da proficiência e estatísticas), e revisão de questões anteriormente respondidas.

A construção do banco de dados exigiu a integração de duas fontes distintas: microdados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) em formato CSV, contendo os parâmetros TRI de cada item (dificuldade b_j , discriminação a_j e acerto casual c_j), bem como informações sobre gabaritos e habilidades avaliadas, e arquivos JSON com textos e alternativas das questões [Yunger 2024, INEP 2024b]. Foi desenvolvido um algoritmo de importação que processa os microdados do INEP, realiza a correspondência com os arquivos JSON e insere as questões no banco de dados utilizando transações. Cada questão é mapeada para sua respectiva habilidade e competência dentro da área de conhecimento, seguindo a estrutura pedagógica do ENEM. O processo de importação garante a integridade referencial, validando a existência de áreas de conhecimento e habilidades antes da inserção, resultando em um banco de dados estruturado com questões do ENEM, seus parâmetros TRI completos e devidamente categorizadas por competências e habilidades.

¹Disponível em <https://app.capivest.com.br>

4.3. Funcionamento do Sistema de Recomendação

O sistema desenvolvido implementa um algoritmo de recomendação adaptativa baseado em TRI que opera em três etapas principais: (1) estimativa inicial de proficiência, (2) seleção de questões adaptativa e (3) atualização contínua da proficiência.

4.3.1. Estimativa Inicial de Proficiência

Quando um usuário se cadastra, é realizada uma avaliação inicial de proficiência em cada área de conhecimento. O processo consiste em apresentar 5 questões pré-selecionadas que cobrem diferentes níveis de dificuldade na escala TRI: 2 questões fáceis (parâmetro de dificuldade $b_j < 0.5$), 2 questões médias ($0.5 \leq b_j \leq 1.5$) e 1 questão difícil ($b_j > 1.5$). Esta distribuição garante que estudantes em diferentes níveis de proficiência tenham questões adequadas ao seu conhecimento, permitindo uma estimativa mais precisa.

Para cada questão, o sistema coleta a resposta (acerto ou erro) e o nível de conforto auto-reportado (escala de 0 a 5). O método de máxima verossimilhança ponderado combina essas informações para estimar a proficiência inicial (θ_0), onde acertos com baixo conforto contribuem menos, refletindo possíveis acertos ao acaso. A busca pelo θ que maximiza a verossimilhança é realizada no intervalo $[-4.0, 4.0]$ com passo de 0.01. A proficiência estimada é convertida para a escala ENEM (500,100) e armazenada.

4.3.2. Algoritmo de Seleção de Questões Adaptativa

A Fig.1 apresenta o algoritmo de seleção que utiliza os princípios da TRI para recomendar questões adequadas ao nível atual de proficiência do estudante, garantindo um desafio progressivo. O algoritmo é parametrizável, permitindo ajustes em três aspectos principais: (1) os intervalos de busca ao redor da proficiência, (2) o número mínimo de questões candidatas para seleção e (3) o número de questões finais a serem retornadas. É importante destacar que o parâmetro de dificuldade b_j representa a dificuldade do item na escala TRI, enquanto θ representa a proficiência do estudante na mesma escala. O sistema busca itens onde $b_j \approx \theta$ para maximizar a informação psicométrica, pois questões cuja dificuldade está próxima da proficiência do estudante são mais informativas para a estimação.

Busca de Questões: A Fig.1 (linhas [5:19]) busca questões em intervalos crescentes ao redor da proficiência. A função `filtrar_questoes_por_bj` retorna questões cujo parâmetro de dificuldade b_j está dentro do intervalo e que ainda não foram respondidas. O processo para quando encontra pelo menos o número mínimo de candidatas necessário.

Seleção de Questões: A Fig.1 (linhas [21:26]) seleciona as questões finais ordenando por discriminação (a_j). Questões com maior discriminação são mais informativas e permitem estimar a proficiência com maior precisão. Se houver poucas candidatas (menor ou igual a `min_candidatas`), ordena todas por discriminação e escolhe as `questoes_finais` melhores. Se houver muitas candidatas, primeiro seleciona uma amostra aleatória de `min_candidatas` (garantindo diversidade) e depois escolhe as

```

1 # Entrada: theta (proficiência), area_conhecimento, intervalos_busca,
2 #          min_candidatas, questoes_finais
3 # Saída:   questoes_finais questões recomendadas
4
5 # 1. Buscar questões em intervalos crescentes ao redor de theta
6 questoes_candidatas = []
7
8 PARA CADA d EM intervalos_busca:
9     questoes = filtrar_questoes_por_bj(area_conhecimento, theta - d, theta + d)
10    questoes_candidatas = questoes_candidatas + questoes
11    SE len(questoes_candidatas) >= min_candidatas:
12        INTERROMPER
13    FIM SE
14 FIM PARA
15
16 # 1.1. Se necessário, buscar todas as questões disponíveis
17 SE len(questoes_candidatas) < min_candidatas:
18     questoes_candidatas = filtrar_questoes_por_bj(area_conhecimento)
19 FIM SE
20
21 # 2. Selecionar questoes_finais melhores por discriminação (a_j)
22 SE len(questoes_candidatas) <= min_candidatas:
23     questoes_selecionadas = ORDENAR(questoes_candidatas, POR a_j
24     ↪ DESC)[:questoes_finais]
25 SENÃO:
26     amostra = ALEATORIO(questoes_candidatas, min_candidatas)
27     questoes_selecionadas = ORDENAR(amostra, POR a_j DESC)[:questoes_finais]
28 FIM SE
29 RETORNAR questoes_selecionadas

```

Figura 1. Algoritmo de seleção de questões adaptativa

questoes_finais com maior discriminação dessa amostra, equilibrando diversidade educacional com qualidade psicométrica.

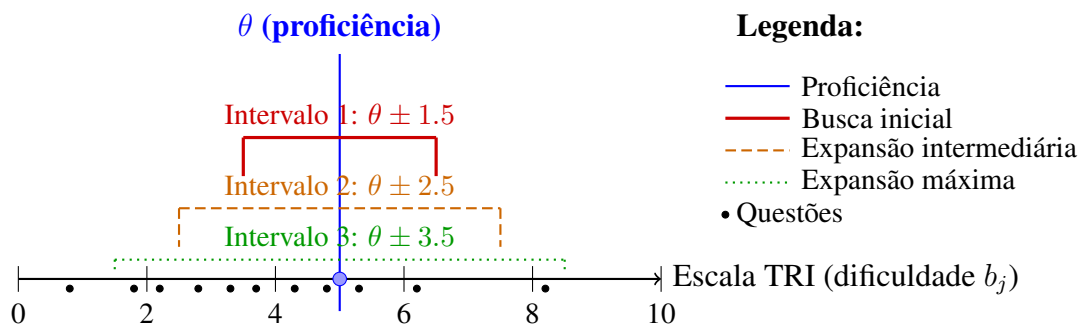


Figura 2. Seleção de questões em intervalos crescentes ao redor da proficiência
 θ (valores de exemplo: $d_1 = 1.5$, $d_2 = 2.5$, $d_3 = 3.5$)

4.3.3. Atualização da Proficiência

A atualização da proficiência do estudante ocorre após a submissão das questões diárias, incorporando dois elementos fundamentais para cada questão: a resposta binária (acerto ou erro) e o nível de conforto auto-reportado pelo estudante em uma escala de 1 a 5, onde 1 representa “muito inseguro” e 5 representa “muito confiante”. O processo de atualização é realizado por meio de três etapas sequenciais que integram os dados das

respostas com os parâmetros psicométricos das questões, utilizando o método de Máxima Verossimilhança (MLE) que incorpora o nível de conforto do estudante. A Figura 3 apresenta o fluxo conceitual completo deste processo.

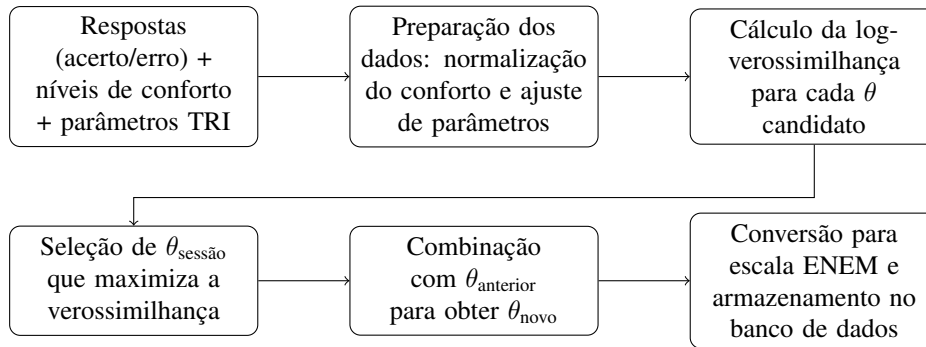


Figura 3. Fluxo conceitual da atualização de proficiência no sistema

A Figura 4 apresenta o algoritmo de atualização de proficiência. O algoritmo utiliza as respostas do estudante (acertos e erros), os níveis de conforto reportados (escala de 1 a 5) e os parâmetros TRI das questões respondidas para estimar uma nova proficiência. O processo busca, de forma sistemática, o valor de proficiência que melhor explica o padrão de respostas observado, considerando tanto a correção das respostas quanto o nível de confiança do estudante em cada questão.

Busca da Proficiência da Sessão: O algoritmo testa diferentes valores de proficiência e seleciona aquele que melhor explica as respostas do estudante. Para cada questão respondida, o sistema realiza três ajustes principais. Primeiro, o nível de conforto reportado pelo estudante é normalizado para uma escala entre 0 e 1. Segundo, o parâmetro de acerto casual é ajustado: quando o estudante reporta baixo conforto, o sistema assume maior probabilidade de acerto ao acaso, refletindo a incerteza na resposta. Terceiro, uma estimativa ponderada é construída combinando a resposta real (acerto ou erro) com o nível de conforto: acertos com baixo conforto têm menor peso, pois podem ser acertos ao acaso, enquanto erros com alto conforto são tratados de forma mais conservadora, indicando que o estudante realmente não dominava o conteúdo. O valor de proficiência que melhor explica todo o conjunto de respostas, considerando esses ajustes, é selecionado como a proficiência da sessão.

Para ilustrar o funcionamento do algoritmo, considere o seguinte exemplo prático. Um estudante com proficiência anterior $\theta_{\text{anterior}} = 0.5$ (escala TRI) respondeu 5 questões em uma sessão. A Tabela 1 apresenta os dados de entrada: respostas (acerto/erro), níveis de conforto reportados e parâmetros TRI de cada questão.

A Tabela 2 detalha o processamento para a Questão 1 como exemplo. O conforto é normalizado ($c_{01} = 0.75$), o acerto casual é ajustado considerando a incerteza ($g_1^{\text{ajustado}} = 0.30$), e a estimativa ponderada é construída ($t_1 = 0.91$), resultando em uma contribuição positiva para a verossimilhança, mas com peso reduzido devido ao conforto moderado. Para a Questão 2, que foi errada com baixo conforto ($c_{02} = 0.25$), o acerto casual é aumentado

```

1 # Entrada: respostas (y_j), niveis_conforto (c_j), parâmetros TRI (a_j, b_j, c_j),
2 #          theta_anterior
3 # Saída:   theta_novo (proficiência atualizada)
4
5 # 1. Buscar theta que maximiza verossimilhança total
6 theta_sessao = None
7 max_verossimilhanca = -INFINITO
8
9 PARA theta EM [-4.0, 4.0] (com passo 0.01):
10     L_total = 0
11     PARA CADA questão j EM questoes_respondidas:
12         # Normalizar conforto para [0, 1]
13         c_01 = (c_j - 1) / 4
14
15         # Ajustar acerto casual baseado no conforto
16         g_j_ajustado = AJUSTAR_ACERTO_CASUAL(c_j, c_01)
17
18         # Calcular probabilidade usando modelo 3PL (discriminação mantém valor
19         ↪ original)
20         P_j = CALCULAR_3PL(theta, a_j, b_j, g_j_ajustado)
21
22         # Construir estimativa ponderada pela percepção de competência combinando
23         ↪ resposta real e conforto
24         t_j = CONSTRUIR_ALVO_SUAVE(y_j, c_01)
25
26         # Calcular log-verossimilhança
27         log_L_j = t_j * LOG(P_j) + (1 - t_j) * LOG(1 - P_j)
28         L_total = L_total + log_L_j
29     FIM PARA
30
31     SE L_total > max_verossimilhanca:
32         max_verossimilhanca = L_total
33         theta_sessao = theta
34     FIM SE
35 FIM PARA
36
37 # 2. Combinar com proficiência anterior (média simples)
38 theta_novo = (theta_anterior + theta_sessao) / 2
39
40 # 3. Converter para escala ENEM e armazenar
41 ARMAZENAR(CONVERTER_PARA_ESCALA_ENEM(theta_novo, area_conhecimento))
42
43 RETORNAR theta_novo

```

Figura 4. Algoritmo de atualização de proficiência

($g_2^{ajustado} = 0.35$), refletindo maior incerteza, e a estimativa ponderada ($t_2 = 0.24$) indica que o erro tem menor impacto na redução da proficiência estimada.

O algoritmo testa valores de θ no intervalo $[-4.0, 4.0]$ e encontra $\theta_{sessao} = 0.62$ como o valor que maximiza a verossimilhança total considerando todas as 5 questões. Este valor reflete que o estudante acertou 3 questões (incluindo uma com alto conforto) e errou 2 (uma com baixo conforto, indicando possível acerto ao acaso).

Combinação com Proficiência Anterior: A nova proficiência é calculada como a média simples entre a proficiência anterior e a proficiência estimada na sessão atual: $\theta_{novo} = (0.5 + 0.62)/2 = 0.56$. Esta abordagem garante estabilidade ao sistema, evitando variações bruscas que podem ocorrer devido a fatores temporais como cansaço, distrações ou estados emocionais momentâneos. Ao combinar igualmente a proficiência histórica com a estimativa da sessão atual, o sistema mantém uma evolução gradual, suavizando flutuações temporais enquanto ainda permite que melhorias consistentes no desempenho

Tabela 1. Exemplo prático: dados de entrada para atualização de proficiência

Questão	Resposta	Conforto	a_j	b_j	c_j
1	Acerto	4	1.2	0.3	0.25
2	Erro	2	0.9	0.8	0.20
3	Acerto	5	1.5	0.5	0.15
4	Acerto	3	1.0	0.6	0.22
5	Erro	1	0.8	1.2	0.18

O estudante acertou 3 questões (questões 1, 3 e 4) e errou 2 (questões 2 e 5). Os níveis de conforto variam de 1 (muito inseguro) a 5 (muito confiante). As questões apresentam diferentes níveis de dificuldade (b_j), discriminação (a_j) e acerto casual (c_j), representando a diversidade do banco de questões do ENEM.

Tabela 2. Exemplo prático: processamento detalhado para duas questões

Etapa	Q1 (Acerto, Conforto=4)	Q2 (Erro, Conforto=2)
Conforto normalizado (c_{01})	0.75	0.25
Acerto casual ajustado ($g_j^{ajustado}$)	0.30	0.35
Probabilidade 3PL (P_j para $\theta = 0.5$)	0.78	0.42
Função $r(c_{01})$	0.70	0.30
Estimativa ponderada (t_j)	0.91	0.24
Log-verossimilhança ($\log L_j$)	-0.25	-0.18

sejam incorporadas ao perfil do estudante. A proficiência final é convertida para a escala ENEM: $\theta_{\text{ENEM}} = 100 \cdot 0.56 + 500 = 556$.

Armazenamento: A Fig.4 (linha [39]) converte a proficiência atualizada para a escala ENEM (500,100) e armazena no banco de dados, atualizando o perfil do estudante na área de conhecimento correspondente. Mais detalhes sobre a implementação completa do algoritmo podem ser encontrados no repositório do código².

4.4. Configuração dos Parâmetros do Algoritmo

No presente estudo, os valores dos parâmetros dos algoritmos utilizados foram os seguintes. Para o algoritmo de seleção adaptativa: `intervalos_busca` = [1.5, 2.5, 3.5], `min_candidatas` = 10 e `questoes_finais` = 5. Estes valores foram escolhidos com base na distribuição real dos parâmetros de dificuldade das questões do ENEM disponíveis no banco de dados, garantindo que o sistema encontre questões adequadas para estudantes em diferentes níveis de proficiência.

Para o algoritmo de atualização de proficiência, a busca pela proficiência que maximiza a verossimilhança foi realizada no intervalo $[-4.0, 4.0]$ com incrementos de 0.01, valores escolhidos para balancear precisão e eficiência computacional, cobrindo adequadamente a distribuição de proficiências na escala TRI padrão. As funções genéricas utilizadas para incorporar o nível de conforto no cálculo foram implementadas com os seguintes valores específicos: a função `AJUSTAR_ACERTO_CASUAL` utiliza a fórmula $g_j^{ajustado} = \min(0.4, c_j + 0.2 \cdot (1 - c_{01}))$, onde 0.4 é o limite máximo e 0.2 é o coeficiente de ajuste, aumentando a probabilidade de acerto casual quando o conforto é baixo, refletindo maior incerteza na resposta; a função `CONSTRUIR_ALVO_SUAVE` utiliza a

²Disponível em <https://github.com/joaocolem/meu-enem-me>

fórmula $t_j = 0.7 \cdot y_j + 0.3 \cdot r(c_{01})$, onde 0.7 é o peso da resposta real, 0.3 é o peso do conforto, e $r(c_{01}) = 0.1 + 0.8 \cdot c_{01}$ mapeia o conforto normalizado $c_{01} \in [0, 1]$ de forma linear para o intervalo $[0.1, 0.9]$ para evitar valores extremos. Os parâmetros de dificuldade e discriminação são mantidos com seus valores originais, preservando as propriedades psicométricas calibradas das questões. Estes valores foram definidos com base em testes iniciais e na necessidade de incorporar a percepção de competência do estudante de forma moderada, evitando distorções nos parâmetros psicométricos e garantindo estabilidade numérica nos cálculos.

5. Resultados e Discussão

5.1. Resultados

Os resultados da avaliação com usuários revelaram percepções positivas em todas as dimensões analisadas. A Tabela 3 apresenta um resumo das estatísticas descritivas para cada dimensão avaliada.

Tabela 3. Estatísticas Descritivas das Dimensões Avaliadas (n=27)

Dimensão	Média	D.P.	Mediana
Usabilidade	4,51	0,53	4,67
Percepção de Adaptação e Dificuldade	3,80	0,78	4,00
Motivação e Engajamento	4,35	0,67	4,50
Recomendação	4,78	0,70	5,00

Aproximadamente 77,8% dos participantes (21 estudantes) classificaram a usabilidade como “Boa” ou “Excelente” (escalas 4 e 5). A percepção de adaptação apresentou 59,3% de respostas positivas (16 estudantes), enquanto o engajamento foi avaliado positivamente por 77,8% dos participantes (21 estudantes).

Destaca-se a alta taxa de recomendação: 92,6% dos participantes (25 estudantes) recomendariam a plataforma para outros estudantes, com média de 4,78 na escala de 1 a 5.

5.1.1. Análise Qualitativa

A análise de conteúdo das respostas abertas revelou quatro temas principais: funcionalidade e variedade de questões (11 menções), revisão e análise de erros (8 menções), metodologia alinhada ao ENEM (8 menções), e adaptação ao nível individual (6 menções). Quanto às melhorias sugeridas, destacaram-se a necessidade de resolução e explicação de questões (3 menções), problemas técnicos (3 menções) e melhorias na usabilidade mobile (2 menções).

5.2. Discussão

Os resultados obtidos fornecem evidências preliminares de que a aplicação rigorosa dos princípios psicométricos da TRI em sistemas de recomendação educacionais pode efetivamente melhorar a experiência de estudo personalizado. A validação parcial da hipótese central através da percepção positiva dos usuários sugere que a abordagem proposta atende a uma necessidade real dos estudantes: receber questões que estejam adequadas ao seu nível de conhecimento, criando um desafio progressivo e motivador. A correlação positiva moderada-alta entre a percepção de adaptação/nível e o engajamento

($r = 0,684$) é particularmente significativa, pois indica que quando os estudantes percebem que as questões estão adequadas ao seu nível, eles se sentem mais engajados com a atividade. Esta relação reforça a importância do algoritmo adaptativo em promover uma experiência que seja simultaneamente desafiadora e acessível, alinhando-se ao princípio da zona de desenvolvimento proximal de Vygotsky. A correlação moderada entre usabilidade e engajamento ($r = 0,623$) sugere que a facilidade de uso da plataforma contribui para manter os estudantes envolvidos, enquanto a correlação entre engajamento e recomendação ($r = 0,462$) indica que estudantes mais engajados são mais propensos a recomendar a plataforma para outros, criando um ciclo positivo de adoção.

A alta taxa de recomendação (92,6%, média de 4,78) e a elevada intenção de uso regular (77,8%) demonstram que os estudantes reconhecem valor prático na abordagem proposta. Este resultado é especialmente relevante considerando que os participantes eram estudantes do primeiro ano do Ensino Médio, que tradicionalmente podem sentir-se intimidados pelo ENEM. A capacidade do sistema de reduzir essa ansiedade inicial, apresentando questões em níveis apropriados, sugere que a personalização baseada em TRI pode ter um impacto positivo na motivação e na percepção de autoeficácia dos estudantes. A percepção de adaptação (média de 3,80) foi a dimensão com menor média entre as avaliadas, embora ainda positiva. Este resultado pode refletir a complexidade de comunicar aos estudantes como o sistema está se adaptando ao seu perfil, ou pode indicar que alguns estudantes ainda não experimentaram suficientemente a adaptação em uma única sessão. Estudos longitudinais permitiriam avaliar se essa percepção melhora com o uso contínuo, à medida que o sistema acumula mais dados sobre o desempenho do estudante.

6. Limitações e Ameaças à Validade

Este estudo apresenta limitações que devem ser reconhecidas para contextualizar adequadamente os resultados e orientar pesquisas futuras. Em primeiro lugar, a amostra utilizada foi de conveniência, composta por 27 estudantes do primeiro ano do Ensino Médio de uma única instituição, o que limita a generalização dos resultados para outros contextos educacionais, demográficos e níveis de escolaridade. Resultados mais amplos e robustos exigem testes com grupos mais diversos, incluindo estudantes de diferentes anos do Ensino Médio, contextos socioeconômicos variados e diferentes regiões do país.

A avaliação transversal, com uma única sessão de uso, impede avaliar como a adaptação do sistema evolui ao longo do tempo e como a percepção dos estudantes muda com o uso contínuo. Estudos longitudinais permitiriam investigar a estabilidade das estimativas de proficiência e o impacto a longo prazo na motivação e no aprendizado. A ausência de grupo de controle ou comparação sistemática com métodos alternativos de recomendação dificulta avaliar a superioridade relativa da abordagem baseada em TRI em relação a outras estratégias de recomendação educacional.

Por fim, o foco exclusivo em medidas de percepção subjetiva, sem incorporar métricas objetivas de aprendizado ou desempenho em avaliações externas, limita a capacidade de inferir sobre o impacto educacional real do sistema, embora forneça insights valiosos sobre a aceitação e a experiência do usuário. Futuras pesquisas deveriam investigar a correlação entre as proficiências estimadas pelo sistema e o desempenho em simulados oficiais do ENEM, bem como avaliar ganhos de aprendizado mensuráveis através de avaliações padronizadas.

7. Conclusão

A integração rigorosa da Teoria de Resposta ao Item com sistemas de recomendação educacionais representa uma fronteira promissora para a personalização do aprendizado, especialmente em contextos de avaliação em larga escala como o ENEM. Este trabalho demonstrou que é viável e eficaz aplicar princípios psicométricos consolidados em sistemas adaptativos de recomendação de questões, criando uma ponte entre a precisão metodológica da TRI e as necessidades práticas de estudantes que buscam uma preparação progressiva e motivadora. A construção do protótipo revelou desafios técnicos significativos, particularmente na integração de microdados oficiais do INEP com textos e alternativas de questões, e na implementação de algoritmos que preservem as propriedades psicométricas dos itens enquanto incorporam informações subjetivas do estudante. A solução desenvolvida, que combina busca em intervalos crescentes, seleção por discriminação e atualização de proficiência com estimativa ponderada pela percepção de competência, representa uma abordagem inovadora para equilibrar rigor psicométrico e experiência do usuário.

Os resultados empíricos obtidos com estudantes do primeiro ano do Ensino Médio sugerem que a abordagem proposta atende a uma necessidade real: a de apresentar o ENEM como um desafio progressivo e alcançável, não como uma barreira intransponível. A alta taxa de recomendação (92,6%) e o elevado engajamento (média de 4,35) indicam que os estudantes reconhecem valor na personalização baseada em TRI, validando a hipótese de que sistemas adaptativos fundamentados em princípios psicométricos podem aumentar a percepção de progresso e a motivação. As implicações deste trabalho estendem-se além do contexto específico do ENEM. A metodologia desenvolvida pode ser adaptada para outros sistemas de avaliação em larga escala que utilizam TRI, e os princípios de integração entre dados oficiais e sistemas de recomendação podem informar o desenvolvimento de ferramentas educacionais mais amplas. O banco de dados construído, com questões categorizadas e parâmetros TRI completos, constitui um recurso valioso para pesquisas futuras em educação e psicometria.

Considerando as limitações identificadas, pesquisas futuras deveriam focar em estudos longitudinais que acompanhem estudantes ao longo de múltiplos anos do Ensino Médio, permitindo avaliar não apenas a percepção imediata, mas o impacto real no aprendizado e no desempenho em avaliações externas. Tais estudos poderiam investigar como a trajetória de proficiência estimada pelo sistema se correlaciona com o desempenho em simulados oficiais do ENEM e se o uso contínuo resulta em ganhos de aprendizado mensuráveis, comparando grupos que utilizam a plataforma com grupos de controle. A validação psicométrica dos parâmetros TRI especificamente para estudantes em diferentes etapas do Ensino Médio constituiria uma contribuição fundamental, considerando que os parâmetros atuais refletem o desempenho de concluintes. Além disso, comparações sistemáticas com métodos alternativos de recomendação, incluindo abordagens baseadas em filtragem colaborativa ou aprendizado de máquina, permitiriam uma avaliação mais precisa da superioridade relativa da abordagem baseada em TRI.

Este trabalho contribui para o campo ao demonstrar que é possível criar sistemas educacionais que sejam simultaneamente rigorosos do ponto de vista metodológico e efetivos do ponto de vista pedagógico, oferecendo uma alternativa às plataformas comerciais que frequentemente simplificam ou distorcem os princípios da TRI. A integração

bem-sucedida entre psicometria e tecnologia educacional abre perspectivas para o desenvolvimento de ferramentas que realmente personalizem o aprendizado, respeitando tanto a complexidade do conhecimento quanto a individualidade do estudante.

Referências

- Andrade, D. F., Tavares, H. R., and Valle, R. C. (2000). *Teoria da Resposta ao Item: Conceitos e Aplicações*. Associação Brasileira de Estatística.
- Andrade, F. H. S. (2020). *Teoria da resposta ao item: influência do tamanho da amostra na estimação dos parâmetros dos itens utilizando os microdados do Enem*. PhD thesis, Universidade de São Paulo. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55136/tde-21092020-170142/en.php>. Acesso em: 28 outubro 2025.
- Andrade, J. M., Laros, J. A., and Gouveia, V. V. (2010). O uso da teoria de resposta ao item em avaliações educacionais: diretrizes para pesquisadores. *Avaliação Psicológica*, 9(3):505–517. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000300009. Acesso em: 28 outubro 2025.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F. M. and Novick, M. R., editors, *Statistical theories of mental test scores*, pages 397–479. Addison-Wesley.
- Braida, H. M. and Rezende, H. V. A. (2025). Simulab: Sistema web para geração de simulados do enem. Master's thesis, Instituto Federal de Educação, Ciência e Tecnologia. Disponível em: https://filipe.braida.com.br/monografias/braida_hugo_maxwel.pdf. Acesso em: 28 outubro 2025.
- Chen, Y. (2017). Recommendation system for adaptive learning. *Journal of Educational Computing Research*, 56(2):159–176. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5766274/>. Acesso em: 28 outubro 2025.
- Corrêa, L. R. C., Saldanha, L. C. D., Lemos, M. O., and Soares, M. C. (2023). Desenvolvimento de plataforma inteligente de avaliação, com aplicação de técnica de ia e utilização de tri. In *Anais do Congresso Brasileiro de Informática na Educação*. Disponível em: https://www.researchgate.net/publication/392368533_DESENVOLVIMENTO_DE_PLATAFORMA_INTELIGENTE_DE_AVALIACAO_COM_APLICACAO_DE_TECNICA_DE_IA_E_UTILIZACAO_DE_TRI. Acesso em: 28 outubro 2025.
- Estuda.com (s.d.). Tri enem: o que é e como afeta sua nota? Disponível em: <https://estuda.com/o-que-e-tri/>. Acesso em: 28 outubro 2025.
- INEP (2021a). Entenda a sua nota no enem: guia do participante. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/guia_participante_enem_2021.pdf. Acesso em: 28 outubro 2025.
- INEP (2021b). Exame nacional do ensino médio – enem: procedimentos de análise. Technical report, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira,

- Brasília, DF. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/procedimentos_de_analise_enem.pdf. Acesso em: 28 outubro 2025.
- INEP (2024a). Estatísticas do enem. Portal do Inep. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>. Acesso em: 28 outubro 2025.
- INEP (2024b). Microdados do enem. Portal do Inep. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>. Acesso em: 28 outubro 2025.
- Silva, F. L., Silva, K. K. A., and Slodkowski, B. K. (2022). A aplicação de sistemas de recomendação no contexto educacional: uma revisão sistemática da literatura. *Revista de Tecnología de la Información y Comunicación en Educación*, 16(2):11–25. Disponível em: https://www.scielo.org.ar/scielo.php?pid=S1850-99592022000200011&script=sci_arttext&tlng=pt. Acesso em: 28 outubro 2025.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA. Tradução de: Cole, M., John-Steiner, V., Scribner, S., & Souberman, E.
- Yunger (2024). Criando uma api open-source para questões do enem. TabNews. Disponível em: <https://www.tabnews.com.br/yunger/criando-uma-api-open-source-para-questoes-do-enem>. Acesso em: 28 outubro 2025.