



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS VII - PATOS
CENTRO DE CIÊNCIAS EXATAS E SOCIAIS APLICADAS
CURSO DE GRADUAÇÃO EM BACHARELADO EM CIÊNCIA DA
COMPUTAÇÃO**

DAVI MIGUEL FRANKLIM DE ALMEIDA

***SCREENING PROGRAMMING: APLICAÇÃO WEB PARA O
DESENVOLVIMENTO DAS HABILIDADES DE RACIOCÍNIO
LÓGICO-MATEMÁTICO***

**PATOS - PB
2024**

DAVI MIGUEL FRANKLIM DE ALMEIDA

***SCREENING PROGRAMMING: APLICAÇÃO WEB PARA O
DESENVOLVIMENTO DAS HABILIDADES DE RACIOCÍNIO
LÓGICO-MATEMÁTICO***

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Computação do Centro de Ciências Exatas e Sociais Aplicadas da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Ciência da Computação.

Orientador: Dr. Jucelio Soares dos Santos

PATOS - PB

2024

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

A447s Almeida, Davi Miguel Franklim de.

Screening programming [manuscrito] : aplicação *web* para o desenvolvimento das habilidades de raciocínio lógico-matemático / Davi Miguel Franklim de Almeida. - 2024.

59 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Ciência da computação) - Universidade Estadual da Paraíba, Centro de Ciências Exatas e Sociais Aplicadas, 2024.

"Orientação : Prof. Dr. Jucelio Soares dos Santos, Coordenação do Curso de Computação - CCEA".

1. Raciocínio lógico - matemático. 2. Teoria de resposta ao item. 3. Teoria clássica dos testes. 4. Confiabilidade. I. Título

21. ed. CDD 004

DAVI MIGUEL FRANKLIM DE ALMEIDA

SCREENING PROGRAMMING: APLICAÇÃO WEB PARA O
DESENVOLVIMENTO DAS HABILIDADES DE RACIOCÍNIO LÓGICO-
MATEMÁTICO

Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso
de Ciência da Computação da
Universidade Estadual da Paraíba,
como requisito parcial à obtenção do
título de Bacharel em Ciência da
Computação

Aprovada em: 18/11/2024.

Documento assinado eletronicamente por:

- **Jucelio Soares dos Santos** (**.475.114-**), em 26/11/2024 07:26:39 com chave **ec42e20cabe011efba512618257239a1**.
- **Vinicius Augustus Alves Gomes** (**.754.334-**), em 26/11/2024 07:54:40 com chave **d6057d2aabe411efb8d306adb0a3afce**.
- **Thiago Alves Laurentino** (**.243.404-**), em 26/11/2024 12:28:09 com chave **0abc1616ac0b11efa7b906adb0a3afce**.

Documento emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QrCode ao lado ou acesse https://suap.uepb.edu.br/comum/autenticar_documento/ e informe os dados a seguir.

Tipo de Documento: Termo de Aprovação de Projeto Final

Data da Emissão: 05/12/2024

Código de Autenticação: 4a3956



*Dedico este trabalho a Deus, minha fonte
de inspiração; à minha amada
família—minha mãe Fátima, meu pai
Francisco, meu irmão Vitor e minha avó
Martina—que sempre me apoiaram
incondicionalmente. Agradeço também aos
meus amigos e à minha namorada, Larissa,
por estarem sempre ao meu lado, enchendo
minha jornada de amor e encorajamento.*

AGRADECIMENTOS

Gostaria de expressar minha sincera gratidão ao meu orientador, Jucelio, por ter sido o melhor orientador que eu poderia ter. Sua orientação, paciência e apoio foram fundamentais em cada etapa deste trabalho.

Agradeço também aos meus colegas e amigos, que me proporcionaram momentos de alegria e motivação, além de apoio nas horas difíceis.

Um agradecimento especial à minha família, que sempre acreditou em mim e me incentivou a seguir em frente, mesmo nos momentos mais desafiadores

Por fim, agradeço a todos que, de alguma forma, contribuíram para a realização deste trabalho. Sua presença e apoio foram essenciais nesta jornada.

"Você tem que lutar para alcançar seus sonhos. Você tem que sacrificar e trabalhar duro para isso."

Lionel Messi

RESUMO

Este estudo investiga e valida o desenvolvimento de um módulo do *Screening Programming*, projetado especificamente para mensurar habilidades de raciocínio lógico e matemático em estudantes de nível superior. Essas habilidades, essenciais para o sucesso acadêmico e profissional, são especialmente importantes em cursos de exatas, onde a resolução de problemas complexos e a realização de análises dedutivas são constantemente exigidas. Para assegurar a confiabilidade e a validade do módulo *Screening Programming* na avaliação dessas habilidades, aplicamos as abordagens da Teoria Clássica dos Testes (TCT) e da Teoria de Resposta ao Item (TRI). A TCT permitiu medir a consistência interna dos itens do teste por meio do coeficiente alfa de Cronbach, um indicador amplamente aceito de confiabilidade, cujos altos valores sugerem uma forte correlação entre os itens, indicando que eles medem de forma coerente as habilidades propostas. Complementando a análise, aplicamos a TRI para explorar as propriedades psicométricas dos itens, avaliando parâmetros fundamentais como a dificuldade, a discriminação e a probabilidade de acerto ao acaso. A dificuldade dos itens reflete o nível de habilidade necessário para respondê-los corretamente, enquanto a discriminação indica a capacidade dos itens de distinguir entre alunos com diferentes níveis de competência. Além disso, incluímos o parâmetro de acerto ao acaso para estimar a chance de um aluno de baixa habilidade acertar uma questão por adivinhação, garantindo que os itens estejam bem estruturados para minimizar esse efeito. Os resultados confirmam a validade e a precisão do instrumento para uso em contextos acadêmicos de nível superior, destacando sua aplicabilidade em avaliações de larga escala e em diferentes instituições. A análise psicométrica mostrou que os itens têm boa capacidade de discriminação e abrangem uma ampla gama de níveis de dificuldade, o que permite ao instrumento avaliar tanto alunos com habilidades avançadas quanto aqueles em fases iniciais de desenvolvimento. A alta consistência interna dos itens reforça a confiabilidade do instrumento, tornando-o uma ferramenta promissora para mensurar habilidades de raciocínio lógico e matemático de maneira precisa e adaptativa.

Palavras-chave: Raciocínio Lógico-Matemático; Teoria de Resposta ao Item; Teoria Clássica dos Testes; Confiabilidade.

ABSTRACT

This study investigates and validates the development of a Screening Programming module specifically designed to measure logical and mathematical reasoning skills in undergraduate students. These skills, essential for academic and professional success, are particularly important in science-related courses, where complex problem-solving and deductive analysis are constantly required. To ensure the reliability and validity of the Screening Programming module in assessing these skills, we applied the Classical Test Theory (CTT) and Item Response Theory (IRT) approaches. CTT allowed us to measure the internal consistency of the test items using Cronbach's alpha coefficient, a widely accepted reliability indicator, whose high values suggest a strong correlation between items, indicating that they consistently measure the proposed skills. To complement this analysis, we applied IRT to explore the psychometric properties of the items, evaluating key parameters such as difficulty, discrimination, and the probability of guessing. Item difficulty reflects the skill level required to answer correctly, while discrimination indicates the ability of the items to differentiate between students with varying levels of competence. Additionally, we included the guessing parameter to estimate the chance of a low-skill student answering a question correctly by mere guessing, ensuring that the items are well-structured to minimize this effect. The results confirm the instrument's validity and precision for use in higher education contexts, highlighting its applicability in large-scale assessments across different institutions. The psychometric analysis showed that the items have good discrimination ability and cover a wide range of difficulty levels, enabling the instrument to assess both advanced students and those in the early stages of skill development. The high internal consistency of the items further strengthens the instrument's reliability, making it a promising tool for accurately and adaptively measuring logical and mathematical reasoning skills.

Keywords: Logical-Mathematical Reasoning; Item Response Theory; Classical Test Theory; Reliability.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de CCI's para diferentes itens.	26
Figura 2 – Exemplo de FII's para diferentes itens.	29
Figura 3 – Modelo de casos de uso.	36
Figura 4 – Modelo de Arquitetura.	38
Figura 5 – Tela de cadastro	39
Figura 6 – Tela de login	40
Figura 7 – Tela de pós-autenticação	41
Figura 8 – Tela de convite	42
Figura 9 – Tela do convite via email	42
Figura 10 – CCI - Todos os itens para habilidade Raciocínio Lógico e Matemático.	50
Figura 11 – FII - Todos os itens para habilidade Raciocínio Lógico e Matemático.	51

LISTA DE TABELAS

Tabela 1 – Análise Psicométrica dos Itens com Base na TCT.	47
Tabela 2 – Parâmetros Psicométricos dos Itens com Base na TRI.	49

LISTA DE ABREVIATURAS E SIGLAS

3PL	Modelo Logístico de Três Parâmetros
CCI	Curva Característica do Item
EAP	<i>Expected A Posteriori</i>
FII	Função de Informação do Item
GUI	Interface Gráfica do Usuário
TCC	Trabalho de Conclusão de Curso
TCL	Termo de Consentimento Livre e Esclarecido
TCT	Teoria Clássica dos Testes
TRI	Teoria da Resposta ao Item
UFCG	Universidade Federal de Campina Grande

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Contextualização	13
1.2	Problema e Proposta de Solução	15
1.3	Objetivos	16
1.4	Questões de Pesquisa	17
1.5	Estrutura do trabalho	18
2	REFERENCIAL TEÓRICO	19
2.1	Raciocínio Lógico-Matemático	19
2.1.1	<i>Definição e Importância</i>	<i>19</i>
2.1.2	<i>Desenvolvimento Cognitivo</i>	<i>21</i>
2.2	Teorias Psicométricas	22
2.2.1	<i>Teoria Clássica dos Testes</i>	<i>22</i>
2.2.1.1	<i>Coeficiente de Ponto Bisserial</i>	<i>23</i>
2.2.1.2	<i>Coeficiente Alpha de Cronbach</i>	<i>24</i>
2.2.1.3	<i>Limitações da TCT</i>	<i>24</i>
2.2.2	<i>Teoria de Resposta ao Item</i>	<i>24</i>
2.2.2.1	<i>Curva Característica do Item</i>	<i>25</i>
2.2.2.2	<i>Estimativa da Habilidade</i>	<i>27</i>
2.2.2.3	<i>Função de Informação do Item</i>	<i>28</i>
3	SCREENING PROGRAMMING	30
3.1	Análise de Requisitos	30
3.1.1	<i>Requisitos Funcionais</i>	<i>30</i>
3.1.2	<i>Requisitos Não-Funcionais</i>	<i>32</i>
3.2	Modelo de Análise	34
3.2.1	<i>Perfil do Usuário</i>	<i>34</i>
3.2.2	<i>Diagrama de Caso de Uso</i>	<i>36</i>
3.2.3	<i>Tecnologias de Suporte</i>	<i>37</i>
3.3	Modelo de Arquitetura	38
3.4	Modelo de Projeto	39

4	METODOLOGIA	43
4.1	Participantes	43
4.2	Instrumento	43
4.3	Procedimentos de Coleta de Dados	44
4.4	Análise de Dados	45
5	RESULTADOS E DISCUSSÕES	46
5.1	Confiabilidade do Instrumento na Mensuração das Habilidades de Raciocínio Lógico-Matemático	46
5.2	Validade e Propriedades Psicométricas dos Itens: Indicadores de Habilidades de Raciocínio Lógico-Matemático	48
6	CONSIDERAÇÕES FINAIS E SUGESTÕES PARA TRABALHOS FUTUROS	52
	REFERÊNCIAS	54

1 INTRODUÇÃO

Neste Capítulo, apresenta-se uma visão geral desta pesquisa, de modo a descrever a contextualização do problema, objetivos e questões de pesquisas.

1.1 Contextualização

O desenvolvimento de habilidades de raciocínio lógico-matemático é amplamente reconhecido como essencial para o desempenho acadêmico e sucesso profissional em áreas que demandam análise crítica, resolução de problemas e tomada de decisão informada, como a Ciência da Computação, a Engenharia e a Matemática (Holovina; Kamenova, 2021; Nurlaelah; Usdiyana; Fadilah, 2024). O raciocínio lógico-matemático envolve a capacidade de análise e síntese, assim como habilidades de dedução e inferência que são fundamentais para que o indivíduo possa lidar com problemas complexos de maneira estruturada e eficiente (Nurlaelah; Usdiyana; Fadilah, 2024). Essas competências permitem que o indivíduo construa esquemas mentais que facilitam a interpretação e manipulação de informações abstratas, aspectos essenciais para a construção de conhecimento em contextos variados (Mitsea; Drigas; Mantas, 2021).

O desenvolvimento dessas habilidades não é imediato, trata-se de um processo gradual e cumulativo que se constrói através de experiências concretas e abstratas (Olteanu, 2022). As interações com problemas estruturados e situações que desafiam a lógica fomentam a construção de esquemas de pensamento que, ao longo do tempo, capacitam o indivíduo a realizar operações mentais mais sofisticadas (Eppe et al., 2022). Essas habilidades cognitivas são indispensáveis para que o indivíduo possa interpretar e resolver problemas de forma autônoma e criativa, características fundamentais para a formação de um pensamento crítico e a promoção da autonomia intelectual (Hendrycks et al., 2021).

Nos últimos anos, o avanço das tecnologias educacionais trouxe novas possibilidades para a avaliação e o desenvolvimento do raciocínio lógico-matemático (Reyes, 2023; Tigrero et al., 2024). Ferramentas digitais, como plataformas de aprendizado e aplicativos interativos, oferecem uma alternativa para a aplicação de testes e avaliações de maneira prática, acessível e escalável (Sillat; Tammets; Laanpere, 2021). Tais tecnologias permitem que educadores e pesquisadores acessem dados sobre o desempenho dos

alunos em tempo real, facilitando a identificação de lacunas no aprendizado e permitindo intervenções pedagógicas direcionadas.

No entanto, embora essas ferramentas digitais tenham potencial para transformar a prática educacional, ainda existem limitações significativas quanto à disponibilidade de instrumentos psicometricamente validados que possam medir o raciocínio lógico-matemático de maneira rigorosa e confiável (Haleem et al., 2022; Barcelos et al., 2018). A falta de validação psicométrica pode comprometer a precisão dos diagnósticos, dificultando a implementação de intervenções pedagógicas baseadas em evidências (Cunha; Neto; Stackfleth, 2016; El-Den et al., 2020). Sem instrumentos que atendam aos critérios de validade e confiabilidade, os dados obtidos a partir de avaliações digitais podem não refletir com precisão as competências cognitivas dos alunos, limitando o potencial de tais ferramentas para fornecer informações significativas que guiem as práticas educacionais (Kim, 2009).

A ausência de uma padronização nas métricas de avaliação também dificulta a comparação entre diferentes contextos e populações, limitando a possibilidade de estudos longitudinais e análises transversais que possam informar políticas educacionais e intervenções pedagógicas eficazes. Essa limitação é especialmente relevante em contextos educacionais, onde uma avaliação precisa das habilidades dos alunos é necessária para identificar suas necessidades e promover um desenvolvimento mais direcionado e fundamentado. Ferramentas bem estruturadas e validadas são essenciais para que educadores possam adaptar o ensino conforme o nível de compreensão e o ritmo de aprendizado dos alunos, promovendo uma educação mais inclusiva e eficaz (Skedsmo; Huber, 2019; Strietholt; Rosén; Gladushyna, 2021).

Portanto, torna-se fundamental o desenvolvimento de instrumentos psicometricamente robustos que possam mensurar o raciocínio lógico-matemático de maneira padronizada e acessível em plataformas digitais (Edelsbrunner; Dablander, 2019). Um instrumento digital que atenda a esses critérios pode não apenas facilitar o acesso a avaliações confiáveis, mas também permitir a coleta de dados comparáveis entre diferentes populações e contextos educacionais. Contribuindo significativamente para práticas educacionais baseadas em dados, promovendo diagnósticos iniciais, monitoramento contínuo do progresso cognitivo e intervenções pedagógicas personalizadas que beneficiem o desenvolvimento dos estudantes e o aprimoramento das práticas educativas (Maquiné; Coelho, 2020).

1.2 Problema e Proposta de Solução

Apesar da importância das habilidades de raciocínio lógico-matemático, há uma escassez de instrumentos de avaliação digital que sejam psicometricamente válidos e confiáveis (Junpeng et al., 2020). Instrumentos tradicionais frequentemente carecem de critérios de consistência interna, precisão e validade que sustentem seus resultados, o que compromete a comparabilidade e a aplicabilidade dos dados obtidos (Monticone et al., 2021; Lima; Osório, 2021). Essa limitação implica em uma dificuldade para educadores e pesquisadores identificarem com precisão os níveis de competência dos indivíduos e desenvolverem intervenções pedagógicas adequadas (Huber; Skedsmo, 2017).

Para abordar essa lacuna, este estudo propõe a seguinte questão: como desenvolver e validar um instrumento digital que avalie o raciocínio lógico-matemático de maneira consistente, precisa e cientificamente embasada? A resposta a esta pergunta envolve a construção de um instrumento que não apenas produza resultados confiáveis, mas também que seja replicável e aplicável em diferentes contextos educacionais, de modo a favorecer análises diagnósticas precisas e intervenções baseadas em dados.

Para responder ao problema identificado, este estudo propõe o desenvolvimento e validação de um módulo de avaliação digital, o *Screening Programming* (Dantas et al., 2023), para mensuração das habilidades de raciocínio lógico-matemático¹. Esse módulo será implementado em formato de aplicação digital, com base em critérios psicométricos rigorosos, e será estruturado em três fases principais:

- **Desenvolvimento de Itens:** A primeira fase consiste na concepção de um conjunto de itens psicometricamente válidos. Os itens serão elaborados para medir competências específicas de raciocínio lógico-matemático, incluindo análise, dedução e resolução de problemas. Esse desenvolvimento será fundamentado na literatura existente sobre avaliação cognitiva, assegurando que cada item represente um nível específico de habilidade.
- **Análise de Consistência Interna:** A segunda fase envolve a aplicação da Teoria Clássica dos Testes (TCT) para avaliar a consistência interna do instrumento

¹ Este módulo foi desenvolvido em parceria com o CNPq, no âmbito do Programa de Iniciação Científica (PIBIC), por meio do projeto "Investigação e Desenvolvimento de um Instrumento Informatizado para Avaliar e Aprimorar Habilidades do Raciocínio Lógico e Matemático," com vigência de 01/09/2023 a 31/08/2024, sob orientação do professor Rodrigo Alves Costa.

(Araújo et al., 2019; Primi, 2012). A consistência interna é um indicador fundamental de confiabilidade, e seu cálculo será realizado por meio do coeficiente alfa de Cronbach, além de outros parâmetros psicométricos para validar a estrutura do instrumento. Essa etapa visa garantir que o instrumento seja confiável e possa ser replicado em diferentes contextos educacionais (Wolcott; Olsen; Augustine, 2022).

- **Avaliação dos Parâmetros Psicométricos:** A terceira fase envolve a análise dos parâmetros psicométricos dos itens, considerando teorias de mensuração, como a Teoria da Resposta ao Item (TRI) (Araújo et al., 2019; Baker, 2001). Essa análise visa avaliar a dificuldade e discriminação dos itens, assegurando que cada um deles mesure com precisão as habilidades de raciocínio lógico-matemático. A avaliação válida deve refletir de forma precisa a competência que se propõe a medir, e, portanto, essa etapa é essencial para o desenvolvimento de um instrumento que seja cientificamente embasado e aplicável em diferentes cenários educacionais (Pasquali, 2017).

A implementação do módulo *Screening Programming* com base em princípios psicométricos rigorosos representa uma contribuição significativa para o campo da avaliação educacional. Esse instrumento, além de oferecer uma medida confiável e válida do raciocínio lógico-matemático, poderá apoiar diagnósticos precisos e intervenções pedagógicas fundamentadas, promovendo uma prática educacional mais eficiente e orientada por dados. Espera-se, assim, que o presente estudo contribua para a criação de instrumentos digitais de avaliação baseados em evidências, beneficiando o desenvolvimento cognitivo de estudantes e fortalecendo práticas educacionais baseadas em dados.

1.3 Objetivos

O presente estudo tem como objetivo principal avaliar a confiabilidade do *Screening Programming* como instrumento de mensuração das habilidades de raciocínio lógico-matemático, examinando sua consistência interna, precisão e validade dos resultados obtidos através da aplicação. Para alcançar este objetivo, foram estabelecidos os seguintes objetivos específicos:

- Conceber e desenvolver um módulo no *Screening Programming* voltado para a avaliação de habilidades de raciocínio lógico-matemático;

- Implementar itens específicos para a mensuração das habilidades de raciocínio lógico-matemático, com base em critérios psicométricos;
- Realizar a validação da consistência interna do instrumento utilizando a TCT, por meio do cálculo do coeficiente alfa de Cronbach e demais parâmetros psicométricos pertinentes;
- Analisar os parâmetros psicométricos dos itens para garantir a precisão e a robustez na mensuração das habilidades de raciocínio lógico-matemático.

1.4 Questões de Pesquisa

Com base nos objetivos propostos para o desenvolvimento e validação do *Screening Programming*, foram formuladas questões de pesquisa que orientam o processo de análise psicométrica e a avaliação da confiabilidade do instrumento. Essas questões têm como foco avaliar a eficácia do banco de itens na mensuração de habilidades específicas de raciocínio lógico-matemático em um contexto educacional, além de verificar a robustez das propriedades psicométricas que sustentam o uso do instrumento em pesquisas e avaliações educacionais.

- **QP1.** O instrumento desenvolvido a partir do banco de itens demonstra boa confiabilidade, permitindo uma mensuração consistente e estável das habilidades de raciocínio lógico-matemático ao longo do tempo e em diferentes contextos de aplicação?
- **QP2.** Os itens que representam os indicadores de habilidades de raciocínio lógico-matemático, como análise, dedução e resolução de problemas, apresentam propriedades psicométricas adequadas, como validade e capacidade de discriminação entre diferentes níveis de habilidade?

Essas questões de pesquisa fundamentam a análise da validade e confiabilidade do *Screening Programming*, orientando a verificação das características psicométricas dos itens e do instrumento como um todo. A investigação proposta avalia não apenas a precisão dos itens em discriminar diferentes níveis de habilidades de raciocínio lógico-matemático, mas também a capacidade do instrumento de oferecer resultados estáveis e consistentes, essenciais para sua aplicação prática e educacional.

1.5 Estrutura do trabalho

Este Trabalho de Conclusão de Curso (TCC) está estruturado em seis capítulos, organizados da seguinte forma: No **Capítulo 1**, Introdução, são apresentados o contexto, o problema de pesquisa, a proposta de solução, os objetivos, as questões de pesquisa e a estrutura do trabalho. O **Capítulo 2**, Referencial Teórico, explora os conceitos fundamentais relacionados ao raciocínio lógico-matemático e às teorias de mensuração, como a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). O **Capítulo 3**, *Screening Programming*, detalha o desenvolvimento do módulo informatizado e sua aplicação na avaliação das habilidades de raciocínio lógico-matemático. No **Capítulo 4**, Metodologia, descreve-se o desenho da pesquisa, os participantes, os instrumentos e as técnicas de análise utilizadas para avaliar a confiabilidade e validade do módulo. O **Capítulo 5**, Resultados e Discussões, apresenta os resultados da aplicação e análise psicométrica, incluindo a consistência interna, a discriminação e a dificuldade dos itens, e discute a eficácia do instrumento na mensuração do raciocínio lógico-matemático. Por fim, o **Capítulo 6**, Considerações Finais e Sugestões para Trabalhos Futuros, sintetiza os principais achados, limitações e implicações do estudo, além de propor direções para pesquisas futuras e aprimoramento do instrumento. Ao final, encontram-se as referências utilizadas no decorrer desta pesquisa.

2 REFERENCIAL TEÓRICO

Este capítulo aborda o raciocínio lógico-matemático, destacando sua importância para resolver problemas complexos e estruturar ideias em contextos acadêmicos e profissionais. Discute-se a relevância dessa habilidade em disciplinas como matemática e programação, além de explorar as bases teóricas da Teoria Clássica dos Testes (TCT) e da Teoria de Resposta ao Item (TRI) para mensuração psicométrica. A TCT é usada para avaliar a consistência interna, enquanto a TRI permite uma avaliação adaptativa e detalhada dos itens. O capítulo também analisa o desenvolvimento gradual dessa competência ao longo da vida, fornecendo o embasamento necessário para a criação de instrumentos de avaliação confiáveis e eficazes.

2.1 Raciocínio Lógico-Matemático

O raciocínio lógico-matemático é uma habilidade cognitiva essencial, que permite ao indivíduo não apenas resolver problemas complexos, mas também aplicar métodos rigorosos de análise e dedução em diversos contextos (Cresswell; Speelman, 2020). Essa competência é essencial em disciplinas como matemática e programação, onde a precisão e a capacidade de conectar ideias de forma estruturada são fundamentais. Compreender e desenvolver o raciocínio lógico-matemático torna-se, assim, uma prioridade educacional, visando formar alunos com pensamento crítico e habilidades sólidas de resolução de problemas (Bronkhorst et al., 2020).

2.1.1 *Definição e Importância*

O raciocínio lógico pode ser definido como a habilidade de estruturar ideias e argumentos de forma coerente, conectando conceitos e relacionamentos por meio de princípios sistemáticos de dedução e inferência (Markovits, 2017). Em termos gerais, a lógica envolve a distinção entre argumentos válidos e inválidos, o que exige do indivíduo uma compreensão clara dos processos de pensamento que sustentam conclusões corretas. Assim, o raciocínio lógico é a base para uma série de competências cognitivas, permitindo ao indivíduo navegar em situações complexas com clareza e precisão (Prado, 2019; Cresswell; Speelman, 2020; Bronkhorst et al., 2020).

No campo da matemática, o raciocínio lógico-matemático permite que o aluno compreenda conceitos abstratos, estabeleça relações entre números e operações, e explore padrões e regularidades. Ele atua como um mecanismo facilitador para a resolução de problemas, pois ajuda o indivíduo a decompor uma questão complexa em partes menores e manejáveis, levando a uma compreensão mais profunda do problema como um todo (Silva; Luna, 2019). Além disso, o pensamento lógico-matemático é amplamente aplicável em outras áreas do conhecimento, como a física, a química e as ciências sociais, onde a capacidade de análise e dedução contribui para a formação de um pensamento crítico e analítico (García-Carmona, 2023; Mitrevski, 2019).

Em termos educacionais, o raciocínio lógico-matemático desempenha um papel essencial no desenvolvimento de habilidades como a análise crítica, a abstração e a resolução de problemas (Cresswell; Speelman, 2020). Essas habilidades ajudam os estudantes a adquirir uma visão sistemática do conhecimento, permitindo que façam conexões entre conceitos aparentemente desconexos e encontrem soluções inovadoras para problemas complexos. O desenvolvimento do raciocínio lógico é, portanto, um objetivo central na educação básica e superior, preparando os alunos para uma série de desafios acadêmicos e profissionais (Firdaus et al., 2015). Estudos indicam que a prática do raciocínio lógico-matemático não apenas melhora a compreensão de conceitos matemáticos, mas também fortalece as habilidades de pensamento analítico e independente dos alunos, promovendo uma aprendizagem que se estende para além das fronteiras de qualquer disciplina específica (Cresswell; Speelman, 2020).

Além de seu valor acadêmico, o raciocínio lógico-matemático é uma habilidade essencial no mercado de trabalho moderno, onde a resolução de problemas complexos e a tomada de decisões informadas são altamente valorizadas (Ritter et al., 2014; Lin, 2023b). Profissões em tecnologia, engenharia e ciências, por exemplo, exigem dos indivíduos uma compreensão profunda de princípios lógicos e uma capacidade analítica aguçada para resolver problemas e otimizar processos (Lund, 2020). A aplicação do raciocínio lógico na programação é particularmente relevante, pois programar requer a capacidade de decompor problemas, identificar padrões, planejar soluções e prever os efeitos de cada decisão no contexto de um sistema maior. Dessa forma, a importância do raciocínio lógico-matemático vai além do ambiente acadêmico, sendo uma competência fundamental para o sucesso profissional em uma sociedade cada vez mais orientada pela tecnologia (Yasin; Nusantara, 2023).

2.1.2 *Desenvolvimento Cognitivo*

O desenvolvimento do raciocínio lógico-matemático é um processo gradual e cumulativo, influenciado tanto por fatores internos (cognitivos) quanto externos (sociais e ambientais) (Lin, 2023a). O raciocínio lógico emerge como parte do desenvolvimento cognitivo humano, ocorrendo em estágios que vão desde o pensamento concreto, predominante na infância, até o pensamento formal e abstrato, típico da adolescência e da vida adulta (Houdé, ; Vygotsky, 2024). Esse desenvolvimento é impulsionado pelo processo de abstração reflexiva, pelo qual o indivíduo constrói relações entre objetos e ideias com base em experiências concretas e, gradualmente, internaliza essas relações em conceitos abstratos (Tallman; O'Bryan, 2024).

A construção do raciocínio lógico-matemático é também um reflexo das interações entre o indivíduo e seu ambiente (Fatimah; Johar; Zubainur, 2020). A aprendizagem ocorre por meio de esquemas mentais que se adaptam continuamente em resposta a novos desafios, sendo enriquecidos e reorganizados à medida que o indivíduo é exposto a situações que exigem pensamento analítico e resolução de problemas (Wang; Yin, 2023). A partir dessa perspectiva, o desenvolvimento do raciocínio lógico não é apenas uma habilidade que surge naturalmente, mas uma competência que pode ser estimulada e fortalecida através de experiências educacionais adequadas, que incentivem a exploração, a experimentação e a resolução de problemas complexos (Ponniah; Safeek, 2024)

Além disso, a teoria sociocultural postula que o aprendizado ocorre em um contexto social e cultural, onde a interação com outros indivíduos desempenha um papel essencial na construção do conhecimento (Tzuriel; Tzuriel, 2021). Através da Zona de Desenvolvimento Proximal, as crianças podem alcançar níveis mais elevados de compreensão e habilidade com o auxílio de um educador ou de pares mais experientes. Essa zona representa, portanto, o potencial de desenvolvimento que pode ser alcançado com o apoio adequado, permitindo ao aluno enfrentar desafios ligeiramente acima de seu nível atual e desenvolver um raciocínio lógico-matemático mais avançado (Zaretsky, 2021).

A educação formal desempenha um papel fundamental no estímulo ao desenvolvimento do raciocínio lógico-matemático (Nite et al., 2023). Quando o ambiente de aprendizagem oferece atividades desafiadoras, mas acessíveis, ele permite que os alunos explorem seus limites cognitivos e adquiram novos conhecimentos e habilidades

de maneira mais eficaz (Chew; Cerbin, 2021). Ao criar ambientes educacionais que incentivem a resolução de problemas, a curiosidade e a colaboração, os educadores podem proporcionar aos alunos oportunidades para desenvolver e consolidar seu raciocínio lógico-matemático (Lin, 2023b). Essa abordagem não só contribui para o desenvolvimento de habilidades cognitivas avançadas, mas também promove uma aprendizagem significativa e duradoura, que prepara os alunos para aplicarem seu conhecimento em contextos práticos e profissionais.

Dessa forma, o desenvolvimento do raciocínio lógico-matemático não é apenas um objetivo educacional, mas uma competência fundamental para a formação de indivíduos capazes de analisar, sintetizar e resolver problemas de forma independente e inovadora (Proença, 2022). Através de um processo gradual de construção e abstração, e com o apoio de ambientes de aprendizagem bem estruturados, os alunos podem desenvolver um raciocínio lógico sólido e aplicar essas habilidades tanto na academia quanto em suas futuras carreiras profissionais (Lin, 2023b).

2.2 Teorias Psicométricas

Nosso objetivo é medir características não observáveis diretamente, como conhecimento em disciplinas educacionais, comportamentos ou até patologias na área da saúde. Chamamos essas características de traços latentes ou habilidades (Primi, 2012). Para avaliá-los, aplicamos duas principais teorias psicométricas: a TCT e a TRI, que nos permitem estimar os parâmetros dos itens e examinar a qualidade do instrumento como um todo.

2.2.1 Teoria Clássica dos Testes

Na TCT, usamos a pontuação total de um teste para medir a habilidade ou o conhecimento de cada indivíduo. Essa pontuação serve como uma estimativa de desempenho, comparável por meio de normas que posicionam o indivíduo em relação a outros ou ao próprio construto avaliado (Araújo et al., 2019). Dessa forma, os escores obtidos representam o nível de proficiência de cada respondente em um instrumento específico.

A TCT considera que a pontuação observada de um indivíduo resulta de duas componentes: a pontuação verdadeira e o erro de medida (Primi, 2012). Nosso foco é

minimizar esse erro para garantir que o escore final reflita precisamente o construto alvo (Pasquali, 2017). Com a pontuação total, aplicamos medidas para avaliar a qualidade dos itens e do instrumento, como o coeficiente de correlação ponto-bisserial e o coeficiente alfa de Cronbach.

2.2.1.1 Coeficiente de Ponto Bisserial

Calculamos o coeficiente de ponto bisserial para mensurar a correlação entre uma variável dicotômica (certo/errado) e uma variável contínua (a pontuação total do teste). Esse coeficiente indica a capacidade de cada item em distinguir indivíduos com diferentes níveis de habilidade no instrumento (Baker, 2001). Itens com coeficientes elevados são mais discriminativos, diferenciando eficazmente entre indivíduos com habilidades altas e baixas (Araújo et al., 2019). A Equação 2.1 descreve o cálculo do coeficiente de ponto bisserial:

$$Ppb = \frac{\bar{X}_A - \bar{X}_T}{S_T} \sqrt{\frac{p}{1-p}} \quad (2.1)$$

Onde:

\bar{X}_A é a média das pontuações dos respondentes que acertaram o item;

\bar{X}_T é a média global das pontuações do teste;

S_T é o desvio padrão das pontuações totais;

p é a proporção de respondentes que acertaram o item.

Itens com coeficientes de ponto bisserial próximos de +1 são altamente discriminativos, o que significa que eles são eficazes em diferenciar entre indivíduos com habilidades altas e baixas. Esses itens contribuem significativamente para a precisão da estimativa da habilidade, pois conseguem identificar com clareza os participantes que possuem maior domínio do traço latente em comparação com aqueles que apresentam habilidades mais limitadas (Araújo et al., 2019). A alta discriminação de tais itens é fundamental para a qualidade do instrumento, uma vez que eles aumentam a capacidade do teste de capturar variações sutis no desempenho dos respondentes. Além disso, ao identificar os itens com maior poder discriminativo, os avaliadores podem ajustar o instrumento para reforçar a sua sensibilidade, garantindo que a mensuração seja representativa e válida para os objetivos pretendidos.

2.2.1.2 Coeficiente Alpha de Cronbach

Utilizamos o coeficiente alfa de Cronbach para avaliar a confiabilidade do instrumento de medida, analisando sua consistência interna. Este coeficiente indica o grau de correlação entre os itens, determinando se eles avaliam de forma coerente o mesmo construto (Araújo et al., 2019). O alfa de Cronbach varia de 0 a 1; valores próximos a 1 indicam uma boa consistência interna. Escores entre 0,70 e 0,80 são considerados adequados, enquanto valores inferiores a 0,70 sugerem ajustes no instrumento (Wolcott; Olsen; Augustine, 2022). A Equação 2.2 define o alfa de Cronbach:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum S_i^2}{S_T^2} \right) \quad (2.2)$$

Onde:

n é o número de itens;

$\sum S_i^2$ é a soma das variâncias dos itens;

S_T^2 é a variância total das pontuações do teste.

2.2.1.3 Limitações da TCT

Apesar de sua ampla aplicação, a TCT apresenta limitações. Como todos os indivíduos respondem ao mesmo conjunto de itens, independentemente do nível de habilidade, isso pode gerar respostas inconsistentes e cansaço (Kanlı, 2020). A TCT também ignora a probabilidade de acerto ao acaso, afetando a precisão dos resultados em testes de múltipla escolha, e limita a comparação entre populações ou versões de testes diferentes, já que a análise depende da amostra utilizada (Baker, 2001).

A teoria também assume que todos os itens têm peso igual na pontuação total, o que pode não refletir a real dificuldade e discriminação de cada item. A TRI aborda essas limitações, oferecendo uma análise mais flexível e detalhada do desempenho individual.

2.2.2 Teoria de Resposta ao Item

Para superar as limitações da TCT, aplicamos a TRI, que complementa a TCT ao fornecer uma análise mais detalhada da probabilidade de um indivíduo acertar itens, considerando respostas a outros itens do teste. A TRI também permite comparações

entre diferentes testes, desde que compartilhem alguns itens, e entre populações diversas (Polat, 2022).

Usamos a TRI para construir instrumentos educacionais, calibrar itens e desenvolver medidas psicométricas. A TRI oferece flexibilidade ao permitir o ajuste dos dados em diferentes modelos estatísticos, comparando habilidades de indivíduos de maneira independente da amostra (Baker, 2001).

Na TRI, prevemos o desempenho de um indivíduo em um item com base em traços latentes. A Curva Característica do Item (CCI) descreve a relação entre a habilidade do indivíduo e a probabilidade de acertar um item (Pasquali, 2017).

2.2.2.1 Curva Característica do Item

A CCI é uma ferramenta essencial para entender como a probabilidade de acerto em um item varia de acordo com o nível de habilidade latente do indivíduo (Vieira; Júnior; Potrich, 2020). Cada item possui uma CCI única que descreve sua qualidade e comportamento em diferentes níveis de habilidade, permitindo que analistas visualizem como a chance de resposta correta se altera para respondentes com habilidades variadas. A CCI é geralmente representada por um gráfico, no qual o eixo x corresponde ao nível de habilidade (θ) do indivíduo, e o eixo y reflete a probabilidade de acerto, variando de 0 a 1 (Araújo et al., 2019). Esse gráfico permite identificar itens mais eficazes para determinados níveis de habilidade e auxilia na construção de instrumentos de medição que discriminem bem entre respondentes com diferentes perfis de habilidade.

Para modelar a CCI, aplicamos o modelo logístico de três parâmetros (3PL), que é amplamente utilizado na TRI. Esse modelo considera três aspectos fundamentais de cada item: a discriminação (a_j), que indica o quanto o item diferencia entre indivíduos com habilidades próximas; a dificuldade (b_j), que reflete o nível de habilidade necessário para que um respondente tenha 50% de chance de acerto; e o parâmetro de acerto ao acaso (c_j), que mede a probabilidade de um indivíduo com habilidade muito baixa acertar o item por sorte (Jiang et al., 2024). A Equação 2.3 define o modelo 3PL:

$$P(\theta) = c_j + (1 - c_j) \frac{1}{1 + e^{-a_j(\theta - b_j)}} \quad (2.3)$$

Onde:

θ representa a habilidade latente do indivíduo;

a_j é o parâmetro de discriminação do item, indicando a inclinação da curva e, portanto, o poder discriminativo do item;

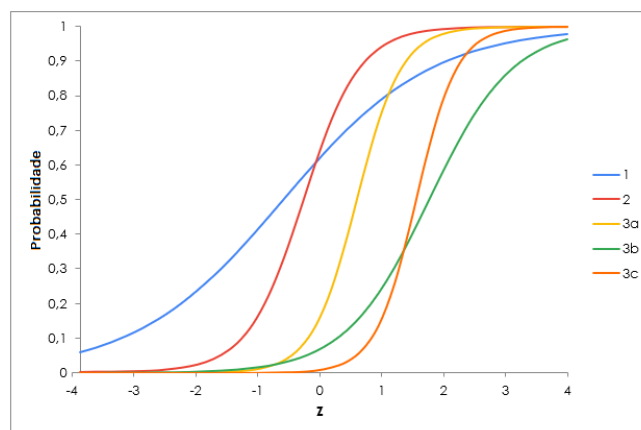
b_j é o parâmetro de dificuldade, determinando o ponto na escala de habilidade onde a probabilidade de acerto é 50

c_j é o parâmetro de acerto ao acaso, representando a probabilidade de acerto independentemente do conhecimento do respondente.

Itens com alto valor de discriminação (a_j) têm inclinações mais acentuadas na CCI, o que permite distinguir mais precisamente entre respondentes com habilidades próximas àquela exigida pelo item. Em contraste, um alto valor de dificuldade (b_j) indica que o item é mais complexo, requerendo um nível de habilidade elevado para que o respondente tenha uma chance razoável de acerto. O parâmetro de acerto ao acaso (c_j), geralmente entre 0 e 0,5, é especialmente relevante em testes de múltipla escolha, pois reflete a probabilidade de um acerto não intencional, como em questões em que o respondente adivinha a resposta correta (Brandão; Guedes et al., 2020).

Para calcular esses parâmetros e gerar as CCIs, utilizamos programas especializados que aplicam funções não lineares. Estes programas transformam os parâmetros em representações gráficas, facilitando a análise visual da relação entre habilidade e probabilidade de acerto para cada item (Araújo et al., 2019). A Figura 1 exibe exemplos de CCIs de diferentes itens, destacando as diferenças na capacidade discriminativa e nos níveis de dificuldade entre eles.

Figura 1 – Exemplo de CCIs para diferentes itens.



Fonte: Elaborada pelo autor (2024).

No gráfico da Figura 1, cada curva corresponde à CCI de um item específico, mostrando como a chance de acerto se altera conforme o nível de habilidade do respondente. O eixo horizontal (x) representa a habilidade, variando de -4 a +4, enquanto o eixo vertical (y) indica a probabilidade de resposta correta, variando de 0 a 1. Neste exemplo, observa-se que o item "3c" possui a maior capacidade discriminativa, distinguindo bem entre respondentes com habilidades próximas. O item "3b" apresenta a maior dificuldade, exigindo um nível de habilidade mais elevado para uma probabilidade de acerto significativa, enquanto o item "1" é o mais simples, com menor capacidade discriminativa.

2.2.2.2 Estimativa da Habilidade

O modelo 3PL gera uma escala chamada habilidade, que varia entre -4 e +4. Estimamos a habilidade com base nas respostas e nos parâmetros dos itens, usando o método Expected A Posteriori (EAP) (Chapman, 2022). O EAP usa estatísticas bayesianas para calcular a habilidade esperada de um indivíduo com base em suas respostas (Kern; Choe, 2021; Chapman, 2022). A Equação 2.4 define o EAP:

$$\theta_{s+1} = \theta_s + \frac{\sum_{i=1}^n a_i [u_i - P_i \theta_s]}{\sum_{i=1}^n a_i^2 P_i(\theta_s) Q_i(\theta_s)} \quad (2.4)$$

Onde:

θ_s é a habilidade estimada do respondente após s iterações;

a_i é o parâmetro de discriminação do item i.

O EAP personaliza as estimativas de habilidade para cada indivíduo, aplicando uma abordagem bayesiana que considera tanto as respostas fornecidas quanto os parâmetros dos itens avaliados. Esse método calcula a habilidade esperada ao combinar as informações dos itens com uma distribuição a priori, o que permite ajustes mais refinados e específicos para cada respondente. Esse processo bayesiano maximiza a precisão das habilidades latentes estimadas ao utilizar uma combinação probabilística, reduzindo o impacto de respostas extremas ou aleatórias e fornecendo uma medida mais confiável do desempenho individual.

2.2.2.3 Função de Informação do Item

A Função de Informação do Item (FII) é uma medida da TRI para avaliar a quantidade de informação psicométrica que cada item fornece em diferentes níveis de habilidade do respondente (Pasquali, 2017). A FII indica em quais pontos do continuum de habilidade o item oferece maior precisão para discriminar entre níveis distintos, sendo especialmente útil para selecionar itens que maximizem a acurácia da estimativa de habilidade. Em um teste, itens com alta FII em um determinado ponto da habilidade são mais eficazes para avaliar indivíduos que possuem habilidade próxima àquele nível, permitindo que o teste seja ajustado para diferentes públicos e contextos de avaliação.

Na prática, a FII auxilia na criação e otimização de instrumentos de medida, permitindo que itens com valores informativos mais altos em determinados níveis de habilidade sejam escolhidos para uma avaliação mais precisa. A FII também é utilizada para entender onde um teste é mais sensível na distinção de habilidades entre respondentes e para ajustar a dificuldade e discriminação dos itens de acordo com o público-alvo. Em um exame adaptativo, por exemplo, o uso da FII permite selecionar itens que se alinham melhor à habilidade do respondente em tempo real, tornando a avaliação mais eficiente.

A Equação 2.5 define a FII:

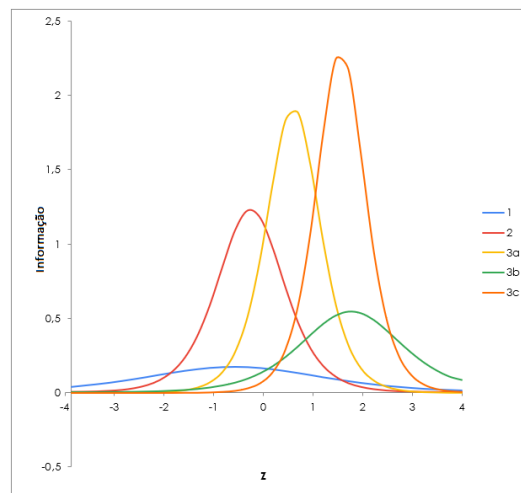
$$I(\theta) = a^2 \frac{Q_i(\theta)(P_i(\theta) - c)^2}{P_i(\theta)(1 - c)^2} \quad (2.5)$$

Onde:

- a representa o parâmetro de discriminação do item, indicando o quanto o item diferencia entre indivíduos de habilidades próximas;
- $P_i(\theta)$ é a probabilidade de um respondente com habilidade θ acertar o item;
- $Q_i(\theta)$ é a probabilidade de erro do respondente, calculada como $1 - P_i(\theta)$;
- c é o parâmetro de acerto ao acaso, que representa a probabilidade de acerto para um indivíduo com habilidade muito baixa.

O gráfico da Figura 2 exemplifica a FII para diferentes itens, destacando as variações de informação em função dos níveis de habilidade dos respondentes.

Figura 2 – Exemplo de FII's para diferentes itens.



Fonte: Elaborada pelo autor (2024).

Na Figura 2, o eixo horizontal (x) indica o nível de habilidade, enquanto o eixo vertical (y) apresenta a quantidade de informação que cada item oferece em distintos pontos dessa habilidade. O item "3a", por exemplo, demonstra uma concentração maior de informação para respondentes com níveis intermediários de habilidade, enquanto o item "2" fornece informações mais relevantes para aqueles com níveis iniciais de habilidade.

3 SCREENING PROGRAMMING

Neste capítulo, apresentamos as etapas realizadas para a construção do sistema *Screening Programming*, uma plataforma projetada para apoiar o desenvolvimento das habilidades acadêmicas de graduandos. Similar ao *Google Classroom*, a aplicação permite o acompanhamento de desempenho e oferece atividades direcionadas. No entanto, o *Screening Programming* foca especificamente em questões avaliativas de nível superior e adota a TRI para ajustar a dificuldade e precisão das avaliações, proporcionando um ambiente de aprendizado mais eficaz para universitários.

3.1 Análise de Requisitos

A análise de requisitos é uma etapa fundamental no processo de desenvolvimento de *software*, responsável por definir o que o sistema deve fazer e as características necessárias para que ele atenda às necessidades dos usuários e às expectativas dos *stakeholders*. Para o desenvolvimento do nosso sistema, focado em mensurar as habilidades de raciocínio lógico matemático de alunos graduandos em cursos de exatas, realizamos uma entrevista semi-estruturada com especialistas da Universidade Federal de Campina Grande (UFCG) e do CognitiveLAB. Esse processo visou garantir que os requisitos do sistema estivessem alinhados com as necessidades técnicas e pedagógicas do projeto.

Essa etapa incluiu diversas reuniões para coleta e análise de requisitos, além da revisão de documentações de projetos semelhantes, o que nos permitiu estabelecer uma base sólida para o desenvolvimento da solução. A análise foi estruturada em requisitos funcionais e não funcionais, ambos essenciais para o sucesso do sistema, assegurando tanto a realização das funcionalidades esperadas quanto o atendimento a atributos como usabilidade, segurança e desempenho.

3.1.1 Requisitos Funcionais

Os requisitos funcionais especificam as funcionalidades que o sistema deve oferecer (Rezende, 2006). Eles detalham os serviços que serão disponibilizados e a forma como essas funções irão interagir com os usuários (Chung et al., 2012). Abaixo estão listadas as principais funcionalidades do sistema, todas projetadas para facilitar o uso por

professores e administradores, proporcionando aos alunos uma experiência educativa enriquecedora.

- Cadastro de professor - O sistema permitirá o cadastro de professores por um administrador, incluindo dados como nome, email, instituição e outras informações básicas. Esse cadastro também possibilitará a personalização das funcionalidades da plataforma para cada professor, além de facilitar o monitoramento de alunos e atividades.
- Cadastro de aluno - Alunos poderão realizar o cadastro antes ou depois de receberem um convite de algum grupo. Será necessário que seja incluídas informações como nome, instituição, ano/série, e outros dados essenciais. Esse cadastro facilitará o acompanhamento do progresso dos alunos nas atividades e permitirá a geração de relatórios personalizados.
- Página de informações sobre o sistema - O sistema oferecerá uma página com detalhes sobre seus objetivos, funcionalidades e instruções de uso, similar a uma página inicial de apresentação. Isso será útil para novos usuários se familiarizarem com a plataforma, sem necessidade de suporte adicional.
- Página de créditos do sistema - Para reconhecer o trabalho dos desenvolvedores e parceiros, o sistema incluirá uma página de créditos com informações sobre as instituições e pessoas envolvidas no projeto, promovendo transparência e fortalecendo parcerias.
- Visualização dos alunos cadastrados - Professores terão uma página onde poderão visualizar e gerenciar todos os alunos cadastrados. Através desta, eles poderão editar dados, excluir cadastros e iniciar atividades específicas para cada aluno, organizando e navegando eficientemente por uma lista de estudantes.
- Visualização e edição de dados pessoais - Professores poderão acessar e atualizar suas próprias informações, como email e senha. Essa funcionalidade garante que dados incorretos possam ser corrigidos e que as informações estejam sempre atualizadas.

- Relatório de atividades e desempenho dos alunos - O sistema gerará relatórios com o histórico de atividades e desempenho dos alunos. Professores poderão acompanhar o progresso de cada aluno, identificando áreas de força e de necessidade de intervenção, para que possam ajustar suas estratégias pedagógicas.
- Área de atividades cognitivas - Uma seção dedicada às atividades será disponibilizada para avaliar habilidades cognitivas, como flexibilidade, controle inibitório e memória de trabalho, inspirada nas funcionalidades do Classroom. Professores poderão organizar atividades e monitorar a execução de tarefas pelos alunos.
- Estimativa de habilidades cognitivas - Após a realização das atividades, o sistema fornecerá uma estimativa das habilidades dos alunos com base em modelos psicométricos, como a TRI. Esse recurso possibilita uma análise objetiva do desempenho, permitindo que professores ofereçam suporte personalizado.
- Visualização dos membros do grupo - Similar ao Classroom, o sistema permitirá que professores e alunos visualizem quem está no grupo ou turma, facilitando o engajamento e o acompanhamento coletivo das atividades.
- Recuperação de senha - O sistema fornecerá uma funcionalidade de recuperação de senha para os usuários. Em caso de esquecimento, o professor ou aluno poderá solicitar a redefinição de senha, recebendo um email com um link seguro para redefinição, garantindo que o acesso seja retomado de forma prática e segura.
- Convites para grupos via email - Alunos receberão convites para participar de grupos diretamente em seu email, enviados pelo professor ou administrador. O email conterá um link que permite ao aluno aceitar o convite e ingressar no grupo, facilitando a integração e participação nas atividades educacionais.

Os requisitos funcionais foram elaborados com atenção para assegurar que o sistema não só funcione de forma eficaz, mas também proporcione uma experiência de uso agradável e intuitiva para professores e alunos.

3.1.2 Requisitos Não-Funcionais

Requisitos Não-Funcionais referem-se às características de qualidade e restrições do sistema que não estão diretamente relacionadas às funções que ele realiza. Em geral,

eles influenciam o desempenho, usabilidade, segurança, escalabilidade, e outros aspectos do sistema que impactam a experiência do usuário e a operação do sistema (Rezende, 2006). A seguir, detalhamos os requisitos não-funcionais.

- Usabilidade - O sistema deve ser intuitivo e acessível para usuários com diferentes níveis de habilidade tecnológica, abrangendo tanto a interface gráfica do usuário (GUI) quanto a apresentação e organização das funcionalidades. A usabilidade é relevante, uma vez que professores e alunos podem não ter formação técnica, sendo necessário que o sistema permita interações simplificadas e uma navegação fácil.
- Desempenho - O sistema deve ser otimizado para garantir tempos de resposta rápidos em todas as operações, como login, carregamento de páginas e execução de atividades. É essencial que essas operações ocorram sem atrasos perceptíveis, mesmo durante períodos de alta demanda de usuários. Um bom desempenho é fundamental para evitar frustrações e garantir uma experiência fluida.
- Confiabilidade - O sistema deve ser altamente confiável e estar sempre disponível para uso. Isso implica que não deve ocorrer frequentemente quedas e que, em caso de falhas, a recuperação deve ser rápida. Serão implementados mecanismos de redundância e backups automáticos para garantir que os dados dos usuários sejam preservados e que o sistema continue operando sem interrupções.
- Segurança - Considerando a sensibilidade dos dados de alunos e professores, a segurança é uma prioridade fundamental. O sistema implementará criptografia para dados sensíveis, como senhas, e utilizará tokenização para minimizar os riscos de vazamento de informações.
- Escalabilidade - O sistema deve ser projetado para suportar um aumento no número de usuários e dados ao longo do tempo. Isso envolve a capacidade de adicionar novos alunos, turmas e atividades sem afetar o desempenho. A escalabilidade é vital para garantir que o sistema possa crescer em resposta à demanda, evitando a necessidade de reescrita ou substituição significativa do código.
- Manutenibilidade - O código do sistema deve ser desenvolvido com boas práticas de programação, incluindo modularidade e documentação adequada, para facilitar

a manutenção e atualização futura. A manutenibilidade é importante para permitir a adição de novas funcionalidades e a correção de bugs de forma rápida e eficiente.

Esses requisitos não-funcionais são fundamentais para assegurar que o sistema não apenas cumpra as necessidades funcionais, mas também proporcione uma experiência estável, segura e escalável ao longo do tempo.

3.2 Modelo de Análise

O modelo de análise tem como objetivo descrever a estrutura lógica e organizacional do sistema, elucidando como os usuários interagem com as funcionalidades, quais papéis são atribuídos a cada um e como as tecnologias utilizadas suportam a execução das tarefas propostas. Este modelo foi elaborado para proporcionar uma visão clara e objetiva do funcionamento do sistema, assim como dos recursos tecnológicos implementados para garantir sua eficiência e usabilidade.

3.2.1 Perfil do Usuário

O sistema foi projetado para atender às necessidades de dois grupos principais de usuários: professores e alunos. Cada grupo possui características e expectativas específicas em relação ao uso do sistema, com diferentes níveis de acesso e funcionalidades que visam facilitar o processo de ensino e aprendizado. Os professores, com conhecimento intermediário em tecnologia e um entendimento da TRI, são responsáveis por gerenciar turmas e acompanhar o desenvolvimento dos alunos. Por outro lado, os alunos, na faixa etária de 6 a 12 anos, interagem com o sistema de forma direta e intuitiva, focados na realização das atividades cognitivas. A seguir, são descritas as particularidades de cada perfil de usuário e como o sistema foi projetado para atender a essas necessidades específicas.

- **Professores:** Esses usuários são responsáveis por gerenciar turmas, cadastrar alunos, supervisionar as atividades e analisar os relatórios gerados pelo sistema. É esperado que os professores possuam um nível básico a intermediário de conhecimento em informática, o que inclui familiaridade com a navegação em sistemas web, edição de informações e uso de ferramentas de comunicação digital. Além

disso, os professores devem ter conhecimento sobre a TRI, uma vez que essa abordagem é utilizada para avaliar as habilidades dos alunos nas atividades cognitivas propostas. Essa compreensão é fundamental, pois as atividades estão diretamente ligadas ao desenvolvimento das habilidades executivas dos alunos.

- **Alunos:** Os alunos, que são o foco principal das atividades, têm idades variadas, sendo o público-alvo crianças de 6 a 12 anos. Dada a faixa etária, espera-se que os alunos tenham apenas um nível básico de familiaridade com o uso de computadores ou dispositivos móveis, tornando essencial que a interface do sistema seja altamente intuitiva e acessível. O design das atividades cognitivas foi pensado para ser desafiador, mas também envolvente e adequado ao desenvolvimento cognitivo dessa faixa etária.

A interação dos usuários com o sistema é otimizada por meio de uma interface intuitiva, que apresenta um menu claro e bem organizado. O fluxo de interação é estruturado em uma sequência de eventos bem definida, visando aprimorar a usabilidade e facilitar a navegação dentro do sistema.

1. O sistema começa com uma tela mostrando os grupos que o aluno está e uma barra de navegação lateral, que permite ao usuário acessar rapidamente áreas principais, como o mural dos alunos, edição de dados, relatórios de desempenho, além de páginas de créditos e informações sobre o sistema. A barra de navegação foi projetada para facilitar o acesso às funcionalidades mais importantes.
2. O professor terá acesso a funcionalidades administrativas, convidar aluno e excluir, bem como o início de atividades para alunos específicos. O fluxo de navegação é simplificado, permitindo que o professor gerencie alunos e monitore seu progresso com poucos cliques. O professor também poderá editar suas próprias informações, garantindo que seus dados estejam sempre atualizados.
3. Os alunos poderão acessar atividades cognitivas diretamente no sistema, respondendo aos itens avaliativos. Após cada atividade, eles terão a oportunidade de visualizar um resumo do seu desempenho, proporcionando um feedback imediato sobre seu progresso.

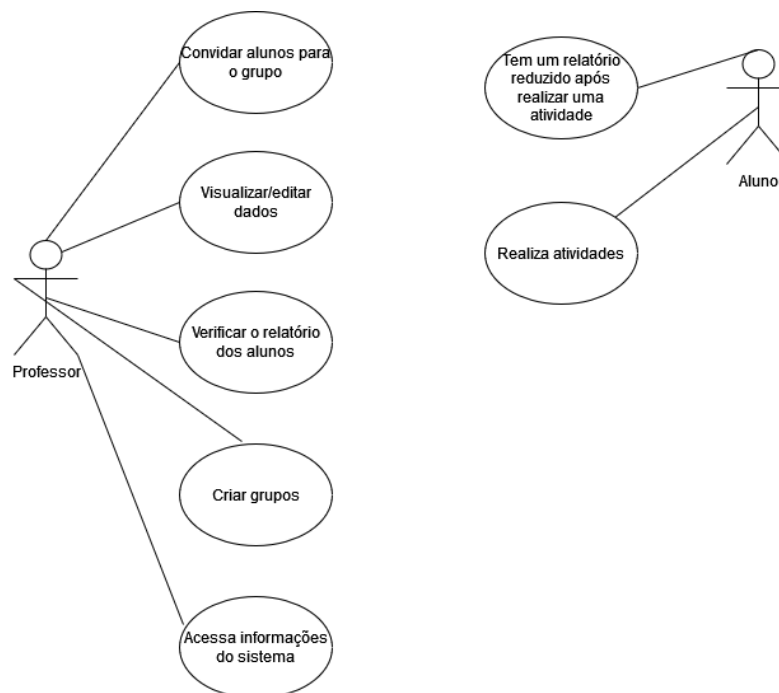
4. Após a conclusão das atividades, o professor poderá acessar relatórios detalhados que mostram os erros e acertos de cada aluno. Esses relatórios são ferramentas úteis para acompanhar o desenvolvimento das habilidades cognitivas ao longo do tempo, permitindo ao professor identificar áreas que podem precisar de mais atenção e ajustes pedagógicos.

Esse fluxo bem definido assegura que o sistema satisfaça as necessidades tanto dos professores quanto dos alunos, ao mesmo tempo que proporciona um ambiente de fácil navegação e gerenciamento de atividades.

3.2.2 Diagrama de Caso de Uso

O Diagrama de Caso de Uso, ilustrado na Figura 3, oferece uma representação visual dos principais atores do sistema e suas interações com as funcionalidades do *Screening Programming*. Esse diagrama é uma ferramenta valiosa para entender como os diferentes usuários — professor e aluno — utilizam as funções disponíveis.

Figura 3 – Modelo de casos de uso.



Fonte: Elaborado pelo autor (2024).

O professor é o principal usuário das funcionalidades pedagógicas, responsável por gerenciar os alunos (incluindo cadastro, edição e exclusão), configurar as atividades cognitivas e monitorar os relatórios de desempenho de cada aluno. Além disso, o acompanhamento e a análise dos instrumentos avaliativos estão entre suas atribuições.

3.2.3 *Tecnologias de Suporte*

A escolha das tecnologias utilizadas no desenvolvimento do *Screening Programming* foi orientada pela necessidade de garantir que o sistema seja escalável, seguro e eficiente, tanto em desempenho quanto em manutenção futura. A seguir, detalhamos as principais ferramentas e *frameworks* empregados no projeto:

- **Front-end:** A interface do usuário foi desenvolvida utilizando HTML, CSS e JavaScript. Essas tecnologias são amplamente reconhecidas no mercado e possibilitam a criação de interfaces dinâmicas e responsivas, assegurando que o sistema seja acessível em computadores e dispositivos móveis. Além disso, essa combinação facilita a integração de bibliotecas e *frameworks* que melhoram a experiência do usuário.
- **Back-end:** O *backend* do sistema foi desenvolvido em Python, utilizando o *framework* Flask. Essa escolha oferece flexibilidade e simplicidade na criação de APIs e na manipulação de dados, permitindo um gerenciamento eficaz das requisições. O backend foi hospedado na plataforma Render, garantindo escalabilidade e eficiência.
- **Banco de Dados:** Para o armazenamento dos dados, optamos pelo PostgreSQL, um banco de dados relacional robusto que oferece desempenho superior e funcionalidades avançadas. A escolha do PostgreSQL permite a integração fácil com o Flask e assegura a integridade dos dados durante a execução das atividades dos alunos.
- **Controle de Versão e Colaboração:** Utilizamos o Git como sistema de controle de versão, juntamente com o GitHub, que serve como repositório remoto e ferramenta colaborativa. Essas tecnologias garantem que o código seja armazenado de forma segura na nuvem, permitindo a colaboração de vários desenvolvedores simultaneamente e mantendo um histórico claro das modificações.

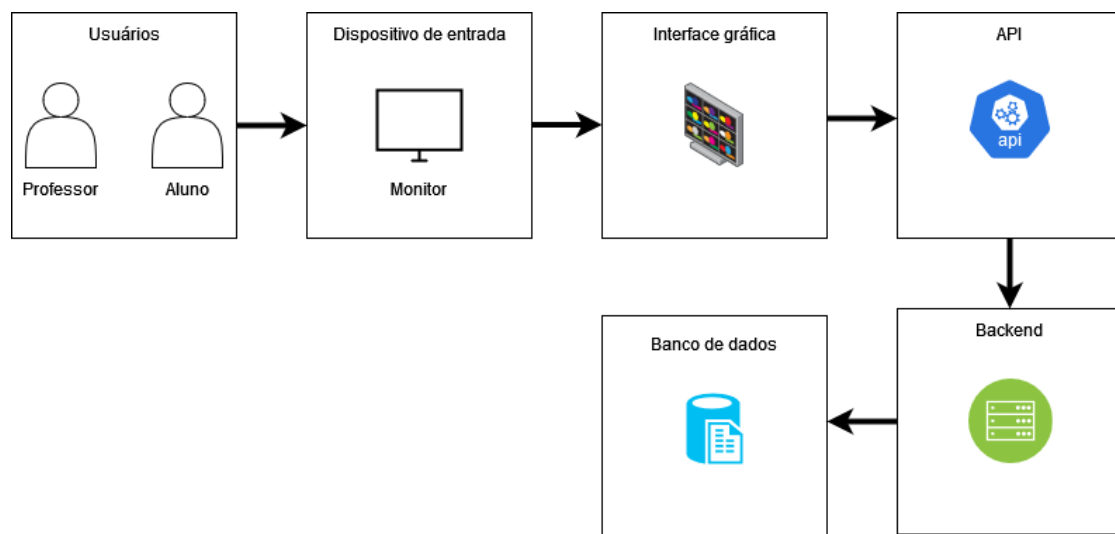
- **Editor de Código:** Para o desenvolvimento do sistema, utilizamos o Visual Studio Code, um editor de código-fonte popular que oferece extensões úteis para o desenvolvimento web. O Visual Studio Code permite a integração direta com o Git e oferece suporte para depuração, além de uma vasta gama de plugins que facilitam a codificação.

A integração dessas tecnologias possibilitou o desenvolvimento de um sistema robusto e eficiente, que atende tanto aos requisitos funcionais quanto aos não-funcionais, assegurando uma experiência de usuário otimizada e segura. O front-end foi hospedado na plataforma Vercel, proporcionando uma entrega rápida e confiável das interfaces.

3.3 Modelo de Arquitetura

Nesta seção, especificamos o funcionamento da arquitetura do *Screening Programming*. A Figura 4 mostra o modelo escolhido para o sistema, que exige dois tipos de usuários: professor e aluno. A interface inclui todas as funcionalidades essenciais para a interação entre os usuários e o sistema.

Figura 4 – Modelo de Arquitetura.



Fonte: Elaborado pelo autor (2024).

Os usuários acessam a aplicação através de dispositivos com acesso à internet, utilizando a interface do *Screening Programming*, que oferece todas as funcionalidades

necessárias, incluindo a geração de relatórios. A arquitetura foi desenvolvida como uma aplicação web para garantir acessibilidade em diversos dispositivos e facilidade de uso.

Todos os dados são armazenados em um banco de dados PostgreSQL, que organiza e processa as informações de forma segura e eficiente, garantindo suporte adequado para as operações do sistema e o gerenciamento das informações dos usuários.

3.4 Modelo de Projeto

A Figura 5 representa a tela de cadastro do sistema. Nessa interface, é possível registrar tanto alunos quanto professores. O sistema identifica automaticamente o tipo de usuário com base no domínio do e-mail fornecido: se o endereço de e-mail tiver o domínio @servidor, o usuário será classificado como professor; se o domínio for @aluno, o sistema o classificará como aluno.

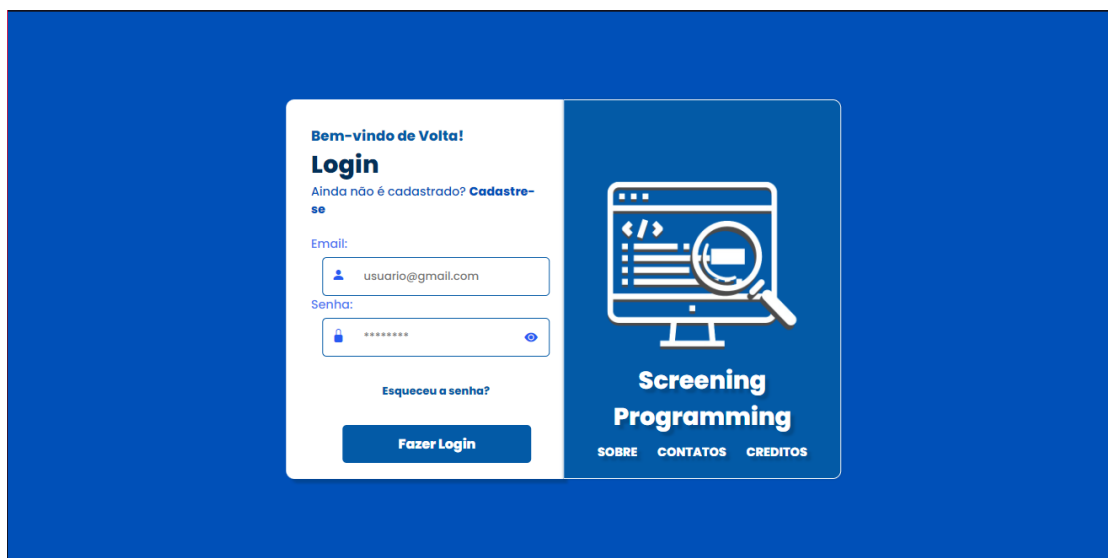
Figura 5 – Tela de cadastro

A imagem mostra a tela de cadastro do sistema 'Screening Programming'. O layout é dividido em duas seções principais. À esquerda, há um formulário branco sobre um fundo azul escuro. O formulário começa com o texto 'Bem-vindo!' e o título 'Criar Conta', seguido pela instrução 'Insira algumas informações e comece conosco!'. Os campos de entrada são: 'Nome Completo:' com um ícone de pessoa e o texto 'Nome Sobrenome'; 'Email:' com o exemplo 'usuario@gmail.com'; 'Data de Nascimento:' com o formato 'dd/mm/aaaa' e um ícone de calendário; 'Senha:' com caracteres ocultos por pontos e um ícone de olho; e 'Confirmar Senha:' com um ícone de olho. Abaixo do formulário, há dois botões: 'Voltar' e 'Cadastre-se'. À direita, há uma seção azul com um ícone de um monitor de computador exibindo código e um ícone de lupa sobre um documento. Abaixo do ícone, o texto 'Screening Programming' é exibido em branco, seguido por links para 'SOBRE', 'CONTATOS' e 'CRÉDITOS'.

Fonte: Elaborado pelo autor (2024).

A Figura 6 apresenta a tela de login do sistema, por meio da qual tanto alunos quanto professores acessam suas contas inserindo suas credenciais. Após o login, o sistema disponibiliza funcionalidades específicas de acordo com o perfil do usuário, adaptando a interface e os recursos conforme o tipo de acesso (aluno ou professor).

Figura 6 – Tela de login



Fonte: Elaborado pelo autor (2024).

A Figura 7 exibe a tela de pós-autenticação, que exige praticamente os mesmos campos para ambos os tipos de usuários, com uma única distinção: professores devem informar sua formação acadêmica, enquanto alunos precisam especificar seu período de curso.

Figura 7 – Tela de pós-autenticação

Pós Autenticação

Por favor, forneça as informações restantes para concluir o cadastro.

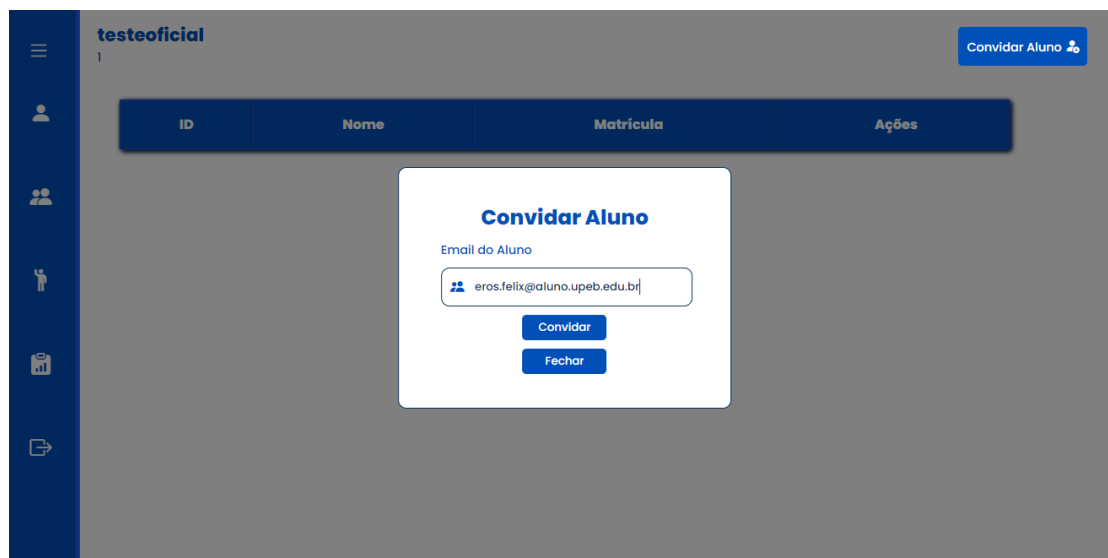
<input type="text" value="Nome Completo"/>	<input type="text" value="Nascimento"/>
<input type="text" value="Outro"/>	<input type="text" value="Período"/>
<input type="text" value="Matricula"/>	<input type="text" value="UEPB"/>
<input type="text" value="Selecione seu estado"/>	<input type="text"/>

Confirmar

Fonte: Elaborado pelo autor (2024).

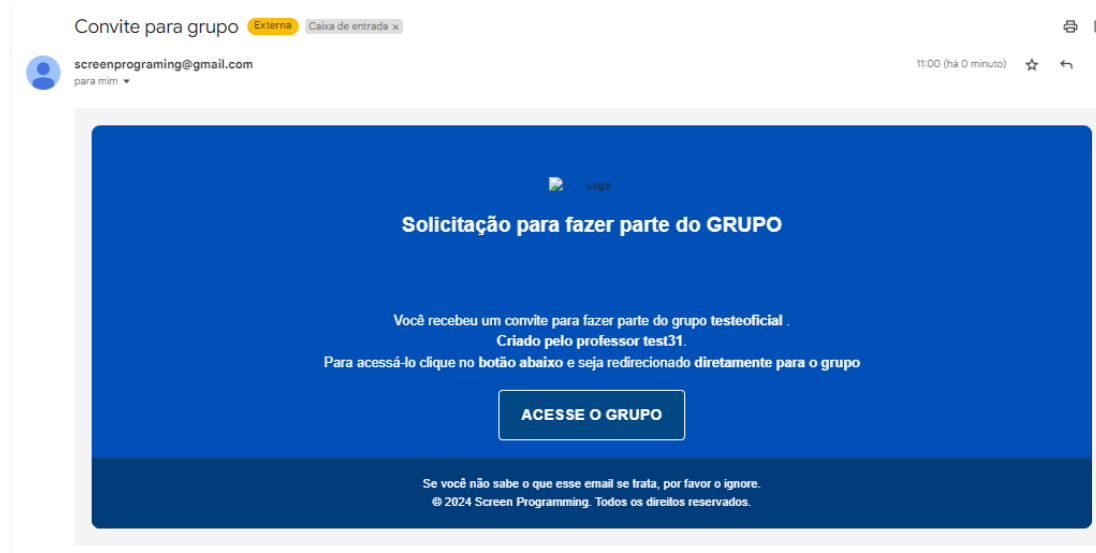
As Figuras 8 e 9 ilustram o fluxo de convite de um docente para que um aluno integre seu grupo. Após selecionar a opção "Convidar", o aluno receberá um e-mail contendo informações relevantes, como o nome do grupo e o nome do professor responsável.

Figura 8 – Tela de convite



Fonte: Elaborado pelo autor (2024).

Figura 9 – Tela do convite via email



Fonte: Elaborado pelo autor (2024).

4 METODOLOGIA

Neste capítulo, são descritos os participantes, o instrumento utilizado, os procedimentos de coleta de dados e as análises realizadas para avaliar a confiabilidade e as propriedades psicométricas do instrumento informatizado de avaliação de raciocínio lógico e matemático. O estudo busca validar um instrumento específico para mensuração de habilidades lógicas e matemáticas em alunos do ensino superior, com ênfase em cursos das áreas de ciências exatas.

4.1 Participantes

A amostra deste estudo foi composta por 80 alunos de graduação matriculados em cursos de instituições de ensino superior voltados para as ciências exatas, incluindo engenharia, física, matemática e computação. Os participantes foram selecionados por conveniência, levando em consideração a disponibilidade e acessibilidade ao estudo, e todos participaram de forma voluntária.

Os critérios de inclusão foram:

- (i) estar regularmente matriculado em um curso de graduação na área de exatas,
- (ii) possuir habilidades básicas de uso de computadores, necessárias para a navegação no ambiente informatizado, e
- (iii) concordar em participar do estudo mediante assinatura do Termo de Consentimento Livre e Esclarecido (TCLE).

Não foram aplicados critérios de exclusão específicos, exceto para casos onde limitações severas comprometessem o uso adequado da plataforma informatizada, como dificuldades significativas com dispositivos eletrônicos. Dessa forma, a amostra final representa um grupo diverso em termos de desempenho acadêmico e habilidades prévias em raciocínio lógico e matemático.

4.2 Instrumento

O instrumento de coleta de dados foi uma aplicação informatizada desenvolvida exclusivamente para avaliar habilidades de raciocínio lógico e matemático. O conteúdo

do teste foi cuidadosamente planejado para abranger: conceitos fundamentais de lógica, álgebra, resolução de problemas, e reconhecimento de padrões matemáticos.

Essas áreas fornecem uma base ampla para avaliar as competências necessárias para o raciocínio lógico-matemático. Cada questão foi elaborada para medir diferentes níveis de complexidade, incluindo desde problemas elementares até aqueles com alta demanda cognitiva, promovendo um desafio adequado para as diversas habilidades dos participantes.

O teste final foi composto por 10 itens, com cada item apresentando quatro alternativas de resposta. Esses itens foram projetados para avaliar facetas específicas da capacidade de raciocínio lógico-matemático dos alunos. Para assegurar a validade do conteúdo, o instrumento passou por uma revisão rigorosa realizada por especialistas nas áreas de educação e psicométrica, que avaliaram: clareza, relevância, e adequação dos enunciados ao público-alvo. Essa avaliação visou garantir que os itens fossem representativos das habilidades que o instrumento pretende medir, em conformidade com as boas práticas de construção de testes educacionais.

A aplicação informatizada foi desenvolvida para oferecer uma experiência de uso interativa e intuitiva, acessível em qualquer dispositivo com conexão à internet. A plataforma registrava automaticamente as respostas e o tempo de resposta de cada participante, permitindo uma coleta de dados precisa e eficiente. O tempo médio para a conclusão do teste foi de aproximadamente 48 minutos, com variações observadas conforme o nível de proficiência e engajamento dos alunos.

4.3 Procedimentos de Coleta de Dados

A coleta de dados foi realizada em um ambiente controlado, onde os participantes puderam acessar a plataforma informatizada em horários previamente agendados. Antes do início do teste, cada participante recebeu orientações sobre o funcionamento da plataforma e os objetivos do estudo. Durante a aplicação, um monitor estava disponível para esclarecer eventuais dúvidas quanto ao uso da plataforma, mas sem interferir no processo de resposta dos participantes.

Após a conclusão do teste, as respostas foram armazenadas de forma segura e confidencial para análise posterior. Todas as etapas do processo de coleta de dados respeitaram as normas éticas e legais para estudos envolvendo seres humanos, incluindo a

obtenção do Termo de Consentimento Livre e Esclarecido (TCLE), bem como a garantia de anonimato e confidencialidade das informações fornecidas.

4.4 Análise de Dados

Realizamos a análise de dados utilizando métodos combinados da TCT e da TRI, garantindo uma avaliação abrangente e detalhada da validade e confiabilidade do instrumento.

Na TCT, calculamos o coeficiente alfa de Cronbach para avaliar a consistência interna dos itens, definindo um valor mínimo de 0,7. Também calculamos a média e o desvio padrão das respostas dos participantes, o que nos permitiu observar a distribuição dos escores e entender a centralidade do desempenho no teste como um todo.

Na TRI, estimamos três parâmetros para cada item: a dificuldade (b), que indica o nível de habilidade necessário para responder corretamente; a discriminação (a), que mostra como o item distingue entre diferentes níveis de habilidade; e a chance de acerto ao acaso (c), que representa a probabilidade de resposta correta sem base em conhecimento. Esses parâmetros ajudaram a identificar quais itens são eficazes para avaliar diferentes níveis de habilidade.

Também calculamos a correlação ponto-bisserial para cada item, o que nos permitiu analisar a relação entre o acerto em um item e o desempenho geral no teste. Esse parâmetro foi essencial para ajustar o instrumento, aprimorando a capacidade dos itens de discriminar entre variados níveis de habilidade.

Para a análise estatística, utilizamos os pacotes Pandas, NumPy e SciPy em Python, que facilitaram a execução precisa dos cálculos necessários para TCT e TRI. Automatizamos os processos por meio de scripts que geraram gráficos e relatórios, permitindo visualizar padrões e tendências de forma clara. Todos os dados foram armazenados em servidores seguros, com backups automáticos, e relatórios visuais, incluindo gráficos de pontuação média, tempos de resposta e discriminação de itens, forneceram uma visão detalhada do desempenho dos participantes.

5 RESULTADOS E DISCUSSÕES

Neste capítulo, apresentamos uma análise detalhada dos resultados obtidos a partir da aplicação do instrumento informatizado para avaliação das habilidades de raciocínio lógico e matemático. Esta seção explora a confiabilidade e a consistência interna do instrumento, analisadas por meio da Teoria Clássica dos Testes (TCT), e as propriedades psicométricas dos itens, avaliadas pela Teoria de Resposta ao Item (TRI). Ambas as abordagens complementares fornecem uma visão abrangente sobre a eficácia e precisão do instrumento para medir as habilidades cognitivas dos participantes.

5.1 Confiabilidade do Instrumento na Mensuração das Habilidades de Raciocínio Lógico-Matemático

A confiabilidade do instrumento foi avaliada através do coeficiente alfa de Cronbach, que resultou em um valor de 0,757. Esse coeficiente indica uma boa consistência interna entre os itens, o que assegura que o instrumento está efetivamente medindo as habilidades que se propõe a avaliar. Segundo a literatura, valores de alfa superiores a 0,70 são considerados aceitáveis para instrumentos em fase de desenvolvimento, como é o caso deste estudo. Desse modo, podemos afirmar que os itens apresentam uma correlação satisfatória entre si, reforçando a adequação do instrumento para a avaliação de habilidades de raciocínio lógico e matemático. Esse nível de confiabilidade é especialmente relevante, pois contribui para diminuir a possibilidade de erros de medição e garantindo uma maior precisão nas avaliações dos participantes.

Além disso, esse resultado está próximo ao limite superior de aceitabilidade, indicando que o instrumento já possui uma estrutura bem definida e que os itens são suficientemente coerentes entre si. Sugerindo que, em aplicações futuras, a confiabilidade do instrumento poderia ser melhorada com a inclusão de novos itens que ampliem a abrangência do conteúdo avaliado, ou com ajustes nos itens existentes para melhorar a homogeneidade do conjunto. Esse nível de consistência interna também apoia o uso do instrumento em contextos educacionais de ensino superior, oferecendo um suporte para pesquisas futuras e aplicações práticas voltadas para a medição de habilidades de raciocínio lógico e matemático.

Para aprofundar a análise psicométrica e explorar a eficácia dos itens individu-

almente, apresentamos na Tabela 1 os valores da proporção de acertos e a correlação ponto-bisserial para cada item do instrumento. A proporção de acertos indica a facilidade de cada item, enquanto a correlação ponto-bisserial fornece uma medida da capacidade discriminativa de cada questão, ou seja, o quanto o item é eficaz em distinguir entre participantes com diferentes níveis de habilidade.

Tabela 1 – Análise Psicométrica dos Itens com Base na TCT.

Id	Proporção de Acertos	Correlação Ponto Bisserial
Mat-001	0,816	0,525
Mat-002	0,575	0,479
Mat-003	0,555	0,210
Mat-004	0,401	0,307
Mat-005	0,773	0,549
Mat-006	0,836	0,635
Mat-007	0,749	0,333
Mat-008	0,686	0,409
Mat-009	0,706	0,688
Mat-010	0,445	0,236

Fonte: Elaborada pelo autor (2024).

A análise dos valores na Tabela 1 permite algumas observações relevantes. Primeiramente, a proporção de acertos varia entre 0,401 (Mat-004) e 0,836 (Mat-006), indicando que os itens abrangem diferentes níveis de dificuldade. Essa diversidade é importante para que o instrumento seja capaz de capturar uma gama ampla de habilidades, incluindo tanto participantes com níveis iniciais quanto avançados em raciocínio lógico e matemático. Itens como o Mat-006, com uma alta proporção de acertos, podem ser considerados mais fáceis, enquanto itens como o Mat-004, com uma proporção mais baixa, apresentam maior desafio aos respondentes.

No que se refere à correlação ponto-bisserial, observamos que os valores variam entre 0,210 (Mat-003) e 0,688 (Mat-009). A correlação ponto-bisserial indica o quanto um item é eficaz para discriminar entre participantes com diferentes níveis de habilidade.

De maneira geral, valores acima de 0,30 são considerados aceitáveis, enquanto valores superiores a 0,50 são considerados bons. Portanto, a maioria dos itens apresenta correlações ponto-bisserial que indicam uma discriminação adequada. No entanto, itens como o Mat-003, com uma correlação ponto-bisserial de 0,210, podem ter uma capacidade discriminativa limitada e poderiam ser revisados para melhorar sua eficácia. Já o item Mat-009, com alta correlação ponto-bisserial (0,688), destaca-se por sua excelente capacidade de discriminação, auxiliando na identificação de alunos com diferentes níveis de habilidade em raciocínio lógico e matemático, o que é valioso para intervenções pedagógicas específicas.

Essas análises ressaltam que o instrumento possui uma estrutura consistente e que a maioria dos itens é eficaz tanto em termos de dificuldade quanto de discriminação. Contudo, a análise também indica oportunidades para ajustes finos em itens específicos, o que poderia potencialmente elevar o nível de confiabilidade e precisão do instrumento. Essa melhoria contínua é fundamental para garantir que o instrumento permaneça adequado e relevante em aplicações futuras, especialmente em estudos que buscam monitorar o desenvolvimento de habilidades de raciocínio lógico e matemático em diferentes contextos educacionais.

5.2 Validade e Propriedades Psicométricas dos Itens: Indicadores de Habilidades de Raciocínio Lógico-Matemático

A análise das propriedades psicométricas dos itens foi realizada utilizando o modelo logístico de três parâmetros (3PL) da TRI, o qual fornece uma visão detalhada sobre as características de cada item, considerando os parâmetros de discriminação (a), dificuldade (b) e acerto ao acaso (c). Esses parâmetros permitem uma avaliação precisa e informada sobre a qualidade e o desempenho dos itens no contexto do teste aplicado. Na Tabela 2, apresentamos os parâmetros dos itens.

Os resultados demonstram que a maioria dos itens possui uma discriminação (a) superior a 1,0, indicando que eles são eficazes em diferenciar participantes com distintos níveis de habilidade. O item Mat-009 apresentou o maior valor de discriminação ($a = 3,867$), destacando-se por sua eficácia em distinguir entre participantes com habilidades altas e baixas em raciocínio lógico e matemático. Esse parâmetro é particularmente relevante em avaliações educacionais, pois itens altamente discriminativos contribuem

para intervenções pedagógicas mais precisas e eficazes.

Quanto à dificuldade (b), observa-se uma ampla variação, com valores entre -1,068 e 1,090. O item Mat-001, com uma dificuldade de -1,032, foi identificado como o mais fácil, enquanto o item Mat-010, com uma dificuldade de 1,090, foi o mais difícil. Essa diversidade permite ao instrumento captar uma gama variada de habilidades entre os participantes, desde os menos habilidosos até aqueles com um nível avançado de raciocínio lógico e matemático.

Tabela 2 – Parâmetros Psicométricos dos Itens com Base na TRI.

Id	a	b	c
Mat-001	2,512	-1,032	0,140
Mat-002	1,365	0,005	0,145
Mat-003	0,696	0,177	0,170
Mat-004	1,223	0,926	0,148
Mat-005	1,715	-1,007	0,155
Mat-006	2,656	-1,068	0,158
Mat-007	0,964	-0,955	0,165
Mat-008	1,704	-0,468	0,169
Mat-009	3,367	-0,521	0,135
Mat-010	0,844	1,090	0,175

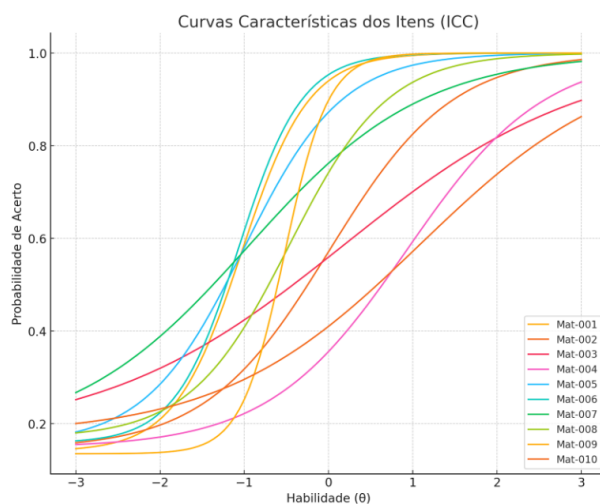
Fonte: Elaborada pelo autor (2024).

No que tange à probabilidade de acerto ao acaso (c), todos os itens apresentaram valores abaixo de 0,40, com o item Mat-008 registrando a maior probabilidade de acerto por chute (c = 0,169). Esses valores indicam que a chance de um participante acertar uma questão por adivinhação é reduzida, o que minimiza a influência do acaso e reforça a validade do instrumento.

Para uma melhor compreensão da qualidade dos itens, apresentamos as CCI na Figura 10, que ilustram a relação entre os diferentes níveis de habilidade e a probabilidade de acerto em cada item. As CCI destacam as variações nos parâmetros de discriminação, dificuldade e acerto ao acaso, fornecendo uma visualização das características psicométricas dos itens.

A análise das CCI indicou que o item Mat-003 foi o menos discriminativo, sugerindo que ele pode não ser tão eficaz em diferenciar participantes com distintos níveis de habilidade. Esse achado indica a necessidade de uma possível revisão ou substituição do item para melhorar sua eficácia. Em contraste, o item Mat-009 apresentou um alto valor de discriminação, destacando-se como um excelente item para avaliar participantes com diferentes níveis de habilidade, o que é desejável em avaliações educacionais abrangentes.

Figura 10 – CCI - Todos os itens para habilidade Raciocínio Lógico e Matemático.

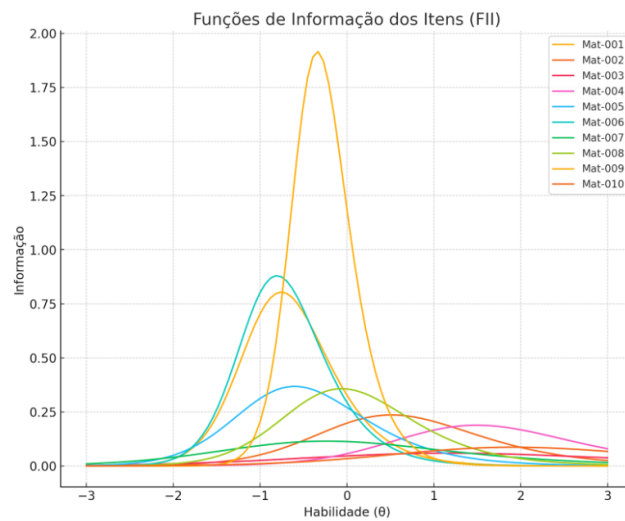


Fonte: Elaborada pelo autor (2024).

A Figura 11 exibe as FII, as quais ilustram a quantidade de informação fornecida por cada item em diferentes níveis do traço latente. A análise das FII permite identificar quais itens são mais eficazes em termos de capacidade de informação, contribuindo para um entendimento mais profundo sobre a eficácia de cada questão na avaliação das habilidades cognitivas dos participantes.

A FII revelou que o item Mat-003 oferece maior informação em níveis medianos de habilidade, tornando-o adequado para avaliar participantes com habilidades intermediárias. Já o item Mat-009 se destaca por fornecer informações em toda a faixa de habilidade, sendo particularmente útil para um instrumento de avaliação abrangente que deseja capturar perfis variados de habilidade.

Figura 11 – FII - Todos os itens para habilidade Raciocínio Lógico e Matemático.



Fonte: Elaborada pelo autor (2024).

Os resultados apresentados evidenciam a eficácia e a confiabilidade do instrumento desenvolvido para medir habilidades de raciocínio lógico e matemático em alunos de nível superior. As análises realizadas oferecem subsídios valiosos para melhorias contínuas no instrumento, garantindo sua adequação e aplicabilidade em contextos acadêmicos. Esse instrumento pode contribuir significativamente para a avaliação das competências dos estudantes, auxiliando na identificação de áreas que necessitam de atenção específica e na formulação de estratégias pedagógicas mais eficazes e direcionadas.

6 CONSIDERAÇÕES FINAIS E SUGESTÕES PARA TRABALHOS FUTUROS

Este estudo propôs e implementou um instrumento informatizado específico para mensurar habilidades de raciocínio lógico e matemático em alunos de graduação, com o objetivo de proporcionar uma avaliação mais detalhada e precisa dessas competências. Por meio da utilização de duas abordagens teóricas complementares — a TCT e a TRI — foi possível obter evidências sobre a confiabilidade e validade psicométrica dos itens que compõem o instrumento. Desta forma:

- **QP1.** A análise pela TCT, particularmente com o uso do coeficiente alfa de Cronbach, demonstrou que o instrumento apresenta alta consistência interna, o que significa que os itens estão bem correlacionados entre si e conseguem medir de forma coerente as habilidades propostas. Esse resultado indica que o instrumento pode ser aplicado repetidamente em contextos similares com resultados consistentes, um aspecto essencial para pesquisas e avaliações acadêmicas de qualidade.
- **QP2.** Além disso, a análise pela TRI permitiu uma avaliação mais detalhada das características dos itens. Essa abordagem forneceu uma visão mais refinada da dificuldade e da capacidade de discriminação dos itens, além de uma estimativa da probabilidade de acerto ao acaso. Esses aspectos tornam o instrumento especialmente útil, pois ele não apenas mede o desempenho dos participantes, mas também ajusta a dificuldade conforme o nível de habilidade dos alunos, oferecendo uma avaliação adaptativa. Essa característica é particularmente valiosa para captar nuances do raciocínio lógico-matemático em diferentes níveis de habilidade, o que amplia as possibilidades de utilização do instrumento em contextos acadêmicos diversificados.

Os resultados obtidos apontam para a eficácia do instrumento como uma ferramenta prática e moderna para avaliação em larga escala. A informatização do processo, além de facilitar a aplicação e a coleta de dados, permite um processamento mais rápido e detalhado das respostas, possibilitando que educadores e pesquisadores tenham acesso a dados de alta qualidade para fundamentar decisões pedagógicas e de pesquisa.

Futuras investigações poderão explorar o uso do instrumento em diferentes contextos educacionais, como em cursos técnicos ou mesmo em programas de pós-graduação,

e com amostras mais diversificadas, para que se possa consolidar seu impacto e versatilidade na área de avaliação educacional. A aplicação do instrumento em grupos maiores e variados poderia não apenas validar ainda mais suas propriedades psicométricas, mas também identificar possíveis ajustes para adaptar o instrumento a diferentes públicos.

Dessa forma, o instrumento desenvolvido neste estudo se revela promissor para futuros estudos e aplicações práticas, contribuindo significativamente para uma mensuração mais precisa e adaptativa das habilidades cognitivas no contexto acadêmico, permitindo uma avaliação mais justa e objetiva das competências dos estudantes.

REFERÊNCIAS

- ARAÚJO, A. L. S. O. et al. Metodologia de pesquisa em informática na educação: Abordagem quantitativa de pesquisa. In: _____. Porto Alegre: SBC, 2019. cap. Teoria de Resposta ao Item. Citado 6 vezes nas páginas 16, 22, 23, 24, 25 e 26.
- BAKER, F. B. *The Basics of Item Response Theory*. [S.l.]: ERIC, 2001. Citado 4 vezes nas páginas 16, 23, 24 e 25.
- BARCELOS, T. S. et al. Mathematics learning through computational thinking activities: A systematic literature review. *J. Univers. Comput. Sci.*, v. 24, n. 7, p. 815–845, 2018. Citado na página 14.
- BRANDÃO, H. M. de L.; GUEDES, L. G. de R. et al. Analysis of student's proficiency evaluation model for adaptative testing. *Brazilian Journal of Development*, v. 6, n. 3, p. 10608–10620, 2020. Citado na página 26.
- BRONKHORST, H. et al. Logical reasoning in formal and everyday reasoning tasks. *International Journal of Science and Mathematics Education*, Springer, v. 18, p. 1673–1694, 2020. Citado na página 19.
- CHAPMAN, R. Expected a posteriori scoring in promis®. *Journal of Patient-Reported Outcomes*, Springer, v. 6, n. 1, p. 59, 2022. Citado na página 27.
- CHEW, S. L.; CERBIN, W. J. The cognitive challenges of effective teaching. *The Journal of Economic Education*, Taylor & Francis, v. 52, n. 1, p. 17–40, 2021. Citado na página 22.
- CHUNG, L. et al. *Non-functional requirements in software engineering*. [S.l.]: Springer Science & Business Media, 2012. v. 5. Citado na página 30.
- CRESSWELL, C.; SPEELMAN, C. P. Does mathematics training lead to better logical thinking and reasoning? a cross-sectional assessment from students to professors. *PloS one*, Public Library of Science San Francisco, CA USA, v. 15, n. 7, p. e0236153, 2020. Citado 2 vezes nas páginas 19 e 20.
- CUNHA, C. M.; NETO, O. P. de A.; STACKFLETH, R. Principais métodos de avaliação psicométrica da validade de instrumentos de medida. *Revista de Atenção à Saúde*, v. 14, n. 47, p. 75–83, 2016. Citado na página 14.
- DANTAS, D. M. et al. Screening programming's reliability to measure predictive programming skills. In: SBC. *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*. [S.l.], 2023. p. 1779–1788. Citado na página 15.

EDELSBRUNNER, P. A.; DABLANDER, F. The psychometric modeling of scientific reasoning: A review and recommendations for future avenues. *Educational Psychology Review*, Springer, v. 31, p. 1–34, 2019. Citado na página 14.

EL-DEN, S. et al. How to measure a latent construct: psychometric principles for the development and validation of measurement instruments. *International Journal of Pharmacy Practice*, Oxford University Press, v. 28, n. 4, p. 326–336, 2020. Citado na página 14.

EPPE, M. et al. Intelligent problem-solving as integrated hierarchical reinforcement learning. *Nature Machine Intelligence*, Nature Publishing Group UK London, v. 4, n. 1, p. 11–20, 2022. Citado na página 13.

FATIMAH, S.; JOHAR, R.; ZUBAINUR, C. Students' logical mathematical intelligence in completing mathematical problems with natural disaster context. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2020. v. 1470, n. 1, p. 012022. Citado na página 21.

FIRDAUS, F. et al. Developing critical thinking skills of students in mathematics learning. *Journal of Education and Learning (EduLearn)*, v. 9, n. 3, p. 226–236, 2015. Citado na página 20.

GARCÍA-CARMONA, A. Scientific thinking and critical thinking in science education: Two distinct but symbiotically related intellectual processes. *Science & Education*, Springer, p. 1–19, 2023. Citado na página 20.

HALEEM, A. et al. Understanding the role of digital technologies in education: A review. *Sustainable operations and computers*, Elsevier, v. 3, p. 275–285, 2022. Citado na página 14.

HENDRYCKS, D. et al. Measuring mathematical problem solving with the math dataset. *arXiv preprint arXiv:2103.03874*, 2021. Citado na página 13.

HOLOVINA, O.; KAMENOVA, D. Critical and logical thinking formation as the educational competence in the modern training system for lawyers. , 2021. Citado na página 13.

HOUDÉ, O. Brain and cognitive development: Logicomathematical intelligence. *Cognitive and Systems Neuroscience*, p. 561. Citado na página 21.

HUBER, S. G.; SKEDSMO, G. Standardization and assessment practices. *Educational Assessment, Evaluation and Accountability*, Springer, v. 29, p. 1–3, 2017. Citado na página 15.

JIANG, Y. et al. A multidimensional irt model for ability-item-based guessing: the development of a two-parameter logistic extension model. *Communications in statistics-simulation and computation*, Taylor & Francis, v. 53, n. 7, p. 3068–3080, 2024. Citado na página 25.

JUNPENG, P. et al. Validation of a digital tool for diagnosing mathematical proficiency. *International Journal of Evaluation and Research in Education*, ERIC, v. 9, n. 3, p. 665–674, 2020. Citado na página 15.

KANLI, E. Assessment of creativity: theories and methods. *Creativity-A Force to Innovation*, IntechOpen, 2020. Citado na página 24.

KERN, J. L.; CHOE, E. Using a response time-based expected a posteriori estimator to control for differential speededness in computerized adaptive test. *Applied Psychological Measurement*, Sage Publications Sage CA: Los Angeles, CA, v. 45, n. 5, p. 361–385, 2021. Citado na página 27.

KIM, Y.-M. Validation of psychometric research instruments: The case of information science. *Journal of the American Society for Information Science and Technology*, Wiley Online Library, v. 60, n. 6, p. 1178–1191, 2009. Citado na página 14.

LIMA, F. F. d.; OSÓRIO, F. d. L. Empathy: assessment instruments and psychometric quality—a systematic literature review with a meta-analysis of the past ten years. *Frontiers in psychology*, Frontiers Media SA, v. 12, p. 781346, 2021. Citado na página 15.

LIN, C. The complete structure of mathematical thinking. In: *Intellectual Development and Mathematics Learning*. [S.l.]: Springer, 2023. p. 79–111. Citado na página 21.

LIN, C. *Intellectual Development and Mathematics Learning*. [S.l.]: Springer, 2023. Citado 2 vezes nas páginas 20 e 22.

LUND, J. Systems engineering knowledge and skills for water and environmental problems. *Civil Engineering and Environmental Systems*, Taylor & Francis, v. 37, n. 4, p. 183–196, 2020. Citado na página 20.

MAQUINÉ, G. O.; COELHO, I. M. W. da S. Ferramentas tecnológicas a serviço da avaliação no contexto de ensino-aprendizagem: Um estudo comparativo. *Informática na educação: teoria & prática*, v. 23, n. 3 Set/Dez, 2020. Citado na página 14.

MARKOVITS, H. The development of logical reasoning. In: *International handbook of thinking and reasoning*. [S.l.]: Routledge, 2017. p. 383–400. Citado na página 19.

- MITREVSKI, B. Teaching critical thinking and problem solving in physics. In: AIP PUBLISHING. *AIP Conference Proceedings*. [S.l.], 2019. v. 2075, n. 1. Citado na página 20.
- MITSEA, E.; DRIGAS, A.; MANTAS, P. Soft skills & metacognition as inclusion amplifiers in the 21 st century. *International Journal of Online & Biomedical Engineering*, v. 17, n. 4, 2021. Citado na página 13.
- MONTICONE, M. et al. Psychometric properties of assessment tools. *Measuring spinal cord injury: a practical guide of outcome measures*, Springer, p. 7–15, 2021. Citado na página 15.
- NITE, S. B. et al. Spatial visualization in informal learning. In: IEEE. *2023 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2023. p. 1–5. Citado na página 21.
- NURLAELAH, E.; USDIYANA, D.; FADILAH, N. The relationship between computational thinking ability and logical mathematical intelligence. *Mosharafa: Jurnal Pendidikan Matematika*, v. 13, n. 1, p. 87–96, 2024. Citado na página 13.
- OLTEANU, C. Programming, mathematical reasoning and sense-making. *International Journal of Mathematical Education in Science and Technology*, Taylor & Francis, v. 53, n. 8, p. 2046–2064, 2022. Citado na página 13.
- PASQUALI, L. *Psicometria: Teoria dos Testes na Psicologia e na Educação*. [S.l.]: Editora Vozes Limitada, 2017. Citado 4 vezes nas páginas 16, 23, 25 e 28.
- POLAT, M. Comparison of performance measures obtained from foreign language tests according to item response theory vs classical test theory. *International Online Journal of Education and Teaching*, ERIC, v. 9, n. 1, p. 471–485, 2022. Citado na página 25.
- PONNIAH, K.; SAFEEK, N. M. M. Exploring theoretical and conceptual frameworks for the enhancement of high-order thinking skills through the study of thirukkural. *International Journal of Academic Research in Progressive Education and Development*, v. 13, n. 2, 2024. Citado na página 21.
- PRADO, J. The development of the reasoning brain and how to foster logical reasoning skills. *development*, 2019. Citado na página 19.
- PRIMI, R. *Psicometria: Fundamentos matemáticos da teoria clássica dos testes. Avaliação Psicológica*, Instituto Brasileiro de Avaliação Psicológica, v. 11, n. 2, 2012. Citado 2 vezes nas páginas 16 e 22.
- PROENÇA, M. C. d. Habilidades matemáticas na resolução de problemas: análise da compreensão de futuros professores. *Bolema: Boletim de Educação Matemática*, SciELO Brasil, v. 36, n. 74, p. 1135–1157, 2022. Citado na página 22.

- REYES, E. S. Technology, creativity and logical-mathematical thinking: a triad to rethink. *Mendive-Revista de Educacion*, v. 12, n. 1, 2023. Citado na página 13.
- REZENDE, D. A. *Engenharia de software e sistemas de informação*. [S.l.]: Brasport, 2006. Citado 2 vezes nas páginas 30 e 33.
- RITTER, F. E. et al. Cognition: Mental representations, problem solving, and decision making. *Foundations for Designing User-Centered Systems: What System Designers Need to Know about People*, Springer, p. 165–200, 2014. Citado na página 20.
- SILLAT, L. H.; TAMMETS, K.; LAANPERE, M. Digital competence assessment methods in higher education: A systematic literature review. *Education Sciences*, MDPI, v. 11, n. 8, p. 402, 2021. Citado na página 13.
- SILVA, S. d. O. A.; LUNA, S. V. d. Correlação entre o raciocínio lógico e o raciocínio matemático em crianças escolarizadas. *Bolema: Boletim de Educação Matemática*, SciELO Brasil, v. 33, p. 1047–1066, 2019. Citado na página 20.
- SKEDSMO, G.; HUBER, S. G. Measuring teaching quality: Some key issues. *Educational Assessment, Evaluation and Accountability*, Springer, v. 31, p. 151–153, 2019. Citado na página 14.
- STRIETHOLT, R.; ROSÉN, M.; GLADUSHYNA, O. The integrity of educational outcome measures in international assessments. *Educational Assessment, Evaluation and Accountability*, Springer, v. 33, p. 1–7, 2021. Citado na página 14.
- TALLMAN, M. A.; O'BRYAN, A. E. Reflected abstraction. In: *Piaget's Genetic Epistemology for Mathematics Education Research*. [S.l.]: Springer, 2024. p. 239–288. Citado na página 21.
- TIGRERO, J. M. A. et al. Digital educational resources in the development of mathematical logical thinking in middle school students. *Revista Iberoamericana de educación*, v. 8, n. 4, p. 12–23, 2024. Citado na página 13.
- TZURIEL, D.; TZURIEL, D. The socio-cultural theory of vygotsky. *Mediated learning and cognitive modifiability*, Springer, p. 53–66, 2021. Citado na página 21.
- VIEIRA, K. M.; JÚNIOR, F.; POTRICH, A. C. G. Measuring financial literacy: Proposition of an instrument based on the item response theory. *Ciência E Nat*, v. 42, p. 38, 2020. Citado na página 25.
- VYGOTSKY, L. The development of thinking and the formation of concepts in the adolescent. In: *LS Vygotsky's Pedological Works, Volume 4: Pedology of the Adolescent II: Pedology of the Transitional Age as a Psychological and Social Problem*. [S.l.]: Springer, 2024. p. 41–179. Citado na página 21.

WANG, Y.-X.; YIN, B. A new understanding of the cognitive reappraisal technique: an extension based on the schema theory. *Frontiers in Behavioral Neuroscience*, Frontiers Media SA, v. 17, p. 1174585, 2023. Citado na página 21.

WOLCOTT, M. D.; OLSEN, A. A.; AUGUSTINE, J. M. Item response theory in high-stakes pharmacy assessments. *Currents in Pharmacy Teaching and Learning*, Elsevier, v. 14, n. 9, p. 1206–1214, 2022. Citado 2 vezes nas páginas 16 e 24.

YASIN, M.; NUSANTARA, T. Characteristics of pattern recognition to solve mathematics problems in computational thinking. In: AIP PUBLISHING. *AIP Conference Proceedings*. [S.l.], 2023. v. 2569, n. 1. Citado na página 20.

ZARETSKY, V. K. One more time on the zone of proximal development. *Cultural-Historical Psychology*, v. 17, n. 2, 2021. Citado na página 21.