



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO PROFISSIONAL EM
MATEMÁTICA EM REDE NACIONAL – PROFMAT**

ELIOENAI DE QUEIROZ ARAÚJO

ANÁLISES PSICOMÉTRICAS DO EXAME CLASSIFICATÓRIO DO IFPI

**FLORIANO
2023**

ELIOENAI DE QUEIROZ ARAÚJO

ANÁLISES PSICOMÉTRICAS DO EXAME CLASSIFICATÓRIO DO IFPI

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática em Rede Nacional (PROFMAT) do Instituto Federal do Piauí/ Campus Floriano, como parte integrante dos requisitos para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Igor Ferreira do Nascimento

FLORIANO
2023

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Araújo, Elioenai de Queiroz

A658a Análises psicométricas do exame classificatório do IFPI / Elioenai de Queiroz Araújo. - 2023.
53 p.: il. color.

Dissertação (Mestrado Profissional em Matemática) - Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Floriano, 2023.

Orientador : Prof Dr. Igor Ferreira do Nascimento.

1. Teoria Clássica do Teste. 2. Teoria de Resposta ao Item. 3. Exame Classificatório do IFPI. 4. Análise de Itens. I.Título.

CDD - 510

Elaborado por Neuda Fernandes Dias CRB 3/1375

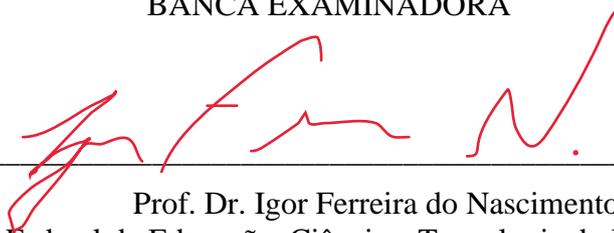
ELIOENAI DE QUEIROZ ARAÚJO

ANÁLISES PSICOMÉTRICAS DO EXAME CLASSIFICATÓRIO DO IFPI

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática em Rede Nacional (PROFMAT) do Instituto Federal do Piauí/*Campus* Floriano, como parte integrante dos requisitos para obtenção do título de Mestre em Matemática.

Aprovada em: 11/09/2023

BANCA EXAMINADORA



Prof. Dr. Igor Ferreira do Nascimento
Instituto Federal de Educação, Ciência e Tecnologia do Piauí – IFPI
Orientador

Roberto Arruda Lima Soares

Assinado de forma digital por Roberto Arruda Lima Soares
Dados: 2023.09.13 17:32:52 -03'00'

Prof. Dr. Roberto Arruda Lima Soares
Instituto Federal de Educação, Ciência e Tecnologia do Piauí – IFPI
Avaliador Interno

Documento assinado digitalmente



RUI MARQUES CARVALHO

Data: 17/09/2023 11:45:09-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Rui Marques Carvalho
Instituto Federal de Educação, Ciência e Tecnologia do Piauí – IFPI
Avaliador Interno

DocuSigned by:

NEI CARLOS DOS SANTOS ROCHA

157F10EA4192470...

Prof. Dr. Nei Carlos dos Santos Rocha
Universidade Federal do Rio de Janeiro – UFRJ
Avaliador Externo

AGRADECIMENTOS

Após dedicar anos a estudos intensos e incansável empenho na concretização deste mestrado, expresso minha profunda gratidão àqueles que me acompanharam e desempenharam papéis cruciais nesta jornada. Por meio destas palavras sinceras, compartilho um fragmento da importância que essas pessoas tiveram e ainda têm nessa conquista, expressando minha mais sincera e profunda gratidão.

Em primeiro lugar, estendo meu agradecimento a Deus pela dádiva da vida e pelas oportunidades maravilhosas que Ele sempre me concede, como esta de empreender e concluir o Mestrado. Os desafios ao longo do percurso foram numerosos, contudo, sua presença constante jamais me permitiu sucumbir.

A meus pais, que invariavelmente desdobraram esforços para prover a melhor educação a seus filhos, especialmente à minha adorada e admirável mãe, Juraci Cavalcante de Queiroz Araújo, e a meu pai, Paulo Henrique de Araújo. Suas orações poderosas e exemplos de vida moldaram minha trajetória de maneira inestimável.

A meus filhos, João Vitor de Queiroz Araújo Batista e João Guilherme de Queiroz Araújo Batista, manifesto minha gratidão pela força e compreensão diante de minhas ausências. Foram inúmeras as ocasiões em que precisei priorizar estudos e leituras incessantes, em detrimento da participação em momentos familiares.

Dirijo meu agradecimento ao ilustre orientador, o Prof. Doutor Igor Ferreira do Nascimento, cuja paciência e dedicação foram imensuráveis. Mesmo em meio a uma agenda repleta de compromissos, sua presença constante para esclarecer minhas dúvidas foi notável. Sua orientação se revelou essencial para a conclusão desta dissertação.

Manifesto gratidão à Professora Janilde, diretora da escola de Caxias, cujo apoio foi inestimável.

Estendo meu reconhecimento ao Amigo Professor Juarez, cuja força e apoio se fizeram presentes ao me hospedar em Floriano.

Agradeço aos Professores Dr. Roberto Arruda e Dr. Ezequias, cuja ajuda e colaboração merecem destaque especial.

Aos estimados professores Ifpianos, Ronaldo Campelo, Fábio Luz, Gildon e Maria Cezar, expromo minha sincera gratidão por exigirem de mim mais do que eu mesmo acreditava ser capaz de realizar. Minha eterna gratidão pelo compartilhamento de conhecimento, tempo e amizade.

Dirijo minha gratidão aos colegas de jornada, alunos Ifpianos Ana Karina, Adriano, Daniel, Dhone, Evanildo, Joenilson, Josiel, Luiz Carlos Araújo, Luís Carlos Barbosa e Valderi. Vocês desempenharam papéis significativos em meu crescimento, e esse papel merece ser recompensado com minha eterna gratidão.

Por fim, à instituição IFPI e à família Ifipiana, manifesto minha gratidão pela oportunidade concedida, pela atenção, desprendimento e simpatia que sempre permearam nossa convivência. Companheiros de trabalho e irmãos na amizade, vocês fizeram parte de minha formação e continuarão presentes em minha trajetória, sem dúvida alguma.

“Os acertos em um teste fornecem um ponto de partida para avaliar o desempenho, mas é o desenvolvimento contínuo das habilidades e competências que realmente determina o sucesso de um participante”.

John Wooden

.

RESUMO

ARAÚJO, Elioenai de Queiroz. **Análises psicométricas do exame classificatório do IFPI**, 2023. 53 f. Dissertação (Mestrado) – Instituto Federal do Piauí – Campus Floriano, Floriano, 2023.

Para ingressar em curso técnico de uma instituição pública de ensino no Brasil os candidatos aos cursos são submetidos a um teste classificatório. Com a expansão do Instituto Federal do Piauí este exame vem abrangendo cidades em todo o território piauiense. Assim, a ideia desse estudo é buscar elementos necessários para explorar os resultados do teste classificatório e, com tais resultados, melhorar o ordenamento dos alunos, permitir dar um resultado maior do que a nota do aluno e ainda fomentar políticas da instituição. Desse modo, o objetivo deste estudo se configura em uma investigação detalhada e examina a validade, confiabilidade e equidade desses exames como meios de avaliar o potencial dos candidatos. A metodologia da pesquisa tem como fundamentos teóricos a Teoria Clássica do Teste e Teoria de Resposta ao Item, bem como elementos da pesquisa quantitativa. Para análise e estudo dos dados usaremos softwares estatísticos como suporte tecnológico. Esperamos com esta investigação construir um modelo específico de análise a partir de dados coletados no exame classificatório do Instituto Federal do Piauí para que em provas futuras os organizadores possam realizar uma melhor avaliação educacional e consequentemente direcionar as políticas públicas da instituição.

Palavras-chave: Teoria Clássica do Teste. Teoria de Resposta ao Item. Avaliação em larga escala; Classificatório. Análise de Itens. Proficiência.

ABSTRACT

ARAUJO, Elioenai de Queiroz. **Psychometric analysis of the IFPI classification examination**, 2023. 53 f. Thesis (Master Degree) – Federal Institute of Piauí – Floriano Campus, Floriano, 2023.

To enter a technical course at a public educational institution in Brazil, candidates for courses are submitted to a qualifying test. With the expansion of the Federal Institute of Piauí, this exam has been covering cities throughout the Piauí territory. So the idea of this study is to look for the necessary elements to explore the results of the classification test and, with such results, to improve the ordering of the students, to allow a result greater than the student's grade and also to promote the institution's policies. Thus, the purpose of this study is a detailed investigation and examines the validity, reliability and fairness of these exams as a means of assessing the potential of candidates. The research methodology has as theoretical foundations the Classical Test Theory and Item Response Theory, as well as elements of quantitative research. For data analysis and study, we will use statistical software as technological support. With this investigation, we hope to build a specific model of analysis from data collected in the classification exam of the Federal Institute of Piauí so that in future tests the organizers can carry out a better educational evaluation and consequently direct the public policies of the institution.

Keywords: Classical Test Theory. Item Response Theory. Large scale assessment; Classificatory. Item Analysis. Proficiency.

LISTA DE GRÁFICOS

Gráfico 1. Análise Gráfica do Item 31	33
Gráfico 2. Curva característica de Todos os Itens	35
Gráfico 3. Medidas de Proficiência	35
Gráfico 4. Gráficos de Dispersão	36
Gráfico 5. De Posicionamento dos itens e Distribuição dos Participantes	37
Gráfico 6. Curva Característica dos 3 Itens âncoras	39

LISTA DE TABELAS E QUADROS

Tabela 1. Numeração da questão, parâmetros, descrição das habilidades necessárias	32
Tabela 2. Percentual de acertos das questões por descritores	33
Tabela 3. Estimativas das Medidas Descritivas	34
Tabela 4. Parâmetro “a”, “b”, “c” das questões 32 ,40 e 43	38
Tabela 5. Apresentação do nível e parâmetros “a”, “b”, “c” das respostas	45
Tabela 6. Classificação dos Tópicos, IFPI 2022.1	48
Tabela 7. Classificação dos descritores, IFPI 2022.1	48
Quadro 1. Código de descritores SAEPI 9º ano.	49

LISTA DE ABREVIATURAS E SIGLAS

AGI	Análise Gráfica do Item
BNCC	Base Nacional Comum Curricular
CCI	Curva Característica do Item
ENADE	Exame Nacional de Avaliação dos Estudantes
ENEM	Exame Nacional do Ensino Médio
EF09MA08	Ensino Fundamental /9ºAno-Matemática/Faixa 8-(Habilidade BNCC)
EF09MA19	Ensino Fundamental /9ºAno-Matemática/Faixa 19-(Habilidade BNCC)
EF09MA32	Ensino Fundamental /9ºAno-Matemática/Faixa 32-(Habilidade BNCC)
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto de Estudo e Pesquisas Educacionais Anísio Teixeira
ML1P	Modelo Logístico de 1 Parâmetro
ML2P	Modelo Logístico de 2 Parâmetro
ML3P	Modelo Logístico de 3 Parâmetro
PROFMAT	Programa De Pós-Graduação Mestrado Profissional Matemática
Q32	Questão 32
Q40	Questão 40
Q43	Questão 43
R32	Item 32
R40	Item 40
R43	Item 43
SAEB	Sistema de Avaliação da Educação Básica
SAEPI	Sistema de Avaliação Educacional do Piauí
TCM	Teoria Clássica da Medida
TCT	Teoria Clássica dos Testes
TRI	Teoria de Resposta ao Item

SUMÁRIO

1 INTRODUÇÃO	14
1.1 OBJETIVOS.....	16
1.1.1 Objetivo Geral	16
1.1.2 Objetivos Específicos	16
1.2 METODOLOGIA.....	16
2 REVISÃO DE LITERATURA.....	18
2.1 CONSIDERAÇÕES INICIAIS SOBRE PSICOMETRIA	19
2.2 TEORIA CLÁSSICA DOS TESTES (TCT).....	21
2.3 TEORIA DE RESPOSTA AO ITEM (TRI).....	22
2.4 CRITÉRIOS PARA AVALIAÇÃO DA QUALIDADE DOS ITENS	24
2.5 O EXAME CLASSIFICATÓRIO DO IFPI.....	25
3 MATERIAL E MÉTODOS	28
3.1 TEORIA DE RESPOSTA AO ITEM (TRI).....	28
3.2 CALIBRAGEM.....	28
3.3 NOTAS.....	29
3.4 ESCALA MÉTRICA.....	30
3.4.1 Principais características de uma escala métrica	31
4 RESULTADOS	32
4.1 IMPORTÂNCIA DA AVALIAÇÃO EDUCACIONAL	40
5 CONSIDERAÇÕES FINAIS.....	42
REFERÊNCIAS	44
ANEXO 1 – TRI.....	48
ANEXO 2 – TCT	49
ANEXO 3 – TABELAS E QUADROS	51

1 INTRODUÇÃO

As mudanças sociais, políticas, científicas e tecnológicas que vêm ocorrendo no mundo, os impactos e os problemas aumentam e causam angústia nas populações da realidade social, refletem-se nas exigências à formação e à qualificação dos profissionais. O conhecimento, a ciência e a pesquisa constituirão o foco de estudo e de práticas que privilegiam sempre a produção do conhecimento. O ensino superior no Brasil ainda é um privilégio de poucos. De acordo com o IBGE (2019), apenas 12,8% da população possui nível superior. Desde o final da década de 1960, o acesso ao ensino superior é realizado por meio de concurso vestibular. O exame era composto apenas por itens objetivos nas áreas de conhecimento consideradas essenciais (ANDRIOLA, 2011).

Os exames classificatórios são avaliações realizadas com o objetivo de classificar candidatos em determinadas categorias, como admissão em instituições de ensino, seleção para empregos ou concursos públicos. Esses exames são projetados para avaliar o conhecimento, habilidades e competências dos candidatos e, assim, determinar sua posição relativa em relação aos demais. Podem abranger diferentes áreas de conhecimento, dependendo do contexto em que são aplicados. Por exemplo, em exames de admissão universitária, podem ser avaliados conhecimentos em disciplinas como matemática, língua portuguesa, ciências, entre outras. Já em concursos públicos, os exames podem abranger conhecimentos específicos relacionados à área de atuação pretendida, além de testes de aptidão e habilidades específicas. Nesse sentido, os resultados servem, na grande maioria, para políticas públicas.

Nesta perspectiva, o estudo em questão sobre análise psicométrica do processo seletivo do Instituto Federal de Educação (IFPI) configura-se de extrema relevância no contexto educacional atual, porque desempenha um importante papel na seleção de candidatos para esta instituição de ensino. Porém, a sua validade, a confiabilidade e a equidade são questionadas frequentemente. Sendo assim, uma análise mais aprofundada torna-se essencial para melhorar esses processos de seleção.

A relevância desta pesquisa é incontestável. Seus resultados podem fornecer esclarecimentos de grande valia para os educadores e administradores responsáveis pela elaboração e administração dos exames. A compreensão das propriedades psicométricas dos exames permitirá aprimorar a formulação de questões, identificar potenciais caminhos a serem seguidos e ajustar o processo de seleção para melhor refletir as habilidades e aptidões dos candidatos. Existem muitos instrumentos de avaliação no país, com destaque para o Sistema de

Avaliação da Educação Básica (SAEB), Exame Nacional do Ensino Médio (ENEM) e Exame Nacional de Avaliação dos Estudantes (ENADE), que avaliam educação básica, ensino médio e ensino superior, respectivamente. Todos eles são coordenados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), órgão responsável por executar as ações educacionais, e vinculado ao Ministério da Educação. Além de mensurar o nível de proficiência do aluno, tais exames são fundamentais para a elaboração de políticas públicas.

No caso do ENEM, a avaliação também é utilizada como exame para ingresso no ensino superior. Embora o número de participantes no ENEM (Exame Nacional do Ensino Médio) de 2021 tenha sido o menor registrado desde 2005, 3,1 milhões de pessoas se inscreveram, enquanto em 2019 foram registrados 5,1 milhões. Durante o governo Dilma Rousseff em 2014, a quantidade era de 8,7 milhões (INEP,2021). Não está claro quais são as causas dessa diminuição no número de participantes, mas especialistas apontam duas vertentes: o período pandêmico e o aumento progressivo da taxa cobrada pelo governo para a realização do exame.

Tendo em vista que as recompensas do Enem são altamente atrativas, uma vez que a aprovação do candidato implica na sua admissão nas Universidades públicas do país, ou na obtenção de bolsa ou financiamento para cursar uma faculdade privada (BRASIL, 2010), é pertinente investigar a eficácia educacional. Nesse sentido, os métodos estatísticos podem ser uma ferramenta importante nesse quesito, o que se espera comprovar neste estudo.

Análise estatística da Folha de São Paulo mostra que entre participantes com mesmo número de acertos, a nota é maior para aqueles que acertam perguntas mais fáceis. Será essa afirmação verdadeira? O Modelo matemático adotado pelo ENEM, a chamada TRI (Teoria da Resposta ao Item), prevê itens calibrados de acordo com o nível de conhecimento daquele tema, dificuldade e probabilidade de acerto casual. A nota dos participantes depende não só do número de acertos, mas também do nível de dificuldade das questões. Entre outros pontos, acertar questões difíceis e errar fáceis é entendido pelo modelo como um possível chute, e há reflexo na nota final. Candidatos que, por exemplo, acertaram 22 das 45 questões de Matemática no ENEM 2019 tiveram notas bem diferentes de um outro candidato que acertou a mesma quantidade de pontos (INEP,2021). É um modelo estatístico que estima a probabilidade de um candidato acertar uma questão, dada sua proficiência na prova.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Aplicar a Teoria Clássica do Teste (TCT) e Teoria de Resposta ao Item (TRI) para os dados do exame classificatório do Instituto Federal.

1.1.2 Objetivos Específicos

- a) Utilizar um modelo logístico de 3 parâmetros da TRI para os dados do classificatório do IFPI;
- b) Investigar a qualidade dos itens da prova, por meio da TCT e TRI;
- c) Melhorar a análise dos resultados do exame classificatório;
- d) Identificar as principais dificuldades dos participantes;

1.2 METODOLOGIA

A metodologia da pesquisa tem como fundamentos teóricos a Teoria Clássica do Teste e Teoria de Resposta ao Item, bem como elementos da pesquisa quantitativa. Para análise e estudo dos dados usaremos o *software* R (2023). Além disso, a relevância estende-se aos próprios candidatos. Um processo de seleção justo e transparente é fundamental para garantir que todos os candidatos, independentemente de sua origem socioeconômica, gênero ou contexto cultural, tenham igualdade de oportunidades para demonstrar seu potencial. Ao analisar a equidade dos exames classificatórios, a pesquisa contribui para a criação de um ambiente mais inclusivo e meritocrático.

Na teoria de resposta ao Item (TRI), foram utilizados modelos matemáticos para mostrar a proficiência do indivíduo, propondo formas de representar a relação entre a probabilidade de um indivíduo acertar ou não um item e parâmetros dos itens, na área específica de conhecimento. O alcance da pesquisa abrange tanto os aspectos técnicos quanto os sociais. Do ponto de vista técnico, a análise psicométrica investiga a validade e confiabilidade dos exames, bem como possíveis melhorias nos processos de elaboração e aplicação. Esses aspectos são essenciais para garantir que os resultados dos exames sejam confiáveis e representativos das habilidades dos candidatos.

Foram analisadas um conjunto de respostas por um grupo de indivíduos que responderam um conjunto de itens e dos indivíduos em uma escala de medida. A TRI possibilita

comparações entre indivíduos de populações diferentes, quando são submetidos a testes ou avaliações que tenham itens comuns e permite comparação entre indivíduos da mesma população, submetidos a testes totalmente diferentes.

2 REVISÃO DE LITERATURA

O presente estudo examina casos práticos de análise psicométrica de exames classificatórios, destacando os desafios enfrentados e as soluções propostas para melhorar a precisão e a justiça dos processos seletivos. Além disso, consideramos as implicações éticas e práticas da aplicação das descobertas da análise psicométrica na tomada de decisões educacionais e profissionais. O Instituto Federal de Educação do Piauí – IFPI mantém vários Cursos Técnicos Integrados ao Ensino Médio onde os alunos participam, fazendo um exame classificatório que é composto por trinta (30) questões de Português e trinta (30) questões de matemática. Sobre os alunos que almejam vagas nos cursos técnicos do IFPI, o estudo de Vogado (2021) investigou se a prova do classificatório contempla os descritores existentes no Sistema de Avaliação Educacional do Piauí (SAEPI), realizada pela Secretaria de Estado de Educação do Piauí - SEDUC – PI.

Segundo Cabrito (2009), o processo de avaliar é algo intrínseco dentro do contexto financeiro e econômico, como um meio de mensurar a qualidade para a atividade produtiva e especulativa. No contexto educacional, o ato de avaliar está associado desde mensuração do nível de aprendizado do aluno para um período de disciplina, até a adoção de políticas públicas.

Segundo Edelen e Reeve (2007), a TRI é um conjunto de técnicas de modelagem que engloba, por meio da análise de itens específicos, a mensuração de determinadas características dos indivíduos avaliados.

Conforme Andrade, Tavares e Valle (2000), essas técnicas permitem representar a relação entre o(s) traço(s) latente(s) pertencente(s) a um indivíduo e a probabilidade de este acertar determinado item. Entende-se por traço latente as características de um indivíduo que não podem ser observadas diretamente, como por exemplo, proficiência em outra língua ou o grau de satisfação de um consumidor. Um item pode ser representado por uma questão de prova, ou até mesmo, por uma pergunta relacionada ao grau de satisfação do indivíduo a um determinado produto ou serviço.

De acordo com Andrade, Tavares e Valle (2000), citado por CAMARGO *et al* (2016, p. 336):

Os modelos decorrentes da TRI são capazes de suprir algumas limitações da Teoria Clássica do Teste (TCT) ou Teoria Clássica da Medida (TCM), que é comumente utilizada na avaliação e seleção dos indivíduos. Na TCT, a avaliação é realizada por meio de escores brutos e padronizados. O escore de um indivíduo depende basicamente do número de acertos alcançados, sem

considerar se esses acertos são oriundos de questões fáceis ou difíceis.

Conforme os autores, A TCT apresenta algumas “falhas” por focar muito no teste e não na aptidão real do aluno, uma vez que o número de acertos é o escore que a pessoa obteve no teste, estando sempre dependente do erro do teste, que pode ser maior ou menor. Assim, por muitas aplicações ao longo do tempo que se faça nunca iremos obter uma análise mesmo sobre a aptidão do real do sujeito, já que sempre haverá algum erro associado.

Duas pressuposições principais embasam a TRI no que se refere às características dos itens: unidimensionalidade e independência local. O primeiro se refere à existência de uma habilidade dominante, que responde por todos os itens do teste. O segundo assume que as respostas a diferentes itens no teste ocorrem de forma independente, ou seja, o desempenho em um item não interfere em outro (ANDRADE; TAVARES E VALLE, 2000; PASQUALI, 2018; SOARES, 2018).

Até o ano de 2008, o ENEM recorria à TCT. A partir de 2009, o exame passou por uma reformulação e a TRI veio em substituição como método de validação dos itens e estimativa do desempenho dos participantes. O modelo usado no ENEM é o de 3 parâmetros, no qual é descrita a probabilidade de um respondente acertar uma questão dependendo da sua proficiência.

Esse modelo, como o nome diz, tem como condição conhecer 3 parâmetros sobre o item: o parâmetro “a” à discriminação do item; o “b” é a sua dificuldade; e o “c” que corresponde à chance de uma pessoa com baixo desempenho responder à questão de maneira satisfatória. Considerando a análise acima, percebe-se a importância do estudo e seguiremos para a aplicação do modelo TRI para itens das questões do exame classificatório do IFPI. O exame de admissão do IFPI é exclusivamente para ordenar os candidatos às vagas de curso técnico da instituição. Por esse motivo, este trabalho tem como objetivo utilizar a TCT e TRI para aprimorar as técnicas de mensuração de proficiência e divulgação dos resultados do exame, identificando, por exemplo, as maiores dificuldades dos participantes.

2.1 CONSIDERAÇÕES INICIAIS SOBRE PSICOMETRIA

De acordo com Pasquali (2009), a Psicometria é um campo científico da psicologia que se dedica à criação e aplicação de instrumentos para a medição de construtos e variáveis psicológicas. Ela se baseia em métodos estatísticos avançados, incluindo análise fatorial, modelagem de equações estruturais e a Teoria de Resposta ao Item, entre outras técnicas

multivariadas. Essas abordagens possibilitam a análise e mensuração da estrutura de construtos psicológicos, mais especificamente, dos processos mentais.

No início do século XX, Charles Spearman desenvolveu um modelo de inteligência que se apoiava em um fator geral, explicando as correlações entre diversos testes de inteligência. Esses avanços levaram ao desenvolvimento de instrumentos para mensurar esses construtos. Vários estudos descrevem a psicometria desempenhando um papel essencial em diversos campos, abrangendo desde a seleção de pessoal até o diagnóstico psicológico, a orientação vocacional e a pesquisa científica. Concentra-se na criação e aplicação de métodos de medição, como testes psicológicos e questionários, para avaliar de forma precisa traços psicológicos, habilidades e características de personalidade, através de instrumentos confiáveis e válidos e, ainda oferece suporte para a tomada de decisões embasadas em dados sólidos em uma variedade de contextos.

Os pesquisadores brasileiros dedicam-se constantemente ao desenvolvimento e validação de instrumentos de avaliação psicológica, visando garantir a confiabilidade e validade das medidas obtidas. Os testes são a principal ferramenta para acessar construtos psicológicos, como inteligência, personalidade e valores. Esses testes buscam explicar o comportamento dos indivíduos à luz das teorias psicológicas e do modo como eles respondem a itens específicos. Os resultados dos testes fornecem informações sobre o perfil psicológico de um sujeito, sendo utilizados em diversas áreas, incluindo a avaliação para dirigir automóveis ou portar armas de fogo.

A qualidade na área educacional, como Davok (2007) menciona, engloba estruturas, processos e resultados educacionais. A busca pela melhoria contínua da qualidade do ensino exige o engajamento de todos os envolvidos no processo, com ênfase na autoavaliação no ambiente universitário (ZANELLA; LOPES; SEIDEL, 2009).

Uma metodologia relevante para avaliar a satisfação é a Teoria de Resposta ao Item (TRI), que se baseia em modelos matemáticos para relacionar variáveis latentes e suas manifestações, possibilitando a criação de medidas padronizadas. A TRI é uma alternativa à Teoria Clássica da Medida (TCM) e tem se tornado popular na área de educação e testes psicológicos. Este estudo concentra-se na implementação das técnicas da Teoria Clássica do Teste e da Teoria de Resposta ao Item para analisar os dados do exame classificatório do Instituto Federal do Piauí, tornando o processo de avaliação mais completo e permitindo a individualização na apresentação dos resultados

Para ingressar num curso técnico numa instituição pública de ensino no Brasil os candidatos aos cursos são submetidos a um teste classificatório. Com a expansão do Instituto

Federal do Piauí este exame vem abrangendo cidades em todo o território piauiense. Assim, a ideia desse estudo é buscar elementos necessários para melhorar a estrutura e a qualidade dessa prova.

Portanto, a Psicometria contribui para a tomada de decisões informadas, fornecendo informações objetivas e confiáveis sobre características psicológicas dos indivíduos. No entanto, é essencial que os resultados obtidos por meio de instrumentos psicométricos sejam interpretados com cautela e levando em consideração o contexto específico em que estão sendo aplicados.

2.2 TEORIA CLÁSSICA DOS TESTES (TCT)

Na área da Psicometria e da avaliação educacional é muito comum utilizar a TCT, ou seja, a Teoria Clássica dos Testes que corresponde a uma estrutura usada para analisar e entender resultados de testes e/ou avaliações, procurando examinar a relação existente entre as pontuações em um teste, as latentes ou verdadeiras e o erro associado de mensuração.

A TCT fornece uma estrutura sólida para a análise e interpretação dos resultados dos testes, permitindo a estimação da confiabilidade, a identificação do erro de medida e a compreensão da relação entre os escores observados e os escores verdadeiros. No entanto, é importante ressaltar que a TCT tem suas limitações e suposições simplificadoras, e outras abordagens, como a Teoria de Resposta ao Item (TRI), têm sido desenvolvidas para superar algumas dessas limitações.

Na concepção de Borgatto e Andrade (2012, p. 148):

a análise clássica dos itens de uma prova baseia-se em seus parâmetros descritivos, os quais auxiliam na interpretação da distribuição das respostas para cada alternativa. As propriedades psicométrica dos itens de uma prova correspondem aos seguintes parâmetros: índice de dificuldade (proporção de participantes que responderam ao item corretamente); índice de discriminação, que mede a capacidade do item de diferenciar os participantes de maior habilidade (27% dos respondentes com pontuações mais altas) daqueles de menor habilidade (27% dos respondentes com pontuações mais baixas), correspondendo à diferença entre a proporção de acertos do primeiro grupo e a do segundo grupo; e correlação bisserial entre a resposta numa dada categoria do item e a pontuação total na prova.

Os autores supracitados ainda asseguram que, no índice de dificuldade, analisa-se o grau de dificuldade de cada item por meio da porcentagem de acerto. Isso significa que quanto menor a porcentagem de acerto maior será o grau de dificuldade. Entretanto, o índice de

discriminação analisa, para determinado item, as porcentagens de acertos dos grupos de estudantes com melhor e com pior desempenho. Um resultado esperado seria que, para um item com boa qualidade, a porcentagem de acerto seja maior para o grupo com melhor desempenho, e quanto maior for a diferença entre as porcentagens de acertos dos dois grupos (com melhor e com pior desempenho), maior será a discriminação do item.

Para Pasquali (2018), algumas limitações da TCT como o não-controle de acertos casuais ou as dificuldades existentes em processos de comparação entre estudantes submetidos a provas ou edições distintas de um mesmo exame (necessários, por exemplo, para avaliar a eficiência de uma política pública adotada), corroboraram para o surgimento e aprimoramento da TRI. A principal inovação que esta metodologia trouxe foi considerar cada questão da prova como unidade básica para as análises estatísticas, desenvolvendo um conceito de escala de proficiência que traz grandes benefícios para as análises de avaliações.

2.3 TEORIA DE RESPOSTA AO ITEM (TRI)

A Teoria da Resposta ao Item (TRI) vem ganhando cada vez mais espaço em testes de larga escala no Brasil, uma vez que já é consagrada em várias partes do mundo. Este método adota uma independência dos itens em relação aos indivíduos nos quais se deseja medir a sua proficiência

De acordo com Rabelo (2013), atualmente a TRI é amplamente utilizada nas avaliações em larga escala em todo o mundo, por exemplo. No Brasil, a primeira aplicação da TRI ocorreu em 1995 por meio do Sistema de Avaliação da Educação Básica (Saeb). Em seguida, foi implementada no Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA), na Prova Brasil e, então, no ENEM.

A Teoria de Resposta ao Item (TRI) é um conjunto de modelos estatísticos utilizados na análise de testes psicométricos. Essa teoria tem como objetivo avaliar o desempenho dos indivíduos em testes com base nas características dos itens apresentados. A TRI parte do pressuposto de que o desempenho dos indivíduos em um teste é influenciado não apenas por suas habilidades ou traços psicológicos, mas também pelas características dos itens que compõem o teste. Cada item é visto como um estímulo que pode ser respondido corretamente ou incorretamente, dependendo das habilidades do indivíduo.

Existem diferentes modelos na TRI, mas o mais comumente utilizado é o modelo de resposta ao item de dois parâmetros (IRT). Nesse modelo, são considerados dois parâmetros

para cada item: a dificuldade do item (representada pelo parâmetro “b”) e a discriminação do item (representada pelo parâmetro “a”). A dificuldade do item indica a habilidade necessária para respondê-lo corretamente, enquanto a discriminação do item indica o quão bem ele é capaz de distinguir indivíduos com diferentes níveis de habilidade. A TRI permite calcular estimativas das habilidades dos indivíduos com base em suas respostas nos itens do teste. Além disso, essa teoria também é usada para equilibrar e calibrar os itens do teste, identificar itens com baixa qualidade ou que possam apresentar viés, e comparar o desempenho de diferentes grupos de indivíduos.

A TRI é amplamente utilizada na prática psicométrica e tem contribuído para o desenvolvimento de testes mais precisos, confiáveis e válidos. Ela fornece uma estrutura teórica sólida para a construção e análise de testes, auxiliando na interpretação dos resultados e na tomada de decisões com base nas informações obtidas a partir dos testes psicométricos.

Segundo Andrade, Tavares e Valle (2000); Pasquali, (2018); Soares, (2018) a TRI se baseia em duas pressuposições principais, que se referem às características dos itens: unidimensionalidade e independência local. A primeira se refere à existência de uma habilidade dominante, que responde por todos os itens do teste. A segunda assume que as respostas a diferentes itens no teste ocorrem de forma independente, ou seja, o desempenho em um item não interfere em outro.

Os mesmos autores destacam que os modelos mais utilizados de TRI são os logísticos. Existem três tipos principais: o modelo logístico de um parâmetro (ML1P) que considera apenas a dificuldade do item; o modelo logístico de dois parâmetros (ML2P) que considera a dificuldade e a discriminação do item, definida como a capacidade que ele possui para diferenciar respondentes com níveis distintos de conhecimento; e o modelo logístico de três parâmetros (ML3P) onde são consideradas a dificuldade, a discriminação e a probabilidade de acerto do item pela casualidade (ou chute).

O modelo utilizado pelo ENEM é o ML3P sendo expresso matematicamente pela equação abaixo:

$$P(X_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}} \quad (2.3.1)$$

Segundo Andrade, Tavares e Vale (2000):

É a probabilidade do indivíduo j com habilidade θ_j acertar o item i , a_i é o parâmetro de discriminação, b_i é o parâmetro de dificuldade e c_i é a probabilidade de acerto ao acaso do item i . O ML2P pode ser obtido fazendo $c_i=0$ e o ML1P fazendo $c_i=0$ e $a_i=1$. A função matemática do modelo, que

associa a probabilidade de acerto de um item com seus parâmetros e com a habilidade do respondente, gera uma curva monótona crescente, denominada curva característica do item (CCI). Além da CCI, uma outra medida frequentemente utilizada na TRI é a Curva de Informação do Item (CII) que permite quantificar a precisão com que se pode estimar o nível de conhecimento de um indivíduo por meio da resposta dada por ele ao item, sendo um instrumento eficaz para descrever e selecionar itens de qualidade.

Atualmente a TRI é amplamente utilizada nas avaliações em larga escala em todo o mundo como, por exemplo, no *Programme for International Student Assessment* (PISA) e no *Test of English as a Foreign Language* (Toefl). No Brasil, a primeira aplicação da TRI se deu em 1995, através do Sistema de Avaliação da Educação Básica (Saeb). Em seguida, foi implementada no Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA), na Prova Brasil e, então, no Exame Nacional do Ensino Médio (ENEM) (RABELO, 2013).

2.4 CRITÉRIOS PARA AVALIAÇÃO DA QUALIDADE DOS ITENS

Segundo Travitzki (2017, p. 264), “os métodos de triagem de itens, que incluem análises de dificuldade e discriminação, desempenham um papel fundamental na avaliação global, permitindo uma avaliação da qualidade dos itens com base em seu desempenho empírico”. No que se refere ao índice de discriminação, Baker (2001) e Bortolotti e Andrade (2007) estabelecem que itens com um valor acima de 0.65 são considerados de alta qualidade.

Quanto à avaliação da dificuldade, a abordagem tradicional da psicometria se baseia no percentual de respostas corretas para classificar a dificuldade de um item. No entanto, deve-se notar que essa classificação varia dependendo do grupo de indivíduos que responde ao item. Por exemplo, um item pode ser considerado fácil para um grupo com habilidades elevadas e difícil para outro grupo com habilidades inferiores (COSTA, LIMA E SOARES, 2020; SOARES, 2018).

Na Teoria de Resposta ao Item (TRI), por outro lado, os parâmetros dos itens são calculados independentemente da amostra de sujeitos utilizada durante o processo de calibração. Essa característica é uma das vantagens da adoção da TRI em avaliações. De acordo com a TRI, tanto o índice de dificuldade quanto a habilidade do respondente seguem uma distribuição normal padronizada. Nesse contexto, o índice de dificuldade é expresso em termos do número de desvios-padrão da média. Embora tecnicamente varie de $-\infty$ a ∞ , é comum encontrar valores entre -3.0 e 3.0, representando teoricamente 99.73% do conjunto (Xie, Davidson, Li e Ko, 2019). Assim, conforme Vendramini e Dias (2005, p. 207), “um item com

$b = -3.0$ é considerado extremamente fácil, $b = 0.0$ indica dificuldade média, e $b = 3.0$ representa um item extremamente difícil”.

No que diz respeito ao parâmetro c , que reflete a probabilidade de acerto ao acaso, espera-se valores próximos ao inverso do número de alternativas disponíveis para marcação (PASQUALI, 2003). No caso dos itens do ENEM, que apresentam cinco alternativas de resposta, valores próximos a 0.2 (1/5) são esperados para este parâmetro. É importante destacar que este parâmetro é estimado conjuntamente com os demais, podendo ter estimativa, então, diferente do esperado.

Quanto a uma avaliação psicométrica global, não existe um critério consensual na literatura para determinar a qualidade de um teste. No entanto, Vendramini e Dias (2005, p. 207) propõem que “testes adequados são aqueles em que os itens demonstram coeficientes de correlação ponto-Bisserial mais elevados para a alternativa correta em comparação com as outras alternativas e apresentam resíduos padronizados do ajuste do modelo inferiores a 2.0”.

2.5 O EXAME CLASSIFICATÓRIO DO IFPI

O Instituto Federal de Educação do Piauí (IFPI) é uma instituição criada em dezembro de 2008, que anteriormente era conhecida como Escola Técnica Federal do Piauí (ETFPI), passa por uma reformulação. A estrutura organizacional do IFPI, estabelecida pela Lei 11.892 de dezembro de 2008, inclui cinco pró-reitorias: Pró Reitoria de Administração, Pró Reitoria de Ensino, Pró Reitoria de Desenvolvimento Institucional, Pró Reitoria de Extensão e Pró Reitoria de Pesquisa e Inovação.

O IFPI é uma instituição de ensino que oferece educação básica, superior e profissional em várias áreas de estudo. Sua atuação é multicampi, especializando-se na oferta de educação profissional e tecnológica em diversas modalidades de ensino, integrando conhecimentos técnicos e tecnológicos às práticas pedagógicas.

A história do IFPI remonta a 1909, quando foi fundada como a Escola de Aprendizes Artífices, localizada em Teresina, Piauí, inicialmente no bairro Pirajá. Posteriormente, em 1934, mudou-se para um antigo casarão na Praça Pedro II, recebendo o nome de Liceu Industrial de Teresina. Em 1937, durante o Estado Novo, a escola foi renomeada como Liceu Industrial do Piauí, refletindo o foco na indústria como motor de desenvolvimento.

Em 1938, a Escola mudou-se para a Praça da Liberdade e, em 1942, com a Lei Orgânica do Ensino Industrial, tornou-se a Escola Industrial de Teresina. Em 1965, pela primeira vez, a marca “Escola Federal” foi introduzida na rede de ensino. Em 1967, a instituição criou seus

primeiros cursos técnicos, incluindo Edificações, Agrimensura e Eletromecânica, que mais tarde se desmembraram em cursos individuais. Em 1994, foi autorizada a transformação em Centro Federal de Educação Tecnológica do Piauí, um marco que se concretizou em 1999. A partir de 2005, o CEFET-PI alinhou-se à política do MEC e expandiu sua oferta educacional, introduzindo o Ensino Médio integrado ao Ensino Técnico em várias áreas. Além disso, em 2007, lançou um programa de Pós-Graduação Lato Sensu, demonstrando um compromisso com a qualificação profissional e educacional da comunidade local e regional.

Ao se transformar em IFPI, a Instituição adquiriu autonomia para criar e extinguir cursos, bem como para registrar diplomas dos cursos por ela oferecidos, mediante autorização do seu Conselho Superior. Além disso, obteve um *status* semelhante ao de uma Universidade Federal em termos de funcionalidade, acesso ao fomento de pesquisa e extensão e todos os programas de apoio dos vários ministérios.

Em 2010, iniciou-se o processo de expansão do IFPI com a inauguração dos seguintes campi: Angical, Corrente, Piripiri, Paulistana, São Raimundo Nonato e Uruçuí. Em 2012, foram inaugurados campi em Pedro II, Oeiras e São João; e, em 2014, houve a inauguração dos campi de Campo Maior, Valença e Cocal. Nesse período, foi criado também o Programa Nacional de Acesso ao Ensino Técnico e ao Emprego (Pronatec), reforçando o desenvolvimento e a interiorização da educação profissional.

O exame classificatório é realizado como uma forma de seleção para o ingresso de alunos em seus cursos técnicos subsequentes ao ensino médio. Ele é uma das formas de acesso aos cursos oferecidos pelo IFPI, e ocorre anualmente e é aberto a candidatos que tenham concluído o ensino médio ou equivalente. Os candidatos realizam uma prova composta por questões de múltipla escolha, que abrangem diferentes áreas do conhecimento, como Língua Portuguesa, Matemática, Ciências Humanas e Ciências da Natureza.

A pontuação obtida no Exame Classificatório é utilizada para a classificação dos candidatos e a distribuição das vagas disponíveis nos cursos técnicos subsequentes do IFPI. Os candidatos com as melhores pontuações têm preferência na escolha das vagas disponíveis. É importante salientar que as informações específicas sobre o Exame Classificatório do IFPI podem variar a cada edição; portanto, é necessário consultar o edital e informações atualizadas divulgadas pelo próprio instituto para obter todos os detalhes sobre o processo seletivo, cronograma, conteúdo programático, critérios de pontuação e formas de inscrição.

1. Dispõe sobre o Exame Classificatório para ingresso na Educação Profissional Técnica de Nível Médio Concomitante e/ou Subsequente;
2. Os cursos da Educação Profissional Técnica de Nível Médio na forma Subsequente;

3. Candidatos que tenham concluído todo o Ensino Fundamental em Escola Pública;
4. Candidatos que tenham concluído Todo o Ensino Médio em Escola Pública Curso: o aluno concluirá, neste Instituto, uma Habilitação Profissional Técnica;
5. Os cursos da Educação Profissional Técnica de Nível Médio na forma Concomitante;
6. Candidatos que tenham concluído Todo o Ensino Fundamental em Escola Pública;
7. Candidatos que cursando o 1ª, 2ª ou 3ª séries do Ensino Médio;
8. Curso: Concluirá uma Habilitação Profissional Técnica neste Instituto, condicionada sua Diplomação à conclusão do Ensino Médio;
9. PROVA;
10. O Caderno de Prova conterà 60 (sessenta) questões;
11. Sendo 30 (trinta) questões de Língua Portuguesa;
12. E 30 (trinta) questões de Matemática;
13. Cada uma delas com 05 (cinco) alternativas, com uma única opção correta.

3 MATERIAL E MÉTODOS

“A abordagem supõe o contato direto e prolongado do pesquisador com o ambiente e a situação que está sendo investigada” (LUDKE: ANDRÉ, 1986, p.11). O campo de investigação foi o Instituto Federal de Educação do Piauí– IFPI e, como instrumento de análise, foi utilizada a Prova de Matemática do Exame Classificatório do IFPI, de 22 de novembro de 2022 por meio do caderno de questões da prova de Ensino Técnico integrado. Os microdados utilizados nesta pesquisa foram obtidos por meio de processo administrativo realizado na plataforma FALABR, seguindo os procedimentos de acesso à informação de acesso restrito. Mais detalhes ver o processo 23546.042775/2022-12. Foi aplicado a TCT e a TRI, 3 parâmetros, nos dados obtidos.

3.1 TEORIA DE RESPOSTA AO ITEM (TRI)

Ao contrário das abordagens tradicionais de avaliação, que se concentram apenas no número de respostas corretas ou incorretas, a TRI leva em consideração as características dos itens, como sua dificuldade e discriminação. Cada item é avaliado em relação à probabilidade de ser respondido corretamente por indivíduos com diferentes níveis de habilidade. Com base nas respostas dos participantes a um conjunto de itens, o modelo TRI estima a posição relativa da habilidade de cada indivíduo em relação à dificuldade dos itens.

A TRI é amplamente utilizada em testes padronizados, como testes de aptidão, exames escolares, avaliações psicológicas e testes de personalidade. Ela fornece informações valiosas sobre o nível de habilidade ou traço latente de um indivíduo, permitindo comparações precisas entre diferentes indivíduos e avaliando seu desempenho em relação a um padrão de referência. A TRI também é útil na construção e aprimoramento de itens de teste, pois permite identificar itens problemáticos, como aqueles que são muito fáceis ou muito difíceis para a população em avaliação. Em resumo, a Teoria de Resposta ao Item é um modelo estatístico avançado que busca fornecer medidas mais precisas e confiáveis das habilidades ou traços latentes dos indivíduos com base em suas respostas a um conjunto de itens de teste.

3.2 CALIBRAGEM

A calibragem de itens é um processo utilizado na psicometria para determinar a dificuldade e a discriminação dos itens de um teste, bem como para realizar a seleção final dos itens que serão utilizados na versão final do teste. Durante a calibragem, são analisadas as

respostas dos participantes a cada item do teste, levando em consideração o desempenho global dos participantes no teste. No processo de calibragem, os itens que não apresentam bons resultados de dificuldade ou discriminação podem ser excluídos do teste. Isso ocorre quando um item é muito fácil ou muito difícil para a amostra de participantes, ou quando não consegue discriminar adequadamente entre os participantes com diferentes níveis de habilidade.

Os critérios para a seleção dos níveis de dificuldade dos itens podem variar dependendo do contexto e do construto que está sendo medido. Normalmente, procura-se uma distribuição de dificuldade que cubra uma ampla gama de habilidades e que permita discriminação entre os participantes. Os critérios para a seleção dos níveis de dificuldade podem ser estabelecidos com base em teorias ou padrões previamente definidos, ou podem ser determinados empiricamente com base nos dados dos participantes.

Os critérios para as âncoras, também conhecidas como itens de referência, são estabelecidos para garantir a ancoragem ou fixação dos níveis de dificuldade do teste. Esses itens são selecionados para terem uma dificuldade conhecida e são utilizados como referência para calibrar os demais itens do teste. As âncoras ajudam a estabelecer um padrão de comparação e a vincular os escores dos itens aos níveis de habilidade ou traço que se pretende medir. Em relação ao pacote R para estimação, existem várias opções disponíveis que podem ser utilizadas na calibragem de itens.

Alguns dos pacotes populares para estimação e análise de itens em R incluem o pacote "ltm" (Item Response Theory models for longitudinal data), o pacote "mirt" (Multidimensional Item Response Theory models) e o pacote "TAM" (Test Analysis Modules). Esses pacotes oferecem funções e métodos para realizar a estimação de parâmetros de itens, como dificuldade e discriminação, utilizando diferentes modelos de Teoria de Resposta ao Item (TRI).

É importante ressaltar que o uso de pacotes R específicos para a estimação de itens depende do tipo de modelo e da abordagem que está sendo utilizada na análise. Cada pacote tem suas próprias funcionalidades e características, portanto, é recomendável consultar a documentação e os recursos disponíveis para escolher o pacote mais adequado às necessidades específicas da análise psicométrica em questão.

3.3 NOTAS

A Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI) são duas abordagens diferentes utilizadas na psicometria para analisar e interpretar os resultados de testes psicológicos. Embora ambas sejam amplamente empregadas, existem diferenças significativas

entre elas. A TCT considera que os escores observados nos testes são compostos pelo escore verdadeiro do indivíduo e pelo erro de medida. Ela se baseia em estatísticas clássicas, como a análise de variância e o coeficiente alfa de Cronbach, para estimar a confiabilidade do teste e interpretar os resultados. A TCT enfatiza a consistência interna dos itens e a relação entre os escores observados e os escores verdadeiros.

Por outro lado, a TRI é baseada em modelos matemáticos mais complexos para analisar a relação entre as respostas dos participantes e os parâmetros dos itens. Ela considera que a probabilidade de um participante responder corretamente a um item depende não apenas do escore verdadeiro do indivíduo, mas também dos parâmetros do item, como sua dificuldade e discriminação. A TRI busca modelar a relação entre os escores observados e os escores latentes dos indivíduos, utilizando modelos como o Modelo de Resposta ao Item (MRI) ou o Modelo de Rasch.

Comparando as duas abordagens, a TRI oferece algumas vantagens em relação à TCT. A TRI permite uma estimativa mais precisa dos parâmetros dos itens, como a dificuldade e a discriminação, e também permite a análise de itens individuais, não apenas do teste como um todo. Além disso, a TRI é capaz de lidar com a heterogeneidade dos itens e a variação das habilidades dos participantes.

No entanto, é importante destacar que tanto a TCT quanto a TRI têm seus usos e aplicações específicas. A escolha entre essas abordagens depende do objetivo do teste, das características dos dados e das necessidades de análise. Em alguns casos, é possível combinar elementos das duas teorias para obter uma análise mais completa e robusta dos resultados.

É importante consultar a literatura especializada, buscar orientação de especialistas em psicometria e considerar as necessidades específicas do teste em questão ao decidir entre a TCT e a TRI. Ambas as abordagens têm suas vantagens e limitações, e a escolha adequada dependerá do contexto e dos objetivos da análise.

3.4 ESCALA MÉTRICA

Uma escala métrica é um tipo de escala de mensuração que possui características específicas que permitem realizar operações matemáticas válidas com os dados obtidos. Essa escala é considerada a mais completa e precisa em termos de mensuração, pois permite a realização de operações de adição, subtração, multiplicação e divisão entre os valores observados.

3.4.1 Principais características de uma escala métrica

Identidade: A escala métrica possui um ponto de origem ou zero absoluto, que representa a ausência total da característica medida. Isso significa que o valor zero na escala reflete a ausência do atributo em questão. Um exemplo de escalas termométricas é a escala Celsius, usada para medir temperatura, em que o valor zero representa o ponto de congelamento da água.

Unidade: A escala métrica possui unidades de medida consistentes e uniformes, permitindo a comparação direta entre os valores observados. As unidades de medida devem ser consistentes e equivalentes em toda a escala. Por exemplo, na escala métrica de comprimento, a unidade pode ser o metro, e cada unidade é igualmente espaçada.

Ordenação: Os valores observados em uma escala métrica podem ser ordenados de forma crescente ou decrescente, permitindo a comparação entre os indivíduos ou as variáveis medidas. É possível identificar relações de maior ou menor em relação à característica medida. Por exemplo, em uma escala de satisfação que varia de 1 a 5, é possível comparar e ordenar os níveis de satisfação dos indivíduos.

Intervalo: A escala métrica possui intervalos constantes entre os valores observados. Isso significa que a diferença entre dois pontos quaisquer na escala tem o mesmo significado em toda a extensão da escala. Por exemplo, em uma escala de 0 a 10 para medir intensidade de dor, a diferença entre 4 e 6 é a mesma que entre 8 e 10.

A escala métrica é amplamente utilizada em várias áreas, como ciências sociais, psicologia, educação e pesquisa em geral. Ela permite a aplicação de técnicas estatísticas mais avançadas e a realização de análises quantitativas mais precisas. É importante distinguir a escala métrica de outros tipos de escalas, como a escala nominal (que apenas categoriza os dados), a escala ordinal (que ordena os dados, mas não permite operações matemáticas válidas) e a escala de intervalo (que possui os mesmos atributos da escala métrica, exceto o ponto de origem absoluto). Ao utilizar uma escala métrica, é fundamental conhecer suas características e propriedades para realizar análises adequadas e interpretar corretamente os resultados obtidos.

4 RESULTADOS

Serão apresentados os resultados para a TCT e a TRI nos dados do exame classificatório dos cursos de médio integrado no ano de 2022, primeiro semestre para alunos do Campus Teresina-Central. Para demonstrar o processo do item pela TCT, a primeira questão do teste será analisada:

Questão 31. Todo número positivo pode ser escrito em notação científica: $k \times 10^n$, em que $1 \leq k < 10$ e n é um número inteiro. Esta notação auxilia quando temos números “muito grandes” ou “muito pequenos”. Com base nessas informações, a notação científica do número

$$A = \frac{0,000168 \times 0,001}{0,024 \times 3,5} \text{ é:}$$

- a) 2×10^{-6}
- b) 2×10^{-5}
- c) 2×10^{-4}
- d) 2×10^{-3}
- e) 2×10^{-2}

Essa questão exige do candidato a habilidade (EF09MA04) e o descritor D24, considerando o conjunto de descritores apresentados na tabela x em anexo. Abaixo apresentaremos a tabela 1.

Tabela 1. Numeração da questão, parâmetros, descrição das habilidades necessárias.

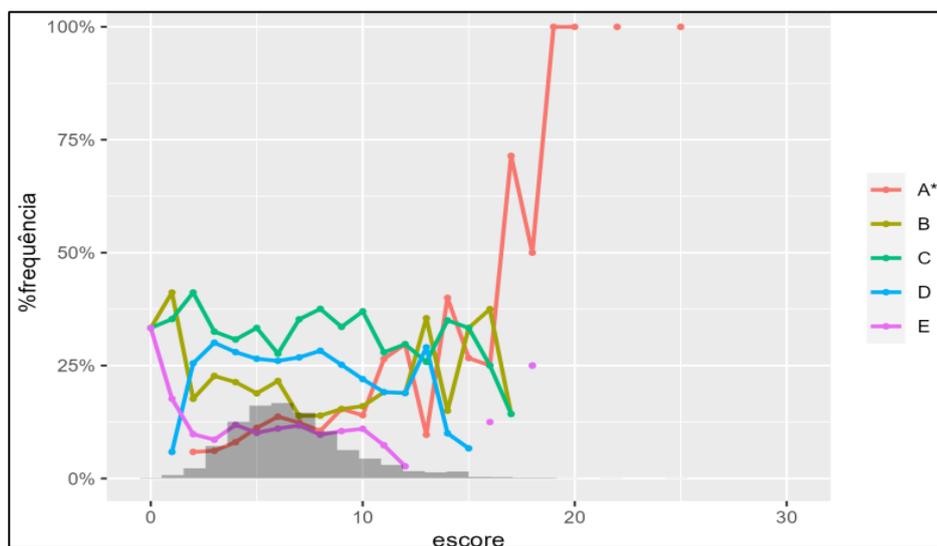
q	Dificuldade	bis	bisp	discr	Cronbach alpha	Descritor	Descrição
R31	9,75%	0,435	0,280	17,25%	0,718511	D24	Efetuar cálculos com potências de expoentes inteiros e aplicar esse conhecimento na representação de números em notação científica.

Fonte: Autor, 2023.

De acordo com a TCT, esse item teve dificuldade de 9,75% e ponto bisserial de 0,280. O gráfico a seguir permite realizar a Análise Gráfica do Item (AGI), e indica que o nível de dificuldade exige um nível de proficiência acima de 20 pontos para maior proporção de acerto observado. Além disso, o parâmetro de discriminação do item indica uma diferença de apenas

17,25% na proporção de acertos entre os grupos A e B. O grupo A é formado pelos candidatos entre as 27% maiores notas no exame, enquanto o grupo B é formado pelos candidatos entre as 27% menores notas no exame. Ou seja, esse item discrimina muito pouco quem sabe e quem não.

Gráfico 1. Análise Gráfica do Item 31



Fonte: Autor, 2023.

Considerando a prova como um todo, a tabela 1 a seguir apresenta o percentual de acerto das questões da prova classificadas por descritores.

Tabela 2. Percentual de acertos das questões por descritores.

Descritor SAEPI	Percentual de acerto
D3	12,2%
D10	12,6%
D27	14,9%
D9	16,0%
D32	18,1%
D11	18,8%
D30	19,1%
D26	20,0%
D13	21,2%
D28	21,3%
D29	23,4%
D20	24,0%
D31	24,1%
D24	25,2%
D33	29,8%

D4	31,6%
D25	34,3%
D14	36,0%

Fonte: Autor.

Com isso, percebe-se que o D3 (Identificar propriedades de triângulos pela comparação de medidas de lados e ângulos) foi o descritor com maior nível de dificuldade, com pouco mais de 12% de acerto dos candidatos. Por outro lado, observa-se que 36% dos respondentes acertaram as questões relacionadas ao descrito D14 (Resolver problema envolvendo noções de volume), sendo considerado do descritor mais fácil da prova.

Estas são as análises que podem ser feitas utilizando a TCT. Como mencionado anteriormente, a TRI possibilita novas análises.

O processo de calibragem da TRU inicia com a exclusão do item 20 por não ter sido acertado por nenhum aluno ou ter sido excluído no certame. Em seguida, durante o processo de estimação, mais 6 itens foram excluídos devido a problemas de estimação, com um limite definido para aqueles com estimativas maiores que 5 ou menores que -5. Este critério foi adotado para evitar o grande número de exclusões de itens. Após as duas etapas, restaram 23 itens para mensurar a proficiência dos alunos. As estimativas dos parâmetros dos itens e os respectivos erros padrão estão apresentados na tabela em Anexo 1.

O quadro a seguir apresenta as medidas descritivas para as estimativas dos 10 itens restantes do processo de calibração.

Tabela 3. Estimativas das Medidas Descritivas.

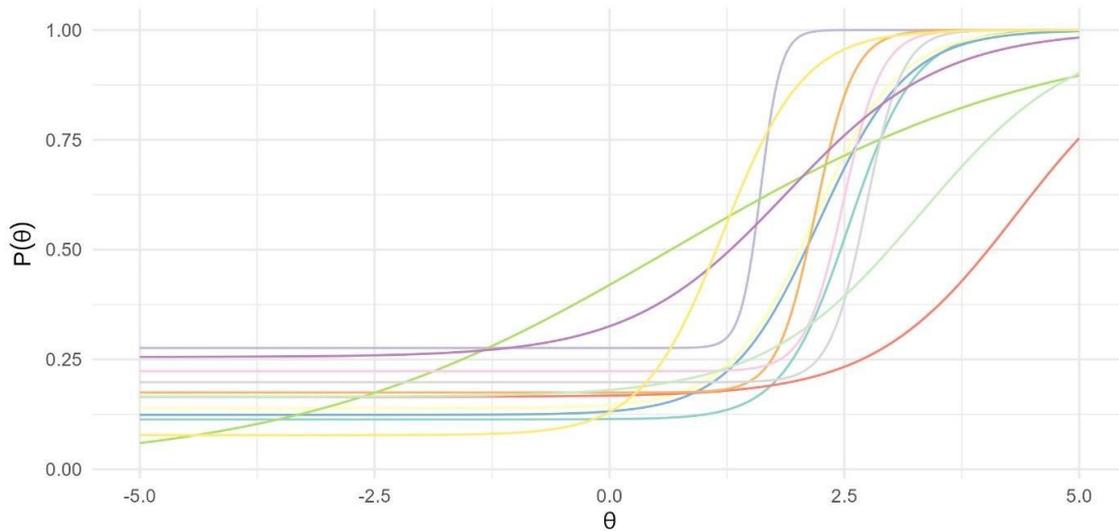
Parâmetro	Média	Desvio	Mediana	Mínimo	Máximo	Amplitude	Assimetria	Curtose
A	2,700	1,989	2,202	0,363	9,373	9,010	1,629	2,978
B	2,362	0,995	2,195	0,669	4,484	3,815	0,489	-0,567
C	0,158	0,081	0,165	0,003	0,303	0,300	-0,158	-0,769

Fonte: Autor.

É possível verificar uma média de 2,700 para o parâmetro de discriminação “a”. No entanto, as estatísticas apresentadas para as estimativas do parâmetro “b” indicam que a prova possui um nível elevado de dificuldade, isto é, os itens estão localizados em regiões que exigem maior nível de proficiência para que seja resolvido corretamente, com média de 2,362.

O gráfico 2 apresenta a curva característica de todos os itens e, embora seja inviável identificar cada um dos itens, percebe-se que todos os itens possuem níveis altos de dificuldade.

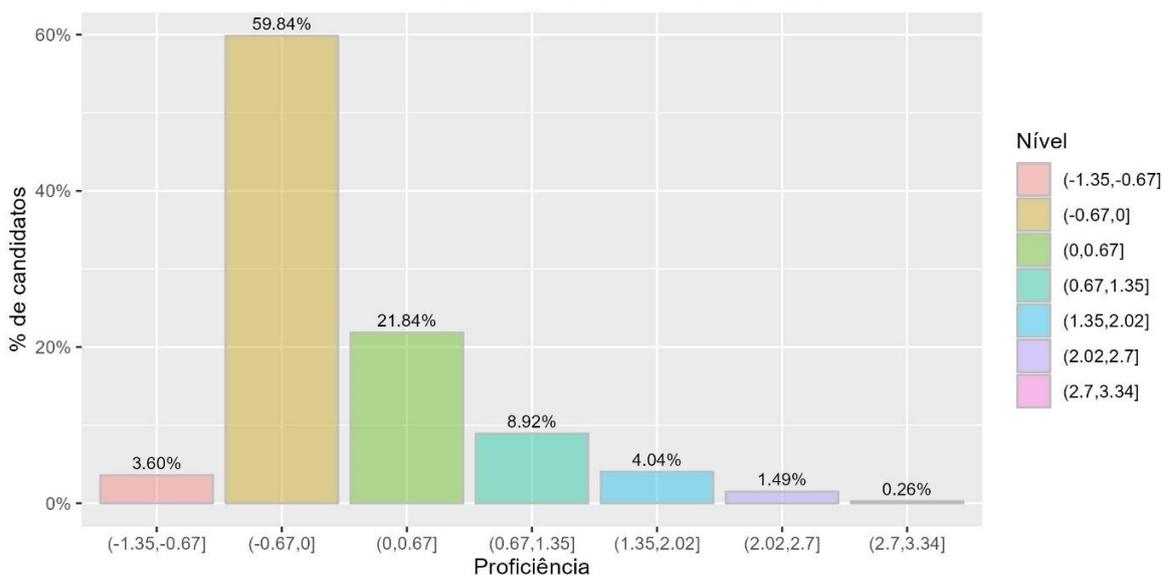
Gráfico 2. Curva característica de todos os itens da prova os itens da prova.



Fonte: Autor, 2023.

Considerando o nível de dificuldade dos itens, tem-se uma distribuição de proficiência com média 0 e desvio padrão 0.67, e por meio do gráfico 3, percebe-se que a classe modal possui alunos de até um desvio padrão abaixo da média, com mais de 50% dos alunos nesta classe. Este resultado indica que não havia itens na prova que permitissem diferenciar a metade dos alunos com menor proficiência.

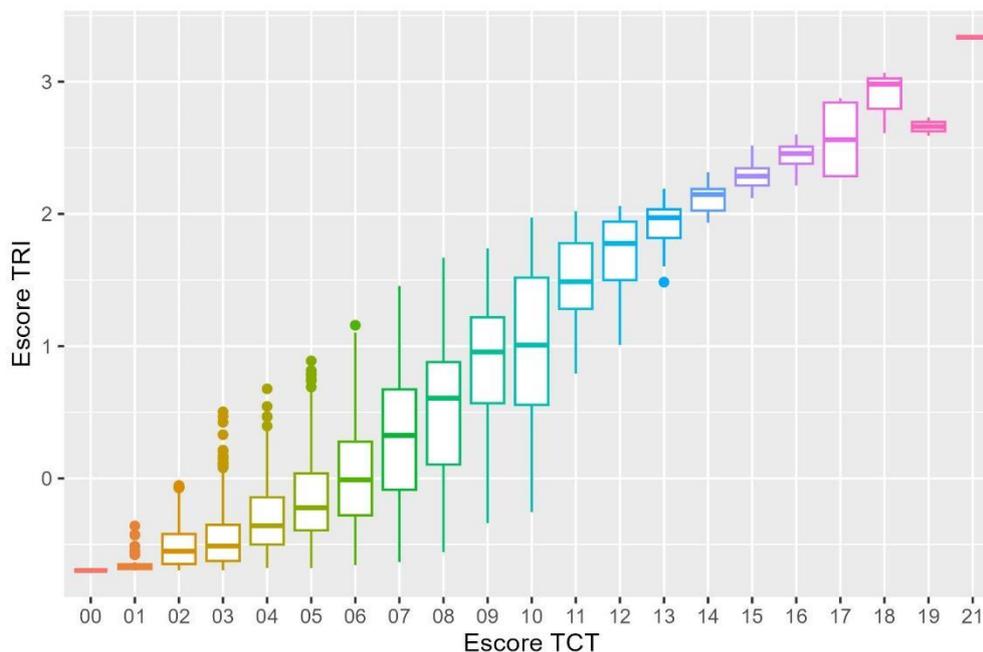
Gráfico 3. Medidas de Proficiência



Fonte: Autor.

O gráfico 4 apresenta a comparação entre os escores gerados pela TCT e pela TRI.

Gráfico 4. Gráficos de Dispersão.



Fonte: **Autor.**

Comparando os resultados da nota apresentada pela TRI e TCT, temos um coeficiente de correlação linear de Pearson de 0.84 (p -valor < 0.05), o que indica uma concordância muito forte entre os dois escores. Considerando que a distribuição de notas não é gaussiana, verificada por meio do teste Shapiro-Wilk (p -valor < 0.05), o coeficiente de Spearman (medida não paramétrica de correlação) estimado foi 0.76, indicando concordância elevada entre as notas. É importante destacar que o coeficiente de correlação de Spearman é baseado no ordenamento dos dados, portanto, há uma correlação elevada entre os ranqueamentos, indicando classificações parecidas utilizando as duas metodologias.

No entanto, é possível perceber que há variabilidade de nota na TRI considerando escores entre 5 e 10. Isso significa que a TRI é capaz de diferenciar pessoas com mesmo escore na TCT, conforme apresentado no gráfico 4. Considerando níveis separados por unidades de desvio padrão, a próxima etapa é criar uma escala de proficiência, em que os níveis são interpretados de acordo com os itens que os contém.

Considerando os critérios da literatura, foi classificado como item âncora aquele localizado em um nível em que a proporção de acerto dos respondentes é superior a 65%, a proporção de acerto na classe imediatamente anterior é inferior a 50% e diferença entre a

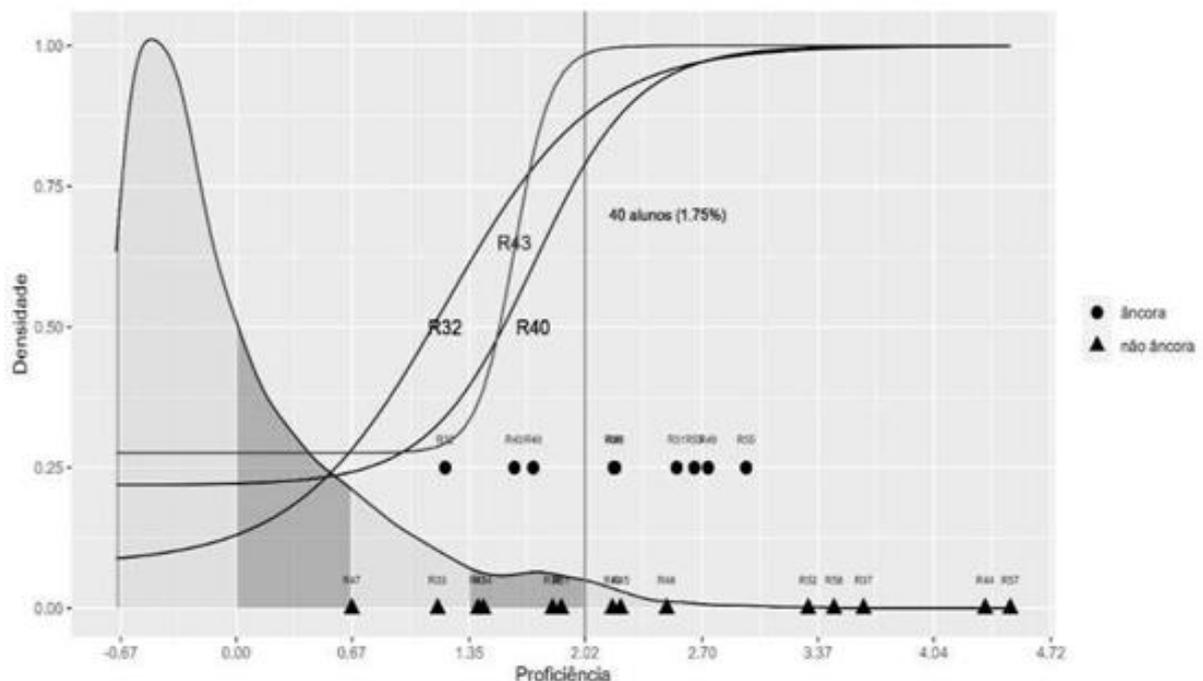
proporção de acerto na classe o nível averiguado e o nível imediatamente anterior maior do que 30%. Considerando este critério, obteve-se 9 itens âncoras. No entanto, os níveis (2.02-2.70] e (2.70-3.34] não atendem ao critério de número mínimo de 50 respondentes (Silva, 2000), fazendo com que exista apenas 3 itens âncoras localizados em 2 níveis, a saber.

O primeiro nível da escala está entre o intervalo de (0.67;1.35], e possui 203 candidatos. Este nível possui um item âncora: Q32. Então, o candidato que está nesse nível é capaz de resolver problemas que envolvam variação proporcional, direta ou inversa, diretamente relacionada às habilidades EF09MA08. No entanto, ao observar o Gráfico 4, podemos perceber que menos de 15% dos participantes dominam esse nível da escala.

O segundo nível da escala está entre o intervalo de (1.35;2.02], e possui 92 candidatos. Este nível possui dois itens âncoras Q40 e Q43. Então, o candidato que está nesse nível é capaz de resolver problemas que envolvam decomposição e cálculo de área, diretamente relacionada às habilidades EF07MA32 e EF08MA19. Apenas 6% dos candidatos dominam os níveis 1 e 2 da escala criada.

O gráfico 5 apresenta o posicionamento dos itens e a distribuição dos participantes, sendo possível perceber que não há itens em níveis menores de proficiência, impossibilitando que o instrumento seja capaz de distinguir as proficiências que lá estão. Isto é, métrica da escala de proficiência não é adequada para níveis mais baixos.

Gráfico 5. De posicionamento dos itens e distribuição dos participantes.



Fonte: Autor

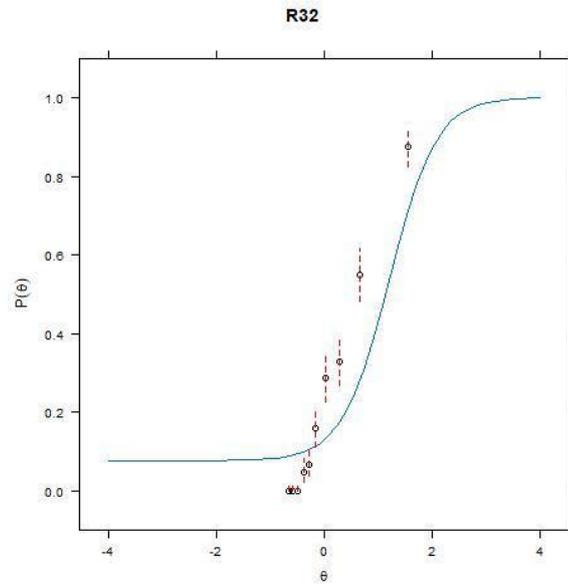
É possível verificar que o item R32 é o item com menor parâmetro de dificuldade, sendo adequado para avaliar alunos entre os níveis 0.67 e 1.35. O item R43 e R40 tem estimativas de dificuldade entre os níveis 1.35 e 2.02. Além disso, é possível perceber as diferenças entre as estimativas dos parâmetros “a”, ”b” e “c” dos itens, conforme apresentado no quadro abaixo:

Tabela 4. Parâmetro “a”, “b”, “c” das questões 32 ,40 e 43.

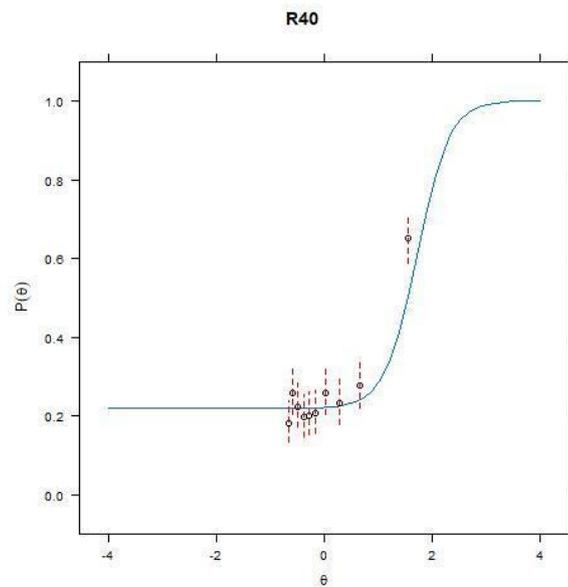
Item	Nível	A	b	c
R32	(0.67 a 1.35]	2,318 (0,45)	1,21 (0,08)	0,078 (0,03)
R40	(1.35 a 2.02]	3,328 (0,76)	1,72 (0,08)	0,219 (0,01)
R43	(1.35 a 2.02]	9,373 (4,49)	1,611 (0,06)	0,276 (0,01)

Fonte: **Autor**

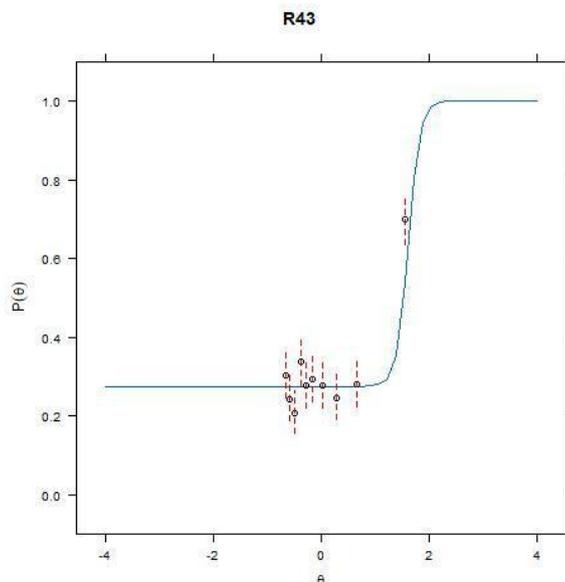
Percebe-se que item R32 possui menor estimativa para o parâmetro de chute, enquanto os itens R40 e R43, estimativas mais elevadas, acima de 20%. Conforme apresentado anteriormente, a ordem das estimativas para a dificuldade dos itens âncoras é R32, R43 e R40, sendo 1.21, 1.611 e 1.72, respectivamente. Por fim, as estimativas para os parâmetros de discriminação ordenam os itens âncoras em R32, R40 e R43, sendo a medida de erro da estimativa do último elevada. Está elevada medida de erro pode ser explicada pelo baixo percentual de participantes em regiões de maior proficiência. A comparação entre a curva característica e o resultado da prova, mostram isto.

Gráfico 6. Curva Característica dos 3 Itens Âncoras.

(a) CCI e dados observados para R32



(b) CCI e dados observados para R40



(c) CCI e dados observados para R43

Fonte: **Autor**

Apesar de o principal objetivo ser a classificação dos participantes, uma prova mais equilibrada, com relação à dificuldade dos itens, permitiria que o instrumento tivesse uma função social, identificando as principais dificuldades e habilidades adquiridas até o momento da avaliação.

4.1 IMPORTÂNCIA DA AVALIAÇÃO EDUCACIONAL

A sociedade capitalista tem interesse que a escola seja excludente e classificatória. Nessa perspectiva a avaliação é uma prática contínua e que sempre esteve presente no contexto da escola e se configura como um assunto polêmico, vista como autoritária, inflexível, ameaçadora e desagradável. Segundo Luckesi (2005, p.63) “Quanto mais ignorância e inconsciência, melhor para os segmentos dominantes da sociedade”. (LUCKESI, 2005, p.63)

Avaliar é o ato de atribuir valor a alguém ou a alguma coisa. As abordagens de avaliação podem gerar controvérsias tanto por parte dos avaliadores quanto dos avaliados, o que pode ser considerado injusto em alguns casos quando realizada de forma mecânica. Quando for necessário tomar decisões, é preciso fazer uma verificação cuidadosa, e analisar toda a situação, pois, uma decisão mal tomada poderá acarretar em sérios problemas.

Quando avaliamos ações, objetos, indivíduos, ou mesmo o desempenho de um aluno, estamos, de fato, atribuindo valores a esses elementos. Essa avaliação pode ser conduzida por meio de uma abordagem construtiva, envolvendo interação dinâmica, ou, alternativamente, pode adotar um caráter autoritário e repressivo. A escolha entre essas abordagens depende de nossa filosofia educacional, de nossas crenças sobre o processo de ensino-aprendizagem e dos objetivos que desejamos alcançar. Saber avaliar faz parte do processo de ensino aprendizagem, portanto, este trabalho contribui para análise de instrumentos de mensuração de proficiência.

De acordo com Freitas, (2008):

“O professor que trabalha de forma consciente e coerente, diversificando instrumentos de avaliação para que seja possível abranger todas as facetas do estudante entenderá a avaliação escolar como uma formalidade do sistema escolar, que retratará o resultado do seu próprio trabalho, por isso precisa ter certos cuidados na sua elaboração e aplicação” (FREITAS, 2008).

O esforço do Brasil em monitorar sistematicamente o desempenho dos alunos é de grande importância. Os processos de avaliação aplicados em todos os anos escolares fornecem informações cruciais que capacitam os professores com uma valiosa ferramenta para mensurar o progresso de seus alunos. No entanto, é fundamental destacar que essas medições não são perfeitas e ainda há a necessidade de assegurar que a avaliação seja capaz de mensurar de forma objetiva e precisa o aprendizado dos alunos.

Luckesi (1996, p. 165) afirma que:

“A avaliação escolar, em sentido lato, deve subsidiar o diagnóstico da situação em que se encontra o aluno, oferecendo recursos para orientá-lo a uma aprendizagem de qualidade, por meio do ensino adequado, pois, “Avaliar significa identificar impasses e buscar soluções” (LUCKESI, 1996, p.165).

Um professor que faz uma prova muito difícil prejudica o aluno, pois não distinguirá, por exemplo, o aluno que não sabe nada do que sabe pouco. Um professor que faz uma prova muito fácil não distinguirá, por exemplo, o aluno que sabe muito do aluno que sabe tudo. Portanto, é fundamental que o docente reflita sobre o processo de avaliação, e as análises psicométricas apresentadas neste trabalho auxiliam o processo de avaliação.

5 CONSIDERAÇÕES FINAIS

Proporcionar um estudo sobre a TRI é também uma maneira de buscar e entender seleções justas e adequadas ao que a sociedade demanda. Em meio a tanta novidade e evolução, ainda se encontra um modelo de avaliação tradicional, decorativa e punitiva. Produzir instrumentos de avaliação de alta qualidade é bastante complexo. Criar itens relevantes para compor uma prova e garantir que estes possuam dificuldade adequada e boa discriminação demanda recursos humanos e materiais.

No entanto, os estudos realizados com base na TCT abordam a questão do desempenho de maneira agregada, enquanto que a análise pela TRI permite a identificação pontual dos conhecimentos, capacidades e habilidades exigidas em cada nível. Nesse sentido, este estudo, fruto de pesquisa bibliográfica, destaca os pontos principais da TCT e da TRI, onde se conclui que uma não substitui a outra, ou seja, a TRI não veio para substituir a TCT, ambas são importantes processos avaliativos.

A TRI se destaca como importante ferramenta de avaliação, pois, possibilita diagnosticar e estimar os desempenhos individuais e coletivos dos alunos, favorecendo a compreensão das dificuldades dos itens e permitindo estabelecer relações de comparação entre as habilidades trabalhadas em relação ao coletivo, como individualmente. Ao desmembrar a proficiência em níveis, evidenciando os conhecimentos exigidos, é possível uma atuação pontual de professores, IES e respectivas autoridades nas questões específicas em que demonstram deficiências na aprendizagem. Essa análise possibilita estudar determinantes baseados em diferentes níveis de proficiência, contribuindo para o avanço nos estudos relacionados a esta área. Em conclusão, a análise psicométrica de exames classificatórios desempenha um papel indispensável na garantia da qualidade, validade e equidade dos processos de seleção educacional e profissional.

Através da aplicação de técnicas psicométricas, como a validação de construto, a análise de itens e a avaliação da confiabilidade, é possível examinar minuciosamente a estrutura e o funcionamento dos exames. Essa abordagem não apenas contribui para a melhoria contínua dos instrumentos de avaliação, mas também ajuda a minimizar possíveis vieses e desigualdades que possam surgir durante o processo de seleção.

Evidenciou-se que a análise psicométrica não se trata apenas de uma etapa técnica, mas também de um componente ético de suma importância. As decisões que envolvem a vida acadêmica e profissional dos candidatos têm implicações duradouras e profundas. Portanto, a

responsabilidade de desenvolver, administrar e interpretar exames classificatórios de maneira justa e embasada em evidências não pode ser subestimada.

Mesmo com avanços significativos na área da análise psicométrica, ainda persistem desafios a serem superados. A adaptação dos métodos psicométricos às novas modalidades de avaliação online, a consideração das diferentes habilidades e competências que os exames podem medir, bem como a busca pela equidade em um mundo cada vez mais diversificado, são temas que merecem atenção contínua e pesquisa aprofundada. Fica evidente que a análise psicométrica é uma ferramenta vital para aprimorar a qualidade e a justiça dos exames classificatórios. Ao continuar a explorar e aperfeiçoar essa abordagem, podemos construir sistemas de avaliação mais transparentes, confiáveis e inclusivos, contribuindo assim para um ambiente educacional e profissional mais igualitário e enriquecedor para todos os indivíduos.

Espera-se que o estudo em questão possa esclarecer alguns conceitos da psicometria como também atuar como um instrumento de difusão de conhecimentos sobre a TRI, contribuindo para uma maior discussão acerca desta teoria no ambiente escolar, em eventos científicos e em trabalhos de caráter acadêmico. E que, a partir dele, possa estimular o desenvolvimento de vários trabalhos futuros com a utilização da TRI, almejando políticas públicas que busquem sempre a melhoria do sistema educacional.

REFERÊNCIAS

- ANDRADE, D. F., TAVARES, H. R.; VALLE, R. C.. **Teoria da Resposta ao Item conceitos e aplicações**. 14º SINAPE: Associação Brasileira de Estatística (ABE). Caxambu, MG, 2000.
- ANDRIOLA, W. B. Descrição dos Principais Métodos para Detectar o Funcionamento Diferencial dos Itens (DIF). **Psicologia: Reflexão e Crítica**, v. 14, n. 3, p. 643–652, 2001. ISSN 0102-7972.
- ANDRIOLA, W. B. Doze motivos favoráveis à adoção do exame nacional do ensino médio (ENEM) pelas instituições federais de ensino superior (IFES). **Ensaio**, v. 19, n. 70, p. 107–126, 2011. ISSN 01044036.
- APA, A. P. A. **Manual diagnóstico e estatístico de transtornos mentais: DSM-5**. 5 ed. Porto Alegre: Artmed, 2014.
- AYALA, R. J. D. **The theory and practice of item response theory**. New York: Guilford Publications, 2013.
- BAKER, F. B. **The basics of item response theory**. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD, 2001.
- BIRNBAUM, Allan. **Some latent trait models and their use in inferring an examinee's ability**. Em F. M. Lord e M.R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp. 395-479). Reading, MA: Addison-Wesley, 1968.
- BOCK, R. D.. **Estimating item parameters and latent ability when responses are scored in two or more nominal categories**. *Psychometrika*, New York, 37 (1), 29-51. doi: <https://doi.org/10.1007/BF02291411>. 1972.
- BORGATTO, A. F; ANDRADE D. F. **Análise clássica de testes com diferentes graus de dificuldade**. *Est. Aval. Educ.*, São Paulo, v. 23, n. 52, p. 146-156, maio/ago. 2012.
- BRASIL, Ministério da Educação. **Base Nacional Comum Curricular: educação é a base**. Educação é a base. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf. Acesso em: 10 jan. 2021.
- BRASIL. **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Enem: relatório pedagógico 2004. Brasília, DF: MEC/ Inep, 2010.
- BRASIL. Ministério da Educação e Cultura. **Sistema Nacional de Avaliação da Educação Básica: Saeb 1995; relatório técnico**. São Paulo: Fundação Carlos Chagas; Rio de Janeiro: Fundação Cesgranrio, 1996.
- CABRITO, Belmiro Gil. Avaliar a qualidade em educação: avaliar o quê? Avaliar como? Avaliar para quê?" **Cadernos Cedes** 29, 2009.

CAMARGO *et al.* Desempenho dos alunos de Ciências Contábeis na prova Enade/2012: uma aplicação da Teoria da Reposta ao Item. **REPeC – Revista de Educação e Pesquisa em Contabilidade**, ISSN 1981-8610, Brasília, v.10, n. 3, art. 6, p. 332-355, jul./set. 2016

COSTA, M. C., Lima, S. H. O., & Soares, D. J. M. **Uma proposta de análise de itens da prova preparatória para o Enade aplicada aos discentes de engenharia civil**. IFMG – Campus Avançado Piumhi, 2020.

DAVOK, D.F. Qualidade em educação. **Avaliação: Revista da Avaliação da Educação Superior**, Sorocaba, v.12, n.3, p.505-513,2007.

DEVELLIS, R. F. **Scale Development: Theory and applications**. 4. ed. Los Angeles: Sage Publications, 2016.

EDELEN, M. O. & Reeve, B. B. **Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement**. *Quality of Life Research*, 2007.

EDELEN, M. O. & Reeve, B. B. **Applying item response theory (IRT)**, 2007. modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, 16(1), pp. 5-18. <http://dx.doi>.

FLETCHER, Phillip R. A. Teoria da resposta ao item: medidas invariantes do desempenho escolar. **Ensaio: avaliação e políticas públicas em educação**, v.1, n. 2, p. 21-28, jan./mar. 1994.

FREITAS, Luiz Carlos de. **Crítica da Organização do Trabalho pedagógico e da Didática**. 7ª Ed. Campinas – SP, Papirus Editora, 1995.

GOMES, Válter. MACHADO, Maria de Lourdes. SARAIVA, Ernani Viana. O ensino superior no brasil: breve histórico e caracterização. In: **anais do IX Colóquio Internacional sobre Gestão universitária na América do Sul**. 25 a 27 de novembro de 2009. Florianópolis: SC.

GOUVEIA, Nicásio; Fernando De Jesus Moreira JuniorII, Angela Pellegrin AnsujiIII. **Utilização da Teoria da Resposta ao Item na Análise do Instrumento da Autoavaliação Institucional na Pós-graduação da UFSM**. RS, 2020.

HUTZ, C. S.; BANDEIRA, D. R.; TRENTINI, C. M. **Psicometria**. Artmed Editora, 2015. (Avaliação Psicológica). ISBN 9788582712368. Disponível em: <<https://books.google.com.br/books?id=cVIIcGAAQBAJ>>. Acesso: 22 ago. 2023.

IBGE. **IBGE - Instituto Brasileiro de Geografia e Estatística**. 2019. 1–5 p. Disponível em: <<https://www.ibge.gov.br/>>.

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Censo da educação superior 2011**. Brasília: Inep; MEC, 2011. Disponível em: <<http://goo.gl/tYbnLd>>.

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo Escolar da Educação Básica 2013**: resumo técnico / Instituto Nacional de Estudos e Pesquisas

Educacionais Anísio Teixeira. – Brasília: O Instituto, 2014.

LOPES, F. L.; VENDRAMINI, C. M. M. Propriedades psicométricas das provas de pedagogia do Enade via TRI. **Avaliação: Revista da Avaliação da Educação Superior**, Campinas, SP, v. 20, n. 1, p. 27-47, mar. 2015.

LUCKESI, C.C. Avaliação da Aprendizagem Escolar: estudos e proposições. São Paulo, Cortez, 1995.

_____, C.C. Avaliação da aprendizagem escolar. 14 ed. São Paulo: Cortez, 2002.

MARQUES, Thiago Valentim. **A teoria da resposta ao item no IFRN**: uma aplicação com os dados do exame de seleção dos cursos técnicos integrados. Monografia (Bacharelado em Estatística) Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Departamento de Estatística. Natal, 2021.

MEC. **Ministério da Educação-história**. Disponível em:<<http://portal.mec.gov.br/institucional/historia>. Acesso em 20 jan. 2023.

MOREIRA JÚNIOR, A. **Educação e tecnologias**: o novo perfil do professor. Editora Vozes, 2010.

MOREIRA JUNIOR, F. J. Aplicações da teoria da resposta ao Item (TRI) no Brasil. **Revista Brasileira de Biometria**, Marília, v.28, n.4, p.137-170,2010.

MOREIRA, F. J., Jr (2010). **Aplicações da teoria da resposta ao item (TRI) no Brasil**. Rev. Bras. Biom., 28(4), pp. 137-170. org/10.1007/s11136-007-9198-0

PASQUALI, L. Provão (ENC) de Psicologia 2000 e 2001: análise dos parâmetros psicométricos. In: PRIMI, R. (Org.). **Temas em avaliação psicológica**. Campinas: Instituto Brasileiro de Avaliação Psicológica, 2002. p. 152-178.

PASQUALI, Luiz, &Primi, Ricardo. (2003). Fundamentos da teoria da resposta ao item: TRI. **Avaliação Psicológica**, 2(2), 99- 110. Recuperado em 08 de abril de 2020, de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000200002&lng=pt&tlng=pt.

PASQUALI, Luiz. (2018). **TRI–Teoria de resposta ao item**: Teoria, procedimentos e aplicações. Curitiba: Editora Appris. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. www.R-project.org/.

PASQUALI, Luiz.. Psicometria. **Revista da Escola de Enfermagem da USP**, 43(spe), 992-999, 2009. doi: <https://doi.org/10.1590/S0080-62342009000500002>.

PRIMI, R.; HUTZ, C. S.; SILVA, M. C. R. da. A prova do ENADE de psicologia 2006: concepção, construção e análise psicométrica da prova. **Avaliação Psicológica**. 2011, v.10, n.3, pp. 271-294.

R Core Team (2023). **_R: A Language and Environment for Statistical Computing_**. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.

RABELO, M.. **Avaliação educacional**: fundamentos, metodologia e aplicações no contexto brasileiro. Rio de Janeiro: SBM, 29, 30-31. 2013.

RASH, G. **Probabilistic models for som intelligence and attainment test**. Copenhagen: Danish Institute for Educational research and St. Paul. MN: Assesment System Corporation, 1960.

SILVA, P. A. da C. **Equalização da Prova Belém na escala SAEB**. Dissertação (Mestrado Profissional em Métodos e Gestão da Avaliação) – Universidade Federal de Santa Catarina, Centro Tecnológico. Florianópolis, 2017.

SOARES, Denilson Junio Marques. Soares, Talita Emidio Andrade. E dos Santos, Wagner. Análise da qualidade psicométrica da prova de matemática do Exame Nacional do Ensino Médio brasileiro de 2018. **Revista Actualidades Investigativas en Educación**, 2021. 21(1), 1-29. Doi. 10.15517/aie.v21i1.42338.

TRAVITZKI, R.. Avaliação da qualidade do Enem 2009 e 2011 com técnicas psicométricas. **Estudos em Avaliação Educacional**. 28(67), 256-288. 2017. doi: <http://dx.doi.org/10.18222/ae.v28i67.3910>.

VENDRAMINI, Claudette Maria Medeiros.E DIAS, Anelise Silva. Teoria de Resposta ao Item na análise de uma prova de estatística em universitários. **Psico- USF**,. 10(2), 201-210. 2005. Doi: <https://dx.doi.org/10.1590/S141382712005000200012>

VENDRAMINI, M.M.V.; SILVA, M.C.; CANALE, M. Análise de itens de uma prova de raciocínio estatístico. **Psicologia em estudo**. Maringá, v. 9, n. 3, p. 487-498, set./dez. 2004.

VOGADO, I. R.. **Uma análise das questões das provas do exame classificatório do IFPI**: Campus Corrente Integrado ao Médio. Trabalho de Conclusão de Curso (Graduação em Matemática – Licenciatura) – Instituto Federal de Educação, Ciência e Tecnologia, Corrente, 2021.

ZANELLA, A.: LOPES, L. F. D.; SEIDEL, E.J. Diagnóstico do ensino aprendizagem e satisfação dos alunos nas disciplinas de estatística da UFSM. **Revista GEPROS: Gestão da Produção, Operações e Sistemas**, Bauru, v.4, n.3,p.123-140, 2009.

ANEXO 1 – TRI

Tabela 5. Apresentação do nível e parâmetros “a”, “b”, “c” das respostas.

Item	Nível	a	b	C
R31	(2.7,3.34]	2,828 (0,88)	2,551 (0,18)	0,114 (0,01)
R32	(0.67,1.35]	2,318 (0,45)	1,21 (0,08)	0,078 (0,03)
R33	não âncora	1,608 (0,36)	1,167 (0,15)	0,028 (0,06)
R34	não âncora	3,129 (0,7)	1,431 (0,08)	0,264 (0,02)
R37	não âncora	1,01 (0,67)	3,634 (1,03)	0,168 (0,03)
R38	não âncora	2,537 (0,55)	1,832 (0,09)	0,108 (0,01)
R39	(2.02,2.7]	2,08 (0,83)	2,187 (0,16)	0,212 (0,02)
R40	(1.35,2.02]	3,328 (0,76)	1,72 (0,08)	0,219 (0,01)
R41	não âncora	2,455 (0,49)	1,4 (0,08)	0,137 (0,02)
R42	não âncora	2,202 (0,56)	2,179 (0,14)	0,139 (0,01)
R43	(1.35,2.02]	9,373 (4,49)	1,611 (0,06)	0,276 (0,01)
R44	não âncora	1,316 (1,17)	4,339 (1,97)	0,165 (0,01)
R45	não âncora	2,113 (0,51)	2,226 (0,14)	0,124 (0,01)
R46	(2.02,2.7]	4,917 (2,21)	2,195 (0,1)	0,175 (0,01)
R47	não âncora	0,497 (0,08)	0,669 (0,19)	0,003 (0,03)
R48	não âncora	4,91 (3,09)	2,493 (0,16)	0,223 (0,01)
R49	(2.7,3.34]	5,166 (3,15)	2,733 (0,2)	0,198 (0,01)
R51	não âncora	1,203 (0,53)	1,883 (0,2)	0,256 (0,07)
R52	não âncora	1,196 (0,7)	3,314 (0,74)	0,165 (0,03)
R53	(2.7,3.34]	2,051 (1,05)	2,654 (0,29)	0,139 (0,02)
R55	(2.7,3.34]	4,092 (2,55)	2,954 (0,28)	0,119 (0,01)
R57	não âncora	0,363 (0,32)	4,484 (1,05)	0,016 (0,2)
R58	não âncora	1,403 (1,1)	3,463 (1,05)	0,303 (0,02)

ANEXO 2 – TCT

Questão	Dificuldade	bis	bisp	discr	Habilidade	Descritor
R31	9,75%	0,434596	0,279642	17,25%	(EF08MA01)	D24
R32	16,41%	0,664829	0,531143	38,29%	(EF06MA09)	D26
R33	17,47%	0,659105	0,541571	39,89%	(EF07MA12)	D29
R34	25,75%	0,590715	0,587197	46,71%	(EF06MA24)	D25
R35	11,90%	0,384195	0,267814	20,08%	(EF09MA03)	D27
R36	27,92%	0,509264	0,529446	45,20%	(EF06MA05)	D24
R37	16,35%	0,438357	0,349627	26,85%	(EF09MA03)	D20
R38	13,57%	0,570627	0,419994	28,04%	(EF09MA08)	D26
R39	20,22%	0,541134	0,475962	36,60%	(EF09MA08)	D20
R40	19,73%	0,570593	0,495997	37,33%	(EF08MA19)	D20
R41	19,92%	0,609072	0,531799	39,75%	(EF09MA05)	D28
R42	13,38%	0,512528	0,37506	26,26%	(EF07MA02)	D28
R43	23,33%	0,578713	0,546194	41,86%	(EF09MA06)	D31
R44	13,07%	0,426924	0,309381	22,69%	(EF08MA06)	D33
R45	13,58%	0,509896	0,375438	25,80%	(EF09MA09)	D31
R46	15,25%	0,470488	0,363982	26,03%	(EF08MA08)	D30
R47	31,58%	0,579149	0,647573	53,61%	(EF09MA09)	D33
R48	16,68%	0,46047	0,370572	29,32%	(EF08MA19)	D13
R49	16,58%	0,400493	0,321454	25,75%	(EF09MA09)	D32
R50	NULA	NULA	NULA	NULA	(EF08MA07)	NULA
R51	0,264888	0,581299	0,586831	0,472095	(EF08MA21)	D14
R52	0,150852	0,468713	0,360951	0,274931	(EF08MA19)	D13
R53	0,128737	0,481824	0,346959	0,23742	(EF09MA06)	D32
R54	0,149617	0,356853	0,273834	0,232388	(EF09MA12)	D11
R55	0,081913	0,371712	0,223441	0,149588	(EF08MA17)	D3
R56	0,099209	0,350948	0,227397	0,163312	(EF09MA13)	D10
R57	0,144552	0,480938	0,363637	0,261665	(EF08MA17)	D9
R58	0,25315	0,458599	0,451741	0,39021	(EF08MA14)	D4

R59	0,241784	0,441509	0,424492	0,375572	(EF09MA14)	nenhum
R60	0,116877	0,378376	0,261848	0,206313	(EF09MA16)	D9

Fonte: Autor.

ANEXO 3 – TABELAS E QUADROS

Tabela 6. Classificação dos Tópicos, IFPI 2022.1

Tópicos	Frequência Observada	Porcentagem
Geometria	6	21,42%
Grandezas e Medidas	2	7,14%
Números e Operações	10	35,71%
Álgebra e Funções	8	28,57%
Estatística e Probabilidade	0	0,00%
Não Contemplado	2	7,14%

Fonte: Autor.

Classificação das questões do classificatório do IFPI, integrado ao médio de 2022.1, referente aos descritores da prova SAEPI,

Tabela 7. Classificação dos descritores, IFPI 2022.1

Descritores	Frequência Observada	Porcentagem
D03	1	3,57%
D04	1	3,57%
D09	2	7,14%
D10	1	3,57%
D11	1	3,57%
D13	2	7,14%
D20	3	10,71%
D24	2	7,14%
D25	1	3,57%
D26	2	7,14%
D27	1	3,57%
D28	1	3,57%
D29	1	3,57%
D30	1	3,57%
D31	2	7,14%
D32	2	7,14%

D33	2	7,14%
Não contemplado	2	7,14%

Fonte: Autor.

Quadro 1. Código de descritores SAEPI 9º ano.

CÓDIGO	DESCRITOR
D01	Identificar a localização/movimentação de objeto em mapas, croquis e outras representações gráficas.
D02	Identificar propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando-as com as suas planificações.
D03	Identificar propriedades de triângulos pela comparação de medidas de lados e ângulos.
D04	Identificar relação entre quadriláteros por meio de suas propriedades.
D05	Reconhecer a conservação ou modificação de medidas dos lados, do perímetro, da área em ampliação e/ou redução de figuras poligonais usando malhas quadriculadas.
D06	Reconhecer ângulos como mudança de direção ou giros, identificando ângulos retos e não-retos.
D07	Reconhecer que as imagens de uma figura construída por uma transformação homotética são semelhantes, identificando propriedades e/ou medidas que se modificam ou não se alteram.
D08	Resolver problema utilizando propriedades dos polígonos (soma de seus ângulos internos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares).
D09	Interpretar informações apresentadas por meio de coordenadas cartesianas.
D10	Utilizar relações métricas do triângulo retângulo para resolver problemas significativos.
D11	Reconhecer círculo/circunferência, seus elementos e algumas de suas relações.
D12	Resolver problema envolvendo o cálculo de perímetro de figuras planas.
D13	Resolver problema envolvendo o cálculo de área de figuras planas.
D14	Resolver problema envolvendo noções de volume.
D15	Resolver problema utilizando relações entre diferentes unidades de medida.
D16	Identificar a localização de números inteiros na reta numérica.
D17	Identificar a localização de números racionais na reta numérica.
D18	Efetuar cálculos com números inteiros, envolvendo as operações (adição, subtração, multiplicação, divisão, potenciação).
D19	Resolver problema com números naturais, envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão, potenciação).
D20	Resolver problema com números inteiros envolvendo as operações (adição, subtração, multiplicação, divisão, potenciação).
D21	Reconhecer as diferentes representações de um número racional.
D22	Identificar fração como representação que pode estar associada a diferentes significados.
D23	Identificar frações equivalentes.
D24	Reconhecer as representações decimais dos números racionais como uma extensão do sistema de numeração decimal, identificando a existência de “ordens” como décimos, centésimos e milésimos.
D25	Efetuar cálculos que envolvam operações com números racionais (adição, subtração, multiplicação, divisão, potenciação).

D26	Resolver problema com números racionais envolvendo as operações (adição, subtração, multiplicação, divisão, potenciação).
D27	Efetuar cálculos simples com valores aproximados de radicais.
D28	Resolver problema que envolva porcentagem.
D29	Resolver problema que envolva variação proporcional, direta ou inversa, entre grandezas.
D30	Calcular o valor numérico de uma expressão algébrica.
D31	Resolver problema que envolva equação do 2º grau.
D32	Identificar a expressão algébrica que expressa uma regularidade observada em sequências de números ou figuras (padrões).
D33	Identificar uma equação ou inequação do 1º grau que expressa um problema.
D34	Identificar um sistema de equações do 1º grau que expressa um problema.
D35	Identificar a relação entre as representações algébrica e geométrica de um sistema de equações do 1º grau.
D36	Resolver problema envolvendo informações apresentadas em tabelas e/ou gráficos.
D37	Associar informações apresentadas em listas e/ou tabelas simples aos gráficos que as representam e vice-versa.

Fonte: **SAEPI, 2021.**