

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
CENTRO DE POLÍTICAS PÚBLICAS E AVALIAÇÃO DA EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO PROFISSIONAL EM GESTÃO
E AVALIAÇÃO DA EDUCAÇÃO PÚBLICA

Wilian Ferreira Rocha

A Dimensão Pedagógica nas Avaliações em Larga Escala: a Teoria de Van Hiele como suporte para produção e análise de itens de Matemática

Juiz de Fora

2026

Wilian Ferreira Rocha

A Dimensão Pedagógica nas Avaliações em Larga Escala: a Teoria de Van Hiele como suporte para produção e análise de itens de Matemática

Dissertação apresentada ao Programa de Pós-graduação Profissional em Gestão e Avaliação da Educação Pública, da Faculdade de Educação, da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Mestre em Gestão e Avaliação da Educação Pública.

Orientadora: Profa. Dra. Liamara Scortegagna

Juiz de Fora

2026

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Ferreira Rocha, Wilian.

A Dimensão Pedagógica nas Avaliações em Larga Escala : a Teoria de Van Hiele como suporte para produção e análise de itens de Matemática / Wilian Ferreira Rocha. -- 2026.

111 f. : il.

Orientador: Liamara Scortegagna

Dissertação (mestrado profissional) - Universidade Federal de Juiz de Fora, Faculdade de Educação/CAEd. Programa de Pós-Graduação em Gestão e Avaliação da Educação Pública, 2026.

1. Avaliação em Larga Escala. 2. Elaboração de Itens. 3. Ensino de Geometria. 4. Teoria de Van Hiele. 5. Teoria de Resposta ao Item. I. Scortegagna, Liamara, orient. II. Título.

Wilian Ferreira Rocha

A Dimensão Pedagógica nas Avaliações em Larga Escala: a teoria de Van Hiele como suporte para produção e análise de itens de Matemática

Dissertação apresentada ao Programa de Pós-Graduação em Gestão e Avaliação da Educação Pública da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Gestão e Avaliação da Educação Pública. Área de concentração: Gestão e Avaliação da Educação Pública.

Aprovada em 20 de julho de 2026.

BANCA EXAMINADORA

Prof.(a)Dr.(a) Liamara Scortegagna - Orientador

Universidade Federal de Juiz de Fora

Prof.(a)Dr.(a) Wagner Silveira Rezende

Universidade Federal de Juiz de Fora

Prof.(a)Dr.(a) Marcus Bessa de Menezes

UFCG

Juiz de Fora, 06/07/2026.



Documento assinado eletronicamente por **Liamara Scortegagna, Professor(a)**, em 20/07/2026, às 11:06, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **MARCUS BESSA DE MENEZES, Usuário Externo**, em 24/07/2026, às 15:28, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wagner Silveira Rezende, Professor(a)**, em 05/08/2026, às 16:41, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **3067139** e o código CRC **815281E3**.

AGRADECIMENTOS

O desenvolvimento de uma dissertação, sobretudo no campo da educação, é algo caro e em alguns momentos bastante difícil. Ainda mais quando os efeitos desse trabalho podem atingir milhares de estudantes, se considerarmos a adoção de processos em uma instituição do tamanho dessa da qual faço parte atualmente. A certeza é que toda a construção desse trabalho contribuiu para meu crescimento pessoal.

Em primeiro lugar, agradeço a Deus por me permitir diariamente contribuir para a vida das outras pessoas e iluminar os meus caminhos. Na filosofia cristã, a vida que vale ser vivida é aquela em que o amor ao próximo é o que nos guia.

Agradeço aos meus pais, Waldomiro e Vilma, que, mesmo um tantinho longe de mim, nessa imensidão que é esse estado de Minas Gerais, me apoiam e fizeram de tudo para que eu pudesse chegar aonde estou hoje.

À minha esposa, Edyenis, um agradecimento mais do que especial. Só ela sabe como foi todo esse processo de construção desse trabalho.

Ao meu sogro João e à minha sogra Wilma, que me aguentaram e me aguentam diariamente.

Aos meus irmãos, Waldeir, Felipe e Tatiane, que muito contribuíram em toda a minha caminhada.

A toda a minha família, de forma especial à minha prima Sarah e ao Eli, por terem me acolhido em Juiz de Fora lá no início.

À minha orientadora, a Professora Dra. Liamara Scortegagna, e ao meu agente de suporte acadêmico, o Professor Me. Vitor Marques, pelas contribuições no desenvolvimento deste trabalho.

Aos membros da banca, o Professor Dr. Marcus Bessa de Menezes e o Professor Dr. Wagner Silveira Rezende, que gentilmente se dispuseram a colaborar com esta pesquisa.

A todos os professores e agentes de suporte acadêmico do PPGP, pelas contribuições e ensinamentos.

A todos os colegas do PPGP, com quem partilhei os momentos de muito trabalho. As experiências e trocas de aprendizados foram grandiosas.

À Fundação CAEd, por ter permitido a realização desse estudo, tão importante para meu aprimoramento profissional.

E a todos e todas que, direta ou indiretamente, contribuíram para a realização dessa pesquisa.

RESUMO

Este trabalho foi desenvolvido no âmbito do mestrado em Gestão e Avaliação da Educação Pública, da Universidade Federal de Juiz de Fora/CAEd. Nele, buscou-se entender como o processo de aquisição do conhecimento geométrico ocorre e como ele se manifestou nos itens aplicados em avaliações em larga escala. Nesse sentido, a pergunta norteadora proposta foi: Como os níveis de pensamento geométrico da Teoria de Van Hiele se manifestam nos itens de Geometria das avaliações em larga escala, qual a relação dessa manifestação com o comportamento empírico desses itens, e de que modo essa análise integrada pode subsidiar a proposição de um Plano de Ação Educacional para a elaboração de itens mais alinhados ao desenvolvimento cognitivo dos estudantes? Para respondê-la, adotou-se o objetivo geral de analisar como os níveis de pensamento geométrico, propostos pela Teoria de Van Hiele, são mobilizados nos itens de Geometria presentes nas avaliações em larga escala, relacionando essa estruturação pedagógica ao comportamento estatístico dos itens, a fim de propor um Plano de Ação para a elaboração de instrumentos mais adequados ao desenvolvimento cognitivo dos estudantes da Educação Básica. Especificamente, objetivou-se caracterizar o processo de elaboração de itens, seus critérios e considerações pedagógicas que podem ser empregadas; investigar os itens do eixo do conhecimento de Geometria quanto ao desenvolvimento das tarefas propostas nos itens e de suas características intrínsecas; e propor critérios para o desenvolvimento de itens que sejam capazes de gerar informações em diferentes pontos da escala. Metodologicamente, utilizou-se nesta pesquisa uma abordagem mista (qualitativa e quantitativa) com delineamento sequencial exploratório, em duas fases. Para a análise, foram utilizados 94 itens públicos do eixo de Geometria (5º e 9º anos do Ensino Fundamental e 3ª série do Ensino Médio), desenvolvidos pela Fundação CAEd com base na matriz de referência do Saeb de 2001. A análise pedagógica dos itens fundamentou-se na Teoria de Van Hiele e na Taxonomia de Bloom, sendo posteriormente cruzada com os parâmetros estatísticos de discriminação, dificuldade e acerto casual (a, b e c) gerados pela Teoria da Resposta ao Item (TRI). Concluiu-se que o rol de itens foi modesto para estabelecer correlações fortes entre os níveis da Teoria de Van Hiele e os parâmetros psicométricos, mas que a elaboração de itens com intencionalidade diagnóstica permite avaliar estudantes estagnados em um nível de conhecimento geométrico inferior. A partir disso, foi feita a elaboração de um Plano de Ação Educacional voltado à Equipe de Instrumentos de Avaliação. O plano propõe a adoção de diretrizes pedagógicas e de um protocolo para a construção de itens que contemplem o *continuum* da aprendizagem geométrica, qualificando a mensuração da proficiência e os diagnósticos educacionais.

Palavras-chave: Avaliação em Larga Escala. Elaboração de Itens. Ensino de Geometria. Teoria de Van Hiele. Teoria de Resposta ao Item.

ABSTRACT

This study was developed within the scope of the Master's program in Public Education Management and Evaluation at the Federal University of Juiz de Fora (UFJF/CAEd). It sought to understand how the process of acquiring geometric knowledge occurs and how it is manifested in items applied in large-scale assessments. In this sense, the guiding question was: How do the levels of geometric thinking of the Van Hiele Theory manifest in the Geometry items of large-scale assessments, what is the relationship between this manifestation and the empirical behavior of these items, and how can this integrated analysis support the proposition of an Educational Action Plan to develop items more aligned with students' cognitive development? To answer this, the general objective was to analyze how the geometric thinking levels proposed by the Van Hiele Theory are mobilized in Geometry items present in large-scale assessments, relating this pedagogical structuring to the statistical behavior of the items, in order to propose an Action Plan for developing instruments more suited to the cognitive development of Basic Education students. Specifically, the study aimed to characterize the item development process, its criteria, and the pedagogical considerations that can be employed; investigate the items of the Geometry knowledge axis regarding the development of the proposed tasks and their intrinsic characteristics; and propose criteria for developing items capable of generating information at different points of the scale. Methodologically, this research utilized a mixed-methods approach (qualitative and quantitative) with an exploratory sequential design, conducted in two phases. For the analysis, 94 public items from the Geometry axis (5th and 9th grades of Elementary Education and 3rd grade of High School) were used, developed by the Fundação CAEd based on the 2001 Saeb reference matrix. The pedagogical analysis of the items was based on the Van Hiele Theory and Bloom's Taxonomy and was subsequently cross-referenced with the statistical parameters of discrimination, difficulty, and guessing (a, b, and c) generated by Item Response Theory (IRT). It was concluded that the set of items was too modest to establish strong correlations between the Van Hiele Theory levels and the psychometric parameters, but that the development of items with diagnostic intent allows for the evaluation of students stagnated at a lower level of geometric knowledge. Based on these findings, an Educational Action Plan was developed for the Assessment Instruments Team. The plan proposes the adoption of pedagogical guidelines and a protocol for the construction of items that encompass the geometric learning *continuum*, thereby enhancing the measurement of proficiency and educational diagnostics.

Keywords: Large-Scale Assessment. Item Development. Geometry Teaching. Van Hiele's Theory. Item Response Theory.

LISTA DE ILUSTRAÇÕES

Figura 1 - Exemplo de item	20
Figura 2 - Processo de elaboração de itens	22
Figura 3 – Curva característica do item.....	36
Figura 4 – Fluxo de seleção dos estudos (Fluxograma Prisma).....	53
Figura 5 – Relação entres descritores do eixo de Geometria da Matriz Saeb 2001..	58
Figura 6 – Possíveis figuras para um item de planificação para o 9EF	89
Figura 7 – Possíveis figuras para um item de planificação para o 5EF	89
Figura 8 – Prismas com planificação usual, menos usual (a) e com sobreposição (b)	90

LISTA DE GRÁFICOS

Gráfico 1 – Curva característica dos itens do descritor D01 - 5EF	100
Gráfico 2 – Curva característica dos itens do descritor D02 - 5EF	101
Gráfico 3 – Curva característica dos itens do descritor D03 - 5EF	101
Gráfico 4 – Curva característica dos itens do descritor D04 - 5EF	102
Gráfico 5 – Curva característica dos itens do descritor D05 - 5EF	102
Gráfico 6 – Curva característica dos itens do descritor D01 - 9EF	103
Gráfico 7 – Curva característica dos itens do descritor D02 - 9EF	103
Gráfico 8 – Curva característica dos itens do descritor D03 - 9EF	104
Gráfico 9 – Curva característica dos itens do descritor D04 - 9EF	104
Gráfico 10 – Curva característica dos itens do descritor D05 - 9EF	105
Gráfico 11 – Curva característica dos itens do descritor D06 - 9EF	105
Gráfico 12 – Curva característica dos itens do descritor D07 - 9EF	106
Gráfico 13 – Curva característica dos itens do descritor D08 - 9EF	106
Gráfico 14 – Curva característica dos itens do descritor D09 - 9EF	107
Gráfico 15 – Curva característica dos itens do descritor D10 - 9EF	107
Gráfico 16 – Curva característica dos itens do descritor D11 - 9EF	108
Gráfico 17 – Curva característica dos itens do descritor D01 - 3EM	108
Gráfico 18 – Curva característica dos itens do descritor D02 - 3EM	109
Gráfico 19 – Curva característica dos itens do descritor D03 - 3EM	109
Gráfico 20 – Curva característica dos itens do descritor D04 - 3EM	110
Gráfico 21 – Curva característica dos itens do descritor D05 - 3EM	110
Gráfico 22 – Curva característica dos itens do descritor D06 - 3EM	111
Gráfico 23 – Curva característica dos itens do descritor D07 - 3EM	111
Gráfico 24 – Curva característica dos itens do descritor D08 - 3EM	112
Gráfico 25 – Curva característica dos itens do descritor D09 - 3EM	112
Gráfico 26 – Curva característica dos itens do descritor D10 - 3EM	113

LISTA DE TABELAS

Tabela 1 – Escala de proficiência (p) de Matemática	39
Tabela 2 – Relação de estudos selecionados na fase dois de triagem	50
Tabela 3 – Itens cedidos pela Fundação CAEd	56
Tabela 4 – Itens e medidas cedidas pela Fundação CAEd, com a descrição da tarefa do item	62
Tabela 5: Síntese do protocolo para produção de itens	87

LISTA DE QUADROS

Quadro 1 – Sistematização dos Achados e Propostas de Ação Institucional	80
Quadro 2 – Eixo 1: Formação e Instrumentalização	82
Quadro 3 - Proposta de Sumário para o Guia Suplementar.....	84
Quadro 4 – Eixo 2: Pesquisa e Expansão Metodológica.....	85

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
BIB	Blocos Incompletos Balanceados
BNCC	Base Nacional Comum Curricular
CAEd	Centro de Políticas Públicas e Avaliação da Educação
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CAT	Teste Adaptativo Computadorizado
CCI	Função Característica do Item
CONSU	Conselho Superior
DOI	Digital Object Identifier
EAP	expected a posteriori
Enem	Exame Nacional do Ensino Médio
FADEPE	Fundação de Apoio e Desenvolvimento ao Ensino, Pesquisa e Extensão
Fundação CAEd	Fundação Centro de Políticas Públicas e Avaliação da Educação
Ideb	Índice de Desenvolvimento da Educação Básica
Inep	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LAME	Laboratório de Avaliação e Medidas Educacionais
MEC	Ministério da Educação
PAE	Plano de Ação Educacional .
PPGP	Programa de Pós-graduação em Gestão e Avaliação da Educação Pública
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROAV	Programa de Apoio à Avaliação Educacional
Saeb	Sistema Nacional de Avaliação da Educação Básica
SIMAVE	Sistema Mineiro de Avaliação da Educação Pública
SisLAME	Sistema de Administração e Controle Escolar
TAI	Teste Adaptativo Informatizado
TCT	Teoria Clássica dos Testes
TRI	Teoria da Resposta ao Item
UFJF	Universidade Federal de Juiz de Fora

SUMÁRIO

1	INTRODUÇÃO.....	13
2	AS AVALIAÇÕES EM LARGA ESCALA E OS DESAFIOS NA PRODUÇÃO DE ITENS	16
2.1	O CENTRO DE POLÍTICAS PÚBLICAS E AVALIAÇÃO DA EDUCAÇÃO (CAEd).....	16
2.2	O CONTEXTO DA PRODUÇÃO DE ITENS E DOS TESTES PADRONIZADOS	19
2.3	AS AVALIAÇÕES, MOMENTOS NO CONTEXTO ESCOLAR E CONSTRUÇÃO DOS TESTES	24
2.4	DELIMITAÇÃO DO CASO	28
3	ASPECTOS PEDAGÓGICOS E ANÁLISE DOS ITENS DO EIXO DO CONHECIMENTO DE GEOMETRIA.....	31
3.1	AS AVALIAÇÕES EM LARGA ESCALA E A PRODUÇÃO DE MEDIDAS.....	31
3.1.1	A produção de medidas dos itens.....	34
3.1.2	A escala de proficiência do Saeb	38
3.1.3	Os itens para avaliações em larga escala	39
3.1.4	A Teoria de Van Hiele.....	41
3.1.5	Aspectos pedagógicos dos itens e das avaliações em larga escala	45
3.2	LEVANTAMENTO BIBLIOGRÁFICO E DOS DADOS DA PESQUISA	46
3.3	ANÁLISE DA MATRIZ E DOS ITENS: OS PRESSUPOSTOS PEDAGÓGICOS E PSICOMÉTRICOS	58
4	PLANO DE AÇÃO: A ELABORAÇÃO DE ITENS BASEADOS NOS NÍVEIS DE VAN HIELE.....	78
4.1	OPERACIONALIZAÇÃO DO PLANO DE AÇÃO	82
4.1.1	Eixo 1: Material Técnico-Formativo e ciclos de discussões.....	82
4.1.2	Eixo 2: Expansão metodológica	85
4.2	PROTOCOLO PARA PRODUÇÃO DE ITENS.....	86
5	CONSIDERAÇÕES FINAIS.....	91
	REFERÊNCIAS.....	93
	APÊNDICE A – SCRIPT R PARA LEVANTAMENTO DE PUBLICAÇÕES	97
	APÊNDICE B – CURVAS CARACTERÍSTICAS DOS ITENS	100

1 INTRODUÇÃO

Esta dissertação foi desenvolvida no contexto do Programa de Mestrado Profissional em Gestão e Avaliação da Educação Pública, focando numa intervenção no ambiente laboral em que o pesquisador atua. Estando a produção de itens e testes no cerne das funções do Analista de Instrumentos de Avaliação na Fundação CAEd, este foi o objeto de pesquisa ao qual se dedica esta dissertação.

Nos últimos anos, tem crescido a demanda por avaliações que traduzam em dados o percurso dos estudantes por meio de avaliações formativas¹. Devido a isso, a utilização das avaliações em larga escala tem ficado cada vez mais presente no contexto escolar, com foco em contribuir com o planejamento pedagógico das escolas. Além disso, as avaliações em larga escala são instrumentos para diagnosticar defasagens e monitorar a educação básica de forma geral. É a partir delas que se produzem medidas que irão validar, justificar e fomentar políticas públicas, mas, além disso, elas consistem em um dos principais instrumentos para que se possa garantir a qualidade da educação, descrita no item VII do artigo 206 da Constituição Federal de 1988 (Brasil, 1988).

Assim, de forma a contribuir no processo de produção de itens, a partir de um viés pedagógico, este trabalho irá focar nos itens de Geometria aplicados em avaliações de Matemática produzidas pela Fundação Centro de Políticas Públicas e Avaliação da Educação (Fundação CAEd) nos últimos quatro anos, parametrizados pela Teoria da Resposta ao Item (TRI) e que possuem *status* de itens abertos². A partir deles, objetivou-se analisar como os níveis de pensamento geométrico, propostos

¹ Existem diversas ações nesse sentido. Nacionalmente, o Ministério da Educação oferece, por meio do CAEd, a Plataforma do Compromisso Nacional Criança Alfabetizada (<https://criancaalfabetizada.caeddigital.net/#!/pagina-inicial>) para acompanhar a aprendizagem dos anos iniciais do ensino fundamental e a Plataforma de Avaliação e Acompanhamento das Aprendizagens (<https://avaliacaoaprendizagensanos finais.mec.gov.br/#!/pagina-inicial>) com foco na avaliação dos anos finais do ensino fundamental. No âmbito estadual, tem-se, por exemplo, iniciativas como as avaliações intermediárias em Minas Gerais e a Prova Paraná no estado do Paraná.

² Necessariamente, os itens selecionados para esta pesquisa compuseram avaliações somativas (que possuem itens sigilosos) dentro do período especificado, mas que, por diversos motivos, passaram a ser públicos. Esses itens são parametrizados em três parâmetros lineares na TRI, na mesma escala Saeb e podem ter composto diversas avaliações somativas em vários estados do Brasil.

pela Teoria de Van Hiele, são mobilizados nos itens de Geometria presentes nas avaliações em larga escala, relacionando essa estruturação pedagógica ao comportamento estatístico dos itens, a fim de propor um Plano de Ação para a elaboração de instrumentos mais adequados ao desenvolvimento cognitivo dos estudantes da Educação Básica.

A Teoria de Van Hiele do pensamento geométrico consiste em cinco níveis de compreensão, que são os níveis de visualização ou reconhecimento; análise; dedução informal ou ordenação; dedução formal e rigor; e em cinco fases de aprendizagem listadas como de questionamento ou informação; orientação direta; explicitação; orientação livre e integração (Kaleff *et al.*, 1994; Liu *et al.*, 2024; Burger e Shaughnessy, 1986; Crowley, 1987; Van Hiele, 1986).

Especificamente, objetivou-se caracterizar o processo de elaboração de itens, seus critérios e considerações pedagógicas que podem ser empregadas; investigar os itens do eixo do conhecimento de Geometria quanto ao desenvolvimento das tarefas propostas nos itens e de suas características intrínsecas; e propor critérios para o desenvolvimento de itens que sejam capazes de gerar informações em diferentes pontos da escala. Ao atingir esses objetivos, é esperado que se tenha respondido à seguinte pergunta norteadora: Como os níveis de pensamento geométrico da Teoria de Van Hiele se manifestam nos itens de Geometria das avaliações em larga escala, qual a relação dessa manifestação com o comportamento empírico desses itens, e de que modo essa análise integrada pode subsidiar a proposição de um Plano de Ação Educacional para a elaboração de itens mais alinhados ao desenvolvimento cognitivo dos estudantes?

A escolha por abordar o eixo do conhecimento de Geometria surge da possibilidade de utilizar a teoria cunhada por Dina Van Hiele-Geldof e Pierre Van Hiele (Hiele, 1986) como um dos suportes pedagógicos para o processo de elaboração de itens para testes de avaliações em larga escala.

A pesquisa adotou uma abordagem mista (qualitativa e quantitativa) com delineamento sequencial exploratório (Gil, 2017). Nessa modelagem metodológica, a investigação estruturou-se em duas etapas interdependentes. A primeira fase, de natureza qualitativa e documental, dedicou-se à análise da matriz e à análise pedagógica dos itens do eixo de Geometria, avaliando suas tarefas, redação e categorizando-os segundo os níveis da Teoria de Van Hiele. Os resultados dessa etapa qualitativa forneceram as bases categóricas para a segunda fase, de natureza

quantitativa. Nesta etapa, buscou-se estabelecer relações estatísticas entre a classificação teórica obtida (os níveis de Van Hiele) e os parâmetros psicométricos empíricos dos itens (advindos da Teoria Clássica dos Testes e da Teoria da Resposta ao Item), permitindo mensurar como a estrutura pedagógica do item impactou seu comportamento estatístico.

Esta dissertação foi estruturada em cinco capítulos. Este primeiro capítulo, a introdução, foi devotado a contextualizar o escopo do tema pesquisado, com sua justificativa, pergunta norteadora e os seus objetivos, além de sua estrutura metodológica.

O capítulo 2 se encarregou de apresentar o contexto do caso de gestão enfrentado no ambiente profissional do pesquisador, de modo a delimitar e justificar a necessidade de intervir nesse processo. Nele foram apresentados o CAEd e a Fundação CAEd, bem como os processos que os trouxeram até a consolidação. Além disso, o ambiente laboral do pesquisador foi detalhado, bem como as particularidades dos processos, concluindo com a delimitação do caso pesquisado.

No capítulo 3, o referencial teórico adotado nesta pesquisa foi estruturado. Nele detalhou-se o processo de produção de medidas e dos testes, bem como sua conceituação. Além disso, a Teoria de Van Hiele foi apresentada, assim como alguns aspectos pedagógicos que estruturaram a análise dos dados. Em seguida, a metodologia de obtenção de dados e de levantamento bibliográfico foi explicitada. Por fim, a análise em torno dos dados obtidos foi conduzida.

O capítulo 4 é reservado ao Plano de Ação Educacional, que foi produzido para contribuir com a Equipe de Instrumentos de Avaliação – Matemática, para orientar o processo de elaboração de itens.

Por fim, o capítulo 5 foi destinado às considerações finais deste trabalho.

2 AS AVALIAÇÕES EM LARGA ESCALA E OS DESAFIOS NA PRODUÇÃO DE ITENS

As avaliações em larga escala se tornaram importantes ferramentas para a construção de políticas públicas, atingindo todos os níveis da educação brasileira. Os sistemas de avaliação nacionais e subnacionais iniciaram-se a partir da década de 1990 no Brasil, sendo um dos principais marcos desse período a criação do Sistema Nacional de Avaliação da Educação Básica (Saeb) em 1993 (Gatti, 2014). Ao longo do tempo, tanto o Saeb quanto os sistemas de avaliação estaduais ganharam a função de monitoramento de políticas públicas educacionais, haja vista a criação de diversos indicadores estaduais e municipais, por influência da criação do Índice de Desenvolvimento da Educação Básica (Ideb), para monitoramento da evolução do sistema educativo (Brooke e Cunha, 2011, p. 26).

O objetivo deste capítulo é apresentar o caso de gestão enfrentado no ambiente de trabalho do pesquisador, levantando fatores que justificam a necessidade de intervenção. Para tanto, este capítulo foi organizado em quatro seções. A primeira seção traz o histórico do CAEd, desde a sua criação até sua consolidação nas avaliações em larga escala no Brasil, além de apresentar o vínculo do pesquisador com essa instituição. A segunda seção se ocupa de apresentar o processo de construção dos itens que compõem as avaliações em larga escala, como ele ocorre na Fundação CAEd atualmente e quais as metodologias são utilizadas na construção dos cadernos de teste. A terceira seção discute as finalidades dos diferentes tipos de avaliação e os apresenta no contexto das avaliações em larga escala, destacando algumas características e particularidades. A quarta seção apresenta a delimitação do caso pesquisado, evidenciando fatores que tensionam a eficiência do fluxo produtivo.

2.1 O CENTRO DE POLÍTICAS PÚBLICAS E AVALIAÇÃO DA EDUCAÇÃO (CAEd)

O CAEd se consolidou no contexto da avaliação da educação pública em todo o país por meio de parcerias em 25 estados brasileiros e o Distrito Federal, além de

parcerias internacionais³. Segundo Tocantins (2018), o CAEd foi criado como um projeto de pesquisa, fruto de um edital lançado pelo MEC em 1997, concebido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), que lançou o Programa de Apoio à Avaliação Educacional (PROAV). O objetivo desse projeto era subsidiar pesquisas em avaliação. Segundo a CAPES, o PROAV tinha como objetivo

criar e consolidar, no País, centros especializados em avaliação educacional nos vários níveis e na formação de recursos humanos para o desenvolvimento de estudos e pesquisas, destinados à solução de problemas e questões relativos à avaliação das ações voltadas para a educação (CAPES, 1997, p. 63).

Ainda segundo Tocantins (2018), a Universidade Federal de Juiz de Fora (UFJF), por meio do Laboratório de Avaliação e Medidas Educacionais (LAME), submeteu um projeto ao PROAV. O LAME realizou o Programa Piloto de Avaliação da Rede Pública de Ensino Fundamental da Microrregião de Juiz de Fora, com avaliação censitária dos estudantes de 33 municípios da microrregião de Juiz de Fora, que culminou na criação do Sistema Mineiro de Avaliação da Educação Pública (SIMAVE), no ano de 2000 (Tocantins, 2018).

Além de programas de avaliação, o LAME produziu o Sistema de Administração e Controle Escolar (SisLAME) em 1999, que tinha foco na gestão das redes de ensino e unidades escolares, informatizando as atividades administrativas. No mesmo ano, o LAME criou o curso de Especialização em Gestão e Avaliação da Educação Pública, voltado para profissionais em exercício na Educação Básica de Minas Gerais (Tocantins, 2018).

Um momento histórico para a instituição foi a mudança organizacional ocorrida em 2001, em que o LAME se tornou o Núcleo de Políticas Públicas e Avaliação de Educação, e que na prática passou a se chamar de Centro de Políticas Públicas e Avaliação da Educação (CAEd) (Tocantins, 2018).

O CAEd se manteve como centro de pesquisa e desenvolvimento tecnológico gerido pela Fundação de Apoio e Desenvolvimento ao Ensino, Pesquisa e Extensão (FADEPE) até 2018, quando passou a ser gerenciado pela Fundação CAEd, que

³ Em maio de 2026, o CAEd/UFJF possuía parcerias firmadas em todos os estados brasileiros, com exceção de Roraima. Veja mais em <https://institucional.caeddigital.net/#!/parcerias> e <https://institucional.caeddigital.net/#!/sobre-o-caed>

obteve seu credenciamento como fundação de apoio à UFJF no mesmo ano⁴, momento no qual passou a dar suporte às atividades desenvolvidas pelo CAEd. Em 2023, foi aprovada a criação do Centro de Políticas Públicas e Avaliação da Educação (CAEd) como um órgão complementar⁵ da Universidade Federal de Juiz de Fora (UFJF), que, de acordo com o artigo 29 do estatuto dessa universidade:

Os Órgãos Suplementares são órgãos integrantes da UFJF, vinculados à Reitoria, criados com a finalidade de dar apoio as atividades de Ensino, Pesquisa, Extensão, Cultura e Desenvolvimento Institucional, Científico e Tecnológico possuindo estrutura e regimento interno próprios. (UFJF, 1998)

Na prática, o CAEd, que antes era vinculado à Faculdade de Educação, passou a ser um órgão com regimento próprio, visando ao desenvolvimento de pesquisa científica, tecnologias educacionais e atividades relacionadas à área da avaliação e gestão da educação, com atuação no desenvolvimento de programas de avaliação em larga escala, tecnologias de avaliação e gestão da educação pública e prestação de serviços vinculados à sua atuação (UFJF, 2023, p.1). Já antevendo a necessidade de formalizar a parceria com a Fundação CAEd para continuidade da execução de suas atividades, no parágrafo único do artigo 2 da Resolução CONSU 83/2023, é indicado que

Parágrafo único. Para a fiel consecução dos objetivos do CAEd, a UFJF firmará convênios, contratos e outros ajustes com fundação de apoio, com a finalidade de obter o apoio indispensável à execução de suas pesquisas (UFJF, 2023, p. 2)

A Fundação CAEd é a fundação de apoio ao órgão complementar CAEd/UFJF, responsável pela operacionalização dos projetos firmados. Em seu site, a fundação se apresenta como instituição responsável por prover apoio especializado para execução de projetos e serviços relacionados à avaliação em larga escala, ao desenvolvimento de tecnologias de avaliação e gestão da educação pública.

Fica a cargo dessa Fundação a produção dos testes que são aplicados pelos diversos programas de avaliação em larga escala parceiros do CAEd. O setor diretamente responsável por produzir os itens e testes é a Equipe de Instrumentos de Avaliação, composta pelas seguintes subequipes: Língua Portuguesa, Matemática, Ciências da Natureza, Ciências Humanas, Revisão Ortográfica e Editoração.

⁴ De acordo com a Resolução CONSU/UFJF N° 27/2018 https://www2.ufjf.br/consu/files/2018/02/Resolucao_27.2018-CONSU_CAED-Credenciamento-Fund.-de-Apoio.pdf

⁵ De acordo com a Resolução CONSU/UFJF N° 83/2023. Disponível em: https://www2.ufjf.br/consu/wp-content/uploads/sites/33/2023/12/Resolucao-83.2023-Consu-UFJF_SEI_Assinada.pdf.pdf

O pesquisador é vinculado a esta equipe, atuando como Analista de Instrumentos de Avaliação, na subequipe de Matemática. Os membros dessa subequipe se ocupam diariamente da elaboração de itens, revisão de itens, validação de revisão ortográfica e validação da editoração do item. Além disso, há também atividades de seleção de itens, montagem e revisão dos cadernos de teste, dentre outras atividades como apoio em formações internas de novos colaboradores, formações em torno da elaboração de itens e validações das plataformas de avaliação digitais.

O ambiente de trabalho das subequipes requer que seus componentes se engajem em discussões sobre a construção pedagógica dos itens que estão sendo avaliados, bem como dos itens que estão sendo selecionados para algum teste.

Pesquisar a elaboração de itens, visando melhorias nesse processo, é crucial para aumentar a qualidade não apenas dos itens, mas também dos testes produzidos e, por consequência, das medidas estatísticas geradas posteriormente. Geralmente, para se produzir um item com qualidade pedagógica, utiliza-se um referencial a partir da progressão do currículo, tomando como ponto de partida a evolução dos pré-requisitos no currículo da etapa e das etapas de escolarização anteriores, e a experiência do profissional responsável pela elaboração, bem como a observação do comportamento estatístico de itens na mesma habilidade e habilidades correlatas.

A próxima seção se incumbirá de caracterizar o principal tipo de item utilizado atualmente nas avaliações em larga escala produzidas pela Fundação CAEd e de mostrar como a sua elaboração ocorre na Equipe de Instrumentos de Avaliação.

2.2 O CONTEXTO DA PRODUÇÃO DE ITENS E DOS TESTES PADRONIZADOS

Para composição dos testes produzidos na Fundação CAEd, são produzidos itens de resposta selecionada, itens de resposta construída, além de itens para testes de fluência e escrita, bem como de itens digitais e aprimorados por tecnologia.

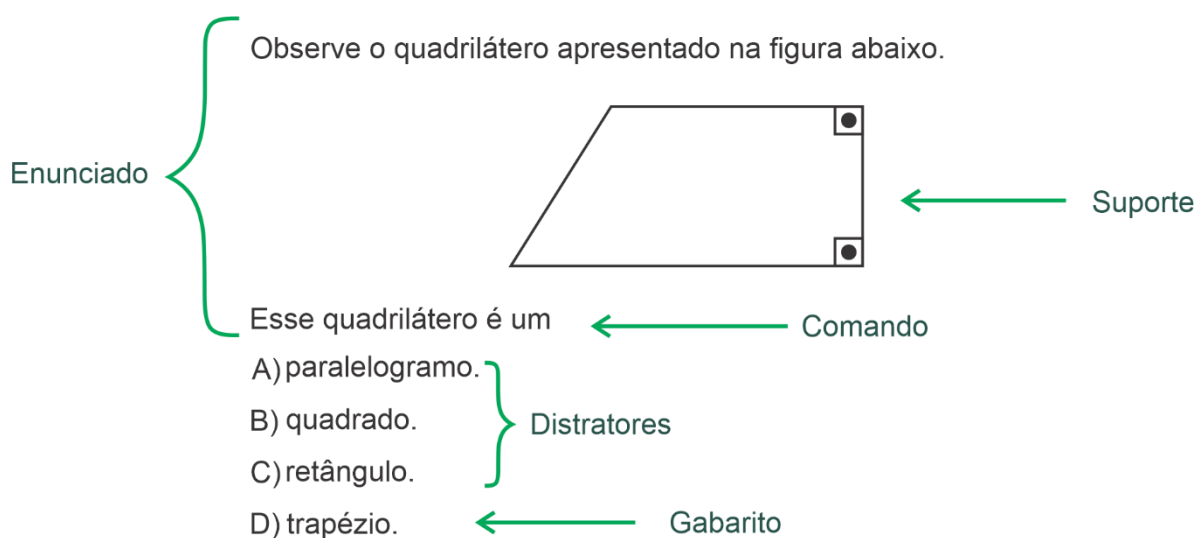
Nas avaliações somativas e formativas, são utilizados principalmente itens com o formato de resposta selecionada. Nesse tipo de item, Hogan (2006) sustenta que eles

apresentam-se ao examinando pelo menos duas, porém não mais do que um pequeno número de opções, entre as quais ele deve escolher uma resposta. Os itens de resposta selecionada são também

denominados itens de resposta múltipla, de múltipla escolha ou de escolha forçada (Hogan, 2006, p. 164).

Nesta pesquisa será utilizado esse formato de item, dado que a estrutura elucidada na seção 3.1.3 e exemplificada pela Figura 1 é a estrutura do item que o CAEd utiliza atualmente⁶.

Figura 1 - Exemplo de item



Fonte: Elaborado pelo autor (2025)

Itens de resposta selecionada possuem características próprias, com tipos de dificuldades diferentes⁷, mas todos eles são compostos por enunciados e alternativas.

Palácios e Oliveira (2022) listam as seguintes características:

- Enunciado: estímulo que faz com que o estudante mobilize recursos cognitivos, a fim de solucionar o problema apresentado. Ele pode se apresentar como uma situação-problema que funcione como um questionamento ou uma contextualização.
 - Suporte: texto, imagem, ou outro recurso
 - Comando: tem a função de apresentar o problema que irá direcionar o estudante a uma resposta, podendo ser elaborado em formato de pergunta ou de sentença a ser completada pelas alternativas. Deve ser claro e objetivo.

⁶ Como descrito em CAEd (2008, p. 19) e na Revista da Rede – SIMAVE – 2024, página 114. Disponível em: https://prototipos.caeddigital.net/arquivos/mg/colecoes/2024/Simave_2024_RR_web.pdf. Acesso em 11 dez. 2025.

⁷ Como descrito no capítulo 4 de Haladyna (2004, p. 67-97).

- Alternativas:
 - Gabarito: a alternativa correta do item, ou seja, é a resposta para o que está problematizado no comando.
 - Distratores: alternativas incorretas. Devem ser plausíveis para quem ainda não desenvolveu a habilidade avaliada, ao mesmo tempo que seu conteúdo precisa, sem sombra de dúvida, ser incorreto.

Existem outras nomenclaturas para cada parte do item, como a adotada pelo Inep, que descreve a estrutura como: texto-base, enunciado e alternativas (Inep, 2010, p. 9).

Embora existam outros modelos de itens de resposta selecionada, o exemplo apresentado na Figura 1 é o modelo mais utilizado em avaliações em larga escala, devido à sua praticidade, à baixa subjetividade ao corrigir, à correção computadorizada e à capacidade de avaliar grande variedade de habilidades em pouco tempo, como preconizado por Palácios e Oliveira (2022, p. 38) Figura 3.

Figura 3.

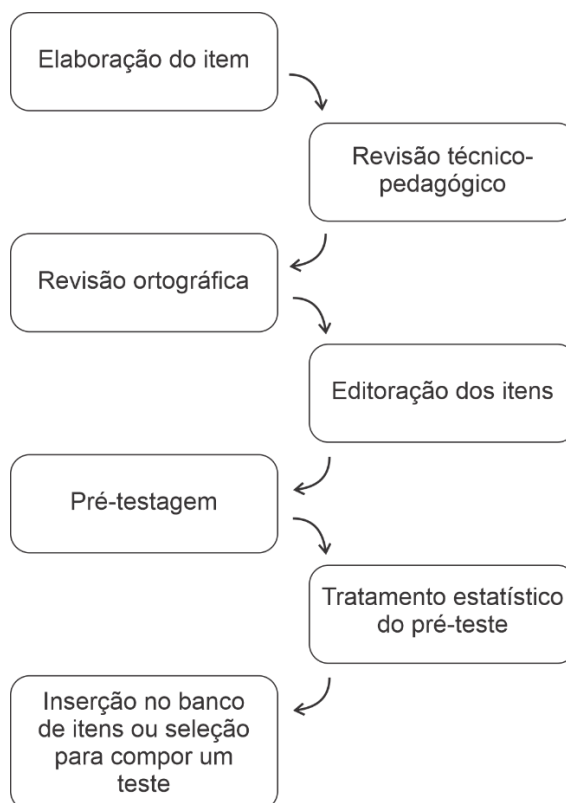
Ainda nesse sentido, Palácios e Oliveira (2022) argumentam que a praticidade desse formato de item conduzirá a uma melhora da aprendizagem, somente se os itens forem elaborados seguindo critérios de qualidade bem estabelecidos, uma vez que itens com falhas produzirão um diagnóstico em desacordo com a realidade dos estudantes e ações baseadas nesse diagnóstico poderão ser inefetivas.

Embora existam características importantes que devem compor o perfil do elaborador de bons itens, o CAEd (2008) sugere que

A excelência na elaboração de itens, contudo, demanda mais do que isso. É preciso imaginação e criatividade na invenção de situações que exijam o conhecimento e as habilidades desejadas. Demanda, principalmente, habilidade e julgamento que só vêm com experiência (CAEd, 2008, p. 18).

Antes de abordar o processo de elaboração de um item, é necessário compreender que o item passa por um processo, pelo qual ele percorre algumas etapas desde a sua concepção até a sua utilização em um teste. No CAEd, o item segue o processo apresentado no esquema da Figura 2.

Figura 2 - Processo de elaboração de itens



Fonte: Elaborado pelo autor (2025), com base em CAEd (2008) considerando os processos internos atuais.

As subequipes da equipe de instrumentos de avaliação da Fundação CAEd possuem os guias internos que orientam os elaboradores quanto às técnicas, regras de padronização, dentre outras orientações. Disponíveis publicamente, têm-se os guias de elaboração e revisão de itens CAEd (2008) e Inep (2010), que serão utilizados, devido à proximidade com o documento interno.

Para iniciar o processo de elaboração de um item, o elaborador precisa selecionar um descritor na matriz de referência e,

a partir de sua experiência e do conhecimento que possui com relação ao desenvolvimento cognitivo dos estudantes que se encontram na etapa de escolarização avaliada, reconhecer o nível de dificuldade desejado na construção do item (CAEd, 2008, p. 20).

Em seguida, o elaborador precisa fazer a construção de cada um dos elementos do item, decidindo a necessidade de utilizar um contexto para o item e a necessidade de utilizar suporte. Definidos esses elementos, o elaborador pode decidir o formato do comando para resposta, que

deve ser claro e objetivo, indicando de maneira precisa a tarefa a ser realizada, afastando qualquer fator complicador para a compreensão do item pelo estudante. Deve contemplar um único descritor da Matriz

de Referência. O comando pode ser interrogativo ou apresentar uma frase incompleta (CAEd, 2008, p. 23).

O exemplo da Figura 1 faz a utilização do formato de comando com frase incompleta.

O passo seguinte é construir as alternativas, atentando-se para o processo de construção da habilidade, para a independência entre cada uma delas e a plausibilidade⁸, além de cuidar para que tenham paralelismo sintático e semântico. Segundo CAEd (2008, p. 24), o gabarito

deve validar a capacidade do estudante em relação à determinada habilidade cognitiva. As demais alternativas, os distratores, produzem informações importantes para a avaliação, na medida em que apontam possíveis caminhos de raciocínio dos estudantes. (CAEd, 2008, p. 24).

Outro ponto a se observar é que o uso de distratores que podem ser facilmente descartados faz a resposta correta surgir do processo de eliminação, e não da execução induzida pelo item e esperada no processo de resolução, devendo ser evitado que se tenha esse tipo de distratores em qualquer item (CAEd, 2008, p. 24).

Ambos os guias Inep (2010, p. 12) e CAEd (2008, p. 26) orientam quanto a pontos a serem considerados na elaboração de um item, como o uso de linguagem clara, objetiva e adequada à etapa a qual o item é destinado, ineditismo do item, não induzir o estudante ao erro e evitar temas que suscitem polêmicas. Além disso, não se deve utilizar expressões como “Assinale a alternativa correta [...]”, “Qual das alternativas [...]”, “A alternativa que indica [...]”, que costumam se apresentar em questões de testes de seleção, que não é o foco de um teste que afere proficiência escolar.

Findado o processo de elaboração, o item começa sua jornada até ser selecionado para compor os testes de alguma avaliação. A primeira etapa é a revisão, na qual especialistas na área de conhecimento verificam se todos os critérios técnicos e pedagógicos foram atendidos. O CAEd (2008) afirma que a lista de critérios

se assemelha a uma lista de controle de qualidade cujas etapas devem ser rigorosamente cumpridas. Dessa forma, após a revisão criteriosa, o item estará pronto para a pré-testagem e os itens que apresentarem um bom comportamento passarão a compor o teste (CAEd, 2008, p. 43).

Findada essa etapa, o item precisa passar por revisão ortográfica, para que especialistas em Língua Portuguesa corrijam possíveis desvios da norma culta,

⁸ Isto é, semelhanças ou similaridade em relação à alternativa correta (CAEd, 2008, p. 28).

devendo retornar aos especialistas da área em que o item foi elaborado para que possam seguir para a próxima etapa.

A partir desse ponto, a subequipe de editoração procede com a formatação e diagramação dos itens e, posteriormente, com a formatação e diagramação dos pré-testes, para que os itens possam ser aplicados, produzindo as medidas estatísticas. Após a pré-testagem, os itens ficam disponíveis para passar por uma análise psicométrica e pedagógica para que sejam selecionados para compor o banco de itens ou para compor um teste de avaliações em larga escala.

É notável que o processo de elaboração de itens requer a participação de muitos profissionais, que precisam mobilizar sua experiência docente, buscar por documentos orientadores e pela literatura especializada, além da própria experiência construída no exercício da função de Analista de Instrumentos de Avaliação.

Em seu estudo desenvolvido na subequipe de Matemática, na então Coordenação de Instrumentos de Avaliação, Souza (2020) identificou que os colaboradores entrevistados durante sua pesquisa se formaram em torno de temas de avaliação em larga escala durante o período em que estão trabalhando na função de analista de instrumentos de avaliação, e que esse processo de formação não era padronizado.

Notadamente, esse processo não sofreu grandes alterações, uma vez que a formação do pesquisador em temas relativos à produção de itens e testes também não ocorreu de forma padronizada. Souza (2020) afirma que os colaboradores entrevistados “alegaram não trabalhar anteriormente com avaliações, sendo o CAEd, ao mesmo tempo, um local de labor e formação” (Souza, 2020, p. 122).

A próxima seção terá como foco central discutir os diferentes tipos de avaliação, de acordo com o momento no contexto escolar, explicitando a clareza nos objetivos da avaliação, sua organização e características.

2.3 AS AVALIAÇÕES, MOMENTOS NO CONTEXTO ESCOLAR E CONSTRUÇÃO DOS TESTES

As avaliações que ocorrem em ambiente escolar podem se classificar em diagnósticas, formativas e somativas, a depender do momento em que ocorrem. As avaliações diagnósticas são aplicadas no início de um percurso de ensino-

aprendizagem, com foco em identificar possíveis defasagens e potencialidades que os estudantes possuem.

As avaliações formativas podem acontecer durante todo o percurso de ensino-aprendizagem, com o principal foco em identificar a evolução dos estudantes durante todo o período letivo, evidenciando se foram consolidados os conhecimentos ou há necessidade de fazer ajustes aos processos adotados até dado ponto nesse percurso.

Por fim, uma avaliação somativa tem o objetivo de verificar o aprendizado dos estudantes ao final de um percurso de ensino-aprendizagem. Ao apresentar a concepção de Bloom *et al.* (1971), Luckesi (2022) segue nesse sentido, afirmando que as adjetivações dadas aos tipos de avaliação “estão comprometidas com o uso que se pode fazer dos resultados da investigação avaliativa, desde que o ato avaliativo simplesmente investiga e revela a qualidade da realidade”.

O autor ainda pontua que o caráter somativo das avaliações pode ser intermediário ou final. A avaliação somativa final objetiva verificar a extensão dos resultados obtidos, enquanto a somativa intermediária visa a resultados parciais que possibilitem reorientação das atividades.

Vale notar que as avaliações diagnósticas e formativas podem ser consideradas como duas classificações de um mesmo instrumento de avaliação, já que uma avaliação diagnóstica pode ter caráter formativo, e que toda avaliação formativa se propõe a fazer um diagnóstico de um dado ponto de um percurso de ensino-aprendizagem.

Quando as avaliações estão inseridas em contextos de larga escala, elas apresentam características similares. Dessa forma, quando uma rede em escala nacional ou subnacional se propõe a aplicar avaliações a seus estudantes, têm-se as chamadas avaliações diagnósticas da rede ou avaliações formativas da rede. As avaliações aplicadas anualmente pelas redes estaduais ou a prova do Saeb são consideradas avaliações somativas (Brasil, 2025, p. 21). A partir dessas avaliações, as redes passam a ter um conjunto de dados robusto, que podem se complementar para dar suporte ao aprendizado.

No ponto de partida do desenvolvimento de um teste, está a necessidade de se possuir com clareza o objetivo desse teste, que inclui também o delineamento das características e público-alvo (Hogan, 2006, p. 160). Nesse sentido, o CAEd (2008) afirma que “os testes de avaliação em larga escala têm como objetivo aferir a

proficiência dos estudantes em determinada área de conhecimento, em períodos específicos de escolarização” (CAEd, 2008, p. 14).

Um passo importante que se faz necessário é definir as habilidades que serão avaliadas, para orientarem a elaboração dos itens. Essas habilidades são definidas pela Matriz de Referência, que segundo o CAEd (2008):

As Matrizes de Referência são compostas por um conjunto de descritores, os quais contemplam dois pontos básicos do que se pretende avaliar: o conteúdo programático a ser avaliado em cada período de escolarização e o nível de operação mental necessário para a habilidade avaliada. Tais descritores são selecionados para compor a Matriz, considerando-se aquilo que pode ser avaliado por meio de itens de múltipla escolha. (CAEd, 2008, p. 14).

É necessário que se note que o conjunto de descritores em uma matriz de referência não serve como orientação quanto ao que é desenvolvido em sala de aula, algo que é orientado pelos currículos nacionais e subnacionais, embora, segundo Bauer (2020), o estabelecimento de metas por rede de ensino e escola, políticas por incentivos por melhorias nos resultados de rendimentos dos estudantes, evidencia

a ampliação das influências das avaliações no currículo e nas práticas escolares, com cada vez mais indicações sobre o que e como ensinar e avaliar, ampliando o papel da avaliação na determinação dos conteúdos e práticas de ensino e de avaliação no interior das salas de aula (Bauer, 2020, p. 10).

Ainda nesse sentido, “a matriz de referência não avalia cada tópico do currículo, mas deve refletir o que se espera como essencial no processo de ensino-aprendizagem para determinado ano ou etapa que esteja sendo avaliado” (Palácios e Oliveira, 2022, p. 17).

Durante a elaboração da matriz de referência, Palácios e Oliveira (2022, p.17) explicam que os especialistas consideram os aspectos “a) o arcabouço de conhecimentos da área a ser avaliada; e (b) o tipo de teste que será elaborado: objetivo e/ou discursivo, ou prático” (Palácios e Oliveira, 2022, p. 17) e seguem os seguintes passos:

- 1) definição do modelo teórico que será adotado, sendo o mais comum um modelo que tem como foco competências, as quais, para serem alcançadas, dependem de um conjunto de habilidades e objetos de conhecimento;
- 2) estudo minucioso do currículo do ano ou etapa a ser avaliado, por disciplina;
- 3) eleição de competências, com suas respectivas habilidades e objetos de conhecimento, bem como as relações possíveis entre esses componentes;
- 4) estudos de validação da matriz de referência proposta por meio de revisão de conteúdo por especialistas da área, bem como análise semântica com o objetivo de verificar a compreensão da matriz por

uma amostra do seu público-alvo e estudo piloto por meio de oficinas de elaboração de itens com base na matriz (Palácios e Oliveira, 2022, p. 17-18).

Após a definição da matriz de referência para o teste, com sua organização e descritores, já se tem o documento que orienta a elaboração dos itens.

Do ponto de vista técnico, um teste de avaliação em larga escala é construído com itens unidimensionais⁹ que compõem os blocos para a montagem dos cadernos de teste. A depender do tipo e da finalidade da avaliação, a sua estrutura e desenho podem variar, de forma a permitir avaliar as habilidades que estão dispostas na matriz de referência da avaliação.

Nas avaliações em larga escala¹⁰, os blocos seguem a metodologia dos Blocos Incompletos Balanceados (BIB), que consiste em distribuir os blocos de itens¹¹ em diversos cadernos, que não possuem a totalidade de blocos, fazendo com que os estudantes recebam individualmente um subconjunto do total de blocos. Essa distribuição ocorre como apresentado por Bekman (2001, p. 121), em que cada caderno possui a mesma quantidade de blocos, estes são utilizados a mesma quantidade de vezes dentro do conjunto total de blocos e cada par de blocos é utilizado o mesmo número de vezes no conjunto total dos cadernos.

Uma vantagem de utilizar os BIB é que, com esses cadernos incompletos, os estudantes fazem testes com uma quantidade limitada de itens, o que torna o teste menos extenso, mas, ainda assim, capaz de produzir melhores estimativas dos parâmetros dos itens e da habilidade dos estudantes.

Assim, “a combinação da Teoria de Resposta ao Item com os Blocos Incompletos Balanceados nos permite obter informações precisas sobre a performance do aluno e da sala de aula simultaneamente” (Bekman, 2001, p. 134).

Na próxima seção, será apresentada a delimitação do caso de pesquisa desenvolvido, expondo a motivação, elementos pesquisados e quais os dados serão analisados no próximo capítulo.

⁹ Que avaliam uma única habilidade (Palácios e Oliveira, 2022, p. 205).

¹⁰ Confira, por exemplo, a Revista da Rede – SIMAVE – 2024, páginas 116-118. Disponível em: https://prototipos.caeddigital.net/arquivos/mg/colecoes/2024/Simave_2024_RR_web.pdf. Acesso em 11 dez. 2025.

¹¹ Esses blocos são balanceados de forma que apresentem, em média, mesmo grau de dificuldade.

2.4 DELIMITAÇÃO DO CASO

Atuando na função de Analista de Instrumentos de Avaliação, o pesquisador se ocupou de diversos processos em torno dos itens, desde a proposição e o estudo das matrizes, passando pela elaboração dos itens, até a finalização dos cadernos de testes.

Nesses processos, os analistas se preocupam com o alinhamento das tarefas exigidas pelos itens ao conhecimento dos estudantes que se pretende avaliar. Assim, a experiência em sala de aula, a sensibilidade para compreender e escrever em linguagem adequada aos estudantes a quem os itens e testes serão destinados e boas referências bibliográficas são importantes competências e ferramentas que o elaborador precisa possuir. O CAEd (2008) afirma que o elaborador de bons itens precisa “entender os processos de desenvolvimento e aprendizagem que caracterizam os estudantes para os quais o item será construído” (CAEd, 2008, p. 18).

Souza (2020) afirma que, no contexto da sua pesquisa, cerca de metade dos itens sobreviviam a todas as etapas da produção de itens, aos procedimentos de revisão e validação. É notável que as formações internas, as melhorias nos processos e o crescimento no quadro de elaboradores próprios dentro da Fundação CAEd fazem com que a taxa de itens que chegam ao fim do processo seja maior.

O exercício de calibrar, pedagogicamente, as tarefas dos itens ainda no processo de elaboração, buscando uma linguagem adequada, é o que torna o processo de construção do item até sua finalização mais fluido, uma vez que a complexidade dos itens é dependente de vários fatores. Conceição (2019) observa que itens considerados como difíceis por meio de alguma metodologia de cálculo de proficiência podem ter baixa complexidade textual e que o contrário também pode ocorrer.

Uma evidência dessa situação é a existência de itens que, embora sejam vistos como itens com tarefas de baixa complexidade, performam como itens de média ou alta complexidade, ou, ainda, ocorrência em sentido contrário. Daí, surge a necessidade de observar esse processo e instrumentalizar a Equipe de Instrumentos de Avaliação para enfrentar essa situação e produzir itens que superem esse desafio.

Nos últimos anos, as diversas áreas têm incorporado avanços tecnológicos e digitais. Por exemplo, os testes aplicados por meios digitais têm ganhado atenção, assim como a utilização de inteligência artificial em vários processos na educação. No

âmbito das avaliações em larga escala, existem movimentos importantes ocorrendo, como a utilização da metodologia de *testlets* no Exame Nacional do Ensino Médio (Enem)¹², pesquisas envolvendo Teste Adaptativo Computadorizado (CAT) no Enem (Jaloto; Primi, 2024), pesquisas envolvendo Teste Adaptativo Informatizado (TAI) para estudantes dos 1º e 2º anos do ensino fundamental (Travitzki *et al.*, 2020) e pesquisas sobre itens aprimorados por tecnologias (Scortegagna, 2023). Além disso, desde 2019 as matrizes de avaliações, como o Saeb, estão sendo transicionadas para as matrizes alinhadas à Base Nacional Comum Curricular (BNCC)¹³.

Nesse sentido, este trabalho foi construído com o intuito de aprofundar-se nas diversas nuances da produção de itens e, para isso, foi escolhido o eixo do conhecimento de Geometria¹⁴. A partir disso, serão analisados os descritores da matriz Saeb de 2001, sendo cinco descritores para o 5º ano do ensino fundamental, onze descritores para o 9º ano do ensino fundamental e dez descritores para as 3ª/4ª séries do ensino médio, totalizando vinte e seis descritores¹⁵. Essa análise será feita observando a progressão que ocorre na BNCC, uma vez que é o documento que rege o ensino básico nacional atualmente.

Os dados obtidos para esta pesquisa são, principalmente, documentos oficiais e os dados estatísticos cedidos pela Fundação CAEd. Foram requisitados cento e quatro itens a essa Fundação, sendo quatro itens de cada um dos vinte e seis descritores. Os critérios para a seleção dos itens foram: itens caracterizados como itens públicos pela Equipe de Instrumentos de Avaliação, que estiveram presentes em avaliações nos últimos quatro anos.

Para complementar esse processo de análise, serão utilizadas as Teorias de Van Hiele e a Taxonomia de Bloom revisada¹⁶, seguindo-se um estudo dos itens,

¹² Como noticiado em: <https://www.gov.br/mec/pt-br/assuntos/noticias/2025/novembro/enem-utiliza-metodologia-de-testlets-para-avaliar-conhecimentos>. Acesso em 11 de dez. de 2025.

¹³ Como informado em: <https://www.gov.br/inep/pt-br/centrais-de-conteudo/noticias/saeb/saeb-utiliza-novas-matrizes-de-referencia-desde-2019>. Acesso em 11 de dez. de 2025.

¹⁴ Na matriz do Saeb de 2001, esse eixo é chamado de “Tema”. Para facilitar a referência, essa pesquisa utiliza “Eixo de Geometria” ou “Eixo de Conhecimento de Geometria”, a mesma nomenclatura utilizada na matriz do Saeb de 2022 e mais próximo do que se utiliza na BNCC (unidade temática).

¹⁵ Será detalhado na seção 3.3.

¹⁶ Essa taxonomia é largamente utilizada na construção de matrizes e currículos, uma vez que permite descrever os objetivos da habilidade/descritor de forma clara. A taxonomia de Bloom revisada

munidos de seus dados estatísticos, para que se possa evidenciar como está o alinhamento das tarefas que os estudantes resolvem nessas avaliações com as expectativas e conhecimentos indicados pelas matrizes e pela base curricular.

Vale ressaltar que a totalidade desses itens foi produzida em uma organização pensada para a matriz de 2001 do Saeb e que um dos desafios que se apresenta nesse contexto é de manter o rigor pedagógico e estatístico ao transicionar para a nova matriz do Saeb que foi proposta para o ensino fundamental em 2022. Como ainda não se tem avaliações com resultados publicados com itens construídos a partir da matriz de 2022, o foco deste trabalho estará em relacionar o referencial levantado com os itens e as medidas estatísticas dos testes.

consiste em duas dimensões, a dimensão dos processos cognitivos e a dimensão do conhecimento. No domínio cognitivo, as categorias são: conhecimento, compreensão, aplicação, análise, síntese e avaliação.

3 ASPECTOS PEDAGÓGICOS E ANÁLISE DOS ITENS DO EIXO DO CONHECIMENTO DE GEOMETRIA

Neste capítulo, o debate ocorre em torno dos itens de matemática do eixo do conhecimento de Geometria. Como suporte, mobilizam-se autores que abordam o processo de construção e validação de itens e testes, autores que tratam de aspectos do ensino de matemática, com o intuito de relacionar aspectos pedagógicos e psicométricos dos itens para identificar padrões e estruturas que contribuam para melhoria da qualidade dos itens. Aqui, serão aprofundados os conceitos de itens, testes, os pressupostos pedagógicos e o campo da Geometria nas avaliações em larga escala.

O objetivo central consiste em estudar o viés pedagógico da elaboração de itens de Matemática para avaliações em larga escala, especificamente no eixo de conhecimento Geometria, ocupando-se, principalmente, da busca por relacionar a matriz de referência adotada com a Teoria de Van Hiele como suporte pedagógico.

Nesse sentido, este capítulo possui três seções. A primeira seção é dedicada à discussão em torno das avaliações em larga escala, definindo termos importantes desse contexto, desde a sua concepção, metodologias de correção e cálculo das medidas estatísticas, a escala de proficiência, além de apresentar a Teoria de Van Hiele, de maneira a servir de base sólida para a análise psicométrica e pedagógica dos itens levantados para esta pesquisa. A segunda seção se ocupa de caracterizar a pesquisa apresentando seu desenho, seus objetivos e questão de pesquisa, além de sua caracterização e métodos de levantamento bibliográfico e de dados. A terceira seção é reservada à análise dos itens, mobilizando a matriz adotada, o referencial teórico, os dados psicométricos e a Teoria de Van Hiele, de forma integrada, a fim de fundamentar o Plano de Ação Educacional (PAE), que será desenvolvido no próximo capítulo.

3.1 AS AVALIAÇÕES EM LARGA ESCALA E A PRODUÇÃO DE MEDIDAS

Para que seja possível compreender e discutir sobre a unidade básica de qualquer avaliação do conhecimento, o item, é necessário discutir o que se entende por teste e testagem. Hogan (2006) apresenta no primeiro capítulo de seu livro, “Introdução à Prática de Testes Psicológicos”, um panorama sobre teste e testagem

psicológica, sob uma perspectiva macro, apresentando as diversas categorias em que os testes se classificam, as principais utilizações desses testes, pressupostos importantes e questões fundamentais no que concerne à testagem, além de um histórico sobre a testagem.

Nesse último ponto, Hogan (2006) aponta o ano de 1965 como o princípio do que, para o autor, é o cenário atual da testagem psicológica. Isto porque foi até esse ponto na história que houve a consolidação do que hoje se conhece como Teoria Clássica dos Testes (TCT) e foi em meados dos anos 1960 que surgiu a Teoria da Resposta ao Item (TRI) ou “teoria moderna do teste”, que é “definida como um novo conjunto de métodos para a investigação de um universo de assuntos referentes a fidedignidade, escalonamento e construção de testes” (Hogan, 2006, p. 22). Ainda, o autor afirma ter sido nessa mesma década que se observou o princípio do ativismo legislativo e judicial envolvendo testes, em grande medida, vindo do governo federal¹⁷.

Hogan (2006) apresenta a existência de diferentes categorias de testes, como os testes de Capacidade Mental, de Conhecimento, de Personalidade, de Interesses, Atitudes e Valores e Neuropsicológicos. Nesse contexto, as avaliações em larga escala se enquadram na categoria dos testes de conhecimento.

Ainda nesse sentido, o autor apresenta questionamentos acerca da fidedignidade, validade e normas do teste, sendo que o primeiro se refere à estabilidade dos seus escores; o segundo, ao que de fato o teste está mensurando; e o terceiro, saber como interpretar os escores de um teste. Sobre isso, ele afirma que “a interpretação dos escores dos testes geralmente depende da utilização de normas, que se baseiam nos escores de grandes grupos de indivíduos que já fizeram o teste” (Hogan, 2006, p. 12-13), concluindo que as tentativas de responder a esses questionamentos formam o núcleo da teoria da testagem e constituem os tópicos com os quais se deve preocupar ao tratar de qualquer teste.

Concernente à testagem, o autor apresenta quatro pressupostos. O primeiro faz referência ao fato de o ser humano possuir traços ou características reconhecíveis, a exemplo da habilidade verbal, da memória, da capacidade de raciocínio quantitativo, dentre outros. Supõe-se que essas características descrevam aspectos potencialmente importantes das pessoas. O segundo pressuposto é que é possível quantificar esses traços, ou seja, é possível criar distinções entre pessoas nas

¹⁷ Governo Federal dos Estados Unidos da América.

características que se estão estudando, e essas pessoas se disporão nesse *continuum*. O terceiro pressuposto questiona se esses traços têm algum grau de estabilidade. Eles podem não ser perfeitamente estáveis, mas não podem flutuar em demasia. O quarto é que o traço razoavelmente estável quantifica com os testes o comportamento que se observa na realidade. É notado que do ponto de vista teórico esse último é o menos importante, mas que do ponto de vista prático é de extrema importância (Hogan, 2006, p. 11-12).

Quanto ao teste, Hogan (2006, p. 28-30) apresenta algumas definições e extrai seis elementos que ele considera “teste” nas ciências comportamentais. A partir desses elementos, ele formula a seguinte definição: “teste é um processo ou um instrumento padronizado que fornece, de maneira quantificada, informações sobre uma amostra de comportamento ou de processos cognitivos” (Hogan 2006, p. 30). Com esse mesmo objetivo de apresentar uma definição para esse conceito, Haladyna (2004) considera o teste como “um instrumento de medida destinado a descrever numericamente o grau ou a quantidade de aprendizado sob condições uniformes e padronizadas” (Haladyna, 2004, p. 4). Entre essas duas definições, Hogan (2006) apresenta uma definição mais abrangente, enquanto Haladyna (2004) já o faz tendo em vista testes de avaliação educacional.

No contexto brasileiro, esse tipo de teste teve sua gênese de forma sistematizada no início da década de 1990, com a criação do Sistema de Avaliação da Educação Básica (Saeb). Em 1995, o Saeb adota o modelo estatístico da TRI (Brasil, 2003, p. 12). Antes, com a TCT, os resultados eram dependentes do conjunto de itens apresentados em um teste a uma população, e a comparação entre diferentes grupos só seria possível se o outro grupo passasse pelo mesmo teste ou um paralelo a este, dificultando efetuar comparações entre grupos com testes diferentes (Klein, 2003, p. 126-127).

Com a introdução da TRI, foi possível superar as limitações da TCT, que analisava apenas percentuais de acertos, passando à análise individual do item, por meio dos parâmetros que indicam propriedades do item. Isso, aliado aos “parâmetros dos itens e as proficiências dos indivíduos serem invariantes” (Klein, 2003, p. 127), permitiu monitorar os sistemas de ensino ao longo do tempo e inclusive entre anos de escolaridade.

Ainda na década de 1990, diversos estados iniciaram alguns projetos avaliativos. Os estados do Ceará e de Minas Gerais esboçavam seus sistemas de

avaliação em 1992, o estado do Paraná em 1995 e São Paulo em 1996. Na década de 2000 é que os estados brasileiros iniciaram um movimento de estruturação de seus sistemas de avaliação (Alavarse *et al.*, 2012).

Embora este trabalho não se ocupe de pormenores da produção de medidas, é importante que se compreenda como se produzem essas medidas a partir da TCT e da TRI. A seguir, são apresentadas algumas características dessas teorias e elucida-se como a compreensão desses elementos auxilia na identificação de melhorias no processo de produção de itens.

3.1.1 A produção de medidas dos itens

A TCT e a TRI são teorias estatísticas do campo da Psicometria, utilizadas como metodologia para medir um traço latente, que, no contexto de uma avaliação externa em larga escala, consiste na capacidade de um estudante resolver um conjunto de itens de determinada área do conhecimento.

Nesse sentido, Pasquali (2011) afirma de forma didática que a TCT se preocupa em explicar o resultado, um escore total T , que, em um teste com itens dicotômicos, é a soma da quantidade de itens respondidos de forma correta. Já a TRI se preocupa com cada um dos itens que compõem o teste e quer saber qual a probabilidade de um item ser acertado ou errado e quais os fatores que afetam essa probabilidade. Além disso, o autor observa que “a TCT se interessa em produzir testes de qualidade enquanto a TRI se interessa em produzir tarefas (itens) de qualidade” (Pasquali, 2011, p. 67).

Ao introduzir o modelo TCT, Pasquali (2011) argumenta que é necessário distinguir o escore bruto ou empírico (T), o escore verdadeiro (V) e o erro (E) e faz as seguintes suposições:

- T : escore bruto ou empírico do sujeito, que é a soma dos pontos obtidos no teste;
- V : escore verdadeiro; magnitude real daquilo que se quer medir (seria o próprio T se não houvesse erro de medida);
- E : erro cometido nessa medida.

O autor postula que o escore empírico é a soma do escore verdadeiro e o erro, isto é,

$$T = V + E$$

e disso, pode-se concluir as expressões $E = T - V$ e $V = T - E$.

Contudo, embora o escore verdadeiro não possa ser diretamente medido, ao aplicar o teste a muitos sujeitos, serão produzidas três distribuições de frequência, dos T, dos V e dos E. Mas ao se considerar um evento aleatório¹⁸ se tem que a média ou expectância matemática do erro é 0, isto é, $E(E) = 0$. Como consequência, tem-se que “o escore verdadeiro é a expectativa do escore empírico” (Pasquali, 2011, p. 67). Isso pode ser observado, já que

$$E(V) = E(T) - E(E)$$

$$E(V) = E(T)$$

Assim, tem-se que o escore verdadeiro V é a esperança matemática do escore empírico T.

Pode-se pressupor ainda que a correlação entre o escore verdadeiro e o erro é zero e que os erros de escore entre um teste e outro teste paralelo¹⁹ ou em duas aplicações repetidas não estão correlacionados (Pasquali, 2011, p. 72-73; Klein, 2013, p. 37). Klein (2013, p. 37) explica que a confiabilidade do teste é definida como o quadrado da correlação entre o escore observado e o escore verdadeiro, que pode ser interpretado como a proporção da variância do escore observado explicada pela variância do escore verdadeiro. O autor ainda afirma que a confiabilidade alta indica que o escore observado é uma boa estimativa do escore verdadeiro. Além disso, depende do tamanho do teste e aumenta quando se aumenta a quantidade de itens no teste, depende dos estudantes testados e, ao acrescentar itens, todo o processo de estimação precisa ser refeito.

Tanto Klein (2013, p. 40) quanto Pasquali (2011, p. 80-82) listam limitações da TCT, citando que os parâmetros nesse modelo dependem do grupo ou da amostra do grupo e dependem do teste utilizado. Essa teoria observa o teste como um todo e não observa um item individual, em que toda a informação é gerada a partir do teste total e não se pode afirmar nada sobre como um estudante pode se comportar diante de

¹⁸ Pasquali (2011, p.48-49) discute “a teoria do erro”, e afirma que um evento empírico é aleatório se sua ocorrência não depende dos eventos que ocorreram anteriormente. Além disso, o autor adiciona que o erro na medida é um evento aleatório, pela teoria do erro, e que a curva normal define uma sequência aleatória de eventos empíricos, normalmente distribuídos em torno de um ponto modal (média) igual a 0 e variância igual a 1. Disto, o autor observa que isto é verdade em uma série com um número infinito de eventos aleatórios, concluindo que em uma distribuição de erros estes se cancelariam, dado que o valor 0 é o que possui maior probabilidade na distribuição.

¹⁹ Dois testes são paralelos se eles estão medindo a mesma coisa, mas com itens (tarefas) diferentes.

um item. Além disso, uma fonte de dificuldades está no conceito de fidedignidade ou precisão da teoria, que é concebida por meio da correlação entre escores obtidos em testes paralelos. Pasquali (2011) explica que “é praticamente impossível satisfazer as condições de definição de formas paralelas” e que é improvável que o erro seja idêntico para todos os estudantes examinados. Diante dessas limitações, psicometristas buscaram alternativas que pudessem contornar diversas dificuldades, esforço que resultou na TRI.

De acordo com Klein (2013),

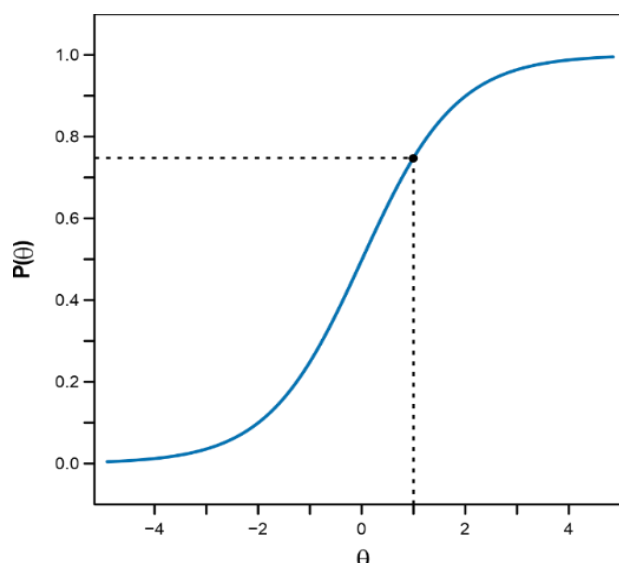
A TRI é um conjunto de modelos estatísticos onde a probabilidade de resposta a um item é modelada como função da proficiência (habilidade) do aluno (variável não observável) e de parâmetros que expressam certas propriedades dos itens, com a propriedade de que quanto maior a proficiência do aluno, maior a probabilidade de ele acertar o item (Klein, 2013, p. 40).

Segundo Pasquali (2011), a TRI trabalha com traços latentes, adotando dois axiomas fundamentais:

- 1) o desempenho do sujeito numa tarefa (item do teste) se explica em função de um conjunto de fatores ou traços latentes (aptidões, habilidades etc.). O desempenho é o efeito e os traços latentes são a causa;
- 2) a relação entre o desempenho na tarefa e o conjunto dos traços latentes pode ser descrita por uma equação monotônica crescente, chamada de CCI (Função Característica do Item ou Curva Característica do Item) [...], onde se observa que sujeitos com aptidão terão maior possibilidade de responder corretamente ao item e vice-versa [...] (Pasquali, 2011, p. 82-83).

A CCI é exemplificada na Figura 3.

Figura 3 – Curva característica do item



Fonte: Elaborado pelo autor (2025)

Nessa curva, um estudante com aptidão $\theta = 1$ tem cerca de 75% de probabilidade de acertar o item que tem essa CCI.

Duas das suposições feitas pela TRI são a unidimensionalidade e a independência local. Quanto à unidimensionalidade, assume-se que há apenas uma aptidão (traço latente, construto ou habilidade, isto é, o θ) responsável pelo desempenho do estudante. Na independência local, considera-se que as respostas de um indivíduo a diferentes itens são estatisticamente independentes, já que as aptidões que afetam o teste são mantidas constantes. Isto quer dizer que a resposta a um item não afeta a resposta a outro item.

Três modelos matemáticos predominam na TRI, que são os modelos de 1, 2 e 3 parâmetros, conhecidos como modelos logísticos²⁰. O modelo de 1 parâmetro avalia somente o parâmetro da dificuldade (b) do item, o modelo de 2 parâmetros avalia a dificuldade (b) do item e a discriminação (a) do item, enquanto o modelo de 3 parâmetros avalia a dificuldade (b) do item, a discriminação (a) do item e, também, o acerto ao acaso (c).

A partir do modelo logístico de 3 parâmetros, pode-se chegar ao de 1 e 2 parâmetros. Segundo Andrade, Tavares e Valle (2000), esses modelos “podem ser usados tanto para a análise de itens de múltipla escolha dicotomizados (corrigidos como certo ou errado) quanto para a análise de itens abertos (de resposta livre), quando avaliados de forma dicotomizada”.

Os autores apresentam que o modelo logístico unidimensional de três parâmetros é dado pela seguinte formulação:

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}$$

e explicam que:

U_{ij} é uma variável dicotômica que assume os valores 1, quando o indivíduo j responde corretamente o item i , ou 0 quando o indivíduo j não responde corretamente ao item i .

θ_j representa a habilidade (traço latente) do j -ésimo indivíduo.

$P(U_{ij} = 1|\theta)$ é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i e é chamada de Função de Resposta do Item – FRI.

²⁰ Modelo logístico é um modelo que envolve uma função em que sua curva de crescimento possui valores de 0 a 1. A função da Figura 3.

Figura 3, é uma função logística.

b_i é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala da habilidade.

a_i é o parâmetro de discriminação (ou de inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item – CCI no ponto b_i .

c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (muitas vezes referido como a probabilidade de acerto casual).

D é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando se deseja que a função logística forneça resultados semelhantes ao de função ogiva normal (Andrade, Tavares e Valle, 2000, p. 9-10).

Ambas, a TCT e a TRI, continuam a ser utilizadas para produção dos parâmetros dos itens e de medida de proficiência dos estudantes. Borgatto e Andrade (2012, p. 148) corroboram essa visão ao afirmarem que uma análise dos itens a partir da TCT auxilia na interpretação da distribuição das respostas para cada alternativa, complementando a análise feita pela TRI, uma vez que o índice de discriminação da TCT mede a capacidade do item de diferenciar os participantes de maior habilidade daqueles de menor habilidade e a correlação bisserial mede a relação entre a resposta dada numa categoria do item e a pontuação total na prova.

Dessa forma, pode-se notar que a compreensão dos parâmetros e medidas gerados em cada uma dessas teorias tem contribuições que nos ajudam a compreender pontos fracos e pontos fortes nos itens que são elaborados, no que diz respeito aos distratores dos itens.

3.1.2 A escala de proficiência do Saeb

Os testes das avaliações em larga escala passam por um processo de produção de medidas após a fase de aplicação. Nesse processo, é utilizada a TCT para analisar a consistência do teste, e, em seguida, se faz o cálculo a partir da TRI. Graças à propriedade da invariância dos parâmetros dos itens e das proficiências dos indivíduos na TRI é que se consegue produzir uma escala de proficiência, que permite monitorar o sistema educacional. Klein (2009) afirma que “tanto os parâmetros dos itens obtidos de grupos diferentes de alunos testados quanto os parâmetros de proficiência baseados em grupos diferentes de itens são invariantes, exceto pela escolha de origem e escala” (Klein, 2009, p. 127). Nesses testes, para se permitir a comparabilidade entre etapas e entre edições das aplicações, se busca utilizar itens comuns (Klein, 2003; Silva e Soares, 2010; Klein, 2013). Ainda, Klein (2013) afirma,

sobre a estimativa das proficiências dos estudantes, que os que “respondem a subconjuntos de itens diferentes de um mesmo banco de itens estão na mesma escala e podem ser comparadas” (Klein, 2013, p. 41).

Klein (2013) afirma que, na estimação da proficiência, se utiliza a distribuição normal com média 0 e desvio padrão 1 como a distribuição a priori para o estimador bayesiano EAP (expected a posteriori) de todos os grupos e anos de avaliação para definição de uma escala original. Posteriormente essa escala pode ser transformada em outra.

A escala de proficiência do Saeb é comum entre as séries e anos “baseada no SAEB/97, quando a média e o desvio padrão da distribuição de proficiências da 8ª série deste ano foram arbitrados respectivamente em 250 e 50” (Klein, 2009, p. 134).

Inep (2020) apresenta uma interpretação da escala de proficiência de forma cumulativa, como apresentado na Tabela 1 – Escala de proficiência (p) de Matemática.

Tabela 1 – Escala de proficiência (p) de Matemática

Nível	5EF	9EF	3EM
Nível 0	$125 < p$	-	-
Nível 1	$125 \leq p < 150$	$200 \leq p < 225$	$225 \leq p < 250$
Nível 2	$150 \leq p < 175$	$225 \leq p < 250$	$250 \leq p < 275$
Nível 3	$175 \leq p < 200$	$250 \leq p < 275$	$275 \leq p < 300$
Nível 4	$200 \leq p < 225$	$275 \leq p < 300$	$300 \leq p < 325$
Nível 5	$225 \leq p < 250$	$300 \leq p < 325$	$325 \leq p < 350$
Nível 6	$250 \leq p < 275$	$325 \leq p < 350$	$350 \leq p < 375$
Nível 7	$275 \leq p < 300$	$350 \leq p < 375$	$375 \leq p < 400$
Nível 8	$300 \leq p < 325$	$375 \leq p < 400$	$400 \leq p < 425$
Nível 9	$325 \leq p < 350$	$p \geq 400$	$425 \leq p < 450$
Nível 10	$p \geq 350$	-	$p \geq 400$

Fonte: Elaborado pelo autor com base nas informações do Inep (2020)

Klein (2019) afirma que o Movimento Todos pela Educação considera que os estudantes que atingem proficiência de 225 no 5EF, 300 no 9EF e 350 no 3EM atingem níveis adequados de proficiência.

3.1.3 Os itens para avaliações em larga escala

Como discutido na seção 2.2, os testes podem ser compostos por itens com formatos variados. A maioria dos testes educacionais possui um ou mais itens com objetivo de medir um domínio do conhecimento, habilidades ou capacidades cognitivas. Um item geralmente possui um comando que irá requerer uma resposta do candidato, e essa resposta pode gerar escores. Para Haladyna (2004), o item é a unidade básica de qualquer teste. O autor afirma ainda que “quanto mais esforço colocamos na construção de melhores itens de teste, provavelmente o teste será melhor” (Thorndike, 1967 apud Haladyna, 2004, p. 3 -Tradução nossa).

O processo de produção de itens para um teste é detalhado por Haladyna (2004). O autor destaca a importância da existência de um guia de elaboração de itens, contendo os formatos que devem ser usados e o que deve ser evitado na sua elaboração, abrangendo orientações sobre a revisão de itens, além de treinamentos quanto a sua elaboração. O autor relaciona a validade de um teste e a elaboração de itens, afirmando que:

podemos definir o que um item deve medir e o tipo de comportamento cognitivo que ele provoca. Podemos escrever o item, que é a etapa de explicitação na validação de construto, e podemos estudar as respostas ao item para determinar se ele se comporta da maneira que esperamos (Haladyna, 2004, p. 3 – tradução nossa).

A mitigação da subjetividade na correção do teste, a possibilidade de testar grande variedade de conteúdos em pouco tempo e o fácil uso do computador para sua correção são pontos destacados por Palácios e Oliveira (2022) para utilizar itens de resposta selecionada em avaliações em larga escala, no entanto, esse tipo de item aumenta a susceptibilidade ao chute, a impossibilidade de avaliar habilidades complexas, a dificuldade na construção de distratores de qualidade e a perda da criatividade dos estudantes na resposta.

Existem diferenças entre o que é considerado um item que é utilizado em um teste padronizado e o que é uma questão ou atividade aplicada em outros contextos, como em sala de aula. Nesse sentido, Palácios e Oliveira (2022, p. 40) afirmam que:

no primeiro caso, precisa haver uma padronização para que o processo não sofra interferência e as informações produzidas não estejam distorcidas, no segundo, o professor ou a professora pode estabelecer critérios de acordo com o contexto, como fornecer apoio diferenciado para estudantes em níveis de desenvolvimento distintos (Palácios e Oliveira, 2022, p. 40)

Palácios e Oliveira (2022, p. 73) afirmam ser importante “que um teste contenha itens com diferentes níveis de dificuldade” e sugerem que o nível de dificuldade dos itens se distribua em cinco faixas, a saber: muito fácil, fácil, mediana, difícil e muito

difícil, sendo 80% nos níveis centrais e 20% nos extremos. Os autores também apontam que “essa distribuição dos itens contribui tanto para alcançar estudantes que possuem diferentes níveis de proficiência, quanto para deixar o teste qualitativamente melhor” (Palácios e Oliveira, 2022, p. 73).

Para que, no contexto da produção de itens de Geometria, se faça a definição de níveis de dificuldade dos itens por meio de um viés pedagógico, é necessário que se apoie no currículo, mas, sobretudo, em como o estudante aprende e desenvolve esse conhecimento.

3.1.4 A Teoria de Van Hiele

A Teoria de Van Hiele foi proposta em 1957 pelos educadores holandeses Pierre van Hiele e Dina van Hiele-Geldof. Esta teoria se tornou uma importante referência para compreender como os estudantes aprendem e progridem na construção de conceitos geométricos.

Em seu livro “Structure and Insight: a theory of mathematics education” ao apresentar as raízes para a sua teoria, Pierre van Hiele (1986) explica que desde jovem interessado pelo ensino, impressionava-se com a dificuldade em compreender os materiais didáticos de álgebra, o propósito dos axiomas e definições e os métodos utilizados para ensinar, dentre outras. Anos depois, como professor, ele afirma ter compreendido e tentado fazer melhor do que havia vivenciado como estudante. Com álgebra, Van Hiele afirma ter sido suficiente analisar as operações e apresentá-las em sequência e os resultados já melhoravam, e que isso está nas raízes da construção da sua teoria de níveis de pensamento. Em Geometria, Van Hiele não obteve o mesmo sucesso, pois os livros didáticos foram feitos para construir a Geometria sobre um conjunto de axiomas, tornando-os muito difíceis e fazendo com que os estudantes não os compreendessem (Van Hiele, 1986, p. 1-2).

Entre 1938 e 1951, Van Hiele trabalhou em uma escola Montessori²¹. Nesse período, ele pôde escrever materiais didáticos e teve contato com as dificuldades dos estudantes, pois eles costumavam falar delas. Foi em 1951, quando Van Hiele se tornou professor de uma escola não Montessori, que ele entrou em conflito com a

²¹ Escola com organização diferente das escolas tradicionais, seguindo o ritmo do estudante. Esse tipo de escola não conta com séries, mas com agrupamentos por idade. Disponível em: <https://www.meimeiescola.com.br/metodomontessoriano>.

abordagem de que o professor deveria ensinar fatos e métodos e estava convencido de que o desenvolvimento da compreensão (*insight*) era o objetivo.

Esse conflito foi catalisador do desenvolvimento de uma dissertação sob orientação de Hans Freudenthal e Martinus Langeveld. Nesse processo, ele teve contato com a Psicologia de Gestalt, que o levou a entender o *insight* como resultado da percepção de uma estrutura e com a teoria dos níveis de pensamento de Piaget. A discordância da teoria de Piaget ajudou Van Hiele a construir sua teoria focada: (1) na aprendizagem (com mais que dois níveis, como preconizado por Piaget); (2) no papel crucial da linguagem para mudança de níveis, tendo em conta que os conceitos são construções humanas mutáveis e que o nível superior pode ser alcançado quando as regras do nível inferior são explicitadas e estudadas, tornando-se uma nova estrutura.

Nota-se que a construção dessa teoria surgiu de um interesse pedagógico do autor em produzir uma sistematização da forma como o estudante compreende Geometria e avança no seu aprendizado.

A Teoria de Van Hiele descreve o desenvolvimento do pensamento geométrico por meio de cinco níveis distintos (Kaleff *et al.*, 1994; Liu *et al.*, 2024; Burger e Shaughnessy, 1986; Crowley, 1987; Van Hiele, 1986). Os cinco níveis²² da Teoria de Van Hiele são:

- Nível 0 - Visualização ou reconhecimento: o estudante raciocina por meio de considerações visuais. As figuras são julgadas pela sua aparência, sem considerar suas propriedades.
- Nível 1 - Análise: o estudante identifica propriedades essenciais das figuras e a visualização das classes das formas. Nesse nível, a análise é informal das partes e atributos dos objetos, por meio de observação e experimentação.
- Nível 2 - Abstração (dedução informal ou ordenação): o estudante nota as relações entre as propriedades de diferentes figuras e entende as inclusões com base nas propriedades.

²² Crowley (1987) esclarece que existem diferentes sistemas de numeração para a teoria de van Hiele, mas que os van Hiele utilizavam a numeração do nível 0, o mais básico, até o nível 4, o mais alto nessa teoria. Essa afirmação pode ser observada em publicações como a de Van Hiele (1986). Essa pesquisa adota o mesmo sistema de numeração desses níveis.

- Nível 3 - Dedução (formal): o estudante compreende o significado das condições necessárias e suficientes para uma figura geométrica, entende o papel dos axiomas e definições e é capaz de construir provas matemáticas.
- Nível 4 - Rigor: o estudante consegue comparar sistemas com base em diferentes axiomas e estudar diversas Geometrias na ausência de modelos concretos.

Crowley (1987, p. 4) afirma que essa teoria possui algumas propriedades, e que elas são particularmente importantes para orientar a tomada de decisões educacionais. A primeira propriedade é de que a Teoria de Van Hiele é “sequencial”, isto é, para ter sucesso em um nível, o estudante precisa ter aprendido as estratégias de níveis anteriores. A segunda é o “avanço”. A autora aponta que “o progresso ou não de um nível a outro, depende mais do conteúdo e metodologia de ensino utilizada do que da faixa etária: nenhuma metodologia de ensino permite que o estudante pule um nível” (Crowley, 1987, p. 4. Tradução nossa). A terceira propriedade é “Intrínseca e extrínseca”. Quanto a essa propriedade a autora disserta que:

Os objetos inerentes a um nível tornam-se objetos de estudo no nível seguinte. Por exemplo, no nível 0, apenas a forma de uma figura é percebida. A figura é, obviamente, determinada por suas propriedades, mas é somente no nível 1 que a figura é analisada e seus componentes e propriedades são descobertos. (Crowley, 1987, p. 4. Tradução nossa).

A quarta propriedade é “linguística”. A autora afirma que uma relação que é correta em um nível pode ser modificada em outro. Ela exemplifica essa propriedade dizendo que o quadrado pode ser classificado como um retângulo ou paralelogramo, mas que um estudante que está no nível 1 ainda não compreende que essas inclusões podem ocorrer. Contudo, destaca que essas noções e linguagens são fundamentais no nível 2.

Por fim, a quinta e última propriedade citada por Crowley (1987, p. 4) é a “incompatibilidade”. A autora explica que o aprendizado pode não ocorrer se a instrução está em um nível e o estudante está em outro. Dessa maneira, se o material, o conteúdo, o professor ou o vocabulário adotado estão em um nível acima do que o estudante está, ele poderá não conseguir acompanhar os procedimentos utilizados.

Essas propriedades indicam a importância da metodologia, da organização e do material adotado para o processo de ensino em sala de aula. E para isso, os Van Hiele propuseram cinco fases sequenciais do aprendizado, que utilizadas para o

ensino, promoveriam o domínio de um nível. As cinco fases de aprendizado²³ são: 1 - investigação, 2 - orientação direta, 3 - explicação, 4 - orientação livre e 5 – integração.

Na fase 1, professor e estudantes se envolvem em uma conversa sobre o tema que será estudado, onde são feitas observações, introdução do vocabulário específico ao nível em que o conteúdo está sendo ensinado. Na fase 2, os estudantes exploram os tópicos, sequenciados cuidadosamente pelo professor, gradualmente revelando as estruturas características do nível explorado. O material deve ser constituído de tarefas curtas, para que elas resultem em respostas específicas. Na fase 3, os estudantes expressam e compartilham os entendimentos sobre as estruturas observadas. Nesse ponto, o papel do professor é mínimo, intervindo apenas para garantir uma linguagem precisa e apropriada. Na fase 4, as tarefas são mais complexas. As tarefas têm mais etapas, podem ser executadas de formas variadas e com resultados variados. Nesse tipo de atividade, os estudantes buscam formas próprias e particulares de resolver as tarefas propostas, exercitando o seu protagonismo, de forma que no processo de investigação, relações entre objetos de estudo se tornem explícitas aos estudantes. Na fase 5, o estudante revê e sintetiza aquilo que foi estudado, objetivando produzir uma visão geral em torno dos objetos e suas interrelações. (Crowley, 1987, p. 5-6)

Ainda nesse sentido, Crowley (1987) afirma que “ao final da quinta fase, os alunos atingem um novo nível de pensamento. O novo domínio do pensamento substitui o anterior, e os alunos estão prontos para repetir as fases de aprendizagem no próximo nível” (Crowley, 1987, p. 5-6 -Tradução nossa).

Com apoio da Teoria de Van Hiele, pode-se compreender os motivos pelos quais é tão importante, ao ensinar um novo conteúdo de Geometria, observar o nível em que o estudante se encontra, quer seja por meio da progressão curricular, quer seja pela resposta aos estímulos apresentados, uma vez que a observação com intenção pode resultar em ações pedagógicas mais rápidas para promover aprendizado ao estudante.

No que diz respeito à produção de itens, essa Teoria permite compreender a construção do conhecimento geométrico, trazendo clareza sobre como ele é consolidado e como isso possibilita a escrita de itens atendendo a esse conhecimento. Para análise pedagógica dos itens e resultados de avaliações, ela permite identificar

²³ Crowley (1987) utiliza atividades envolvendo o losango para ilustrar o funcionamento das fases de aprendizado.

o tipo de tarefa que o estudante possui habilidades para resolver. Dessa maneira, o planejamento pedagógico pode contar com um forte aporte para guiar o processo de ensino adaptando-o às necessidades do estudante.

3.1.5 Aspectos pedagógicos dos itens e das avaliações em larga escala

As avaliações da aprendizagem escolar, sobretudo aquelas que são desenhadas para avaliar toda uma rede de ensino, carregam consigo pressupostos da produção de uma imagem da qualidade do ensino ofertado por essa rede. A expectativa é que os itens que compõem as avaliações desse tipo tenham traduzido os descritores da matriz em um teste que consiga fazer um diagnóstico da realidade.

De acordo com Rodrigues (2006), a análise qualitativa dos itens em um teste pode ser feita em termos pedagógicos, de conteúdo e forma, e quantitativa em termos psicométricos. A autora sustenta que “a construção de instrumentos, que objetivam medir o nível de competência dos alunos, requer um conhecimento sistemático das habilidades específicas a serem alcançadas no processo ensino-aprendizagem” (Rodrigues, 2006, p. 46). Na visão de Ribeiro e Groenwald (2024), o processo de análise psicométrica dos itens é fundamental para o contexto de desenvolvimento e aplicação de avaliações educacionais, permitindo que se identifique se os testes avaliam de forma justa e precisa.

Nota-se que as orientações no sentido pedagógico na produção dos itens se limitam a informar ao elaborador a necessidade de adequação do item à etapa avaliada, alinhamento do item ao descritor, plausibilidade dos distratores e paralelismo sintático (CAEd, 2008, p. 26-28; Haladyna, 2004, p. 99). Foi nesse sentido que o CAEd (2008) afirma que o elaborador

[...] deve ser capaz de, a partir de sua experiência e do conhecimento que possui com relação ao desenvolvimento cognitivo dos estudantes que se encontram na etapa de escolarização avaliada, reconhecer o nível de dificuldade desejado na construção do item (CAEd, 2008, p. 20).

Possuir pleno conhecimento em torno de uma habilidade para desenvolver o item de que o elaborador precisa é uma condição necessária, mas não suficiente para elaboração de bons itens.

Segundo Rodrigues (2006), no sentido da validade dos testes, é necessário que o teste meça aquilo que de fato ele se propõe a medir e que um dos procedimentos específicos para a validade do teste é

a análise teórica dos itens, incluindo a análise semântica por sujeitos da própria população de interesse e a análise do conteúdo do teste por peritos nas áreas do conhecimento. Os itens que não alcançarem tais critérios deverão ser retirados do conjunto de itens (Rodrigues, 2006, p. 49).

Nesse processo, Rodrigues (2006) definiu, na metodologia de seu trabalho, três principais pontos para análise pedagógica de itens, focados especificamente no conteúdo e na forma, dos quais são citados, a seguir, dois deles:

1. Analisou-se **pedagogicamente a prova** como um todo, observando-se:
 - 1.1. a distribuição e a proporção de conteúdos abrangidos do total esperado;
 - 1.2. o nível de dificuldade dos itens que compuseram cada tema abrangido;
 - 1.3. a distribuição das competências exigidas para a resolução do item.
2. Analisou-se **pedagogicamente cada item**, por meio das seguintes observações:
 - 2.1. construção do enunciado, sua linguagem, ilustrações e nível de complexidade;
 - 2.2. plausibilidade dos distratores;
 - 2.3. coerência do gabarito;
 - 2.4. adequação entre o propósito do descritor;
 - 2.5. adequação ou não para a série avaliada (Rodrigues, 2006, p. 49).

Sobre os itens do Saeb que analisou, Rodrigues (2006, p. 67) afirma ter identificado problemas pedagógicos quanto a adequação à etapa avaliada ou ao desenvolvimento cognitivo esperado para a etapa. Outra preocupação da autora foi com a linguagem adotada, em à qual os itens foram satisfatórios.

Essa preocupação é descrita por Haladyna (2004, p. 99), no sentido de que a linguagem utilizada no item não interfira na testagem da habilidade que está sendo avaliada. Conceição (2019) também compartilha dessa preocupação, concluindo que itens que são psicometricamente difíceis podem possuir complexidades textuais menos elevadas, e que esse deveria ser um tópico presente nas orientações dos guias de elaboração de itens.

3.2 LEVANTAMENTO BIBLIOGRÁFICO E DOS DADOS DA PESQUISA

Para este estudo, foram feitos alguns movimentos para levantamento de dados e da bibliografia que pudessem dar suporte à pesquisa. O primeiro deles foi o levantamento bibliográfico na busca por livros que pudessem contribuir na construção de um referencial teórico que apresentasse as características do processo de

desenvolvimento de um teste para avaliações em larga escala, bem como as suas particularidades. Os autores de livros adotados são Palácios e Oliveira (2022), além de Haladyna (2004), Hogan (2006), Pasquali (2011) e Andrade, Tavares e Valle (2000), já consolidados no contexto de psicometria.

Com relação à elaboração de itens para aplicação no nível da Educação Básica, foram selecionados guias de elaboração e revisão de itens. Como essas são publicações técnicas, o foco ficou em utilizar os guias produzidos pelo Inep, responsável pela avaliação Saeb, e pelo CAEd, responsável pela aplicação de avaliações em diversos estados brasileiros, além de ser local laboral do pesquisador. Com o intuito de angariar esses documentos, foi feita a busca nos sites institucionais dessas duas instituições e, também, em um mecanismo de busca na internet pelo termo “guia de elaboração de itens”. Foram selecionados os guias de elaboração Inep (2010), Inep (2012) e CAEd (2008), este último, específico para o conteúdo de matemática. O banco de dissertações desse mestrado também foi uma fonte de informações relevantes sobre esse tema.

Além disso, foi feito um levantamento bibliográfico com objetivo de buscar artigos que estudassem os aspectos pedagógicos relacionados aos itens de avaliações em larga escala. Nesse sentido, essa etapa metodológica foi estruturada com base nas orientações de Gil (2017) para pesquisa em base de dados e seguiu-se o processo de triagem recomendado pela diretriz internacional PRISMA 2020 (Principais Itens para Relatar Revisões Sistemáticas e Meta-análises) como descrito em Page *et al.* (2022). A primeira versão dessa diretriz foi publicada em 2009 e passou por um processo de atualização que ocorreu entre 2018 e 2020, culminando na diretriz PRISMA 2020. Essa última versão:

foi elaborada para revisões sistemáticas de estudos que avaliam os efeitos de intervenções em saúde [...]. No entanto, os itens da lista de checagem são aplicáveis a publicações de revisões sistemáticas que avaliam outras intervenções (como intervenções sociais ou educacionais) (Page *et al.*, 2022, p. 3).

A adoção dessa diretriz para mapear o panorama das pesquisas sobre elaboração de itens se deu para garantir rigor, reprodutibilidade e transparência no levantamento bibliográfico. Neste trabalho, essa diretriz foi adotada de maneira

adaptada, uma vez que alguns dos itens dessa diretriz²⁴ não são aplicáveis ao escopo dessa pesquisa.

Assim sendo, o levantamento bibliográfico adotou a diretriz PRISMA 2020 para fazer uma revisão sistemática da literatura, obedecendo às fases de identificação, triagem, elegibilidade e inclusão.

Na fase de identificação, foi feito, inicialmente, um levantamento por meio do Google Acadêmico (com os filtros padrão) utilizando o termo “elaboração de itens” que retornou 1790 resultados. A inclusão dos termos “larga escala”, “matemática” e “pedagógico” reduziu a quantidade de registros para cerca de 650 resultados, o que ainda demandaria um longo processo de estudo de cada um dos resultados de pesquisa. Devido à limitação do Google Acadêmico em permitir exportação dos resultados de pesquisa para tratamento adicional, foi utilizada a API *OpenAlex*, de modo a produzir um arquivo com os resultados de pesquisa. Para a extração dos registros, utilizou-se a API para a base de dados *OpenAlex*, operacionalizada por meio de um *script*²⁵ desenvolvido em linguagem R. O *script* foi configurado para gerar um arquivo separado por vírgulas (.csv), utilizando o termo “elaboração de itens”, contendo os seguintes metadados ordenados por relevância: título, autores, ano, fonte, DOI e resumo. De forma a manter a integridade dos resultados obtidos na busca na base de dados *OpenAlex*, mantiveram-se campos ausentes como valores nulos, sem exclusão de registros, totalizando aproximadamente 1 759 registros. A abordagem por meio da API *OpenAlex* se mostrou consistente em comparação²⁶ com o Google Scholar, uma vez que a quantidade de resultados (incluindo duplicatas) ficou com pequena diferença de uma para a outra.

De posse do arquivo gerado na etapa anterior, seguiu-se para a fase da triagem dos registros brutos, por meio do que Gil (2017) define como leitura exploratória que

²⁴ Há um total de 27 itens nessa diretriz. Como o produto desse trabalho não é a revisão sistemática de literatura, buscou-se atender aos pontos que eram aplicáveis, expondo os resultados apresentados pelos referenciais incluídos.

²⁵ Esse script foi feito com auxílio de inteligência artificial, com modificações pontuais do pesquisador. O script é apresentado no APÊNDICE A – SCRIPT R PARA LEVANTAMENTO DE PUBLICAÇÕES.

²⁶ O OpenAlex se apresenta como alternativa para análise de grande volume de publicações e outras aplicações. Uma tese recente propõe um framework (GSM-ALEX_metrics) para analisar em grande escala os periódicos indexados no Google Scholar Metrics. (Gavron, 2025)

tem o “objetivo verificar em que medida a obra consultada interessa à pesquisa” (Gil, 2017, p. 49). Nessa etapa, retiraram-se os resultados duplicados, resultando em 1 562 resultados e aplicou-se como critério de inclusão adicional a presença do termo “larga escala”, que afunilou a quantidade de resultados para 19. A escolha desse critério deu-se devido ao escopo da pesquisa estar inserido em um ambiente de construção de instrumentos de avaliação nessa categoria. A Tabela 2 apresenta a relação de trabalhos após a aplicação dos critérios da triagem, os quais foram submetidos à etapa de avaliação de elegibilidade.

Tabela 2 – Relação de estudos selecionados na fase dois de triagem

Título	Ano	Ligação ²⁷	Autores
Elaboração de itens para avaliações em larga escala	2018	DOI*	Edson Ferreira Costa
Importância da leitura e gerenciamento do Banco Nacional de Itens na elaboração de avaliação institucional	2023	DOI	Atair Silva de Sousa
Análise da adequação dos itens do Teste de Progresso em medicina	2022	DOI	Edlaine Faria de Moura Villela, Miguel Ângelo Hyppolito, Júlio César Moriguti, Valdes Roberto Bóllela
Proposta de análise de itens das provas do SAEB sob a perspectiva pedagógica e a psicométrica	2006	DOI	Margarida Rodrigues
Publicação em cultivo in vitro de plantas: qualidade para o avanço científico e tecnológico	2014	DOI	Jean Carlos Cardoso
Protocolo fonoaudiológico de introdução e transição da alimentação por via oral para pacientes com risco para disfagia (PITA)	2010	DOI	Aline Rodrigues Padovani
Coesão e coerência na avaliação externa de língua portuguesa: gramática e linguística textual na análise de itens de testes de proficiência do Espírito Santo	2022	DOI	Higor Everson Araujo Pifano
Uma análise do desempenho dos estudantes no exame nacional do ensino médio e as contribuições para o ensino-aprendizagem de física	2017	RI-UFJF*	Carlos, Pablo Rafael de Oliveira
Avaliações em larga escala em ciências humanas no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF): a construção de um guia de orientação do trabalho de analistas e auxiliares de instrumentos de avaliação em questão	2019	RI-UFJF*	Tocantins, Andreia Cristina Teixeira
Avaliação educacional: o ENEM e a Teoria de Resposta ao Item - TRI	2022	DOI	Caroline Chinelato Silveira de Almeida
Matriz de referência para avaliação cognitiva na dimensão sociocultural da educação física escolar	2016	DOI	José Airtón F. Portes
Avaliações em larga escala de língua portuguesa	2020	DOI	Begma Tavares Barbosa, Hilda Linhares Micarello, Rosângela Veiga Júlio Ferreira
Uma análise de qualidade do uso de grandes modelos de linguagem para geração automática de itens avaliativos em português	2025	DOI	João Nogueira, João Augusto Pilato de Castro, Lucas Larcher, Rosângela Veiga Júlio Ferreira, Begma Tavares Barbosa, Jairo Francisco de Souza

²⁷ As ligações marcadas com “*”, não foram obtidas nos resultados gerados pela API OpenAlex e precisaram ser obtidas manualmente.

Estudo estatístico e pedagógico de itens de Matemática do CAEd na área de números e operações, álgebra e funções do Proeb	2020	RI-UFJE*	Souza, Tiago Gioseffi de
O gênero textual itens avaliativos de prova: uma reflexão sobre o discurso e o ensino	2011	DOI	Mara Lúcia Castilho
Recomendações de acessibilidade para pessoas com distintas condições de baixa visão ou daltonismo em exames de larga escala	2025	DOI	Andréa Poletto Sonza, Daner Silva Martins, Lídia Ruiz, Maria Elisabete Rodrigues Freire Gasparetto
A apropriação do gênero item por docentes: uma abordagem lexical	2019	DOI	Bruno de Assis Freire de Lima
Instrumento para avaliação de qualidade em creches: documentos oficiais brasileiros como fonte de pesquisa	2022	DOI	Priscila Carvalho de Castilho
Desvendando habilidades críticas de Matemática por meio da Psicometria: O uso da Teoria Clássica dos Testes e Teoria de Resposta ao Item em Avaliações Educacionais	2024	DOI	Bruno Thayguara de Oliveira Ribeiro, Cláudia Lisete Oliveira Groenwald

Fonte: Elaborado pelo autor (2026).

A terceira fase consistiu na análise da elegibilidade das obras obtidas ao concluir a segunda fase. Para isso, foi feita leitura seletiva que, para Gil (2017), é o processo de leitura em que ocorre a seleção daquilo que de fato contribui para a pesquisa. Dessa forma, é preciso se ter em mente os objetivos da pesquisa, para que as obras selecionadas contribuam para a solução do problema pesquisado. A diretriz Prisma 2020 recomenda que sejam apresentados os motivos pelos quais os estudos não se adequam ao contexto pesquisado. E, para que isso aconteça de forma clara, é necessária a definição de um critério.

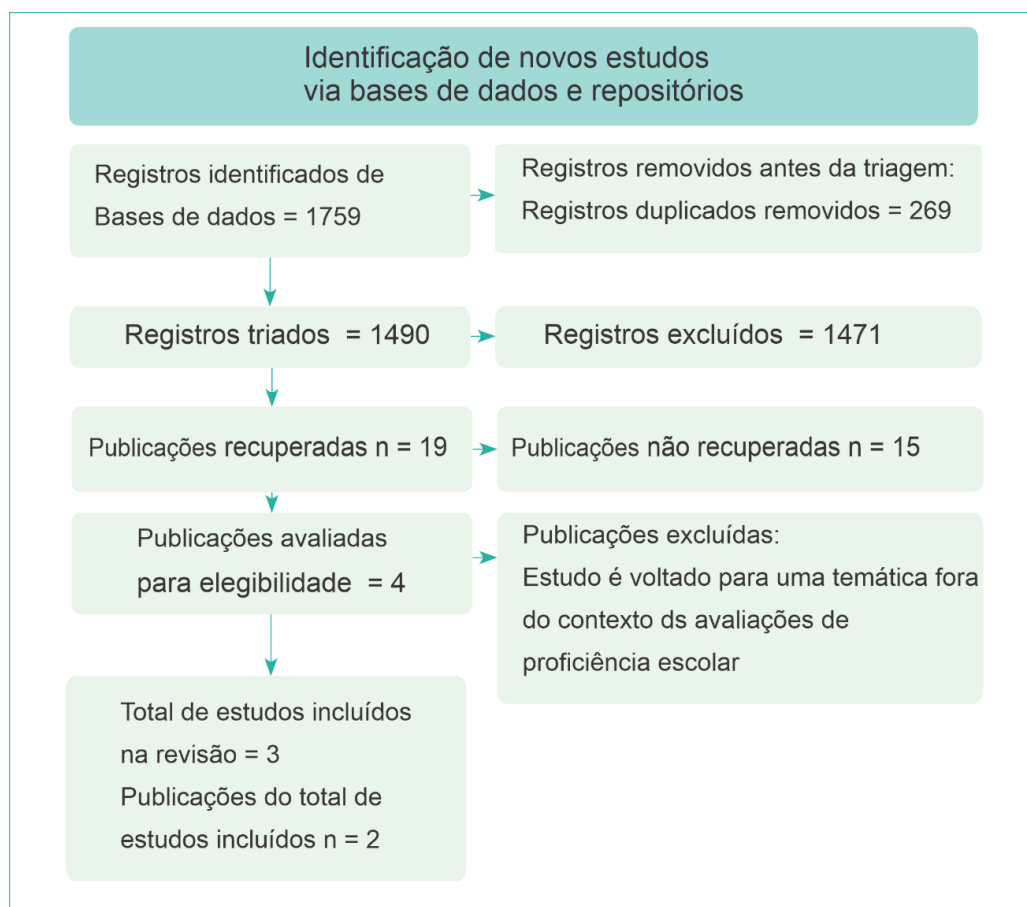
Diante disso, buscou-se identificar estudos que abordassem os aspectos pedagógicos da construção de itens aplicados em avaliações de proficiência escolar, na disciplina de matemática. Dessa forma, foi feita a leitura e a busca pelos termos “matemática” e “pedagógico” ou “pedagógica”, com o objetivo de identificar relevância para o contexto desta pesquisa.

Entre esses estudos, quatro atenderam aos critérios adotados nesta terceira fase: Almeida (2022), Souza (2020), Rodrigues (2006) e Ribeiro e Groenwald (2024), sendo os dois primeiros dissertações de mestrado e os outros dois artigos em revistas científicas. Devido ao foco em itens do Enem, o estudo de Almeida (2022) não foi considerado para a inclusão.

Nessa última fase, o estudo desenvolvido por Souza (2020) já havia sido selecionado anteriormente no banco de dissertações do PPGP e já compunha o referencial bibliográfico. Assim, no escopo final para análise e síntese, por atenderem estritamente a todos os critérios delimitados, foram incluídos os trabalhos de Rodrigues (2006) e Ribeiro e Groenwald (2024).

O detalhamento de todo o processo de identificação, triagem, elegibilidade e inclusão, desde a recuperação dos 1.759 registros brutos até a seleção final, encontra-se sumarizado na Figura 4.

Figura 4 – Fluxo de seleção dos estudos (Fluxograma Prisma)



Fonte: Elaborado pelo autor (2026), com base em Page *et al.* (2022).

O processo de elaboração de itens é marcado por uma complexidade que exige constante julgamento e adaptação. Diante disso, buscou-se um suporte pedagógico que garantisse intencionalidade e maior consistência a essa prática, culminando na adoção da Teoria de Van Hiele nesta pesquisa.

Antes de aprofundar alguns pontos em torno da seleção dos itens e outras decisões metodológicas, é necessário abordar o uso de Inteligência Artificial Generativa (IAG) nessa pesquisa. Em conformidade com as diretrizes de integridade da Portaria CNPq nº 2 664/2026, que institui a Política de Integridade na Atividade Científica do CNPq, declaro²⁸ que o uso de IAG Google Gemini, modelos PRO e Flash, ocorreu no período de dezembro de 2025 a julho de 2026 em momentos específicos.

Descrevo que houve os seguintes momentos em que o uso de IAG ocorreu:

- Apoio no Levantamento Bibliográfico (em dezembro de 2025);

²⁸ O uso de primeira pessoa do singular se deu devido à necessidade do pesquisador de declarar pessoalmente a forma de uso dessa ferramenta.

- Revisão do levantamento bibliográfico e apoio à leitura (de maio a julho de 2026);
- Construção de elementos para análise psicométrica (em julho de 2026);
- Revisão ortográfica preliminar (em julho de 2026).

O “apoio no levantamento bibliográfico” ocorreu com a utilização do modelo PRO do Gemini na produção do script em linguagem R para fazer a geração de um arquivo .csv, como explicitado anteriormente. A análise do conteúdo desse arquivo ocorreu sem nenhum auxílio de IAG.

A “revisão do levantamento bibliográfico e apoio à leitura” ocorreu devido à necessidade de revisitar os dados levantados contidos no arquivo bruto que estava arquivado junto aos documentos utilizados para essa pesquisa. Nessa revisão, os procedimentos de seleção foram refeitos para que todos os itens da diretriz PRISMA fossem verificados. Em uma das etapas dessa revisão, além da leitura e busca por palavras-chave nos resumos dos artigos, foi utilizada a ferramenta NotebookLM²⁹, que se apoia nessa mesma IAG para gerar as respostas, com o intuito de produzir um melhor contexto de análise dos 19 artigos recuperados. Os artigos incluídos na revisão foram lidos na íntegra pelo pesquisador.

A “construção de elementos para análise psicométrica” se refere aos procedimentos de construção das CCI dos itens, agrupados por descritor em cada etapa. Esse procedimento incluiu a produção do script em linguagem R, fazendo a leitura do arquivo com os dados de parâmetros dos itens e a geração dos gráficos.

A “revisão ortográfica preliminar” foi feita para que o texto final da dissertação fosse disponibilizado à banca para avaliação, com o menor número de erros gramaticais possível. A IAG foi instruída a não oferecer reescrita ou ajustes no texto, fornecendo indicação dos erros gramaticais, desvios de norma-padrão e necessidade de ajustes pontuais, de modo a manter o estilo de escrita do pesquisador. Dessa forma, o pesquisador se declara responsável, de modo integral, pelo texto final desta dissertação.

Com a homologação da BNCC, a principal avaliação da Educação Básica, o Saeb, começou um movimento de transição de suas matrizes a partir de 2019. Devido

²⁹ Essa plataforma é integrada ao ecossistema Google e utiliza a IAG Gemini sob um paradigma ancorado nas fontes fornecidas pelo usuário, por meio de arquivo ou de link, em que o processamento e raciocínio analítico é feito com base nessas fontes.

a fatores como a pandemia de COVID-19, até o ano de 2023³⁰ a matriz de referência de 2001 continuou sendo a matriz para o 5º ano e 9º ano do ensino fundamental e para as 3ª/4ª séries do ensino médio para Língua Portuguesa e Matemática. De acordo com a Portaria Nº 435/2025 do Inep, a aplicação do Saeb em 2025 contará com aplicações mistas utilizando a matriz alinhada à BNCC para as etapas do 5º ano (5EF) e 9º ano do ensino fundamental (9EF) e, para as 3ª/4ª séries do ensino médio (3EM), continuará com a matriz de 2001.

Dessa maneira, esta pesquisa adota uma análise a partir da matriz de 2001, uma vez que os dados produzidos e publicados pelo CAEd até 2025 se referem a avaliações que tiveram como base as matrizes Saeb para as etapas indicadas anteriormente.

Para que seja feita uma análise psicométrica e pedagógica dos itens aplicados em testes de avaliações em larga escala, foi requisitada à Fundação CAEd a disponibilização de quatro itens em cada um dos descritores. Os critérios de inclusão foram: itens componentes de instrumentos de avaliações em larga escala produzidas pelo CAEd parametrizados em três parâmetros lineares pela TRI na escala Saeb, considerados como itens abertos; já os de exclusão, foram os itens que não pertencem ao eixo do conhecimento Geometria e que ainda não possuem medidas estatísticas produzidas, ou que não são de propriedade³¹ do CAEd. Além disso, foi requisitada à equipe de Medidas Educacionais do CAEd uma planilha contendo os dados e apresentando as medidas psicométricas dos itens na TRI.

A requisição desse material foi feita por e-mail ao CAEd/UFJF, que culminou na mobilização das equipes da Fundação CAEd, responsáveis por esses dados para sua disponibilização. Os dados solicitados foram disponibilizados mediante assinatura de um Termo de Autorização permitindo o uso dos itens e das medidas estatísticas para esta pesquisa. O acesso foi feito via servidor SFTP (Secure File Transfer Protocol), configurado com exclusividade para acesso aos dados solicitados.

³⁰ Veja mais informações em <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/matrizes-e-escalas>.

³¹ Existem itens fornecidos pelo Inep para serem itens âncora, utilizados para calibrar os itens na mesma escala do Saeb. Klein (2009, p. 138) discute as características e critérios de seleção dos itens-âncora.

O motivo de selecionar itens em uma mesma escala é para garantir que as medidas dos itens sejam comparáveis, tanto entre etapas quanto entre anos de aplicação. As medidas estatísticas advindas da TCT não foram disponibilizadas, uma vez que traçar relações entre essas medidas para aplicações e testes diferentes não é possível. Os descritores e quantitativo de itens por descritor cedidos pela Fundação CAEd estão explicitados na Tabela 3. Nessa tabela, é apresentada a matriz Saeb de 2001 do Eixo de Geometria, com a descrição de cada descritor e em cada etapa. Além disso, foi inserida a quantidade de itens cedidos para cada descritor dessa matriz.

Tabela 3 – Itens cedidos pela Fundação CAEd

Etapa	Descritor	Descrição	Quantidade
5EF	D01	Identificar a localização/movimentação de objeto em mapas, croquis e outras representações gráficas.	4
5EF	D02	Identificar propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações.	4
5EF	D03	Identificar propriedades comuns e diferenças entre figuras bidimensionais pelo número de lados, pelos tipos de ângulos.	4
5EF	D04	Identificar quadriláteros observando as posições relativas entre seus lados (paralelos, concorrentes, perpendiculares).	4
5EF	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.	4
9EF	D01	Identificar a localização/movimentação de objeto em mapas, croquis e outras representações gráficas.	4
9EF	D02	Identificar propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando-as com as suas planificações.	3
9EF	D03	Identificar propriedades de triângulos pela comparação de medidas de lados e ângulos.	4
9EF	D04	Identificar relação entre quadriláteros por meio de suas propriedades.	3
9EF	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.	3
9EF	D06	Reconhecer ângulos como mudança de direção ou giros, identificando ângulos retos e não-retos.	3
9EF	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.	4

9EF	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).	2
9EF	D09	Interpretar informações apresentadas por meio de coordenadas cartesianas.	3
9EF	D10	Utilizar relações métricas do triângulo retângulo para resolver problemas significativos.	3
9EF	D11	Reconhecer círculo/circunferência, seus elementos e algumas de suas relações.	3
3EM	D01	Identificar figuras semelhantes mediante o reconhecimento de relações de proporcionalidade.	3
3EM	D02	Reconhecer aplicações das relações métricas do triângulo retângulo em um problema que envolva figuras planas ou espaciais.	4
3EM	D03	Relacionar diferentes poliedros ou corpos redondos com suas planificações ou vistas.	4
3EM	D04	Identificar a relação entre o número de vértices, faces e/ou arestas de poliedros expressa em um problema.	3
3EM	D05	Resolver problema que envolva razões trigonométricas no triângulo retângulo (seno, cosseno, tangente).	3
3EM	D06	Identificar a localização de pontos no plano cartesiano.	4
3EM	D07	Interpretar geometricamente os coeficientes da equação de uma reta.	4
3EM	D08	Identificar a equação de uma reta apresentada a partir de dois pontos dados ou de um ponto e sua inclinação.	4
3EM	D09	Relacionar a determinação do ponto de interseção de duas ou mais retas com a resolução de um sistema de equações com duas incógnitas.	4
3EM	D10	Reconhecer, dentre as equações do 2º grau com duas incógnitas, as que representam circunferências.	4

Fonte: Elaborado pelo autor (2026), com base nos dados disponibilizados pelo CAEd (2026).

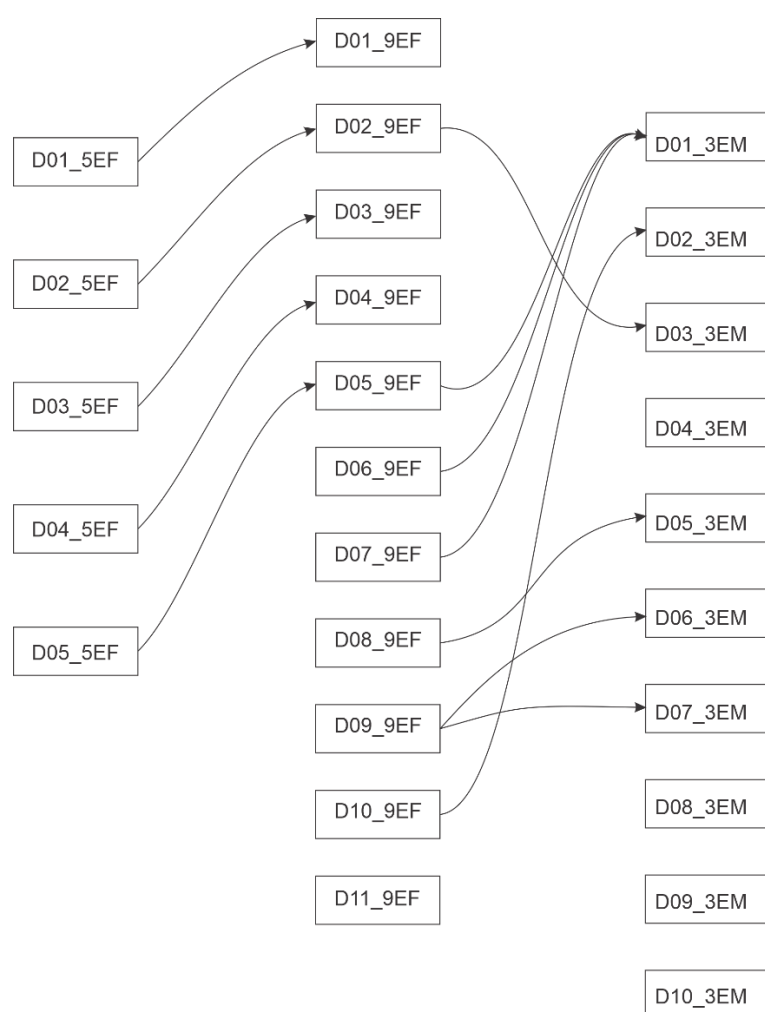
Para alguns descritores, a Fundação CAEd não disponibilizou a quantidade de quatro itens, por não haver esse quantitativo de itens disponíveis, atendendo aos critérios apresentados.

Na próxima seção, munidos dos dados disponíveis para este estudo e apoiados no que foi apresentado no Capítulo 2 e na seção 3.1, será produzida uma análise da matriz de referência do Saeb 2001, no eixo do conhecimento de Geometria, e a análise dos itens a partir de suas tarefas, de forma a subsidiar a discussão entre os níveis de Van Hiele e a observação empírica dos itens, culminando na proposição do plano de ação.

3.3 ANÁLISE DA MATRIZ E DOS ITENS: OS PRESSUPOSTOS PEDAGÓGICOS E PSICOMÉTRICOS

Os descritores da matriz do Saeb, do ponto de vista da construção de itens, apresentam uma estrutura que pressupõe um controle sobre o nível de dificuldade do item produzido, uma vez que há descritores que avaliam o mesmo objeto do conhecimento. A Figura 5 ilustra uma relação entre os descritores³² da matriz do Saeb de 2001.

Figura 5 – Relação entres descritores do eixo de Geometria da Matriz Saeb 2001



Fonte: Elaborado pelo autor (2026)

³² Nessa pesquisa, utilizou-se uma codificação composta de [descriptor] _ [etapa] para facilitar a referência a um determinado descritor de uma determinada etapa, já que a matriz não possui essa diferenciação.

A matriz no eixo de Geometria apresenta descritores em que os objetos de conhecimento se relacionam. Note, por exemplo, que os descritores do 5EF possuem uma relação direta com os descritores do 9EF, uma vez que eles possuem a mesma redação ou uma redação muito próxima entre uma etapa e a outra. Na progressão entre essas etapas, identifica-se um processo de direcionamento e complexificação esperada na etapa seguinte. No descritor D02_5EF, o enfoque era em poliedros e corpos redondos, já no D02_9EF se avança para figuras bidimensionais e tridimensionais de forma geral. Entre D03_5EF e D03_9EF, há uma mudança no enfoque, uma vez que no 5EF era esperada a identificação de propriedades de figuras bidimensionais de forma mais ampla, enquanto no 9EF é esperado esse processo focado em triângulos.

Quando se analisa a transição dos descritores do 9EF para a 3EM, nota-se que no ensino médio os descritores requerem um amadurecimento que culmina na formalização algébrica e trigonométrica da Geometria.

Pode-se observar que alguns descritores do 9EF dessa matriz se desdobram, tornando-se aplicações de descritores dos 3EM. Esse é o caso para o descritor D10_9EF, que trata das relações métricas no triângulo retângulo, atuando como um requisito básico para os descritores do ensino médio: o D02_3EM expande essas relações para problemas envolvendo figuras espaciais, e o D05_3EM, que introduz a abstração das razões trigonométricas (seno, cosseno e tangente).

Outra convergência ocorre no estudo de proporções. As noções de ampliação e redução em malhas que foram dos descritores 5EF_D05 (em nível mais elementar) e D05_9EF e a identificação de propriedades na homotetia no descritor D07_9EF se unificam e se formalizam no D01_3EM, que exige a identificação de figuras semelhantes estritamente pelo reconhecimento das relações matemáticas de proporcionalidade.

Em relação à Geometria Espacial, a habilidade de relacionar figuras tridimensionais às suas planificações no D02_9EF possui continuidade direta no D03_3EM, que trata de planificações e vistas. Nesse campo, se avança no ensino médio ao introduzir o D04_3EM, que requer a compreensão da relação de Euler.

O salto de complexidade mais expressivo concentra-se na introdução da Geometria Analítica. O que começou com a localização em mapas nos descritores D01_5EF e D01_9EF, e a interpretação de coordenadas cartesianas no D09_9EF, tornam-se a base para a localização de pontos no descritor D06_3EM. É no 3EM em

que se espera que o pensamento geométrico seja traduzido para a álgebra, com os descritores D07_3EM, D08_3EM e D09_3EM. Seguindo nesse sentido, o descritor D11_9EF, que avalia o reconhecimento dos elementos básicos da circunferência, torna-se a base fundamental para a identificação matemática de sua equação do 2º grau no plano cartesiano, avaliado pelo D10_3EM.

Além dessas conexões apresentadas, há outras que são implicitamente relacionadas a descritores do 3EM, como é o caso dos descritores D03, D04, D06 e D08 do 9EF. Um exemplo dessa dinâmica é descritor D06_9EF (reconhecer ângulos como mudança de direção ou giros e identificar ângulos retos), que, embora não intimamente relacionado a outro no ensino médio, contribui para que o estudante execute tarefas como as avaliadas pelo D10_9EF (Utilizar relações métricas do triângulo retângulo para resolver problemas significativos) e é o pré-requisito elementar que permite ao estudante operar o D05_3EM.

Esta progressão das habilidades geométricas não acontece por simples acúmulo de conteúdos, mas na forma como o estudante aprende os novos conteúdos. Observar essa trajetória a partir da Teoria de Van Hiele e com apoio da taxonomia de Bloom nos ajuda a entender de que forma esse aprendizado ocorre nessa progressão. Sob o prisma da taxonomia de Bloom, pode-se examinar um descritor ou habilidade a partir de seu verbo ou ação prescrita, de forma a entender o objetivo esperado, e, sob o prisma da Teoria de Van Hiele, é possível compreender o processo em que a consolidação do conhecimento geométrico ocorre.

No 5EF, os estudantes ainda processam a Geometria com predominância no Nível 0 (Visualização) da Teoria de Van Hiele, e iniciam a transição para o nível 1 (Análise). Nessa etapa, a Geometria é compreendida a partir da aparência global, pela contagem de elementos visíveis ou a partir de propriedades mais elementares, sem formalismos. No descritor D04_5EF, por exemplo, o estudante identifica quadriláteros observando diretamente as posições relativas de seus lados na figura. O quadrilátero é um objeto visual apenas, desconsiderando as interseções entre as suas classificações, como o fato de que todo quadrado é um retângulo. A demanda cognitiva das tarefas exigidas concentra-se nos processos de nível inferior da taxonomia de Bloom: Lembrar e Compreender, o que é refletido nos verbos utilizados nos descritores da matriz (Identificar e Reconhecer). O estudante precisa resgatar da memória a forma de um poliedro (D02_5EF) ou de um polígono (D03_5EF) e associá-la à imagem apresentada.

No 9EF é esperado que o estudante já tenha consolidado o conhecimento geométrico até o Nível 1, além de estar concluindo a transição para o Nível 2 (Abstração) da Teoria de Van Hiele. Os objetos geométricos precisam deixar de ser uma figura e passar a ser compreendidos a partir de um conjunto de propriedades interligadas. O descritor D04_5EF e o D04_9EF ilustram isso. Enquanto no 5EF o foco era visual, no 9EF o foco fica nas propriedades. Além disso, com a introdução das transformações homotéticas (D07_9EF) e das relações métricas (D10_9EF), o estudante é levado a deduzir informações implícitas por meio de propriedades e definições. Nessa etapa, a exigência cognitiva se eleva, cobrando do estudante a aplicação e algumas análises. As ações indicadas pelos verbos dos descritores passam a exigir o uso de procedimentos em situações específicas ("Utilizar [...]", "Interpretar[...]", "Resolver problema utilizando propriedades"). O estudante já precisa conseguir informações sobre aquilo que não está explícito.

Quando se chega ao 3EM, há um salto para um processo mais abstrato em alguns descritores e o estudante chega até o nível 3 (Dedução) da Teoria de Van Hiele, de forma que objetos geométricos passam também a ser traduzidos em linguagem algébrica. O estudante passa a interpretar os coeficientes de uma reta (D07_3EM) ou reconhecer a equação de uma circunferência (D10_3EM), o que exige que o estudante deduza formas geométricas pela manipulação de variáveis e axiomas em um plano cartesiano. As exigências cognitivas atingem as categorias de análise e avaliação da Taxonomia de Bloom. Os verbos na matriz requerem, por exemplo, as habilidades de relacionar e interpretar. Nessa etapa, o estudante mobiliza a álgebra para justificar, avaliar o espaço geométrico.

A partir dessa análise, o estudo das tarefas dos itens cedidos para esta pesquisa será feito, visando observar como essa progressão ocorre nos itens aplicados nos testes das avaliações em larga escala produzidas pelo CAEd. A descrição das tarefas dos itens foi feita pelo próprio pesquisador a partir do ponto 2 descrito por Rodrigues (2006). A Tabela 4 apresenta a listagem dos itens, as suas medidas estatísticas e a descrição das tarefas de cada item. Essa tabela apresenta a lista de itens completa, conforme dados cedidos pelo CAEd, com o código do item – um identificador único para cada item – a etapa, os parâmetros a , b e c , e uma descrição das tarefas para cada item, elaborada pelo pesquisador.

Tabela 4 – Itens e medidas cedidas pela Fundação CAEd, com a descrição da tarefa do item

ITEM	Etapas	Descritor	Parâmetro A	Parâmetro B	Parâmetro C	Descrição da tarefa do item
M060197I7	5EF	D01	0,01803	225,3	0,18651	Determinar a localização de um objeto ao final de uma movimentação dados o ponto de partida em um mapa e os movimentos executados, utilizando referencial diferente do estudante.
M073478I7	5EF	D01	0,01151	271,3	0,12964	Determinar a localização de uma pessoa ao final de uma movimentação dados o ponto de partida em uma planta baixa e os movimentos executados, utilizando referencial diferente do estudante.
M090345I7	5EF	D01	0,01719	223,3	0,06447	Identificar a localização de um objeto em uma planta baixa, utilizando referencial diferente do estudante.
M050030B1	5EF	D02	0,01580	161,7	0,06861	Identificar a um par de objetos com formato tridimensional, que corresponde a planificações dadas de um corpo redondo e de um paralelepípedo.
M050346H6	5EF	D02	0,01350	233,4	0,21615	Identificar uma figura tridimensional que corresponde a planificação de um poliedro dado.
M050430H6	5EF	D02	0,01353	156,8	0,13434	Identificar uma figura tridimensional que corresponde a planificação de um dado corpo redondo.
M073633I7	5EF	D02	0,00776	233,3	0,12052	Determinar a planificação que corresponde a um poliedro dado.
M050350C2	5EF	D03	0,01567	287,5	0,17159	Identificar dentre quatro quadriláteros aquele que corresponde a um quadrado, a partir de seu formato.
M060283B1	5EF	D03	0,01538	270,1	0,25991	Nomear um polígono a partir da quantidade de lados.

M050241H6	5EF	D04	0,01805	153,1	0,04088	Identificar dentre quatro quadriláteros aquele que corresponde a um quadrado, a partir da posição relativa de seus lados.
M050459H6	5EF	D04	0,01397	224,7	0,21963	Identificar dentre quatro quadriláteros aquele que corresponde a um retângulo, a partir da posição relativa de seus lados.
M050672H6	5EF	D04	0,01608	264,4	0,29774	Identificar dentre quatro quadriláteros aquele que corresponde a um retângulo, a partir da posição relativa de seus lados.
M040078H6	5EF	D05	0,00996	227,4	0,36390	Identificar dentre quatro figuras desenhadas em uma malha quadriculada, aquela que representa uma ampliação de uma dada figura desenhada em malha quadriculada.
M050050ES	5EF	D05	0,00733	391,2	0,33675	Determinar o fator de ampliação do perímetro entre duas figuras desenhadas em uma malha quadriculada.
M050214B1	5EF	D05	0,01888	233,5	0,15444	Determinar o fator de redução do perímetro entre uma figura uma malha quadriculada.
M050217I7	9EF	D01	0,00817	205,5	0,11934	Identificar a localização de um objeto em uma planta baixa, utilizando referencial diferente do estudante.
M090054I7	9EF	D01	0,01402	338,6	0,25383	Determinar a localização de um objeto ao final de uma movimentação dados o ponto de partida em um mapa e os movimentos executados, utilizando referencial diferente do estudante.
M090511H6	9EF	D01	0,01541	231,7	0,17042	Identificar a localização de um objeto em uma organização retangular de objetos.
M070022H6	9EF	D02	0,01374	274,5	0,14016	Identificar a planificação de um corpo redondo.
M070063I7	9EF	D02	0,00941	300,8	0,12147	Identificar a planificação de um poliedro.
M090275H6	9EF	D02	0,01885	352,3	0,16826	Identificar a planificação de um poliedro.

M080717E4	9EF	D03	0,01511	279,6	0,13717	Dentre quatro triângulos dados, identificar a partir da medida de seus lados aquele que possui uma certa propriedade.
M090200G5	9EF	D03	0,02698	316,5	0,18801	Classificar um triângulo dado a partir da medida de seus lados e seus ângulos.
M090370I7	9EF	D03	0,02453	365,1	0,26046	Classificar um triângulo dado a partir da medida de seus ângulos.
M090706H6	9EF	D03	0,03469	371,5	0,20528	Classificar um triângulo a partir da medida de seus lados e seus ângulos, dados a partir de suas características.
M060023H6	9EF	D04	0,02951	320,6	0,16873	Nomear um quadrilátero a partir de suas propriedades.
M090137H6	9EF	D04	0,01494	350,0	0,12162	Nomear um quadrilátero a partir de suas propriedades.
M090192G5	9EF	D04	0,01938	136,0	0,08350	Identificar dentre quatro quadriláteros aquele que corresponde a um quadrado.
M090203EX	9EF	D04	0,01300	321,4	0,27479	Classificar, a partir de suas propriedades comuns, quatro quadriláteros dados.
M051656E4	9EF	D05	0,00646	320,8	0,15025	Determinar o fator de ampliação da área entre duas figuras desenhadas em uma malha quadriculada.
M090513E4	9EF	D05	0,01911	347,1	0,17418	Determinar o fator de redução da área entre duas figuras desenhadas em uma malha quadriculada.
M090609I7	9EF	D05	0,00849	194,6	0,17398	Identificar dentre quatro figuras desenhadas em uma malha quadriculada, aquela que representa uma aplicação de uma dada figura desenhada em malha quadriculada.
M090125C2	9EF	D06	0,02730	313,1	0,32549	Determinar a posição do indicador de um objeto, dado a medida do ângulo de giro.
M090293A8	9EF	D06	0,04087	333,4	0,19700	Determinar o ângulo de giro do ponteiro de um relógio.
M090514E4	9EF	D06	0,01109	330,5	0,29316	Identificar tipo de ângulo de mudança de direção.

M090048I7	9EF	D07	0,01997	316,5	0,21598	Identificar dentre quatro representações, aquela que representa uma redução homotética de um triângulo.
M090291C2	9EF	D07	0,02688	327,9	0,17495	Determinar a medida de um lado de uma figura dada, transformada homoteticamente meio de uma aplicação com razão k.
M090665I7	9EF	D07	0,03761	363,3	0,16100	Determinar a constante de proporcionalidade de uma transformação homotética.
M090764H6	9EF	D07	0,02532	314,9	0,07954	Determinar a constante de proporcionalidade de uma transformação homotética.
M073878I7	9EF	D08	0,01966	340,3	0,28678	Determinar a medida de um ângulo, dados os outros ângulos do polígono apresentado.
M080577E4	9EF	D08	0,02797	311,2	0,26615	Determinar a medida de um ângulo interno de uma composição de polígonos regulares.
M090259H6	9EF	D08	0,01924	384,8	0,10261	Determinar a medida de um ângulo interno de uma composição de polígonos regulares.
M091435H6	9EF	D08	0,02241	325,8	0,17079	Determinar a medida de um ângulo interno de um polígono regular.
M051810I7	9EF	D09	0,01248	285,4	0,18013	Identificar a localização de um par ordenado, modificado em cada coordenada por meio de movimentações , em um plano cartesiano.
M052148I7	9EF	D09	0,02055	320,5	0,14286	Identificar par ordenados em que um objeto está representado.
M090145H6	9EF	D09	0,01695	319,4	0,21421	Identificar o ponto que indica a localização de um par ordenado, em um plano cartesiano.
M090280H6	9EF	D09	0,01526	275,9	0,30208	Identificar o ponto que indica a localização de um par ordenado, em um plano cartesiano.
M090378A9	9EF	D10	0,02874	326,6	0,17782	Determinar a medida do comprimento de uma escada dada as medidas dos catetos de um triângulo retângulo.

M090801H6	9EF	D10	0,01487	305,0	0,34940	Determinar a medida do comprimento de um objeto, dada as medidas dos catetos de um triângulo retângulo.
M091441H6	9EF	D10	0,03714	364,2	0,15670	Determinar a medida do comprimento de um objeto, dada as medidas dos catetos de um triângulo retângulo.
M091442H6	9EF	D10	0,01606	339,2	0,28772	Determinar a medida da altura de uma construção, dada as medidas das projeções dos catetos de um triângulo retângulo.
M090515H6	9EF	D11	0,01904	262,7	0,26278	Identificar o segmento que representa um dos elementos de uma circunferência apresentada sobre malha quadriculada.
M090522E4	9EF	D11	0,02563	317,7	0,22353	Identificar o segmento que representa um dos elementos de uma circunferência.
M090573H6	9EF	D11	0,04227	323,7	0,21670	Identificar o segmento que representa um dos elementos de uma circunferência apresentada sobre malha quadriculada.
M090574H6	9EF	D11	0,01174	273,5	0,17494	Identificar semicírculo dentre quatro figuras apresentadas.
M00081141	3EM	D01	0,00832	295,2	0,06404	Identificar, entre cinco opções apresentadas, o triângulo que possui relação de semelhança com um triângulo dado como referência.
M100070I7	3EM	D01	0,02289	332,1	0,14401	Reconhecer um par de triângulos semelhantes em um conjunto de figuras planas dadas
M120005D3	3EM	D01	0,00643	463,1	0,15429	Reconhecer um par de triângulos semelhantes em um conjunto de figuras planas dadas
M120352H6	3EM	D01	0,01028	379,9	0,23463	Reconhecer um par de triângulos semelhantes em um conjunto de figuras planas dadas
M110512I7	3EM	D02	0,01332	362,2	0,14055	Determinar a altura de um objeto a partir das projeções dos catetos de um triângulo retângulo.

M120354H6	3EM	D02	0,02082	412,7	0,15004	Determinar o comprimento da projeção do cateto, dadas as medidas do cateto e da hipotenusa do triângulo retângulo.
M121171I7	3EM	D02	0,02966	365,0	0,14508	Determinar o comprimento da hipotenusa, dada as medidas dos catetos de um triângulo retângulo.
M121172I7	3EM	D02	0,02836	366,6	0,22390	Determinar o comprimento da hipotenusa, dada as medidas dos catetos de um triângulo retângulo.
M110100H6	3EM	D03	0,01129	283,3	0,07006	Corresponder uma figura geométrica tridimensional a sua planificação.
M120355H6	3EM	D03	0,01493	270,4	0,08255	Corresponder uma figura geométrica tridimensional a sua vista superior.
M120356H6	3EM	D03	0,01324	274,7	0,04985	Corresponder uma figura geométrica tridimensional a sua planificação.
M120375I7	3EM	D03	0,01397	266,4	0,12770	Corresponder uma figura geométrica tridimensional a sua vista superior.
M00081139	3EM	D04	0,01807	381,8	0,18330	Determinar a quantidade de faces de um poliedro, dada a quantidade vértices e arestas.
M00086026	3EM	D04	0,01969	370,4	0,19501	Determinar a quantidade de faces de um poliedro, dada a quantidade vértices e arestas.
M101479I7	3EM	D04	0,02483	406,7	0,17351	Determinar a quantidade de vértices de um poliedro, dada a quantidade faces e arestas.
M120015EX	3EM	D04	0,02669	372,8	0,18809	Determinar a quantidade de arestas de um poliedro, dada a quantidade faces e vértices.
M100458H6	3EM	D05	0,03252	381,7	0,20624	Determinar a diagonal de um retângulo, dadas: a altura, o ângulo de inclinação e o seno desse ângulo.
M110034I7	3EM	D05	0,03049	371,8	0,23205	Determinar a altura de um prédio, utilizando um triângulo retângulo, dadas a distância ao prédio e o ângulo de inclinação e o cosseno desse ângulo.
M120094I7	3EM	D05	0,01399	394,5	0,26594	Determinar a altura de um objeto, utilizando um triângulo retângulo, dadas a distância ao objeto e o ângulo de inclinação e o cosseno desse ângulo.

M120790I7	3EM	D05	0,02213	350,9	0,22278	Determinar a altura de um objeto, utilizando um triângulo retângulo, dada a medida da hipotenusa, o ângulo de inclinação e o seno desse ângulo.
M100385H6	3EM	D06	0,01339	240,1	0,09100	Identificar o ponto que indica a localização de um objeto por meio de um par ordenado, em um plano cartesiano.
M120164I7	3EM	D06	0,01010	334,6	0,06310	Identificar o ponto que indica a localização de um par ordenado, em um plano cartesiano.
M120214I7	3EM	D06	0,01146	298,3	0,03599	Identificar o ponto que indica a localização de um par ordenado, em um plano cartesiano.
M120372I7	3EM	D06	0,02587	364,8	0,15755	Identificar o par ordenado que indica a localização de um ponto no plano cartesiano, que tem suas coordenadas modificadas por translações.
M120006D3	3EM	D07	0,00410	421,8	0,04340	Identificar, no plano cartesiano, o coeficiente linear de 3 equações da reta.
M120117B1	3EM	D07	0,01707	384,9	0,19485	Identificar, no plano cartesiano, o coeficiente linear de 2 equações da reta.
M120702H6	3EM	D07	0,06071	408,9	0,18927	Interpretar o sinal dos coeficientes de uma equação da reta na forma reduzida, dada a representação da reta no plano cartesiano.
M120318ES	3EM	D08	0,01694	407,5	0,23763	Determinar a equação de uma reta, dado um ponto e a inclinação da reta.
M120695I7	3EM	D08	0,01601	380,9	0,27299	Determinar a equação de uma reta, dados dois pontos.
M120785H6	3EM	D08	0,06064	380,2	0,13174	Determinar a equação de uma reta, dados dois pontos.
M121403H6	3EM	D08	0,02316	396,4	0,22947	Determinar a equação de uma reta, dado um ponto e a inclinação da reta.
M120216A9	3EM	D09	0,00692	317,8	0,12617	Determinar, dentre os pontos apresentados, aquele que representa a solução de um sistema de equações.

M120680ES	3EM	D09	0,00886	345,4	0,05798	Determinar, dentre os pontos apresentados, aquele que representa a solução de um sistema de equações.
M121490H6	3EM	D09	0,03105	383,5	0,22351	Determinar, dentre os pontos apresentados, aquele que representa a solução de um sistema de equações.
M120168A9	3EM	D10	0,01146	463,2	0,21556	Reconhecer, dentre equações de segundo grau, aquele que representa uma circunferência.
M120297B1	3EM	D10	0,01039	429,6	0,22825	Reconhecer, dentre equações de segundo grau, aquele que representa uma circunferência.
M121701H6	3EM	D10	0,01466	394,6	0,18584	Reconhecer, dentre equações de segundo grau, aquele que representa uma circunferência.

Fonte: Elaborado pelo autor (2026), com base nos dados disponibilizados pelo CAEd (2026).

Partindo para o processo de análise empírica dos itens apresentados na Tabela 4, nota-se que as tarefas descritas pelos itens do descritor D01_5EF e D01_9EF exigem, nessa etapa, a mobilização de conhecimentos no nível 0 (Visualização) da Teoria de Van Hiele, uma vez que o estudante precisa de habilidades de observação e localização no espaço. A resolução dessas tarefas não demanda o reconhecimento de propriedades matemáticas, mas sim a capacidade de operar de forma empírica sobre o espaço visual para se orientar em mapas ou plantas baixas. Embora sejam de nível 0 na Teoria de Van Hiele, os conceitos avaliados nesse descritor são estudados durante os quatro anos iniciais do Ensino Fundamental, aparecendo na matriz do 9EF como descritor comum à etapa do 5EF.

Nos itens, é notável que o avanço na complexidade da tarefa se dá de acordo com a quantidade de mudanças de referencial e quantidade de movimentações. Do ponto de vista estatístico, os itens apresentam parâmetro B adequado à etapa no 5EF, de acordo com a Tabela 1, e enquanto no 9EF eles se apresentam, em geral, mais fáceis, tendo em vista que dois dos itens possuíam estrutura próxima ao visto no 5EF e exibem parâmetro B mais de um desvio padrão abaixo do nível considerado adequado, como visto na subseção 3.1.2. Além disso, os itens do 5EF se mostraram mais discriminativos que os itens do 9EF e com menor probabilidade de acerto ao acaso, como é possível observar por meio das CCI dos itens desse descritor, apresentadas no APÊNDICE B – CURVAS CARACTERÍSTICAS DOS ITENS.

Para o descritor D02_5EF, a exigência cognitiva concentra-se predominantemente no nível 0 na Teoria de Van Hiele. Nessa etapa, é esperado que o estudante associe uma figura tridimensional a uma de suas planificações. O estudante avalia a equivalência entre as representações bidimensionais e tridimensionais a partir do aspecto global das formas, reconhecendo objetos comuns, sem necessitar deduzir ativamente propriedades estruturais. As tarefas buscam que ele utilize propriedades mais elementares. Enquanto isso, no 9EF o descritor D02_9EF avalia a mesma habilidade, mas exigindo que o estudante observe, além do formato das faces poligonais, se essas faces se conectam às faces corretas e/ou se uma determinada organização gera alguma sobreposição. Foi exatamente essa relação que foi observada nos itens, tanto no 5EF quanto no 9EF, que performaram com dificuldade adequada, para a sua etapa. O avanço proposto pela habilidade leva o estudante do 9EF para processos que exigem conhecimentos geométricos no nível 1 (Análise) de Van Hiele. Nesse caso, os itens se mostraram mais discriminativos,

uma vez que o tipo de processo esperado demanda maior compreensão dos elementos que compõem o objeto apresentado no item.

Desde os primeiros anos de escolarização, o estudante tem contato com figuras geométricas bidimensionais. No início o foco fica nas diferenças proeminentes, notáveis pela observação. Depois, o estudante começa a conhecer novas figuras e conhecer algumas propriedades e características. No 5EF, o descritor D03_5EF evidencia justamente o início da transição do Nível 0 da Teoria de Van Hiele para o nível 1. As tarefas dos itens, e o parâmetro de dificuldade, demonstram esse momento, exigindo do estudante a comparação da inclinação dos lados, a quantidade de lados e se ele, a partir da figura de um polígono dado, ele é capaz de identificar a nomenclatura desse polígono.

Ao transicionar para o 9EF, o D03_9EF foca no estudo dos triângulos. As propriedades e características dos triângulos, medidas dos lados e os ângulos passam a ser centrais. As tarefas dos itens passam a ter exigências nos níveis 1 (Análise) e 2 (Abstração) da Teoria de Van Hiele, exigindo compreensão de propriedades dos triângulos, bem como relações entre eles.

As tarefas associadas ao descritor D04_5EF consolidam a exigência do nível 1 de Van Hiele para o 5º ano, uma vez que demanda dele a necessidade de reconhecimento das figuras a partir das posições relativas de seus lados. Assim, para solucionar o item corretamente, o estudante precisa abandonar o julgamento puramente visual e focar nas propriedades geométricas e nas relações entre os lados. As tarefas dos itens demonstram a exigência dessas relações, embora todos apresentem apoio de malha quadriculada, além dos riscos de o estudante reconhecer o quadrilátero apenas pela sua forma. Nota-se que, dentre os itens do 5EF cedidos, o item que apresentou parâmetro B mais baixo foi nessa habilidade, e isso, provavelmente, deve-se ao fato de requerer a identificação do quadrado (devido à elementaridade).

No 9EF, o descritor D04_9EF foca nas propriedades dos quadriláteros. Embora apenas um exija observar as propriedades para identificar propriedades comuns e fazer a classificação dos quadriláteros, há de se observar as propriedades para que seja possível nomeá-lo. Novamente, a identificação de um quadrado se mostra demasiadamente elementar para os estudantes, sobretudo nessa etapa, uma vez que o parâmetro B de um dos itens desse descritor (M090192G5) se apresenta em um nível que, segundo Inep (2020), é um nível de defasagem.

O salto na exigência de conhecimentos geométricos do D3 e D4 entre o 5EF e 9EF, demonstrou que pode tornar o item mais discriminativo e que o tipo de tarefa influencia na dificuldade. Os itens M060023H6 e o M090203EX possuem parâmetros B idênticos, mas o primeiro é mais discriminativo que o segundo. Já o M090137H6 é menos discriminativo, mas apresentou o parâmetro dificuldade mais alto. Nesses 3 itens o estudante está trafegando entre os níveis 1 e 2 de Van Hiele.

Por fim, as tarefas atreladas ao descritor D05_5EF e D05_9EF, embora apresentem a mesma redação, precisam ser observadas com cuidado. Para o 5EF, não é um problema para o estudante analisar ampliações e reduções, verificando modificações nos lados e na área das figuras apresentadas na malha, mas a determinação do fator de ampliação ou redução se torna um tanto difícil para a etapa, haja vista que o item com parâmetro B mais elevado, tem exatamente essa tarefa. Mas, embora esse seja o caso, o outro item com tarefa análoga, apresenta-se com parâmetro B próximo a média da escala. Uma possibilidade é o tipo de figura comparada: em um, triângulos, no outro, uma figura composta que pode ser decomposta em retângulos.

Nos descritores do 9EF restantes, exige o domínio principalmente dos níveis 1 e 2 de Van Hiele. A exemplo do D05_9EF, a determinação do fator numérico de ampliação ou redução da área em malhas quadriculadas, obriga o estudante a quantificar e comparar dimensões de forma empírica, o que pressupõe o abandono somente de percepções visuais, demandando uma análise mais formal e abstrata na resolução. Notou-se também que há presença de itens nessas habilidades que performaram com alto poder discriminativo, devendo-se, provavelmente, estar relacionado ao tipo de tarefa e exigência de conhecimentos geométricos.

No D06_9EF, a determinação de posições a partir do ângulo de giro exige a análise da amplitude do movimento como uma grandeza mensurável. Esse descritor se mostra como uma evolução dos descritores D01_5EF e D01_9EF, uma vez que movimentos agora são descritos em termos mudança de direção e giros.

Embora o D09_9EF possua característica de buscar do estudante uma interpretação, os itens se mostram em um processo de identificação de pontos. É compreensível que o verbo interpretar, para itens de resposta selecionada, se apresente em um espaço nebuloso entre a identificação e a resolução de problemas, já que para se resolver um problema, há necessidade de identificar os elementos que o constituem, interpretar as relações entre eles para modelar o problema para em

seguida chegar a uma solução/resultado. Assim, embora o centro do que foi avaliado pelas tarefas dos itens esteja na identificação, a inserção de um contexto³³ em que o estudante precisa angariar as informações sobre os pontos a serem identificados, faz com que o item tenha um caráter interpretativo.

No descritor D11_9EF, é avaliado o reconhecimento de segmentos específicos de uma circunferência (como o raio, o diâmetro, a flecha...), e as relações entre elas caracterizando o nível de reconhecimento e análise (níveis 0 e 1 de Van Hiele). Por outro lado, os itens focam apenas nível 0 de Van Hiele. Desses itens, três pedem a identificação de raio e diâmetro.

Os descritores D07_9EF, D08_9EF e D10_9EF exigem o domínio do nível 2 de Van Hiele (Abstração). O D07_9EF requer a compreensão de transformações homotéticas, com a determinação de constantes de proporcionalidade e o cálculo da medida de lados transformados.

As tarefas do D08_9EF, solicitam a determinação de ângulos internos desconhecidos a partir de composições de polígonos regulares, obrigando o estudante a deduzir valores através de premissas e propriedades pré-estabelecidas (como a soma dos ângulos internos).

Por fim, no descritor D10_9EF, o estudante precisa resolver problemas envolvendo as relações métricas no triângulo retângulo, por meio de análise da situação e aplicar uma regra lógica e algébrica para encontrar uma medida desconhecida.

Quando se avança para a análise dos itens da 3ª série do Ensino Médio (3EM), observa-se o teto de complexidade da matriz, com tarefas que exigem predominantemente a mobilização dos níveis 2 (Abstração) e 3 (Dedução) de Van Hiele devido à introdução da formalização algébrica e trigonométrica.

No descritor D01_3EM, as tarefas são de identificar ou reconhecer triângulos semelhantes, que embora pareçam estar no nível 0 ou 1 de Van Hiele, demandam o domínio do nível 2 de Van Hiele. A resolução deixa de se apoiar em um aspecto puramente visual de semelhança e passa a requerer o estabelecimento de relações matemáticas estritas de proporcionalidade entre os lados. Do ponto de vista estatístico, a dificuldade varia consideravelmente entre os itens cedidos, destacando-se o item M120005D3 com um parâmetro B bastante elevado (463,1), o que pode ser

³³ Como é o caso dos itens M051810I7 e M052148I7.

justificado pelo fato de a verificação dessa proporcionalidade ser complexa devido à disposição das figuras planas dadas.

Para os descritores D02_3EM e D05_3EM, o triângulo retângulo passa a ser utilizado como um modelo geométrico, caracterizando exigências de nível 3 de Van Hiele (Dedução). No D02_3EM, as tarefas exigem a aplicação de relações métricas em situações espaciais ou planas, enquanto no D05_3EM, exige-se a determinação de medidas (como alturas e diagonais) utilizando as razões trigonométricas de seno e cosseno. Os itens desses descritores performam com alta complexidade, apresentando parâmetros B majoritariamente elevados, com destaque para o item M120354H6, evidenciando que a manipulação de projeções e ângulos de inclinação impõe uma carga cognitiva alta para a etapa. Novamente, esse alto índice de exigência remete a itens com alta discriminação e dificuldade mais elevada, para a maioria dos itens analisados desses descritores.

No bloco de Geometria espacial, os descritores D03_3EM e D04_3EM apresentam naturezas analíticas distintas. O D03_3EM avalia a correspondência entre figuras tridimensionais e suas planificações ou vistas superiores, situando-se em uma transição entre os níveis 1 e 2, onde se espera que o estudante mobilize relações entre faces e superfícies de figuras tridimensionais às suas correspondentes projeções. Por sua vez, o descritor D04_3EM exige diretamente o nível 3 de Van Hiele, uma vez que as tarefas requerem alto nível de abstração das características do sólido geométrico para cobrar a aplicação direta da Relação de Euler. Os itens desse descritor performam em um nível de dificuldade muito homogêneo em um ponto da escala mais avançado.

Na matriz, o bloco de Geometria Analítica, que engloba os descritores D06_3EM a D10_3EM, leva o estudante a um processo de análise informal e dedução formal, onde o espaço e a figura passam a ser descritos por meio de estruturas algébricas.

Embora o descritor D06_3EM faça parte desse grupo a identificação de pontos no plano cartesiano, segue os moldes do que o 9EF faz, mas, até certo ponto, mais simples, uma vez que o verbo identificar pressuponha uma complexidade menor do que o verbo interpretar. Os itens apresentam parâmetro B um pouco elevado, devido à complexidade, mas em um ponto adequado na escala.

No D07_3EM, o foco fica na interpretação geométrica dos coeficientes linear e angular da reta. As tarefas exigem que o estudante traduza o comportamento da reta

para a sua estrutura algébrica reduzida. As exigências desse nível refletem em parâmetros B elevados, como no item M120006D3 (421,8). Em direção semelhante, o D08_3EM requer a determinação da equação da reta a partir de pontos ou inclinações dadas, enquanto o D09_3EM exige relacionar o ponto de interseção de retas com a resolução de um sistema de equações. Nesses casos, o suporte geométrico atua como representação gráfica de uma operação estritamente analítica, de onde o estudante precisa angariar informações para indicar corretamente a resposta.

O descritor D10_3EM eleva a circunferência de uma representação à expressão de uma equação de 2º grau. Essa abstração exige do estudante o reconhecimento, dentre equações do 2º grau com duas incógnitas, aquela que representa uma circunferência, demandando uma análise formal da estrutura matemática da equação.

Nessa nossa análise, os itens permitem medir o que se deseja medir, que é a proficiência do estudante, como salientado por Rodrigues (2006). Além disso, no que cabe observar no conjunto de dados, de forma geral, os itens se apresentaram com níveis de dificuldade adequados às etapas, cabendo em algumas habilidades considerar uma abordagem diferente para melhorar esse alinhamento.

No processo de estudo dos itens e suas medidas, foi possível identificar que havia itens em que a abordagem era muito similar a outros itens de habilidades que demandariam um nível de conhecimento geométrico mais elevado. Isso pode ser resultado da necessidade de avaliar defasagens ou dificuldade para elaborar itens em níveis cognitivos variados.

As tarefas apresentadas nos itens são compatíveis com o tipo de processo que é esperado do estudante em cada etapa, mas carecem de tarefas que abordem níveis mais elementares da Teoria de Van Hiele em casos em que é esperado do estudante um nível de desenvolvimento mais avançado.

Não foi possível perceber uma relação explícita entre as medidas estatísticas e os níveis da Teoria de Van Hiele. Contudo, foi possível notar que o nível de exigência delineado no processo de análise da matriz reflete-se diretamente nos itens. Tarefas que exigem níveis cognitivos mais altos, como nos níveis mais altos da Teoria de Van Hiele, tendem a ser mais difíceis e mais capazes de separar os estudantes que desenvolveram uma habilidade dos que não desenvolveram. Isso foi observado em itens que exigem níveis mais altos de Van Hiele, indicando que essa relação reflete

em itens com maior poder discriminativo, indicado pelo parâmetro A. Com relação ao parâmetro da dificuldade do item (parâmetro B), essa correlação direta não pôde ser identificada, mas uma maior carga cognitiva já é indicativa de que os itens possam performar como mais difíceis. Para trazer alguma resposta nesse sentido, há a necessidade de se estudar um número maior de itens de cada habilidade.

Uma consequência dessas afirmações é que, se os níveis da Teoria de Van Hiele se relacionarem com a discriminação do item, então um item feito para avaliar conhecimentos no nível 0 pode fornecer menos informação sobre o que os estudantes de fato sabem? Isso, de fato, pode ser verdade, mas como as avaliações em larga escala são instrumentos para avaliar o conhecimento de milhares de estudantes, com diferentes quantidades de habilidades consolidadas, faz-se necessário o uso de itens com abordagens diferentes e que avaliem habilidades cognitivas diferentes.

Com isso, a produção de itens que permitam obter a melhor informação do que o estudante sabe é necessária e, nesse sentido, tarefas que avaliam diferentes fases no processo de aprendizagem vão conferir ao instrumento de avaliação, como um todo, maior capacidade de obter essas informações.

Nesse ponto, é crucial esclarecer que a Teoria de Van Hiele não é uma taxonomia³⁴ como indica o significado da palavra, mas uma estrutura, como visto na seção 3.1.4, que permite identificar onde o estudante está na consolidação do conhecimento geométrico e, com isso, ser capaz de entender que tipo de conhecimento deve ser cobrado dele para que ele consiga avançar. No contexto de avaliações em larga escala, uma contribuição que essa Teoria fornece é uma lente que permite compreender onde a habilidade descrita na matriz está posicionada no percurso cognitivo e, a partir disso, identificar o que se deve esperar desses estudantes na elaboração do item.

A Teoria de Van Hiele pode permitir que, ainda no processo de elaboração, se produzam itens que exijam conhecimentos em diferentes níveis, o que é particularmente interessante quando se quer avaliar etapas intermediárias (como 6EF, 7EF e 8EF). As nuances e compreensão do processo a partir desses níveis trazem aportes pedagógicos e predição de fatores associados a dificuldades de

³⁴ Segundo o Dicionário Online Aulete, taxonomia é: Ciência da classificação. (<https://www.aulete.com.br/taxonomia>)

aprendizagem dos estudantes, que podem melhorar a compreensão dos resultados estatísticos das avaliações.

Por fim, nota-se que a classificação de descritores e de itens nos níveis dessa Teoria esteve, em geral, dentro do que era o objetivo do descritor esperar do estudante. A Teoria de Van Hiele não foi produzida pensando em ser aproximada do processo de avaliação educacional em larga escala, logo, é esperado que a teoria não se ajuste completamente a esse novo ambiente, uma vez que a tomada de decisões para o ensino em sala de aula e a tomada de decisões em avaliações em larga escala ocorrem de forma diferente. Uma das propriedades importantes da Teoria de Van Hiele é ela ser sequencial e isso se apresenta como vantagem para a produção de itens que visam avaliar as habilidades dos estudantes em um continuum.

Aqui, foi feita uma aproximação dos itens a essa Teoria e que ela se apresenta como possibilidade de aumentar o leque de opções para a construção de itens, já que, se uma habilidade se apresenta a uma etapa de escolaridade para ser avaliada até o nível 2, permite que sejam pensados itens em níveis mais elementares, que por sua vez, podem ser utilizados para avaliar estudantes em etapas em que um nível mais elementar seja o foco da avaliação ou avaliar estudantes em defasagem.

Apoiada nessas considerações, a implementação da abordagem desenvolvida nessa seção demonstra potencial de aumentar a robustez do banco de itens que possui itens desenvolvidos dessa forma. Com objetivo de partilhar com a Equipe de Instrumentos de Avaliação os achados desta pesquisa, foi construído o plano de ação apresentado no próximo capítulo.

4 PLANO DE AÇÃO: A ELABORAÇÃO DE ITENS BASEADOS NOS NÍVEIS DE VAN HIELE

Durante o desenvolvimento deste trabalho, construiu-se uma argumentação em torno da qualidade dos itens e da importância de que estes variem pedagogicamente em dificuldade e tipo de tarefa, permitindo que estudantes com proficiências em diferentes pontos da escala consigam demonstrar seus conhecimentos. Uma analogia válida é a de que, para um instrumento medir com precisão, ele precisa estar adequadamente graduado. Uma balança para medir a massa de uma pessoa, por exemplo, não traz informações precisas se for utilizada para medir ingredientes em uma cozinha; da mesma forma, uma balança de precisão culinária não consegue mensurar a massa de um adulto, limitando-se a indicar que o peso excede a sua capacidade máxima. A informação não deixa de existir, mas torna-se inconclusiva ou insuficiente.

É nesse sentido que Palácios e Oliveira (2022) sugerem que os itens de um teste se distribuam entre muito fáceis, fáceis, medianos, difíceis e muito difíceis. Dessa forma, o instrumento obtém informações precisas para estudantes alocados em diferentes estratos da escala.

Assim, este trabalho reuniu os referenciais necessários para compreender os elementos constituintes dos itens, a produção e a montagem do teste, além dos pressupostos pedagógicos para a sua análise. Apresentaram-se, também, as metodologias utilizadas para a produção das medidas psicométricas, elucidando como a proficiência dos estudantes é calculada e posicionada na escala.

Além disso, utilizou-se um conjunto de itens como dados empíricos de avaliação dos conhecimentos geométricos dos estudantes. A partir deles, observou-se como esses saberes se apresentaram com relação à expectativa (a matriz da avaliação) e à medida de proficiência de estudantes nos itens corrigidos utilizando a metodologia da TRI (em três parâmetros lineares). Nessa análise, buscou-se identificar se os níveis da Teoria de Van Hiele poderiam qualificar as demandas cognitivas das tarefas dos itens e como, a partir de dados produzidos no passado, essas demandas possivelmente se manifestaram.

Na seção 3.3, percebeu-se que o comportamento estatístico dos itens se relaciona com a exigência cognitiva das tarefas e que o nível associado na Teoria de Van Hiele pode ajudar a descrever esse comportamento. Com isso, a observação dos

níveis dessa teoria oferece uma possibilidade de enriquecer o banco de itens da instituição.

No processo de estudo dos itens e suas medidas, foi possível perceber as necessidades institucionais primárias para utilização da metodologia abordada neste trabalho: a padronização formativa e a necessidade de expansão investigativa. Novamente, como indicado por Souza (2020), a lacuna formativa é apontada como fator impeditivo para um banco de itens mais robusto e para a produção de testes mais consistentes. Em seu trabalho, Souza (2020) apresentou fatores que levavam à eliminação de itens e, aqui, buscou-se compreender, a partir das medidas estatísticas e dos itens, como a diversidade de níveis de Van Hiele é empregada nos itens e quais convergências se observam.

Para intervir nessas lacunas, o Quadro 1 sistematiza os achados gerais e os organiza em dois eixos de atuação do Plano de Ação Institucional: o eixo de Formação (focado no desenvolvimento de materiais instrumentais) e o eixo de Pesquisa (focado na escalabilidade do método). O segundo eixo passa a ser indissociável do primeiro, uma vez que as avaliações são produzidas utilizando mais de um eixo do conhecimento, enquanto este trabalho focou no eixo do conhecimento de Geometria.

Quadro 1 – Sistematização dos Achados e Propostas de Ação Institucional

Eixos de Atuação	Achados Gerais	Proposta de Ação
Eixo 1: Formação e Instrumentalização	A análise dos itens evidenciou que há relação entre a variação dos parâmetros e os Níveis de Van Hiele, o que pode ser explicado pelos saltos de exigência cognitiva (como a transição do raciocínio visual para o dedutivo formal). Constatou-se que, em alguns descritores, há ausência de construção de itens (no rol de itens cedidos) com exigências cognitivas variadas.	Material Técnico-Formativo e ciclos de discussões: Desenvolvimento e entrega de um Guia de Elaboração Suplementar à equipe de Matemática. O material servirá para instrumentalizar os elaboradores com as análises apresentadas neste trabalho, visando fundamentar a construção de itens apoiados nos Níveis de Van Hiele e permitir a avaliação do estudante por meio de tarefas com diferentes exigências cognitivas. Além disso, propõem-se ciclos de discussões para apropriação e inclusão desse material no fazer diário da equipe. Um checklist será produzido para diagnosticar e recuperar itens que não se enquadram no nível esperado, visando reduzir a subjetividade na definição da adequação do item às competências cognitivas do estudante e a taxa de retrabalho técnico geradas por essa inadequação.
Eixo 2: Expansão Metodológica	O cruzamento entre pressupostos psicométricos (TRI) e pressupostos pedagógicos cognitivos (Teoria de Van Hiele) demonstrou ser eficaz para auditar e avaliar os itens de avaliações em larga	Expansão Metodológica: Aplicar a metodologia de análise apresentada neste trabalho aos demais eixos do conhecimento da matriz. A ação visa a instituição de grupos de trabalho na equipe para mapear as habilidades de

	escala. No entanto, essa abordagem restringiu-se apenas ao eixo de Geometria, deixando os demais eixos da matriz sem um referencial cognitivo de validação padronizado. Essa assimetria evidencia a necessidade de institucionalizar um processo de expansão para os outros eixos do conhecimento.	Álgebra e Números, utilizando referenciais teóricos específicos de cada área da Educação Matemática (como a Teoria dos Campos Conceituais de Vergnaud). Essa ação visa padronizar a análise pedagógica na produção de todos os blocos do teste, culminando na futura criação de novos materiais de apoio.
--	---	--

Fonte: Elaborado pelo autor (2026)

A partir desses eixos e com auxílio da ferramenta 5W2H (What, Why, Who, Where, When, How, How much), a próxima seção será dedicada a apresentar a operacionalização da execução dessas ações.

4.1 OPERACIONALIZAÇÃO DO PLANO DE AÇÃO

Para garantir que as ações apontadas no Quadro 1 saiam do escopo teórico e sejam efetivamente incorporadas ao fluxo produtivo da Fundação CAEd, estruturou-se a viabilidade do plano por meio da ferramenta de gestão 5W2H. Com essa metodologia é possível mapear o plano de ação, especificando as responsabilidades, os prazos, os métodos e o impacto institucional de cada eixo.

4.1.1 Eixo 1: Material Técnico-Formativo e ciclos de discussões

Esse eixo visa à produção de um guia para servir de apoio para formação inicial de novos colaboradores para a equipe de instrumentos de avaliação, da subequipe de Matemática, de forma a permitir um aprofundamento nesse momento de formação sobre os itens e as avaliações em larga escala. Souza (2020) afirmou que esse processo de formação ocorre em ambiente laboral e, para que isso ocorra de forma proveitosa e demandando menor tempo de formação, é importante que se tenha apoio nos Guias de elaboração, mas também em uma referência no contexto pedagógico, permitindo alinhamento entre os colaboradores.

O Quadro 2 apresenta o detalhamento do plano de ação no eixo 1, mapeando os processos pelos quais essa ação precisa passar para sua implantação.

Quadro 2 – Eixo 1: Formação e Instrumentalização

Diretriz (5W2H)	Detalhamento da Execução
O que será feito? (What)	Produção de um guia para suportar a discussão em torno da aplicação dos níveis de Van Hiele na produção de itens e realização de um ciclo de reuniões para discutir esse material.
Por que será feito? (Why)	Necessidade de suprir a carência de itens com exigências cognitivas variadas no banco e padronizar o julgamento pedagógico da equipe, reduzindo a subjetividade.

Quem será o responsável? (Who)	O autor desta pesquisa (como Analista de Instrumentos de Avaliação), sob a validação da coordenação da subequipe de Matemática da Fundação CAEd.
Onde será implementado? (Where)	Na Equipe de Instrumentos de avaliação, subequipe de Matemática.
Quando será executado? (When)	A partir do início do terceiro trimestre de 2026, após a aprovação da supervisão de instrumentos de avaliação. O guia complementar e o uso do checklist ocorrerão no primeiro mês de execução que, para se encaixar às demandas da subequipe, podem ocorrer mensalmente.
Como será conduzido? (How)	1. Consolidação das análises empíricas desta dissertação em um manual prático (com exemplos de itens de níveis 0 a 3 de Van Hiele). 2. Apresentação da ferramenta de elaboração à supervisão da equipe. 3. Realização de ciclos de discussões mensais com os analistas, aplicando o guia e o checklist como método de diagnóstico de itens diretamente no processo de produção.
Quanto custará? (How much)	Não haverá custos adicionais para a instituição, uma vez que a proposta pode ser acomodada nos processos formativos da subequipe.

Fonte: Elaborado pelo autor (2026).

Nesta ação, a proposta é que os elaboradores identifiquem que, embora seja possível determinar se a exigência cognitiva de um item está alinhada com o desenvolvimento do estudante a partir da progressão curricular, a utilização dos níveis de Van Hiele, como evidência empírica desse desenvolvimento, facilita notar em que ponto ele se encontra.

Além disso, essa teoria já apresenta níveis definidos e tem, segundo Crowley (1987), a propriedade de ser sequencial. Ao relacionar esses níveis e propriedades com os objetivos educacionais das matrizes e currículos, torna-se possível estruturar a complexidade que o item terá.

Esse Guia vai abordar temas que são de conhecimento dos Analistas que compõem a equipe atualmente, pelo fato de que o mesmo material pode ser utilizado no processo de formação de novos Analistas. Nele, o objetivo é que a vivência em sala de aula seja um fator contribuinte ao processo de avaliação, mas que essa subjetividade seja apoiada em estudos consolidados, resultando em itens mais consistentes. Problemas estruturais em itens que foram pensados para avaliar um

estudante de nível de conhecimento geométrico mais elementar, mas que avaliam outros níveis, são o que se deseja evitar com este Guia.

Se a intenção é avaliar um estudante de 2º ano do Ensino Fundamental, o Guia mostrará que faz total sentido inserir um triângulo como distrator ao pedir a identificação de um quadrado, pois nesse estágio (Nível 0 - Visualização), a criança julga a figura apenas pelo seu aspecto global. Por outro lado, o Guia baliza o limite da cobrança para o 5º ano, evitando que a avaliação exija processos que fujam ao escopo do nível de maturidade da etapa.

A organização desse Guia, descrita em uma proposta de sumário, é apresentada no Quadro 3. Distribuído em capítulos, esse Guia fará a apresentação da proposta, traçando os objetivos esperados com essa metodologia. Na sequência, o material apresentará cada um dos níveis do conhecimento geométrico da Teoria de Van Hiele, com exemplos práticos, apoiados pela bibliografia utilizada neste trabalho. Posteriormente, esse guia mostrará as relações que os níveis de Van Hiele podem ter com as medidas estatísticas obtidas com os itens. Por fim, os últimos dois capítulos são focados na prática com os itens, sendo que o primeiro deles tratará de um checklist para analisar itens em produção e, o outro, do processo de visita ao banco para correções de problemas existentes.

Quadro 3 - Proposta de Sumário para o Guia Suplementar

SUMÁRIO
1. APRESENTAÇÃO
2. A LENTE CRIADA PELA TEORIA DE VAN HIELE
3. OS NÍVEIS DE PENSAMENTO GEOMÉTRICO
3.1. Nível 0 (Visualização)
3.2. Nível 1 (Análise)
3.3. Nível 2 (Abstração)
3.4. Nível 3 (Dedução Formal)
4. O COMPORTAMENTO EMPÍRICO
5. CHECKLIST TÉCNICO-PEDAGÓGICO
6. ESTUDOS DE CASO

Fonte: Elaborado pelo autor (2026)

Essa ação é particularmente efetiva, já que mune o elaborador com um processo pedagógico pesquisado diretamente com os estudantes, mapeando os processos cognitivos de aprendizagem. As avaliações da proficiência escolar pretendem avaliar a aprendizagem, e o conhecimento dos níveis de Van Hiele torna a

identificação da adequação de uma tarefa a um estudante com um determinado desenvolvimento de conhecimento geométrico mais rápida e alinhada às expectativas.

4.1.2 Eixo 2: Expansão metodológica

Nesse eixo, o objetivo é permitir que a abordagem que foi construída para a Geometria reflita numa agenda de (re)padronização do processo de construção de itens. Uma vez que vantagens estratégicas são demonstradas no eixo de Geometria, a expansão para outros eixos é necessária.

O eixo apresentado no Quadro 4, foi estruturado para que, ao se adotar a metodologia apresentada por este trabalho, seja possível expandir os procedimentos para os outros eixos da matriz de avaliação do Saeb.

Quadro 4 – Eixo 2: Pesquisa e Expansão Metodológica

Diretriz (5W2H)	Detalhamento da Execução
O que será feito? (What)	Estruturar uma agenda contínua de pesquisa aplicada, focada na análise pedagógico-cognitiva dos demais eixos da matriz.
Por que será feito? (Why)	Para garantir a escalabilidade da metodologia, é necessário que não se limite ao eixo de Geometria; todo o instrumento de avaliação em larga escala deve operar sob referenciais consistentes de aprendizagem.
Quem será o responsável? (Who)	O autor desta pesquisa (como Analista de Instrumentos de Avaliação), junto à Equipe de Instrumentos de avaliação - Matemática com a cooperação da Equipe de Medidas Educacionais da Fundação CAEd.
Onde será implementado? (Where)	Na Equipe de Instrumentos de avaliação, subequipe de Matemática.
Quando será executado? (When)	Operacionalização a partir de fevereiro de 2027, ou após a conclusão das etapas do eixo I. As ações propostas dependem de alinhamento entre diferentes equipes internas da Fundação CAEd, o que pode atrasar o início dessa ação.
Como será conduzido? (How)	1. Criação de grupos de estudos internos para definir os referenciais teóricos adequados a cada eixo (ex: Teoria dos Campos Conceituais de Vergnaud para Álgebra). 2. Replicação do delineamento metodológico desta pesquisa (análise documental cruzada com TRI). 3. Produção progressiva de novos guias suplementares abrangendo toda a matriz de matemática.
Quanto custará? (How much)	Os custos para a instituição serão mínimos, já que a proposta pode ser acomodada nos processos formativos da subequipe. Por envolver mais de uma equipe e demandas de pesquisa, há a possibilidade de redução de produtividade nas atividades

	semanais, que pode ser compensada pelo retorno do investimento.
--	---

Fonte: Elaborado pelo autor (2026).

Especificamente para essa ação, a matriz Saeb adotada pode ser a alinhada com a BNCC, uma vez que, por ser uma ação de longo prazo, já será possível fazer a mesma análise proposta neste trabalho a essa iteração da matriz. Além disso, como a nova matriz inclui as diretrizes para o 2º ano do Ensino Fundamental, o escopo dessa ação pode ser ampliado para abranger essa etapa.

Assim, com essas duas ações, espera-se contribuir com o processo formativo dos colaboradores que atualmente compõem a equipe de instrumentos de avaliação – Matemática, bem como daqueles que venham a integrá-la futuramente. Esses processos desenvolvimento profissional são fundamentais para manter o alinhamento entre as avaliações e as novas demandas educacionais que se apresentam.

4.2 PROTOCOLO PARA PRODUÇÃO DE ITENS

Na seção 3.3, identificou-se que, para cada habilidade, considerando a expectativa de aprendizagem do estudante em seu percurso escolar, é possível determinar o tipo de tarefa e como ela se apresenta nos níveis da Teoria de Van Hiele. Dessa forma, essa seção se desenvolve no sentido de apresentar elementos para um protocolo para a criação de itens, levando em conta os níveis dessa teoria. O conteúdo dessa seção será estruturado no guia desenvolvido para o Eixo I³⁵.

Esse protocolo é composto por dois pontos: 1. Construção de famílias de itens; e 2. Existência de apoios. Com esses elementos, julga-se que o processo de elaboração fica munido de critérios que permitem a construção de itens que se enquadram em diferentes níveis de Van Hiele.

O primeiro ponto desse protocolo, denominado “Construção de famílias de itens”, foi inserido para que se produzam itens em diferentes níveis, quando possível. Espera-se, por exemplo, que o elaborador tome um descritor e identifique o nível de Van Hiele atingido por um estudante que desenvolveu as habilidades e os pré-requisitos para aquele descritor. O elaborador produz um item, com uma tarefa que seja adequada a esse nível, o qual será considerado o item-alvo desse descritor. A

³⁵ Discutido na seção 4.1.1.

partir desse item, o elaborador irá produzir outros itens, podendo ser um em nível superior e outro em nível imediatamente inferior.

O segundo ponto, denominado “Existência de apoios”, pode ser visto tanto como um elemento a dar suporte ao primeiro ponto, quanto como um elemento a ser observado para facilitar a determinação do nível de Van Hiele associado. Para esse segundo ponto, o “apoio” é um componente do item que contribui para que o estudante consiga identificar propriedades e características dos objetos estudados. Na produção de um item, incluir um apoio ou não produz por vezes nesse item uma variação no processo de resolução.

Pode-se, por exemplo, em habilidades de identificar propriedades de um polígono ou relações entre polígonos, incluir uma malha quadriculada para contribuir com o processo de resolução, já que ela permite identificar características explicitadas pela estrutura rígida da malha. A explicitação de medidas, mesmo aquelas que podem ser inferidas a partir das propriedades de um determinado polígono, também contribui para facilitar a execução da tarefa pelo estudante. A exibição de polígonos, poliedros ou outros objetos geométricos em sua apresentação padrão é um facilitador, uma vez que é esperado que o estudante que está no nível 0 da Teoria de Van Hiele consiga identificar esses objetos.

É possível sumarizar esses processos na Tabela 5.

Tabela 5: Síntese do protocolo para produção de itens

	Ação	Objetivo	Implicações nos Níveis de Van Hiele
Família de itens	Produzir item de uma habilidade e elaborar outros itens com objetivo de construir uma tarefa em níveis abaixo ou acima do nível esperado.	Produzir itens que permitam avaliar o desenvolvimento de conhecimentos geométricos de diferentes níveis de uma habilidade ou descritor.	Intencionalidade na transição entre os níveis.
Existência de apoios	Inserção ou retirada de elementos. Uso de malhas quadriculadas,	Ajustar a complexidade da tarefa, de forma a induzir a identificação	Atingir os níveis esperados para

	figuras rotacionadas, explicitação de medidas etc.	das propriedades de diferentes maneiras.	uma habilidade, de forma efetiva.
--	--	--	-----------------------------------

Fonte: Elaborado pelo autor (2026).

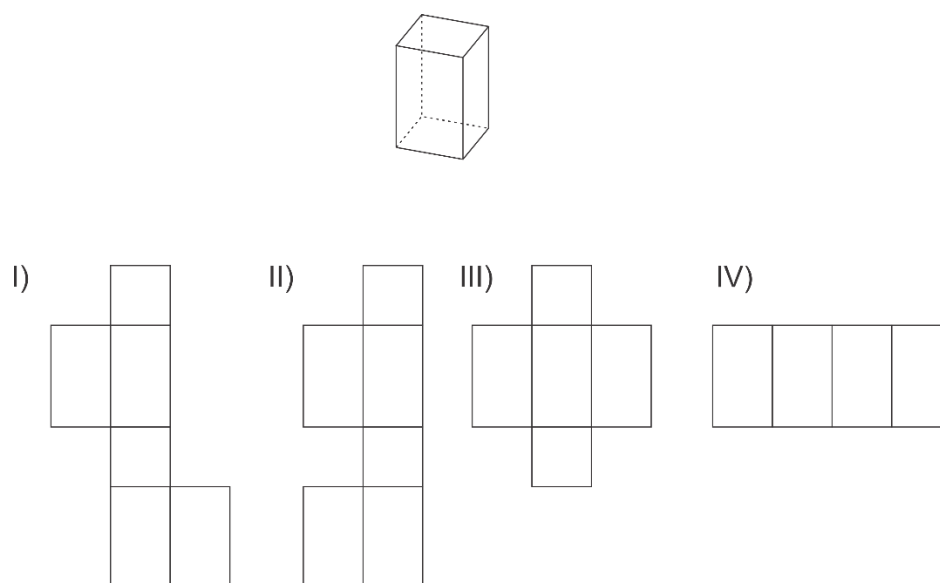
O descritor D02_9EF (Identificar propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando-as com as suas planificações) pode exemplificar os processos para seguir nesse protocolo.

A estrutura dos itens para essa habilidade tem pouco texto e foca na observação e identificação das propriedades das faces que compõem a figura tridimensional e as relações com os polígonos que as descrevem. No caso dos corpos redondos, o objetivo é semelhante, adicionadas as especificidades relativas às suas superfícies laterais curvas e bases circulares.

A habilidade D02_9EF encontra relação direta com a D02_5EF (Identificar propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações) e D03_3EM (Relacionar diferentes poliedros ou corpos redondos com suas planificações ou vistas). Esse conjunto de descritores se mostra perfeito para demonstrar as vantagens de produzir itens com dificuldades cognitivas diferentes, buscando atingir determinado nível de Van Hiele.

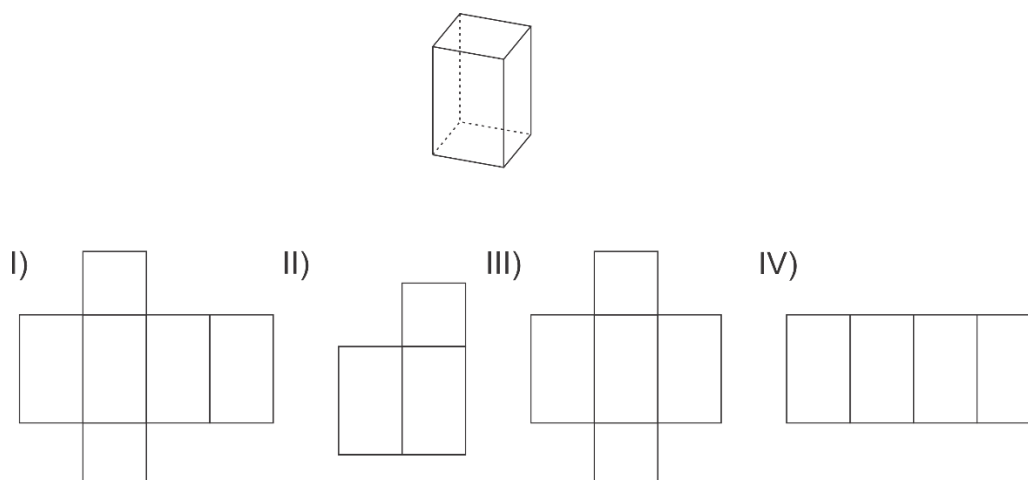
Para o 9EF, é esperado haver elementos que exijam que o estudante identifique se há garantia de que os polígonos que representam as faces estão conectados de forma correta às outras faces correspondentes ou se estão em posição trocada. Na Figura 6 II), duas das faces se sobrepõem se essa organização for utilizada como a planificação dessa figura geométrica tridimensional. Com essa construção, é esperado que esse item esteja avaliando estudantes que já consolidaram habilidades de nível 1 na Teoria de Van Hiele. Um item em nível 0, mas que já exige alguma habilidade de análise por parte do estudante, pode utilizar a estrutura apresentada na Figura 7. Note que, com essa organização, o estudante pode resolver o item apenas pela identificação da quantidade de faces do poliedro apresentado. Se o objetivo é produzir um item em que ele precise fazer apenas o reconhecimento do formato das faces, pode-se incluir corpos redondos ou planificações de outros prismas, facilitando a identificação do formato das bases.

Figura 6 – Possíveis figuras para um item de planificação para o 9EF



Elaborado pelo autor (2026).

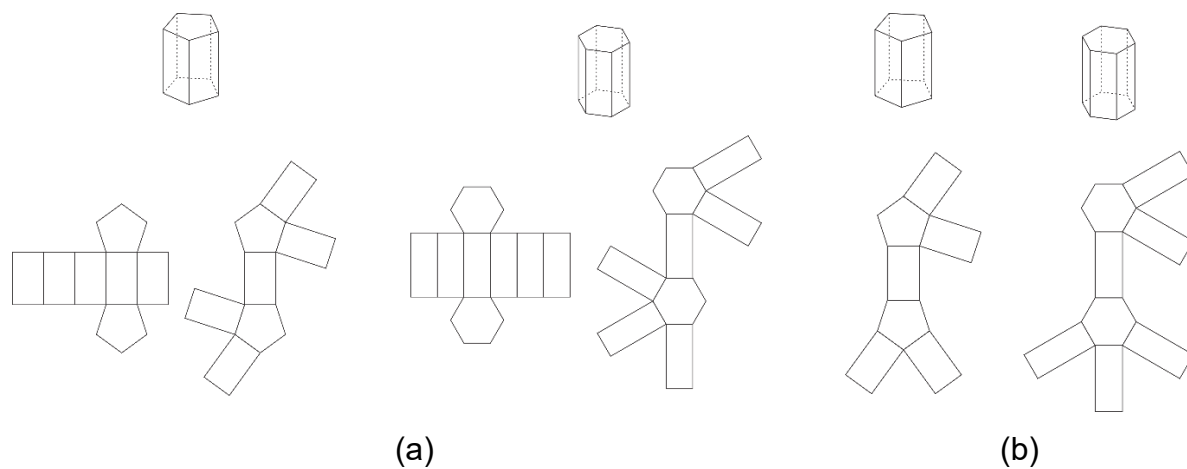
Figura 7 – Possíveis figuras para um item de planificação para o 5EF



Elaborado pelo autor (2026).

Já para produzir itens em um nível acima, isto é, nível 2 de Van Hiele, o uso de estruturas fora do padrão convencional e de sobreposições é recomendado. O processo de análise precisa ser mais minucioso e a exigência cognitiva é típica do que é esperado no 3EM. A Figura 8 apresenta uma maneira de produzir itens seguindo o mesmo processo anterior, utilizando planificações usuais, em que as faces laterais ficam justapostas; menos usuais, em que as faces laterais se conectam às bases e duas possibilidades em que se propõe uma sobreposição.

Figura 8 – Prismas com planificação usual, menos usual (a) e com sobreposição (b)



Elaborado pelo autor (2026).

Dessa forma, com o uso desse protocolo, é esperado que o processo de elaboração de itens ganhe um elemento adicional, a fim de contribuir com a identificação de fatores que permitam determinar o nível de exigência cognitiva esperada e compatibilizar essa exigência ao item produzido.

5 CONSIDERAÇÕES FINAIS

Este trabalho investigou a relação entre a dimensão pedagógica e o comportamento estatístico na elaboração de itens de Matemática para avaliações em larga escala, com foco em Geometria. A principal intenção com este trabalho foi buscar uma ferramenta que permitisse reduzir a subjetividade na estimativa das exigências cognitivas dos itens, permitindo, com isso, diversificar abordagens na produção dos itens. Para isso, a Teoria de Van Hiele foi adotada como suporte para orientar o processo de construção e revisão dos itens.

Os objetivos desta pesquisa foram alcançados ao mostrar como os níveis de pensamento geométrico se manifestam nas tarefas e podem afetar os resultados empíricos. A análise da Matriz de Referência do Saeb (2001) e da amostra de 104 itens revelou que o aumento da dificuldade estatística, medida pela Teoria da Resposta ao Item (TRI), acompanha os saltos cognitivos exigidos pela Teoria de Van Hiele. Isso demonstra que a dificuldade de um item não é aleatória; ela reflete a exigência de passagem do reconhecimento visual (Nível 0) para a análise de propriedades (Nível 1) e, posteriormente, para a dedução (Níveis 2 e 3).

Ao observar as diferentes etapas de ensino, notou-se que no 5º ano as tarefas se concentram principalmente na percepção espacial, enquanto no 9º ano, ocorre uma transição que exige do estudante o foco nas propriedades das figuras, sendo o momento em que os itens passam a cobrar deduções informais. Já na 3ª série do Ensino Médio, a Geometria se alia à álgebra, exigindo dedução formal e resultando nos itens mais difíceis da escala.

Além disso, o indicativo de relação entre os níveis da Teoria de Van Hiele e a dificuldade empírica do item, bem como a forma como ele diferencia os estudantes que consolidaram as habilidades tem potencial para contribuir na construção de itens que identifiquem, de forma assertiva, o desenvolvimento de competências do eixo do conhecimento de Geometria.

Esses achados indicam que o erro do estudante muitas vezes ocorre porque ele está estagnado em um nível de raciocínio inferior, o que justifica a necessidade de construir itens com essa intencionalidade diagnóstica.

A partir desses resultados, propôs-se um Plano de Ação Institucional focado na formação e na instrumentalização da Subequipe de Matemática da Equipe de Instrumentos de Avaliação da Fundação CAEd. A execução das ações propostas se

apresenta como uma solução viável e de fácil implantação, dado que o processo formativo está intrínseco às atividades da equipe. As soluções são de baixo custo permitindo a padronização do trabalho dos elaboradores e a melhoria da validade dos testes produzidos pela instituição.

Por fim, devido às limitações técnicas do estudo que tem foco exclusivo no eixo de Geometria e nos parâmetros da matriz do Saeb de 2001, é necessário que pesquisas futuras expandam essa metodologia de análise conjunta (pedagógica e psicométrica) para os outros eixos do conhecimento matemático, utilizando aportes teóricos específicos desses eixos. Além disso, quando houver dados relacionados às novas matrizes alinhadas à Base Nacional Comum Curricular (BNCC), a expansão desse estudo permitirá a garantia de que a instituição continue a aprimorar a qualidade de todo o seu banco de itens de forma padronizada e fundamentada.

REFERÊNCIAS

ALAVARSE, O. M.; BRAVO, M. H.; MACHADO, C.. Avaliação como diretriz das Políticas Educacionais dos Governos Federal, Estaduais e Municipais: o caso brasileiro. **III Congresso Ibero Americano de Política e Administração da Educação**. Zaragoza, Espanha. 2012.

ANDERSON, L. W.; KRATHWOHL, D. R. (ed.). **Taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy of educational objectives**. Nova York: Longman, 2001.

ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. da C. Teoria da Resposta ao Item: conceitos e aplicações. **14º SINAPE, Associação Brasileira de Estatística**. 2000.

BAUER, A. “Novas” relações entre currículo e avaliação? Recolocando e redirecionando o debate. **Educação em Revista**, v. 36, p. e223884, 2020. DOI: 10.1590/0102-4698223884. Disponível em: <https://www.scielo.br/j/edur/a/LPNd9sV7qbfkGBjv8zcSvWK>. Acesso em: 17 jun. 2025

BEKMAN, R. M. Aplicação dos blocos incompletos balanceados na Teoria de Resposta ao Item. **Estudos em Avaliação Educacional**, São Paulo, n. 24, p. 119–136, 2001. DOI: 10.18222/eae02420012202. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2202>. Acesso em: 15 nov. 2025.

BORGATTO, A. F.; ANDRADE, D. F. de. Análise clássica de testes com diferentes graus de dificuldade. **Estudos em Avaliação Educacional**, São Paulo, v. 23, n. 52, p. 146–156, 2012. DOI: 10.18222/eae235220121934. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/1934>. Acesso em: 15 nov. 2025.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Relatório Nacional Saeb 2003**. 266 p. Brasília: Inep, 2006.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Guia de Elaboração e Revisão de Itens**. Brasília, DF: Inep, 2010.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Guia de Elaboração e Revisão de Itens – Provinha Brasil**. Brasília, DF: Inep, 2012.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Escalas de proficiência do SAEB. Brasília, DF: INEP, 2020.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Matrizes de referência de matemática do Saeb – BNCC**. Brasília, 2022.

BROOKE, N.; CUNHA, M. A. de A. A avaliação externa como instrumento de gestão educacional nos estados. **Estudos & Pesquisas Educacionais**, Fundação Victor Civita, São Paulo, n. 2, nov. 2011.

BURGER, W. F.; SHAUGHNESSY, J. M. Characterizing the van Hiele levels of development in geometry. **Journal for Research in Mathematics Education**, 17(1),

31–48, 1986. DOI: 10.2307/749317. Disponível em: <https://www.jstor.org/stable/749317>. Acesso em: 12 mai. 2025.

CAEd. **Guia de Elaboração de Itens**. Juiz de Fora: CAEd/UFJF, 2008.

CAPES. INFOCAPES. Boletim Informativo, Brasília, v. 5, n. 4, p. 1-78, out./dez. 1997. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/1997INFOCAPESn41997.pdf>. Acesso em: 1 jun. 2025.

CONCEIÇÃO, J. A. **Complexidade textual de itens de avaliação em larga escala: proposta de um mecanismo analítico**. 130f. Tese (Doutorado em Estudos da Linguagem) - Centro de Ciências Humanas, Letras e Artes, Universidade Federal do Rio Grande do Norte, Natal, 2019.

CROWLEY, M. L. **The van Hiele model of the development of geometric thought**. In M. M. Lindquist; A. Schulte (Eds.), *Learning and teaching geometry, K-12* (p. 1-16). National Council of Teachers of Mathematics. (1987).

GAVRON, E. M. **Google Scholar Metrics e OpenAlex: construção de indicadores para medir impacto e visibilidade de periódicos**. 2025. Tese (doutorado em Ciência da Informação) - Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Programa Pós-Graduação em Ciência da Informação, Florianópolis, 2025.

GIL, A. C. *Como elaborar projetos de pesquisa*. 6. ed. São Paulo: Atlas, 2017.

HALADYNA, T. M. **Developing and validating multiple-choice test items**. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2004.

VAN. HIELE, P. M. **Structure and Insight: A Theory of Mathematics Education**. Orlando: Academic Press; 1986.

HOGAN, T. P. **Introdução à Prática de Testes Psicológicos**. Tradução de Luís Antônio Fajardo Pontes. Rio de Janeiro: LTC, 2006.

JALOTO, A.; PRIMI, R. Enem de próxima geração com menos itens e alta confiabilidade usando CAT. **Estudos em Avaliação Educacional**, São Paulo, v. 35, p. e10142, 2024. DOI: 10.18222/eae.v35.10142. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/10142>. Acesso em: 11 dez. 2025.

KALEFF, A. M. M. R.; HENRIQUES, A. de S; REI, D. M.; FIGUEIREDO, L. G. **Desenvolvimento do Pensamento Geométrico – O Modelo de Van Hiele**. *Bolema: Boletim de Educação Matemática*, v. 9, n. 10, p. 21–30, 1994. Disponível em: <https://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema/article/view/10671>. Acesso em 5 maio 2025.

KLEIN, R. Utilização da teoria da resposta ao item no Sistema Nacional de Avaliação da Educação Básica (SAEB). **Ensaio: aval. pol. públ. educ.**, Rio de Janeiro, v.11, n. 40, p. 283-296, jul. 2003. Disponível em http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0104-40362003000300003. Acesso em 5 maio 2025.

KLEIN, R. Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB). **Revista Meta: Avaliação**, [S.l.], v. 1, n. 2, p. 125-140,

sep. 2009. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/38/17>. Acesso em: 05 jun. 2026.

KLEIN, R. Alguns aspectos da teoria de resposta ao item relativos à estimação das proficiências. **Ensaio: avaliação e políticas públicas em educação**, v. 21, n. 78, p. 35-56, 2013. Disponível em: <https://www.scielo.br/j/ensaio/a/FxGm4KDdQ56hF8JMbwxjzJJ>

KLEIN, R. Uma solução para a divergência de diferentes padrões no SAEB. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 27, n. 103, p. 229–249, abr. 2019. Disponível em: <https://www.scielo.br/j/ensaio/a/HfJtDVXJxR5FYSLzGFGFMbJ>

LIU, J.; ZHOU, L.; BHARAJ, P. K.; ZHOU, D.; LO, J.-J. Using international large-scale assessment for learning: Analyzing U.S. students' geometry performance in TIMSS. *JRAMathEdu* (Journal of Research and Advances in Mathematics Education), [S. l.], v. 9, n. 3, 2024. Disponível em: <https://journals2.ums.ac.id/index.php/jramathedu/article/view/4019>. Acesso em: 25 maio. 2025.

Page M. J. *et al.* A declaração PRISMA 2020: diretriz atualizada para relatar revisões sistemáticas. **Revista Panamericana de Salud Publica**. n. 46, 2022. DOI: 10.26633/RPSP.2022.112. Disponível em: <https://www.scielosp.org/j/rpsp/i/2022.v46/> Acesso em: 15 dez. 2025.

PALÁCIOS, C.; OLIVEIRA, L. K. M. de. (Org). **Avaliação da Educação básica e seus instrumentos**. Ed. dos Autores. 1ª edição. Rio de Janeiro. 2022.

PASQUALI, L. **Psicometria**: teoria dos testes na psicologia e na educação. Petrópolis: Vozes, 2011. 397p.

RIBEIRO, B. T. de O.; GROENWALD, C. L. O. Desvendando habilidades críticas de Matemática por meio da Psicometria: O uso da Teoria Clássica dos Testes e Teoria de Resposta ao Item em Avaliações Educacionais. **Revista Cearense de Educação Matemática**, v. 3, n. 7, p. 1–18, 2024. DOI: [10.56938/rceem.v3i8.4176](https://doi.org/10.56938/rceem.v3i8.4176). Disponível em: <https://www.sbembrasil.org.br/periodicos/index.php/rceem/article/view/4176>. Acesso em: 15 dez. 2025.

RODRIGUES, M. M. M. Proposta de análise de itens das provas do SAEB sob a perspectiva pedagógica e a psicométrica. **Estudos em Avaliação Educacional**, São Paulo, v. 17, n. 34, p. 43–78, 2006. DOI: 10.18222/eae173420062117. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2117>. Acesso em: 15 dez. 2025.

SCORTEGAGNA, L. Itens digitais no contexto de avaliações em larga escala. **Pesquisa e Debate em Educação**, Juiz de Fora: UFJF, v. 13, p. 1-23, e40259. 2023. DOI: 10.34019/2237-9444.2023.v13.40259. Disponível em: <https://periodicos.uff.br/index.php/RPDE/article/view/40259>.

TRAVITZKI, R.; ALAVARSE, O. M.; DE RIZZO MENEGHETTI, D.; DE TOLEDO CATALANI, Érica M. Teste adaptativo informatizado da Provinha Brasil – Leitura: resultados e perspectivas. **Estudos em Avaliação Educacional**, São Paulo, v. 31, n. 78, p. 525–553, 2020. DOI: 10.18222/eae.v0ix.7216. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/7216>. Acesso em: 11 dez. 2025.

TOCANTINS, Andreia Cristina Teixeira. **Avaliações em Larga Escala em Ciências Humanas no Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz De Fora (CAEd/UFJF): A Construção de um Guia de Orientação do Trabalho de Analistas e Auxiliares de Instrumentos de Avaliação em Questão**. Dissertação (mestrado profissional) - Universidade Federal de Juiz de Fora, Faculdade de Educação/CAEd. Programa de Pós-Graduação em Gestão e Avaliação da Educação Pública. P.183. 2018.

UNIVERSIDADE FEDERAL DE JUIZ DE FORA. **Estatuto da UFJF**. Disponível em: <https://www2.ufjf.br/ufjf/wp-content/uploads/sites/3/2015/10/estatuto.pdf>. 1998. Acesso em: 02 dez. 2025.

UNIVERSIDADE FEDERAL DE JUIZ DE FORA. Conselho Superior. **Resolução n. 27/2018**. Aprova a concordância com o pedido de credenciamento da Fundação CAEd como fundação de apoio da UFJF. 2018. Acesso em: 02 dez. 2025.

UNIVERSIDADE FEDERAL DE JUIZ DE FORA. Conselho Superior. **Resolução n. 83/2023**. Aprova a criação do Centro de Políticas Públicas e Avaliação da Educação (CAEd) como um órgão suplementar da Universidade Federal de Juiz de Fora (UFJF) e aprova o Regimento Interno deste órgão suplementar, em consonância com o Estatuto da Universidade Federal de Juiz de Fora. 2023. Disponível em https://www2.ufjf.br/consu/wp-content/uploads/sites/33/2023/12/Resolução-83.2023-Consu-UFJF_SEI_Assinada.pdf.pdf. Acesso em: 02 dez. 2025.

SILVA, W.; SOARES, T. M. Eficácia dos processos de linkagem na avaliação educacional em larga escala. **Estudos em Avaliação Educacional**, São Paulo, v. 21, n. 45, p. 191–211, 2010. DOI: 10.18222/eae214520102033. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/2033>. Acesso em: 22 mai. 2025.

SOUZA, T. G. de. **Estudo estatístico e pedagógico de itens de Matemática do CAEd na área de números e operações, álgebra e funções do Proeb**. Dissertação (mestrado profissional) - Universidade Federal de Juiz de Fora, Faculdade de Educação/CAEd. Programa de Pós-Graduação em Gestão e Avaliação da Educação Pública. 158 f. 2020.

APÊNDICE A – SCRIPT R PARA LEVANTAMENTO DE PUBLICAÇÕES

O script em linguagem R apresentado a seguir foi produzido com auxílio da inteligência artificial generativa Gemini (em 12 de dezembro de 2025), recebendo adequações pontuais por parte do autor. Esse script recebe o termo de busca e as páginas inicial e final, constrói a requisição de integração com a API OpenAlex e extrai os metadados de até 200 resultados por página. Essas informações são salvas em uma variável temporária e, ao final da execução, são exportadas para o arquivo “relatório_com_resultados.csv”.

```
#install.packages(c("httr", "jsonlite", "dplyr", "purrr", "tibble",
"stringr", "readr"))

library(httr)
library(jsonlite)
library(dplyr)
library(purrr)
library(tibble)
library(stringr)
library(readr)

# -----
# Função auxiliar: reconstrói o resumo a partir do índice invertido
# -----
reconstruir_resumo <- function(abstract_inverted_index) {

  if (is.null(abstract_inverted_index)) {
    return(NA_character_)
  }

  palavras <- names(abstract_inverted_index)

  # Coletar todas as posições e converter para inteiro
  todas_posicoes <- suppressWarnings(
    as.integer(unlist(abstract_inverted_index))
  )

  if (length(todas_posicoes) == 0 || all(is.na(todas_posicoes))) {
    return(NA_character_)
  }

  resumo_vec <- character(max(todas_posicoes, na.rm = TRUE) + 1)

  for (palavra in palavras) {

    idxs <- suppressWarnings(
      as.integer(abstract_inverted_index[[palavra]])
    )

    idxs <- idxs[!is.na(idxs)]

    if (length(idxs) > 0) {
      resumo_vec[idxs + 1] <- palavra
    }
  }
}
```

```

    paste(resumo_vec, collapse = " ")
  }

# -----
# Função principal: coleta dados do OpenAlex por páginas
# -----
coletar_openalex_por_paginas <- function(
  termo_busca,
  pagina_inicial,
  pagina_final,
  per_page = 200,
  delay = 0.2
) {

  base_url <- "https://api.openalex.org/works"
  todas_paginas <- list()

  for (page in pagina_inicial:pagina_final) {

    query <- list(
      page = page,
      filter = paste0("title_and_abstract.search:", termo_busca),
      sort = "relevance_score:desc",
      per_page = per_page,
      include_xpac = "true"
    )

    resp <- GET(base_url, query = query)
    stop_for_status(resp)

    json <- fromJSON(
      content(resp, as = "text", encoding = "UTF-8"),
      simplifyVector = FALSE
    )

    message(
      paste("Página", page, "-", length(json$results), "registros")
    )

    df_pagina <- bind_rows(
      lapply(json$results, function(item) {

        titulo <- if (!is.null(item$title)) item$title else NA_character_

        ano <- if (!is.null(item$publication_year))
          item$publication_year else NA_integer_

        doi <- if (!is.null(item$doi)) item$doi else NA_character_

        fonte <- if (!is.null(item$host_venue) &&
          !is.null(item$host_venue$display_name)) {
          item$host_venue$display_name
        } else {
          NA_character_
        }

        autores <- if (!is.null(item$authorships)) {
          paste(
            vapply(

```

```

        item$authorships,
        function(a) {
          if (!is.null(a$author$display_name))
            a$author$display_name
          else
            NA_character_
        },
        character(1)
      ),
      collapse = ", "
    )
  } else {
    NA_character_
  }

  resumo <- if (!is.null(item$abstract_inverted_index)) {
    reconstruir_resumo(item$abstract_inverted_index)
  } else {
    NA_character_
  }

  data.frame(
    titulo = titulo,
    ano = ano,
    doi = doi,
    fonte = fonte,
    autores = autores,
    resumo = resumo,
    stringsAsFactors = FALSE
  )
})
)

todas_paginas[[as.character(page)]] <- df_pagina
Sys.sleep(delay)
}

bind_rows(todas_paginas)
}

df <- coletar_openalex_por_paginas( #variável em que serão carregados os
  resultados
  termo_busca = "elaboração de itens", #palavra chave a ser pesquisada
  pagina_inicial = 1,
  pagina_final = 9 #a função tem parâmetro definido para 200 resultados por
  página. Insira a quantidade de páginas totais. Considerando que a pesquisa
  comum tenha sido feita, a quantidade de páginas permitirá o código de fazer
  uma execução otimizada. Definir pagina_final maior do que a quantidade de
  resultados, fará com que o script faça requisições adicionais e retorne
  "Página ## - 0 registros"
)

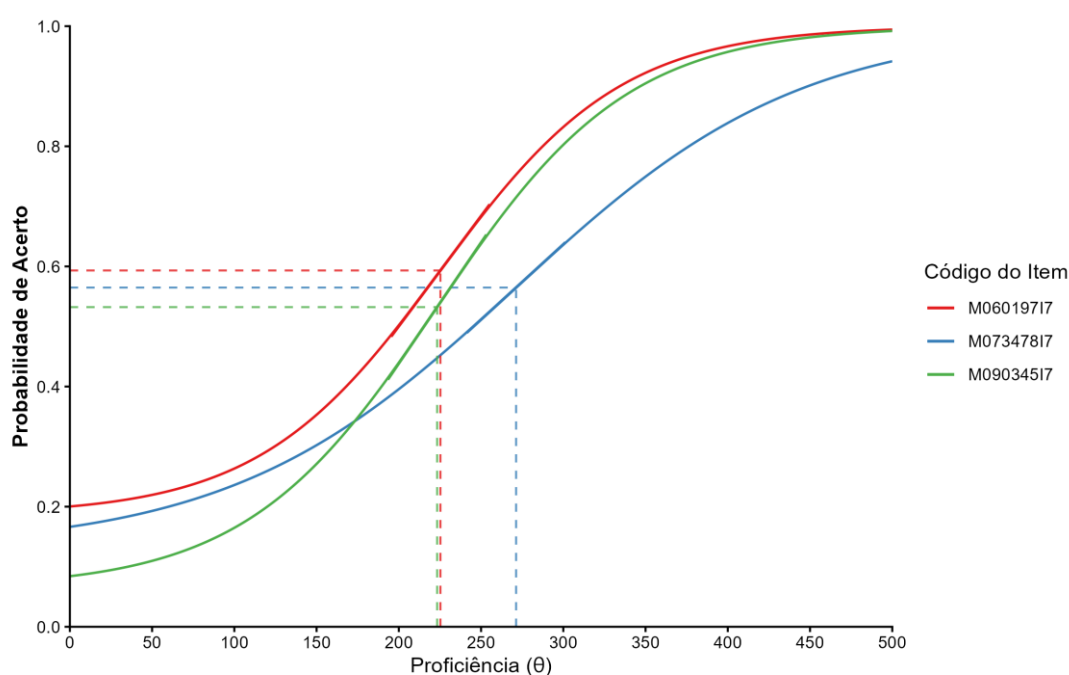
write_csv(df, "relatório_com_resultados.csv")

```

APÊNDICE B – CURVAS CARACTERÍSTICAS DOS ITENS

Utilizando os dados da Tabela 3, foram gerados os gráficos com as Curvas Características dos Itens (CCI), agrupados por descritor em cada etapa. Em cada gráfico, utilizou-se uma cor diferente para cada item, além da marcação de dois segmentos tracejados que se intersectam no ponto de inflexão da curva e de um segmento³⁶ indicando a reta tangente a esse ponto. É exatamente na projeção desse ponto de inflexão no eixo horizontal (θ) que se identifica o parâmetro B (dificuldade), correspondendo a uma probabilidade específica de acerto no eixo vertical (P). O parâmetro A (discriminação), por sua vez, é proporcional à inclinação dessa reta tangente.

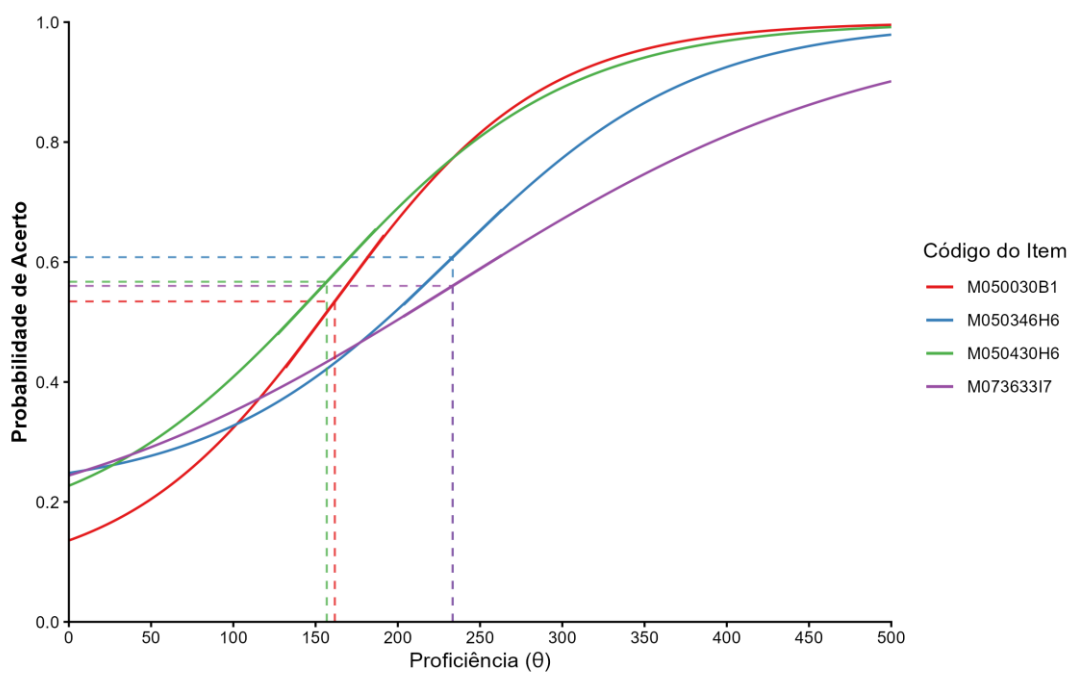
Gráfico 1 – Curva característica dos itens do descritor D01 - 5EF



Fonte: Elaborado pelo autor (2026).

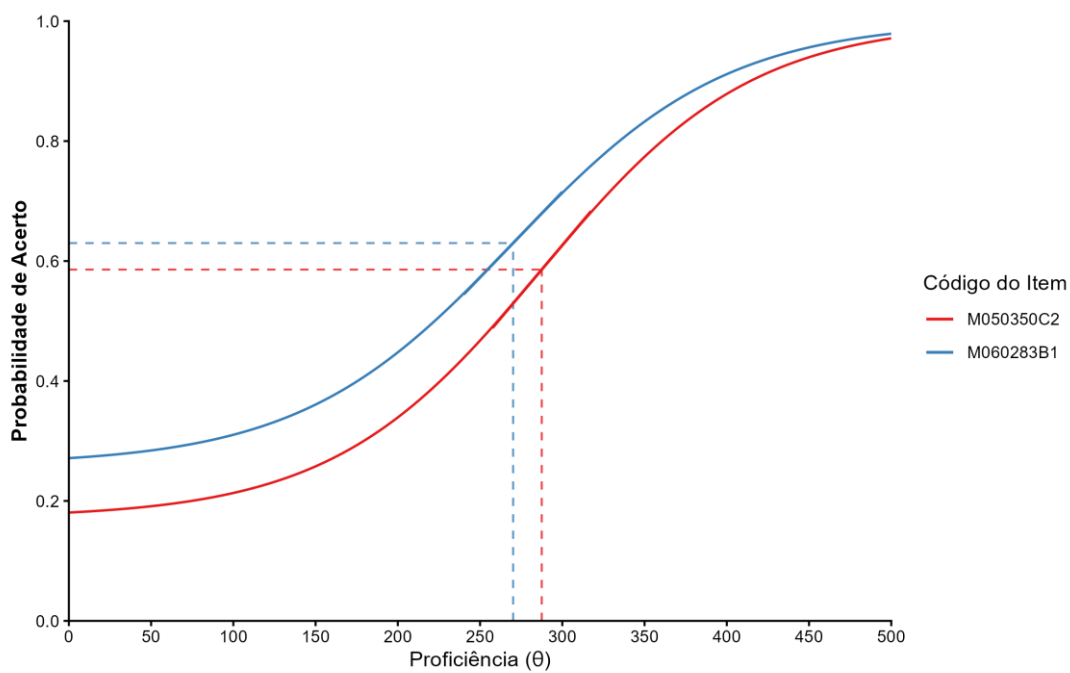
³⁶ Em alguns gráficos esse segmento parece imperceptível, devido à curva não variar muito em relação a essa reta.

Gráfico 2 – Curva característica dos itens do descritor D02 - 5EF



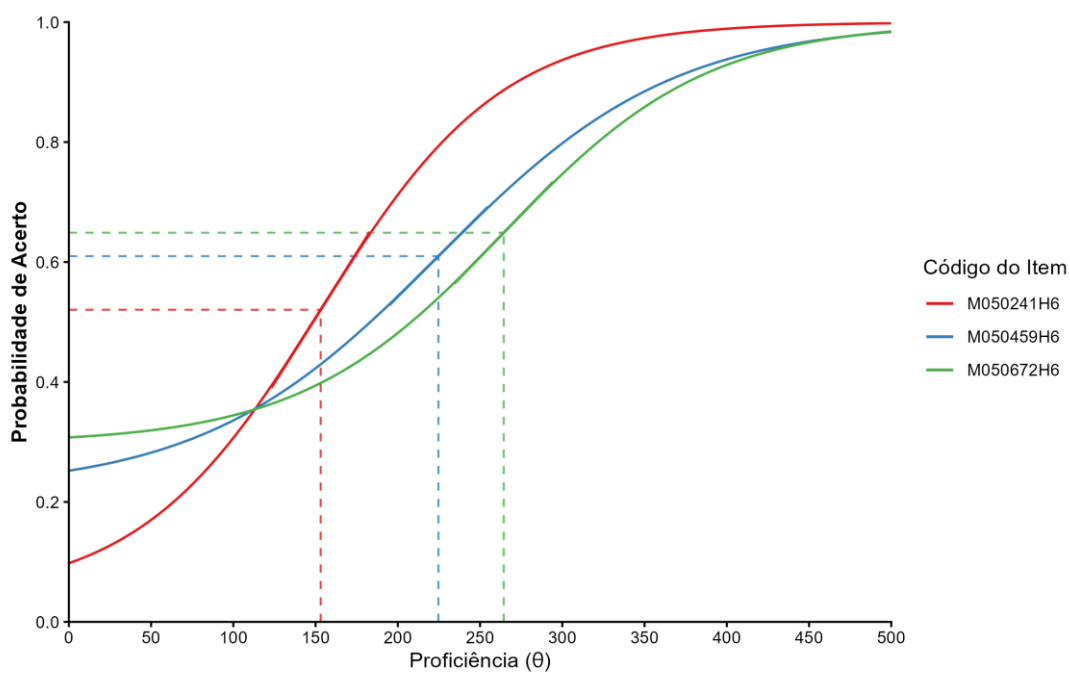
Fonte: Elaborado pelo autor (2026).

Gráfico 3 – Curva característica dos itens do descritor D03 - 5EF



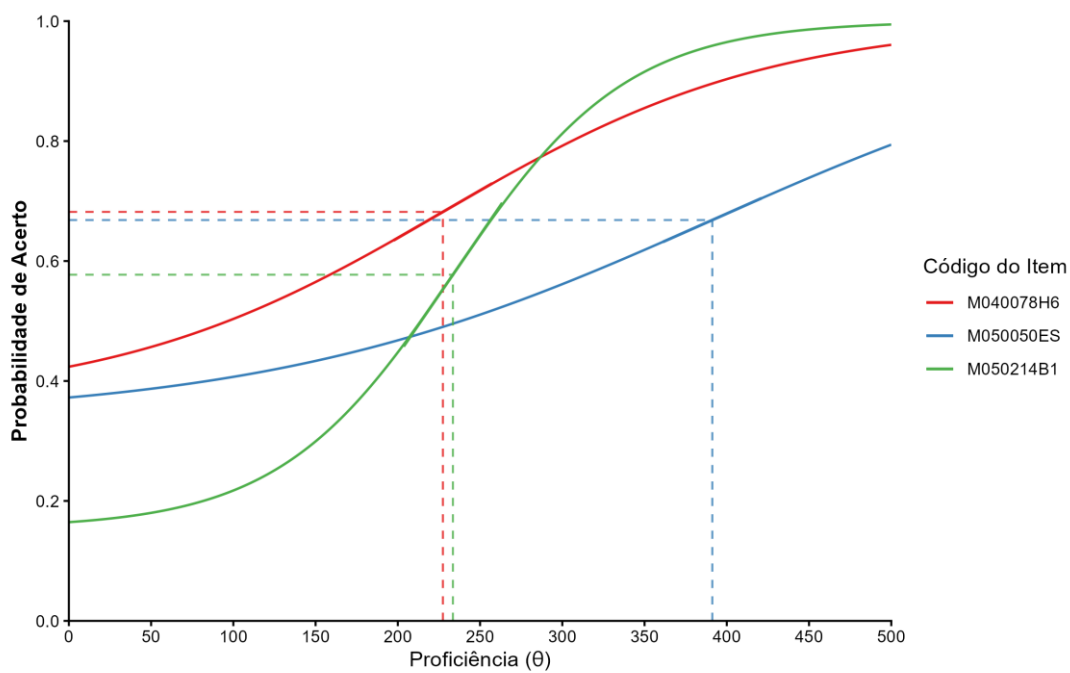
Fonte: Elaborado pelo autor (2026).

Gráfico 4 – Curva característica dos itens do descritor D04 - 5EF



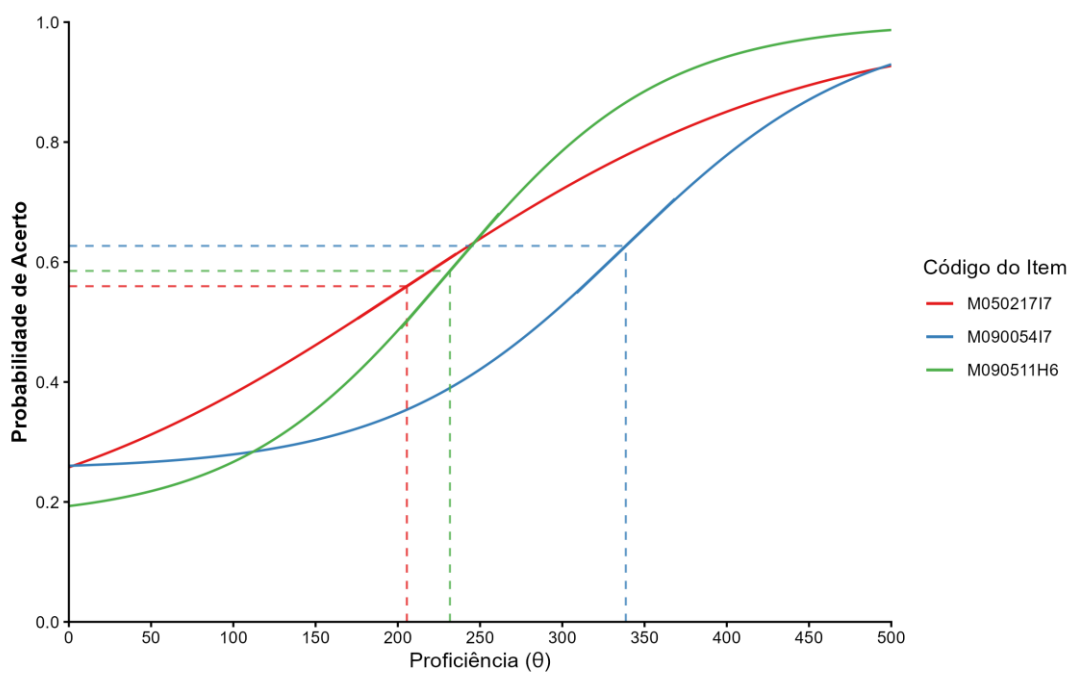
Fonte: Elaborado pelo autor (2026).

Gráfico 5 – Curva característica dos itens do descritor D05 - 5EF



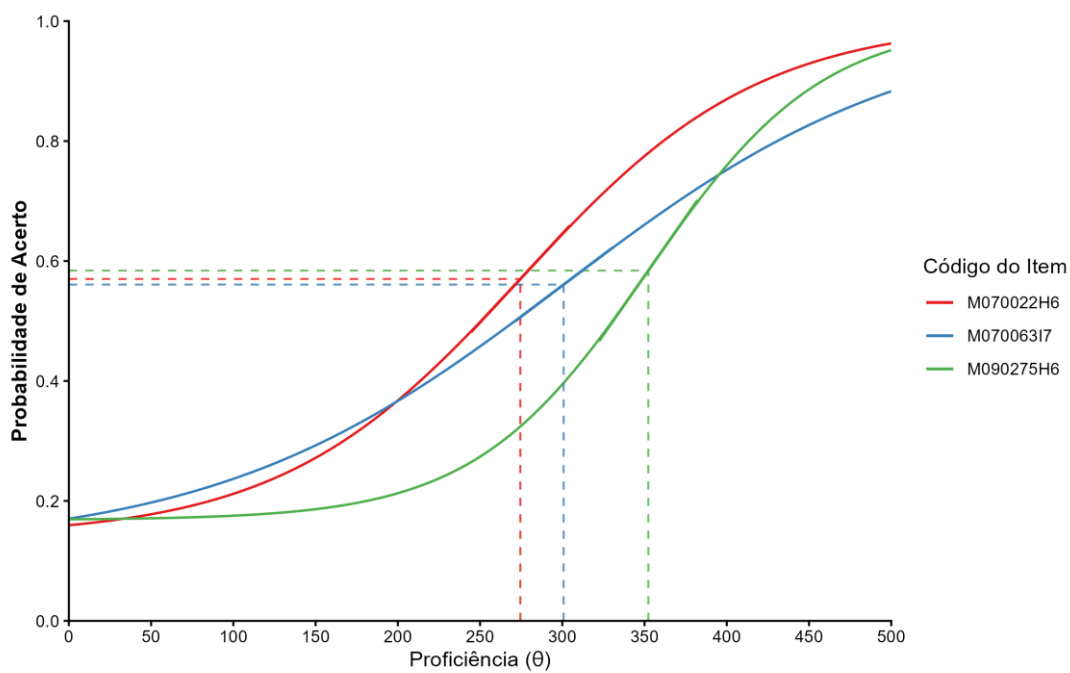
Fonte: Elaborado pelo autor (2026).

Gráfico 6 – Curva característica dos itens do descritor D01 - 9EF



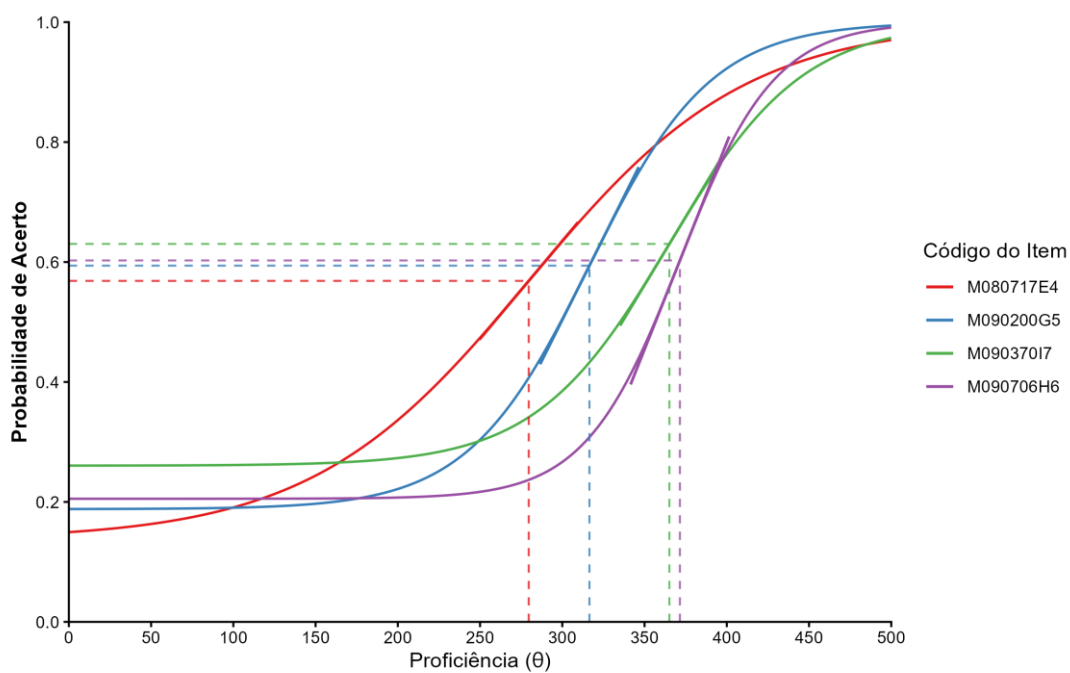
Fonte: Elaborado pelo autor (2026).

Gráfico 7 – Curva característica dos itens do descritor D02 - 9EF



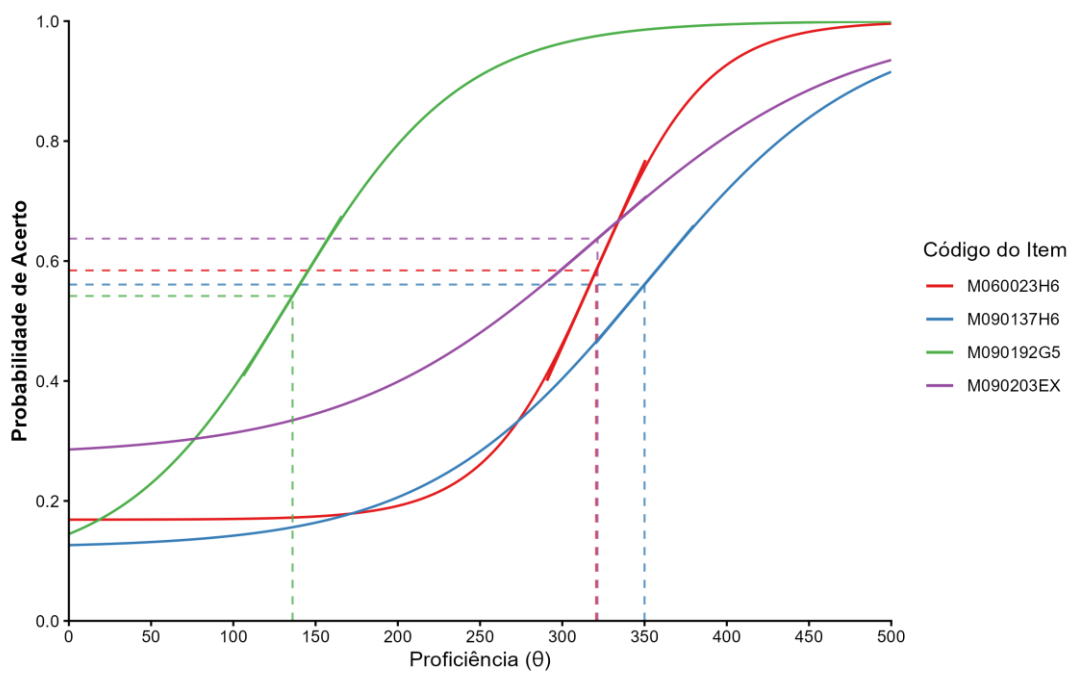
Fonte: Elaborado pelo autor (2026).

Gráfico 8 – Curva característica dos itens do descritor D03 - 9EF



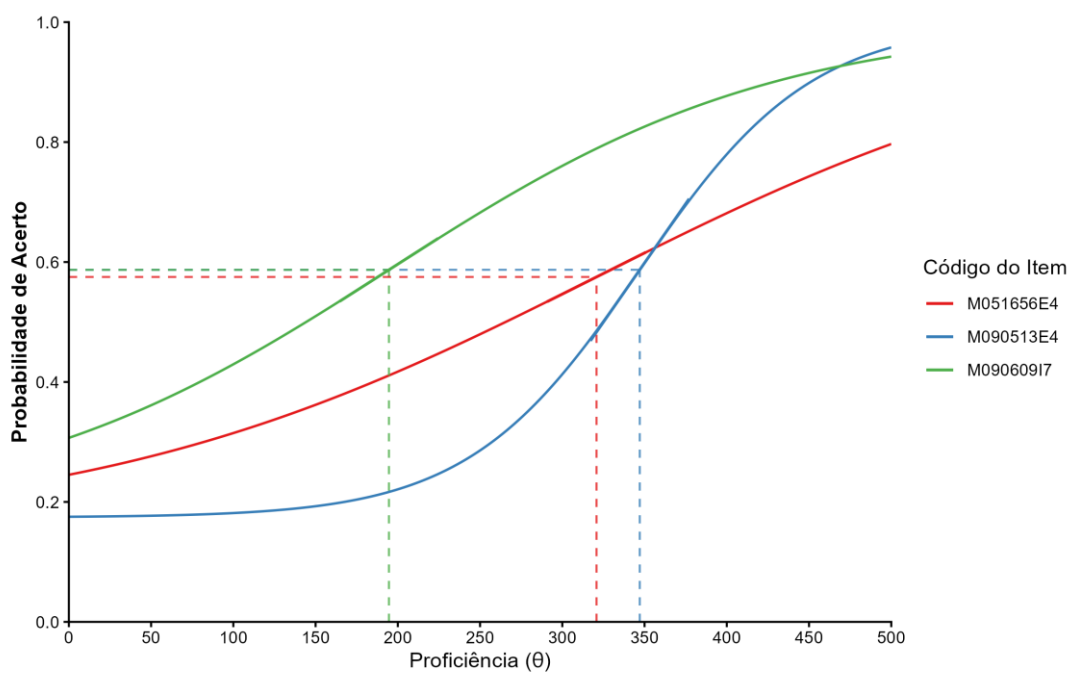
Fonte: Elaborado pelo autor (2026).

Gráfico 9 – Curva característica dos itens do descritor D04 - 9EF



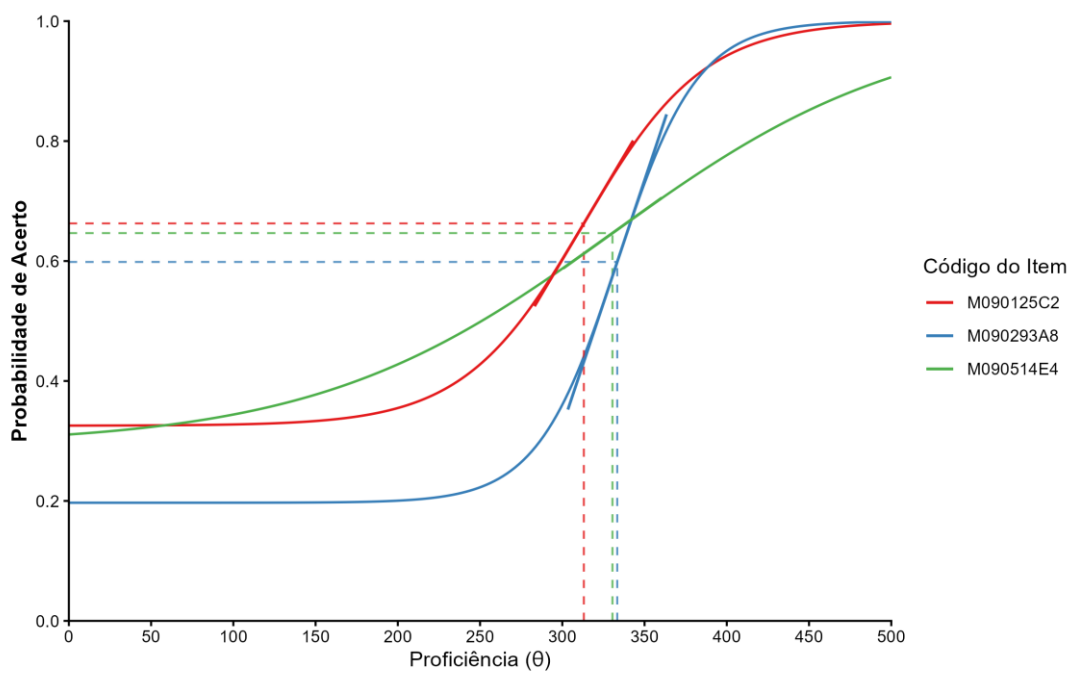
Fonte: Elaborado pelo autor (2026).

Gráfico 10 – Curva característica dos itens do descritor D05 - 9EF



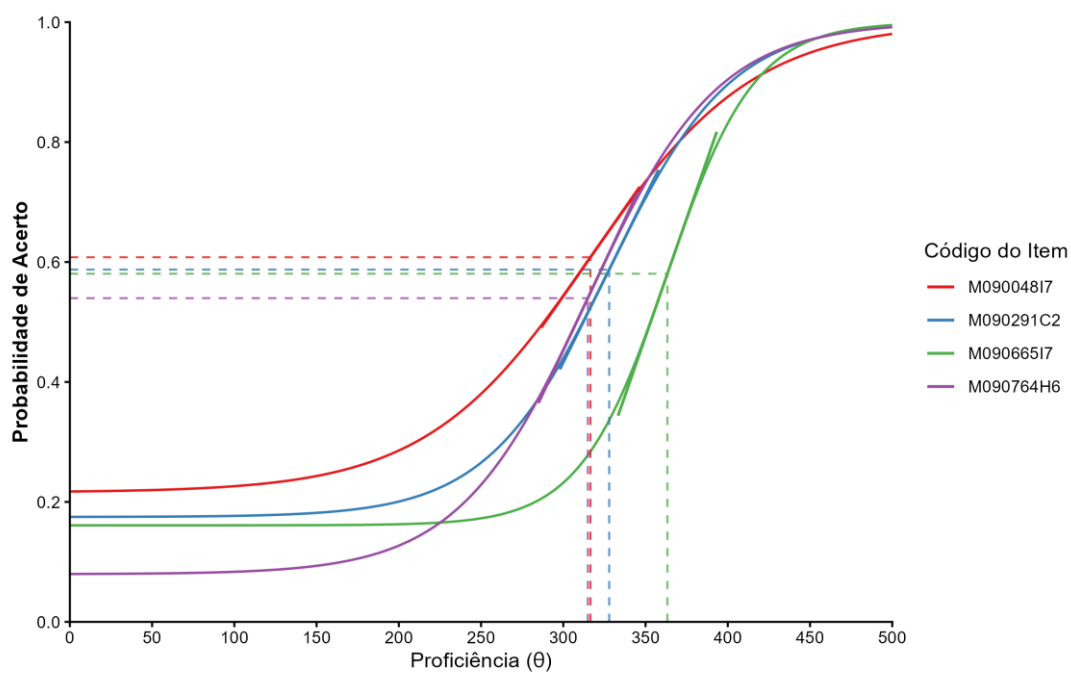
Fonte: Elaborado pelo autor (2026).

Gráfico 11 – Curva característica dos itens do descritor D06 - 9EF



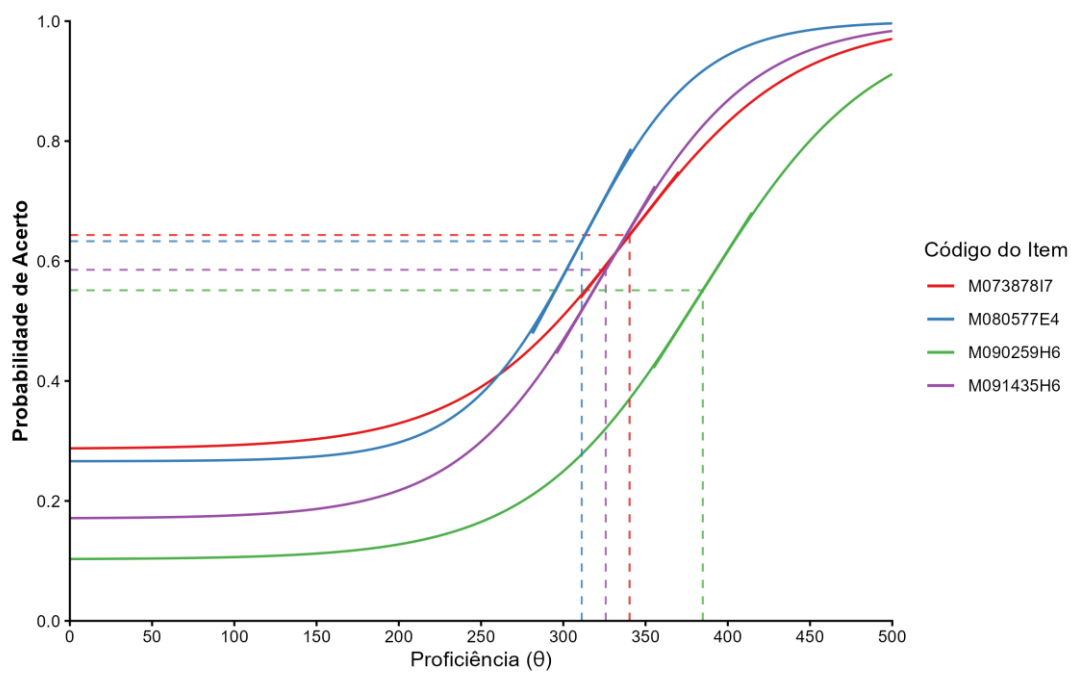
Fonte: Elaborado pelo autor (2026).

Gráfico 12 – Curva característica dos itens do descritor D07 - 9EF



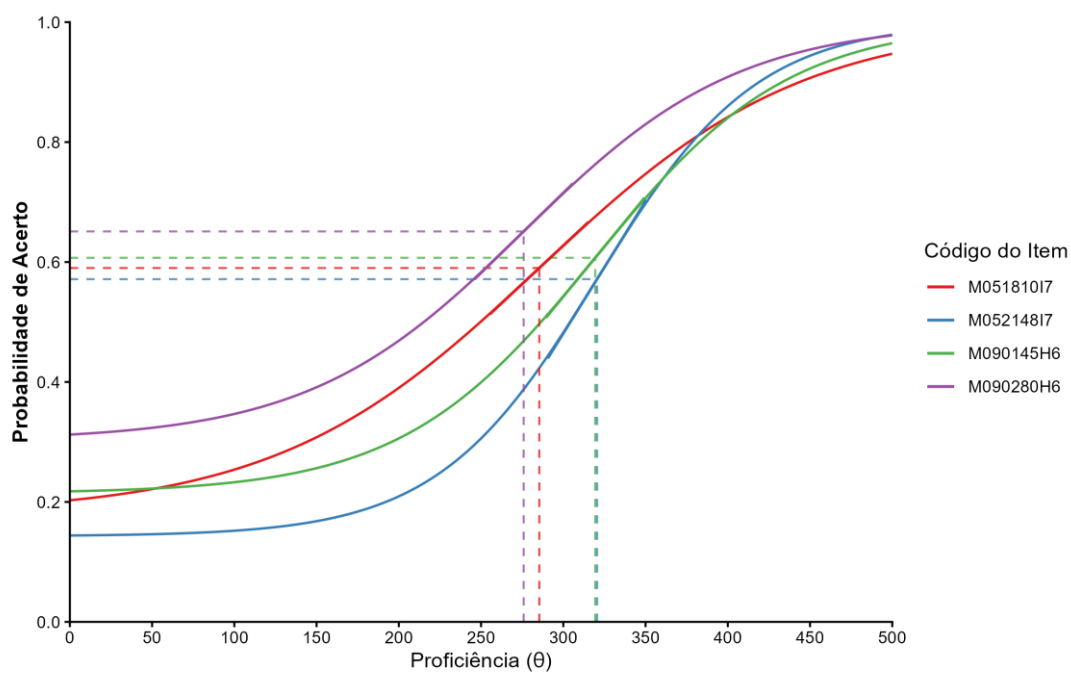
Fonte: Elaborado pelo autor (2026).

Gráfico 13 – Curva característica dos itens do descritor D08 - 9EF



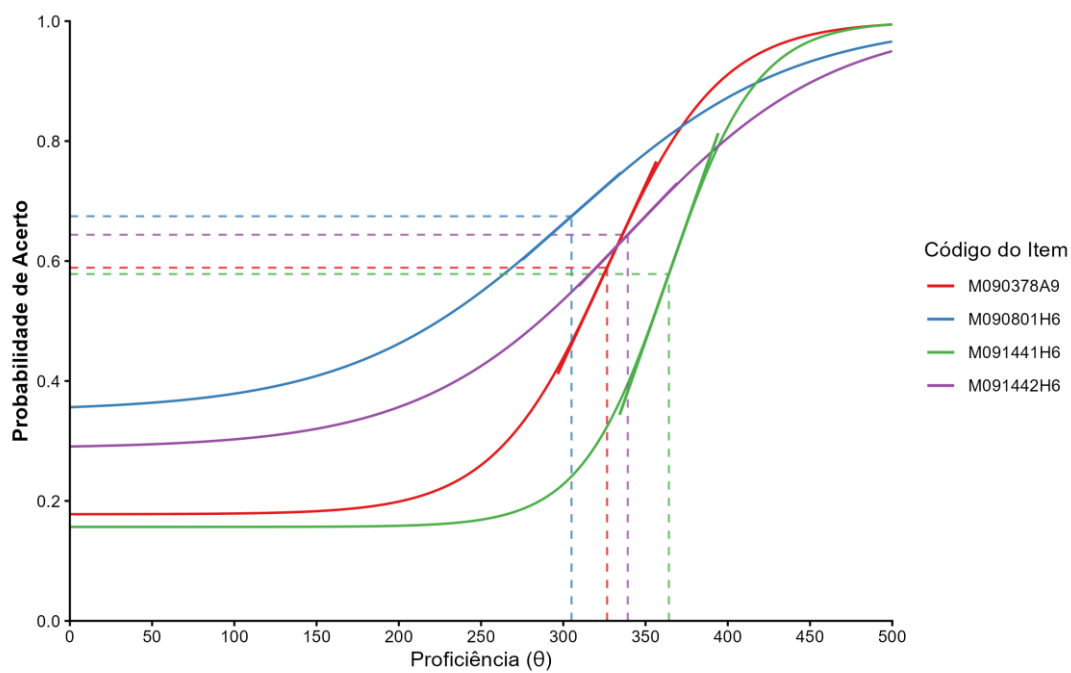
Fonte: Elaborado pelo autor (2026).

Gráfico 14 – Curva característica dos itens do descritor D09 - 9EF



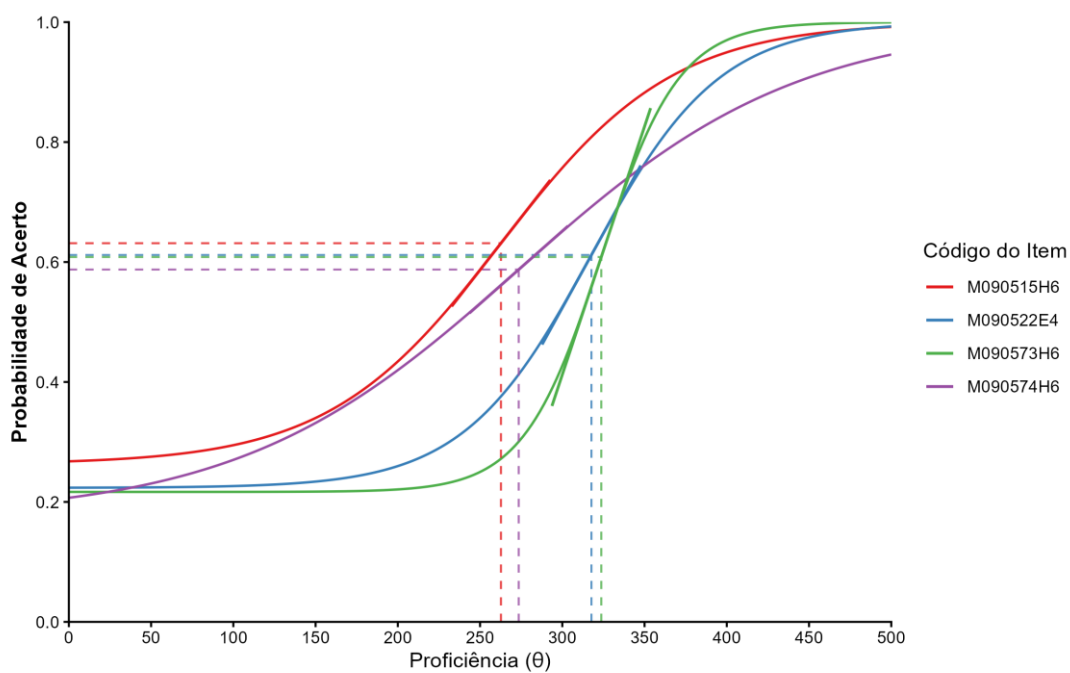
Fonte: Elaborado pelo autor (2026).

Gráfico 15 – Curva característica dos itens do descritor D10 - 9EF



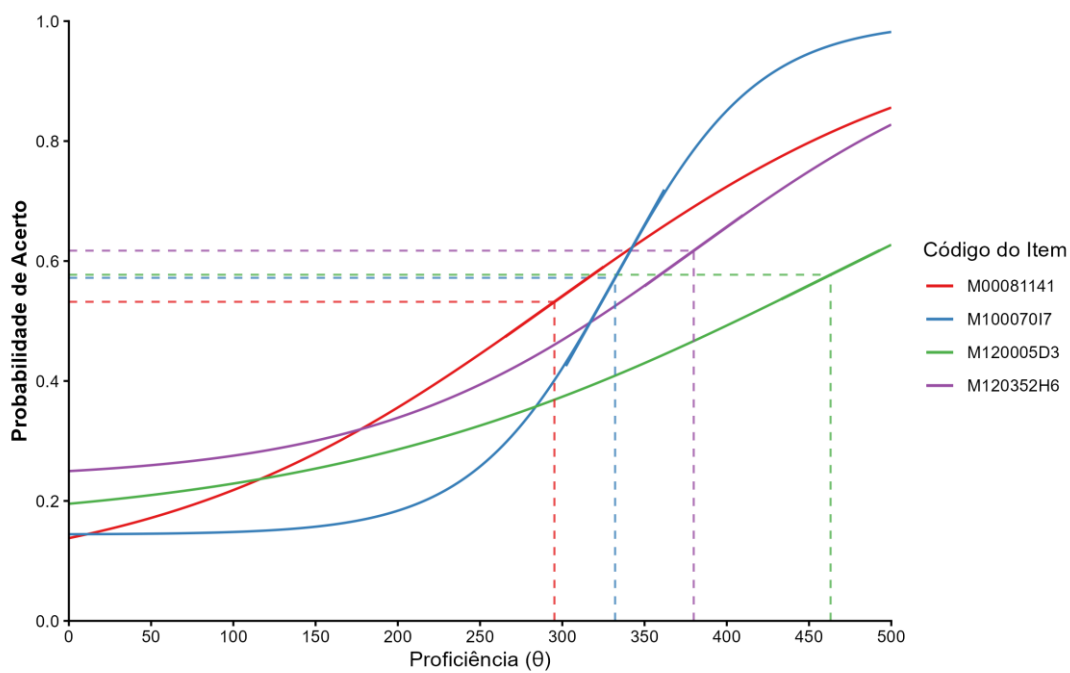
Fonte: Elaborado pelo autor (2026).

Gráfico 16 – Curva característica dos itens do descritor D11 - 9EF



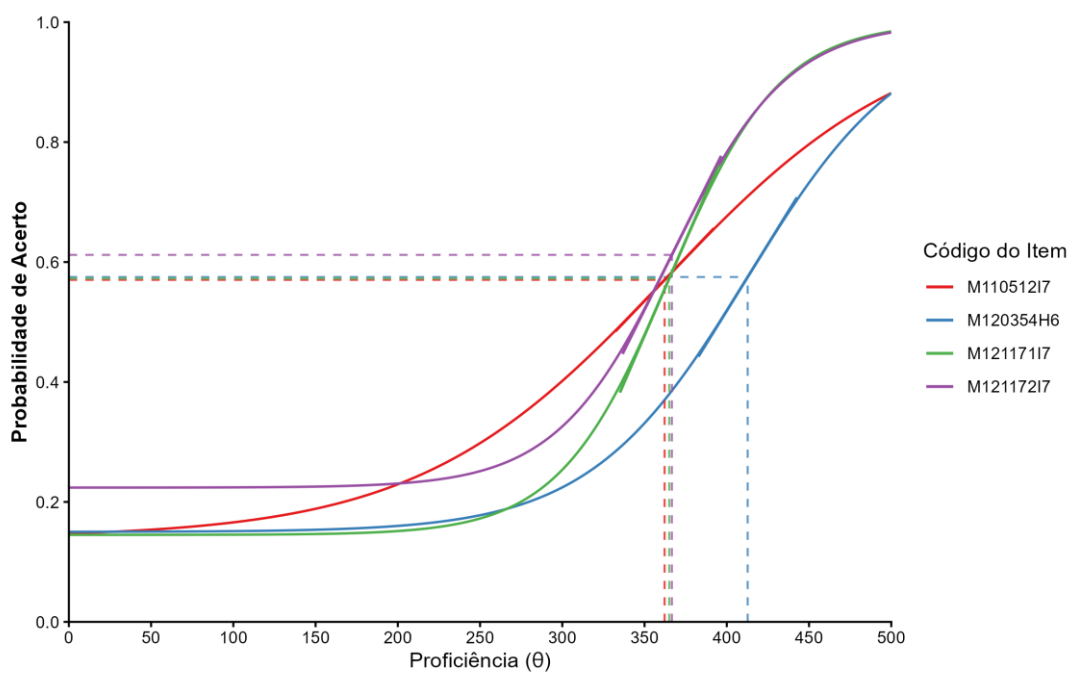
Fonte: Elaborado pelo autor (2026).

Gráfico 17 – Curva característica dos itens do descritor D01 - 3EM



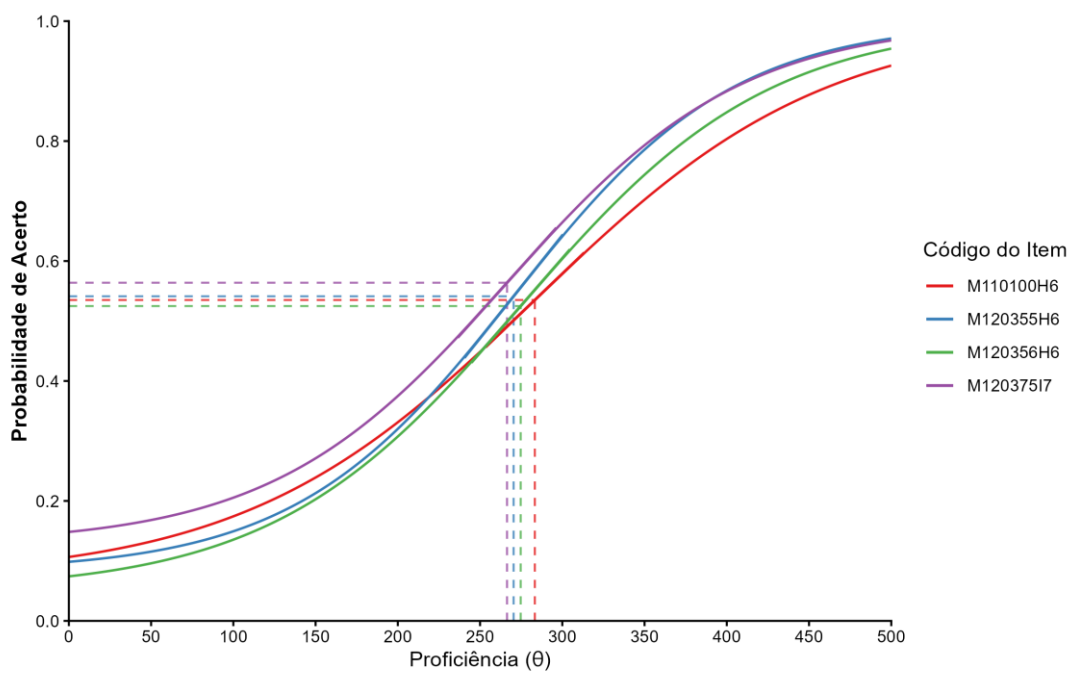
Fonte: Elaborado pelo autor (2026).

Gráfico 18 – Curva característica dos itens do descritor D02 - 3EM



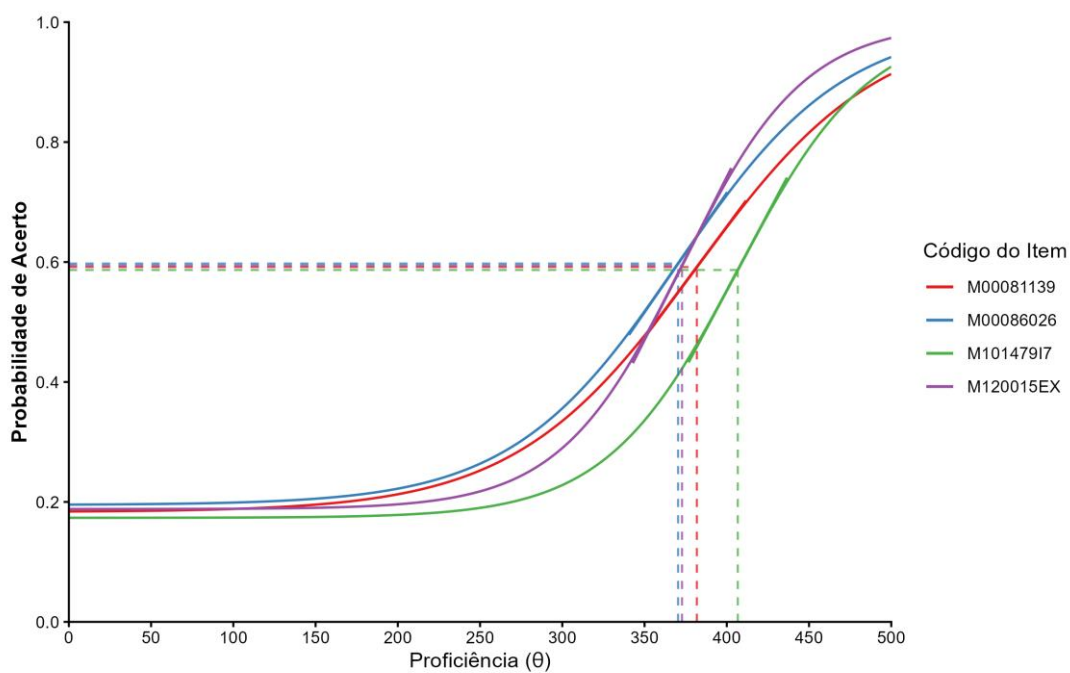
Fonte: Elaborado pelo autor (2026).

Gráfico 19 – Curva característica dos itens do descritor D03 - 3EM



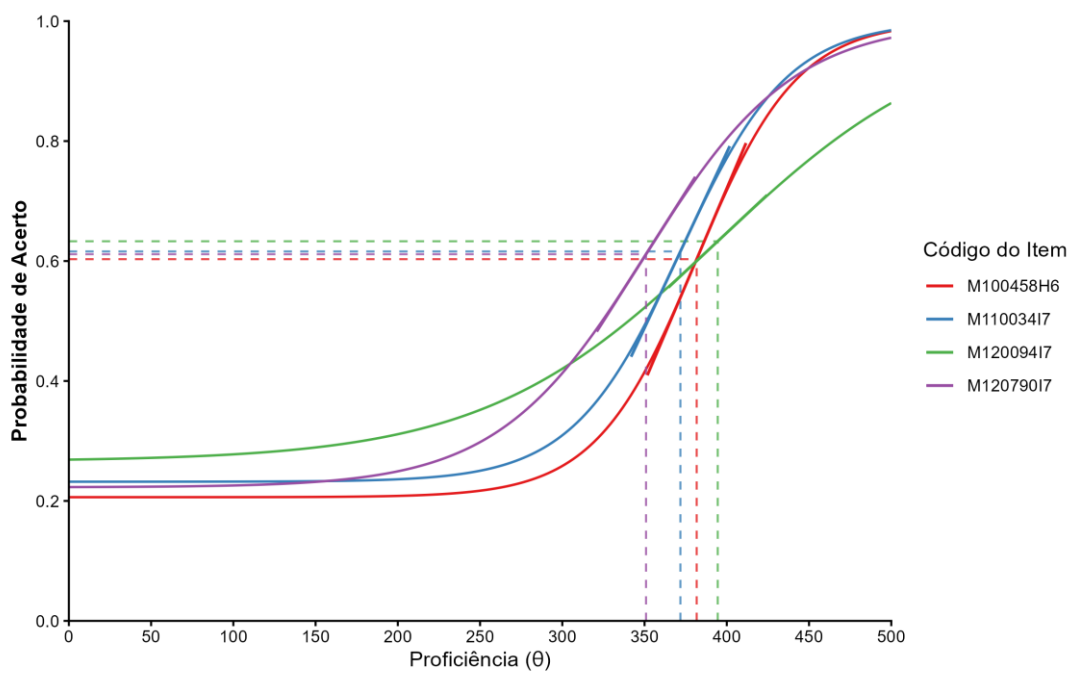
Fonte: Elaborado pelo autor (2026).

Gráfico 20 – Curva característica dos itens do descritor D04 - 3EM



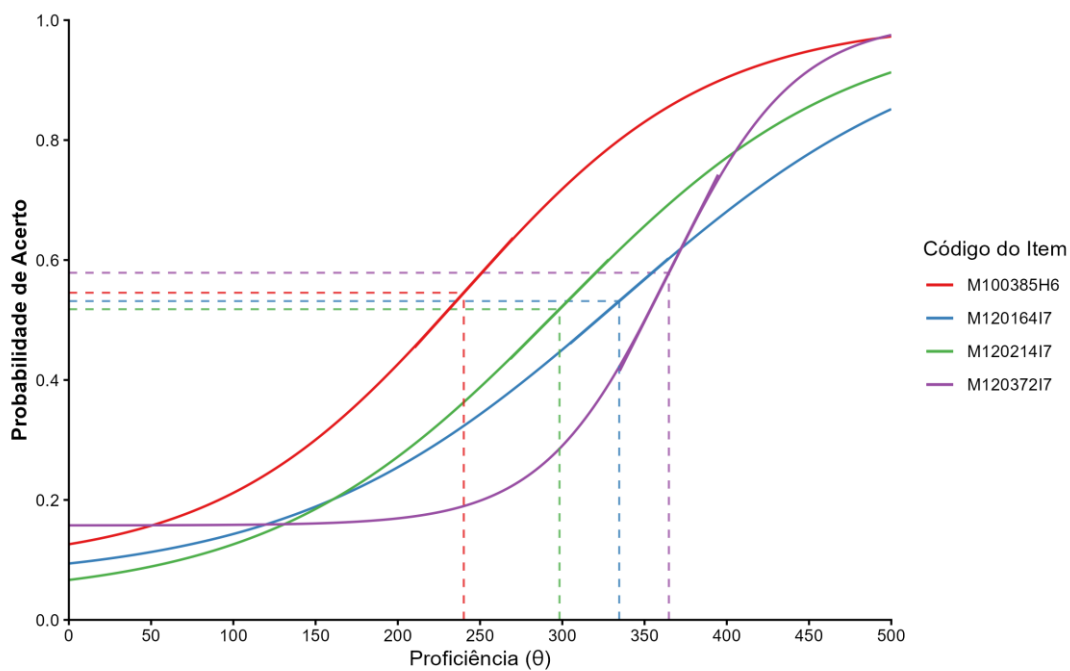
Fonte: Elaborado pelo autor (2026).

Gráfico 21 – Curva característica dos itens do descritor D05 - 3EM



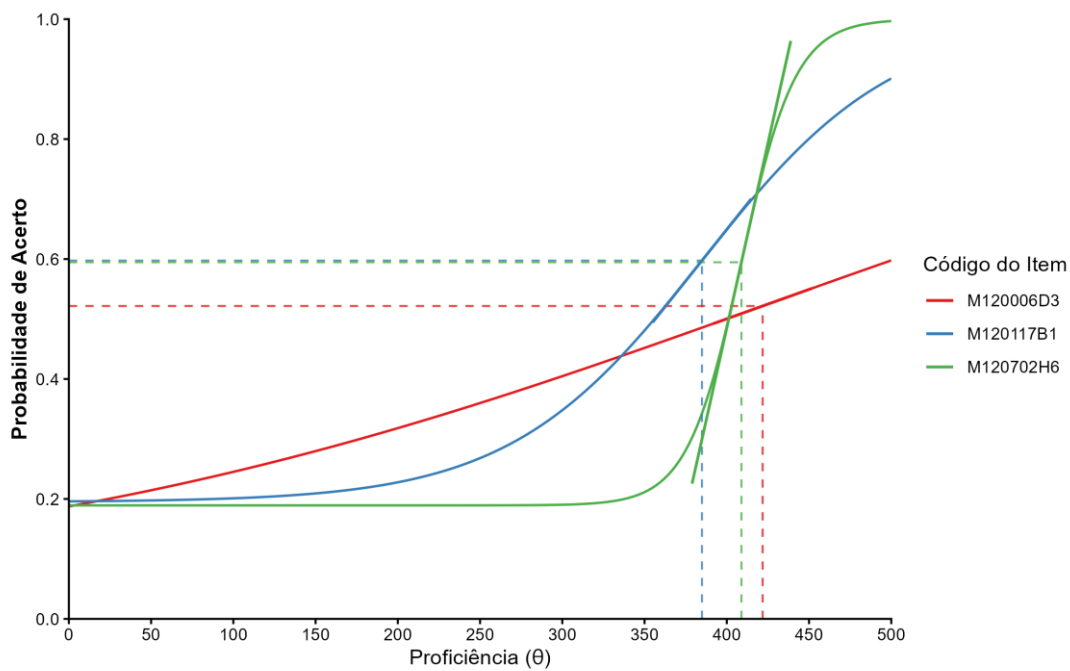
Fonte: Elaborado pelo autor (2026).

Gráfico 22 – Curva característica dos itens do descritor D06 - 3EM



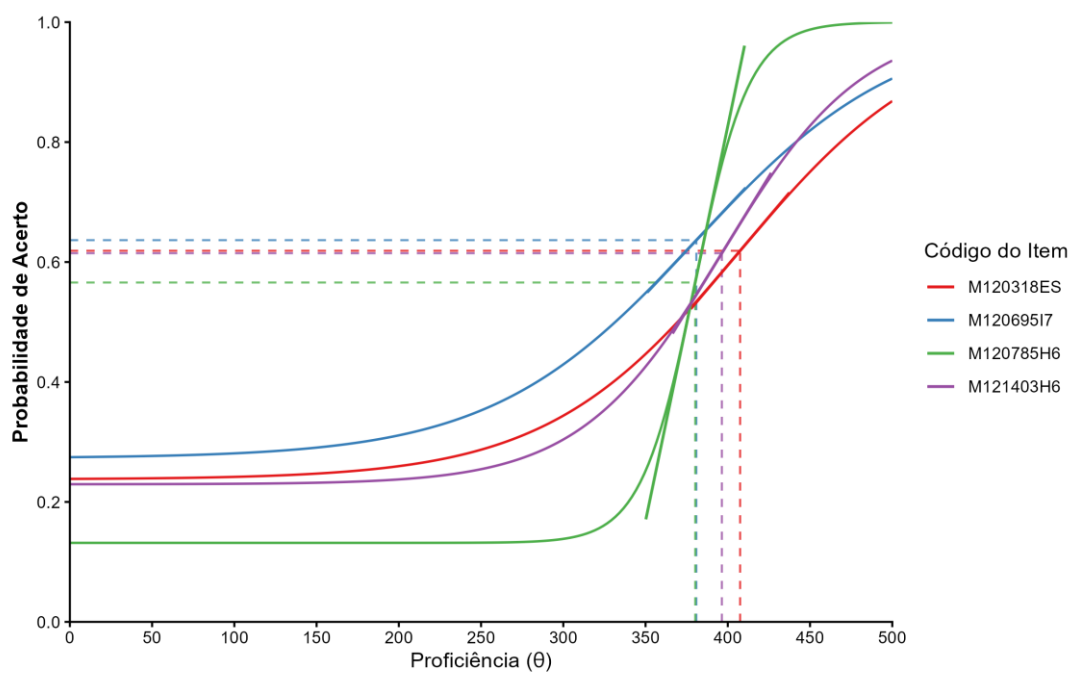
Fonte: Elaborado pelo autor (2026).

Gráfico 23 – Curva característica dos itens do descritor D07 - 3EM



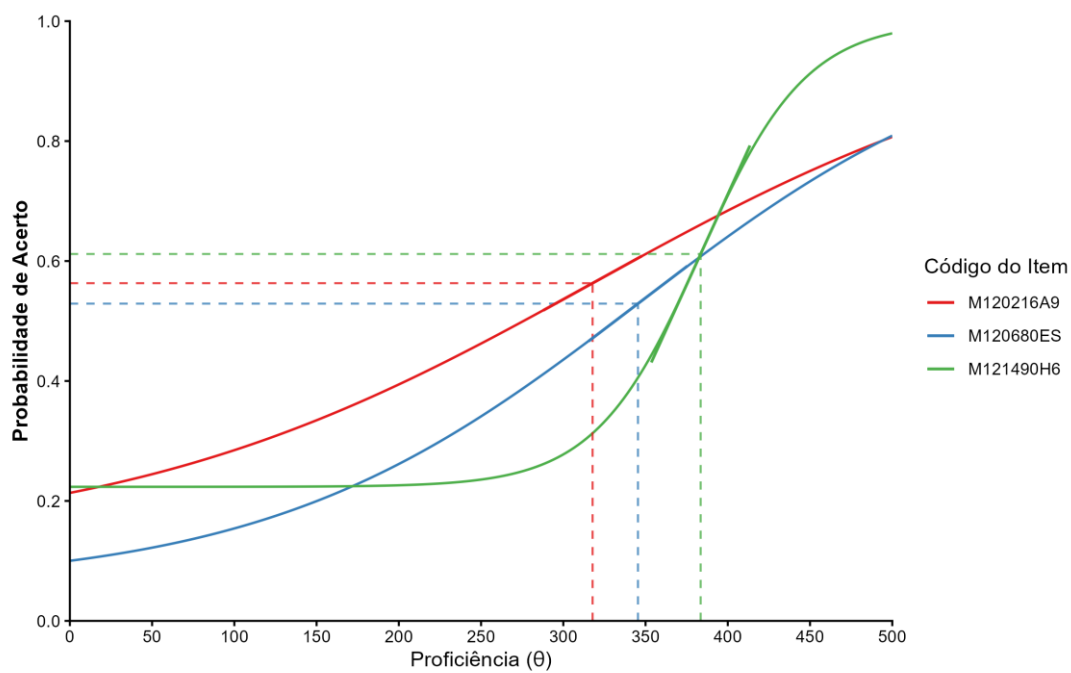
Fonte: Elaborado pelo autor (2026).

Gráfico 24 – Curva característica dos itens do descritor D08 - 3EM



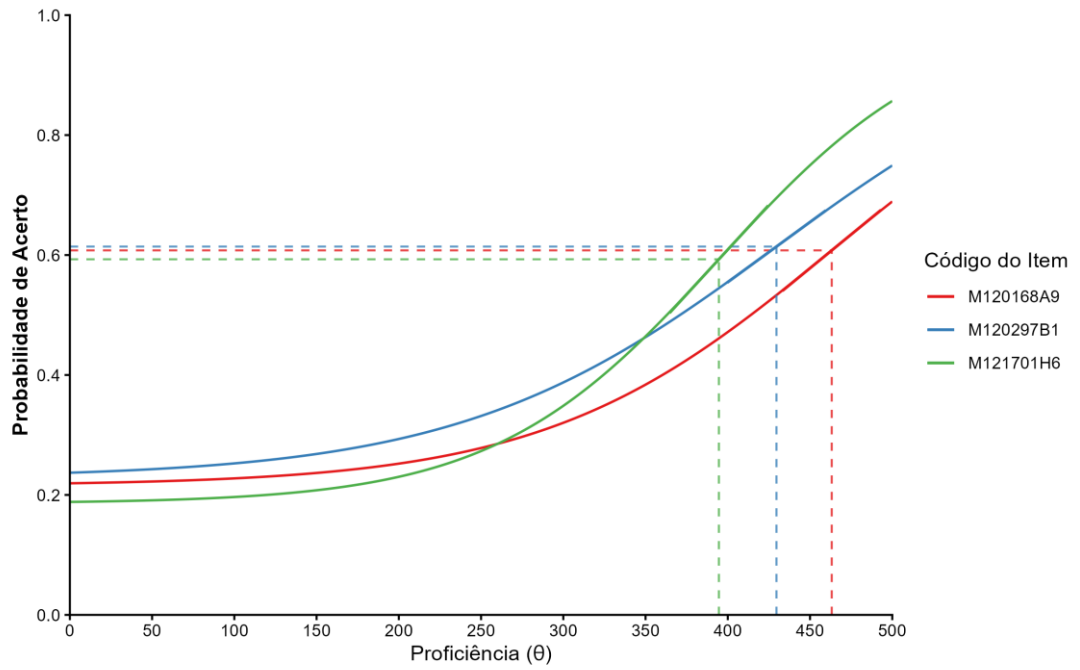
Fonte: Elaborado pelo autor (2026).

Gráfico 25 – Curva característica dos itens do descritor D09 - 3EM



Fonte: Elaborado pelo autor (2026).

Gráfico 26 – Curva característica dos itens do descritor D10 - 3EM



Fonte: Elaborado pelo autor (2026).