

GABRIELA THAMARA DE FREITAS BARROS

**PROCEDIMENTOS PARA A CONSTRUÇÃO DE
INDICADORES POR MEIO DA TEORIA DE RESPOSTA AO
ITEM: A CRIAÇÃO DE UMA MEDIDA DE NÍVEL
SOCIOECONÔMICO FAMILIAR**

Dissertação de Mestrado submetida ao Programa de Pós Graduação em Métodos e Gestão em Avaliação da Universidade Federal de Santa Catarina, como requisito final para obtenção do título de Mestre.

Orientador: Prof. Dr. Adriano Ferreti Borgatto

Florianópolis
2016

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática da Biblioteca Universitária da UFSC

BARROS, Gabriela

Procedimentos para a construção de indicadores por meio da Teoria de Resposta ao Item: a criação de uma medida de nível socioeconômico familiar / Gabriela Barros ; orientador, Prof. Dr. Adriano Ferreti Borgatto - Florianópolis, SC, 2016. 163 p.

Dissertação (mestrado profissional) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Métodos e Gestão em Avaliação.

Inclui referências

1. Métodos e Gestão em Avaliação. 2. Indicadores Contextuais. 3. Nível Socioeconômico. 4. Teoria de Resposta ao Item. 5. Educação. I. Ferreti Borgatto, Prof. Dr. Adriano . II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Métodos e Gestão em Avaliação. III. Título.

GABRIELA THAMARA DE FREITAS BARROS

**PROCEDIMENTOS PARA A CONSTRUÇÃO DE
INDICADORES POR MEIO DA TEORIA DE RESPOSTA AO
ITEM: A CRIAÇÃO DE UMA MEDIDA DE NÍVEL
SOCIOECONÔMICO FAMILIAR**

Esta Dissertação foi julgada adequada para obtenção do Título de Mestre, e aprovada em sua forma final pelo Programa de Pós-Graduação em Métodos e Gestão em Avaliação.
Florianópolis, 22 de agosto de 2016.

Prof. Renato Cislighi, Dr.
Coordenador do Programa

Banca Examinadora:

Prof. Adriano Ferreti Borgatto, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof. Pedro Alberto Barbeta, Dr.
Universidade Federal de Santa Catarina

Prof. Dalton Francisco de Andrade, Dr.
Universidade Federal de Santa Catarina

Adolfo Samuel de Oliveira, Dr.
Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

AGRADECIMENTOS

Aos meus pais e familiares, que sempre estiveram comigo, me oferecendo incondicional apoio, incentivo e amor. Tudo isso, sem dúvida, foi fundamental para eu conseguir chegar até aqui.

Aos meus amigos que sempre estiveram torcendo por mim, me acolhendo quando necessário e compreendendo os momentos em que estive ausente.

Aos meus colegas de trabalho pelo apoio e paciência nos momentos que mais precisei.

Ao professor Joaquim José Soares Neto pelo apoio e por sempre destacar a importância da formação acadêmica na minha profissão.

Ao professor José Francisco Soares por sugerir o tema estudado ressaltando a importância de estudar a desigualdade educacional e por me ensinar muito sobre o tema.

Aos professores Pedro Alberto Barbeta, Dalton Francisco de Andrade e Adolfo Samuel de Oliveira por aceitarem fazer parte da banca e pelas importantes contribuições a este trabalho.

Por último, mas não menos importante, meu orientador, Prof. Adriano Ferreti Borgatto, por ter aceitado o desafio de me orientar.

Muito Obrigada!

“A tarefa não é tanto ver aquilo que ninguém viu,
mas pensar o que ninguém ainda pensou sobre
aquilo que todo mundo vê.”

(Arthur Schopenhauer)

RESUMO

A produção e discussão de indicadores contextuais no âmbito da educação é um tema de bastante importância, visto que a partir deles é possível desenvolver políticas públicas que permitam melhorias no sistema educacional. O presente trabalho evidencia a importância de uma medida de nível socioeconômico na avaliação dos sistemas de ensino. Ao verificar as desigualdades existentes entre os diversos públicos, que frequentam as escolas brasileiras de educação básica, é possível desenvolver políticas que permitam minimizar o impacto da origem social dos alunos em sua trajetória escolar. Este trabalho sugere procedimentos para a construção de indicadores por meio da Teoria de Resposta ao Item (TRI), a partir da criação de uma medida de nível socioeconômico familiar utilizando os questionários aplicados, em 2013, aos concluintes do Ensino Médio do Enem e alunos participantes do Saeb. O estudo foi feito a partir das respostas de 6.273.995 estudantes, utilizando inicialmente vinte e quatro questões. Procurou-se estudar um conjunto de modelos da TRI, de natureza politômica e dicotômica, a fim de utilizar aquele que melhor se ajustou aos dados. Para tanto, foram testados os modelos politômicos de respostas nominais, de respostas graduais de Samejima e o de crédito parcial generalizado, bem como o modelo dicotômico de dois parâmetros. Os testes mostraram que o modelo de respostas graduais foi o modelo que melhor se ajustou aos dados, ensejando, assim, uma medida mais informativa para toda a escala, quando comparada com os outros modelos. Verificou-se, ainda, se os pressupostos deste modelo estavam sendo atendidos, como, por exemplo, a invariância e a unidimensionalidade, o que, convergiu para um modelo final com dezessete questões. Buscou-se validar o Indicador de Nível Socioeconômico (Inse) cruzando a medida calculada com informações como localização rural e urbana e dependência administrativa e correlacionando-o com outras medidas de nível socioeconômico ou relacionadas ao construto. Por fim, utilizando os parâmetros das questões e a probabilidade calculada pelo modelo para diversos valores de Inse, dividiu-se a escala de nível socioeconômico utilizando métodos estatísticos de agrupamento o que resultou em nove níveis de Inse e esses níveis foram caracterizados posicionando as questões e suas alternativas a partir de suas probabilidades.

Palavras-chave: Indicadores Contextuais. Nível Socioeconômico. Teoria de Resposta ao Item.

ABSTRACT

The production and discussion of contextual indicators in education is a topic of great importance seeing that from them it is possible to develop public policies to improve the educational system. This study highlights the importance of socioeconomic level measurement in the evaluation of education systems, so that by checking the inequalities between the various public attending basic education in Brazilian schools can develop policies to minimize the impact of origin social students in their school life. This work suggests procedures for the construction of indicators using item response theory (IRT) from creating a family Socioeconomic level measure using questionnaires, in 2013, given to Enem and student participants Saeb 2013. The study was conducted from the responses of 6.273.995 students and using initially twenty-four issues. A set of IRT models, polytomous and dichotomous nature, were studied in order to determine the best fit to the data. Therefore, the polytomous models nominal responses, the gradual responses from Samejima and the generalized partial credit, as well as the dichotomous model of two parameters was tested. The tests showed that the model of gradual responses was the model that best fit the data, giving rise thus a more informative measure. It was verified the assumptions of this model, as invariance and one-dimensionality, which converged on a final model with seventeen questions. To validate the Socioeconomic Level Indicator (Inse) calculated crossed the measure with information such as rural and urban location and administrative dependence and correlating it with other socioeconomic measures or related to construct. Finally, using the parameters of the probability calculated by the model for various Inse values, we divided the socioeconomic scale using cluster statistical methods which resulted in nine levels Inse and these levels were characterized by positioning issues and their alternatives from their probabilities.

Keywords: Contextual indicators. Socioeconomic level. Item Response Theory.

LISTA DE GRÁFICOS

Gráfico 1. Curva de Informação.....	48
Gráfico 2. Screeplot.....	66
Gráfico 3. Comportamento diferencial do item.....	70
Gráfico 4. Curva característica questão dicotômica.....	78
Gráfico 5. Curva característica questão politômica.....	78
Gráfico 6. Curva de Informação da questão com baixa informação	79
Gráfico 7. Curva de Informação da questão com alta informação ..	79
Gráfico 8. Curva de Informação do Teste.....	80
Gráfico 9. Curva de informação do teste e erro padrão.....	80
Gráfico 10. Screeplot – Questões do 5º ano do EF Saeb.....	95
Gráfico 11. Screeplot – Questões do 9º ano do EF Saeb.....	95
Gráfico 12. Screeplot – Questões da 3ª série do EM Saeb.....	95
Gráfico 13. Screeplot – Questões da 3ª série do EM Enem.....	96
Gráfico 14. Curva de Informação do Instrumento dos três modelos	104
Gráfico 15. Curva de Informação e erro padrão do conjunto de questões.....	119
Gráfico 16. Inse por Dependência Administrativa.....	125
Gráfico 17. Inse por Localização.....	125
Gráfico 18. Inse por Turno Saeb.....	126
Gráfico 19. CCI Questão Q20 destacado os pontos de máximo.....	129
Gráfico 20. Ponto de máxima probabilidade das questões e suas alternativas, por Inse.....	130
Gráfico 21. Dendograma.....	135
Gráfico 22. Distribuição dos Alunos em 7 grupos de Inse.....	137
Gráfico 23. Distribuição dos alunos em 9 grupos do Inse.....	142

LISTA DE ESQUEMA

Esquema 1. Relações entre pares de questões Índice Q3 primeira calibração.....	108
Esquema 2. Relações entre pares de questões Índice Q ₃ segunda calibração.....	109

LISTA DE ILUSTRAÇÕES

Ilustração 1. Comparação gráfica de curvas características dos itens	48
Ilustração 2. Comparação gráfica das curvas de informação.....	49
Ilustração 3. Desenho amostral.....	65
Ilustração 4. CCI questões relacionadas para serem fixadas	105
Ilustração 5. CCI questões com parâmetros estimados livremente	106
Ilustração 6. <i>Screepplot</i> comparativo 5º ano do Ensino Fundamental Saeb	116
Ilustração 7. <i>Screepplot</i> comparativo 9º ano do Ensino Fundamental Saeb	116
Ilustração 8. <i>Screepplot</i> comparativo 3ª série do Ensino Médio Saeb	117
Ilustração 9. <i>Screepplot</i> comparativo 3ª série do Ensino Médio Enem	117
Ilustração 10. Curva de Informação por questão.....	119
Ilustração 11. Curva característica por questão.....	121

LISTA DE QUADROS

Quadro 1. Comparativo das questões para cálculo Inse, Saeb e Enem	58
Quadro 2. Reagrupamento das questões para ajuste modelo dicotômico	101
Quadro 3. Comparativo questão freezer Saeb e Enem	107
Quadro 4. Opções de agrupamentos por meio de métodos hierárquicos	136
Quadro 5. Correspondência nomes grupos	140
Quadro 6. Descrição dos Níveis.....	140

LISTA DE TABELAS

Tabela 1. Quantitativo de alunos nos Microdados e Censo Educacional 2013	56
Tabela 2. Quantitativo de alunos por avaliação/exame.....	86
Tabela 3. Porcentagens de respostas por questão e avaliação/exame	87
Tabela 4. Correlação policórica entre questões, 5º ano do EF Saeb	92
Tabela 5. Correlação policórica entre questões, 9º ano do EF Saeb	92
Tabela 6. Correlação policórica entre questões, 3º série do EM Saeb	93
Tabela 7. Correlação policórica entre questões, 3ª série do EM Enem	93
Tabela 8. Cargas fatoriais do 1º fator por grupo.....	94
Tabela 9. Parâmetros primeira calibração modelo nominal	97
Tabela 10. Parâmetros segunda calibração modelo nominal	98
Tabela 11. Parâmetros e erro padrão terceira calibração modelo nominal	98
Tabela 12. Parâmetros primeira calibração modelo respostas graduais.....	99
Tabela 13. Parâmetros primeira calibração modelo de crédito parcial generalizado.....	100
Tabela 14. Parâmetros primeira calibração modelo dicotômico ...	102
Tabela 15. Índice Q_3 , 5º ano do Ensino Fundamental Saeb	110
Tabela 16. Índice Q_3 , 9º ano do Ensino Fundamental Saeb	110
Tabela 17. Índice Q_3 , 3º série do Ensino Médio Saeb.....	110
Tabela 18. Índice Q_3 , 3º série do Ensino Médio Enem	111
Tabela 19. Parâmetros questões	112
Tabela 20. Parâmetros das questões – Calibração Final	113
Tabela 21. Cargas Fatoriais do 1º fator por grupo no modelo final	115
Tabela 22. Autovalores modelo inicial e final	117
Tabela 23. Quantitativo de alunos selecionados para cálculo do Inse	123

Tabela 24. Estatísticas Descritivas Inse por Avaliação/Exame	124
Tabela 25. Correlação Inse e medidas do Atlas de Desenvolvimento Humano.....	127
Tabela 26. Probabilidade por alternativa da questão Q20.....	128
Tabela 27. Probabilidade por alternativa das questões para diversos valores de Inse	132
Tabela 28. Probabilidade média por questões/alternativas nos grupos	138

SUMÁRIO

1 INTRODUÇÃO	25
1.1 CONTEXTUALIZAÇÃO	25
1.2 PROBLEMA DE PESQUISA	29
2 OBJETIVOS	31
2.1 OBJETIVO GERAL	31
2.2 OBJETIVOS ESPECÍFICOS	31
3 REVISÃO DA LITERATURA	33
3.1 NÍVEL SOCIOECONÔMICO	33
3.2 O NÍVEL SOCIOECONÔMICO E A TEORIA DE RESPOSTA AO ITEM	34
3.3 A TEORIA DE RESPOSTA AO ITEM - MODELAGEM	42
3.4 TEORIA DE RESPOSTA AO ITEM - CONTRIBUIÇÕES	50
4. METODOLOGIA	55
4.1 CONTEXTO DE PESQUISA	55
4.2 INSTRUMENTOS E QUESTÕES	57
4.2.1 Mapeamento das questões	57
4.3 SELEÇÃO DA AMOSTRA	64
4.4 TEORIA DE RESPOSTA AO ITEM	65
4.4.1 Pressuposto da Unidimensionalidade	65
4.4.2 Pressuposto da Independência local	67
4.4.3 Pressuposto da Invariância - DIF	68
4.4.4 Modelos da TRI	70
4.4.4.1 Modelo Dicotômico	71
4.4.4.2 Modelos Politômicos	72
4.4.4.2.1 Modelo de respostas nominais	72
4.4.4.2.2. Modelo de respostas graduais – Samejima	73
4.4.4.2.3 Modelo de crédito parcial e modelo de crédito parcial generalizado	75
4.4.5 Escolha do modelo	76
4.4.6 Ajuste dos modelos	76
4.4.6.1 RMSEA – Raiz do Erro Quadrático Médio de Aproximação	77

4.4.6.2 <i>Análise dos parâmetros dos itens</i>	77
4.4.6.3 <i>Análise gráfica – Curva Característica do Item (CCI)</i>	78
4.4.6.4 <i>Curva de Informação (CI)</i>	79
4.4.7 Estimação do nível socioeconômico	81
4.4.8 Segmentação e interpretação da escala de Inse	83
5. RESULTADOS	86
5.1 ANÁLISE DESCRITIVA.....	86
5.2 TEORIA DE RESPOSTA AO ITEM	91
5.2.1 Análise de correlação e dimensionalidade	91
5.2.2 Modelagem dos dados	96
5.2.2.1 <i>Modelos Politômicos</i>	96
5.2.2.1.1 Modelo de respostas nominais	97
5.2.2.1.2 Modelo de respostas graduais – Samejima	99
5.2.2.1.3 Modelo de crédito parcial generalizado	100
5.2.2.2 Modelo dicotômico	101
5.2.3 Escolha do modelo	103
5.2.3.1 <i>Curva de Informação (CI)</i>	103
5.2.4 Ajuste do modelo	104
5.2.4.1 <i>Análise do Comportamento Diferencial dos Itens – DIF</i>	105
5.2.4.2 <i>Independência local e unidimensionalidade</i>	107
5.2.4.2.1 Independência local - Índice Q_3	107
5.2.4.2.2 Questões e parâmetros das questões	111
5.2.4.2.3 Unidimensionalidade	114
5.2.4.2.4 RMSEA – Raiz do Erro Quadrático Médio de Aproximação	118
5.2.5 Modelo Proposto	118
5.3 ESTIMAÇÃO DO NÍVEL SOCIOECONÔMICO DOS ALUNOS	122
5.4 O INSE E O PARÂMETRO DAS QUESTÕES.....	128
5.5 O INSE E A INTERPRETAÇÃO DA ESCALA.....	134
6. CONSIDERAÇÕES FINAIS	144
REFERÊNCIAS	146

APÊNDICE A – PARÂMETROS E ERRO PADRÃO MODELO NOMINAL – ESCOLHA MODELO	153
APÊNDICE B – PARÂMETROS E ERRO PADRÃO MODELO DE RESPOSTAS GRADUAIS – ESCOLHA MODELO	155
APÊNDICE C – PARÂMETROS E ERRO PADRÃO MODELO DE CRÉDITO PARCIAL GENERALIZADO – ESCOLHA MODELO.....	156
APÊNDICE D – PARÂMETROS E ERRO PADRÃO MODELO DICOTÔMICO DE DOIS PARÂMETROS – ESCOLHA MODELO.....	157
APÊNDICE E – PARÂMETROS E ERRO PADRÃO MODELO DE RESPOSTAS GRADUAIS – MODELO FINAL	158

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

No âmbito da educação, monitorar o aprendizado dos alunos nas disciplinas como português e matemática, por exemplo, já é uma prática bastante consolidada no Brasil e no mundo. O processo de avaliar é uma prática atual e é pautada por uma exigência da sociedade, pois é uma forma de monitorar o ensino, proporcionando uma educação de qualidade e melhoria constante. Avaliação educacional é um tema que não engloba apenas avaliar o desempenho dos alunos nas disciplinas que são ensinadas, ou seja, não inclui olhar apenas diretamente a aprendizagem. Com isso, é importante também conhecer em que condições esse aprendizado acontece, buscando outros temas que não estejam relacionados com a aprendizagem e que são importantes para formação do aluno como indivíduo.

No Brasil, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) é uma autarquia federal vinculada ao Ministério da Educação (MEC) cuja missão, entre outras, é promover avaliações sobre o Sistema Educacional Brasileiro. Estas avaliações possuem como objetivo subsidiar a formulação e implementação de políticas públicas, bem como produzir informações para o público em geral. Dentre os diversos produtos entregues para sociedade pelo Inep, cabe destacar o Exame Nacional do Ensino Médio (Enem) e o Sistema Nacional de Avaliação da Educação Básica (Saeb).

O Exame Nacional do Ensino Médio (Enem) iniciou-se em 1998 e tinha como finalidade avaliar o desempenho dos estudantes ao final da educação básica, buscando contribuir para a melhoria da qualidade da educação oferecida e a partir de 2009 passou a ser utilizado como mecanismo de seleção para o ingresso no Ensino Superior (Inep, [2009?]). É atualmente utilizado também para o acesso a programas oferecidos pelo Governo Federal, tais como o Programa Universidade para Todos – ProUni, e a cada ano, aumenta a sua aceitação na sociedade e nas universidades.

Outro produto entregue pelo Inep para a sociedade é o Sistema Nacional de Avaliação da Educação Básica (Saeb) que avalia a educação básica brasileira buscando contribuir para a melhoria da qualidade a partir de subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas. Iniciou-se em 1990 com a participação de uma amostra de escolas públicas urbanas e

foram avaliadas as 1ª, 3ª, 5ª e 7ª séries do Ensino Fundamental e atualmente compreende:

- Avaliação Nacional da Educação Básica (Aneb): avaliação amostral dos alunos da rede pública com 10 a 19 alunos em turmas regulares e da rede privada com 10 ou mais alunos em turmas regulares do país, em áreas urbanas e rurais, matriculados no 5º ano e 9º ano do Ensino Fundamental. Alunos das redes privadas, municipais, estaduais e federal com 10 ou mais alunos em turmas regulares na 3ª série do Ensino Médio;
- Avaliação Nacional do Rendimento Escolar (Anresc ou Prova Brasil): avaliação censitária que abrange as escolas com no mínimo 20 alunos matriculados nas séries/anos avaliados, envolvendo os alunos do 5º ano e 9º ano do Ensino Fundamental das escolas públicas das redes municipais, estaduais e federal; e
- Avaliação Nacional da Alfabetização (ANA): avaliação censitária envolvendo os alunos da 3ª série do Ensino Fundamental das escolas públicas com divulgação para as escolas que tiveram pelo menos dez alunos no momento da avaliação.

O sistema procura também oferecer dados e indicadores que possibilitem maior compreensão dos fatores que influenciam o desempenho dos alunos nas áreas e anos avaliados (Inep, [2015?]). Com isso, além das provas é aplicado aos alunos tanto no Enem como na Aneb e Anresc um questionário contextual que aborda diversos temas relacionados à educação e também sobre escolaridade dos pais, posse de bens, contratação de serviços e renda familiar. As respostas aos questionários são disponibilizadas no site do Inep no formato de microdados. Atualmente a ANA não aplica questionário aos seus alunos.

Franco (2004) comenta sobre a importância das avaliações de larga escala para as políticas educacionais, relaciona as medidas cognitivas e medidas contextuais e reforça que o procedimento de análise estatística exige não só a mensuração de variáveis de interesse direto, como por exemplo, a proficiência dos alunos, mas também um amplo conjunto de variáveis que estejam relacionadas com a variável de interesse. Assim, como a avaliação relaciona construtos, não há avaliação sem boas medidas sobre a escola, professores, direção e corpo discente. O autor ainda ressalta que esses indicadores estão mais diretamente relacionados com a ação de uma política educacional do que

com a proficiência e que uma avaliação da educação pode envolver apenas medidas contextuais.

Os questionários contextuais são instrumentos utilizados em avaliações e exames que coletam informações sobre diversos assuntos relacionados à educação, permitindo que estudos sobre estes apontem para temas de grande interesse para os formuladores de políticas educacionais. Franco (2004) destaca que a obtenção de medidas contextuais é um procedimento complexo e que o campo das medidas sociais está muito pouco desenvolvido no Brasil.

De acordo com Siedenberg (2003), indicadores podem ter além de uma função meramente informativa, também uma função avaliativa, normativa ou decisória, uma vez que servem de referência para implementação de objetivos, metas e estratégias. O Brasil, em um sentido amplo, apresenta uma série de medidas de acompanhamento do desenvolvimento socioeconômico que norteiam a política pública do país. Dentre os principais indicadores, destacam-se, por exemplo, o Produto Interno Bruto - PIB, a renda per capita, o Índice de Desenvolvimento Humano - IDH e o Coeficiente de Gini, fornecendo muitas vezes informações agregadas até para o município. Utilizam-se, também, medidas indiretas e agregadas como as coletadas no Censo Demográfico, por exemplo, média dos anos de instrução das pessoas responsáveis pelos domicílios, proporção de pessoas responsáveis analfabetas e proporção de domicílios com coleta regular de lixo.

No âmbito da educação, recentemente algumas iniciativas estão sendo propostas no que se refere a indicadores contextuais na esfera da escola ou até mesmo aluno, como por exemplo, o indicador de infraestrutura da escola e de nível socioeconômico familiar do aluno. Esses indicadores são apresentados com o objetivo de contextualizar o ambiente escolar em que os alunos estão inseridos e a forma como se dá o aprendizado. Franco (2004) questiona como pensar em políticas públicas sem considerar essas medidas contextuais e esse questionamento se deve a algumas reflexões, como por exemplo, de que existe desigualdade entre as escolas, pois alunos de níveis socioeconômicos distintos frequentam escolas distintas com condições de aprendizado diferentes.

Uma medida contextual que sempre se destaca quando se discute educação é o nível socioeconômico familiar dos alunos, isto porque, segundo Buchmann (2002) essa medida determina fortemente os resultados educacionais. Desta forma, estudar a forma como o aprendizado dos alunos se dá e suas relações sem considerar o contexto

socioeconômico familiar desses alunos pode levar à resultados não fidedignos.

O Inep, em 2014, pela primeira vez em suas avaliações divulgou um indicador de nível socioeconômico das escolas considerando informações coletadas no exame Enem e nas avaliações Aneb e Anresc. O objetivo deste indicador foi contextualizar, permitindo a análise dos resultados da escola e a comparação com outros estabelecimentos com perfil parecido. A primeira divulgação foi feita na ANA e em seguida o indicador das escolas foi apresentado com os resultados da Prova Brasil 2013 e Enem por escola 2013 (Inep, [2014?]).

Também, no que se refere à medida contextual nível socioeconômico familiar, já existem tentativas estaduais de utilizar esse construto em políticas públicas. Por exemplo, a Secretaria de educação de São Paulo com o objetivo de atender ao “Plano de Metas Compromisso Todos pela Educação”, estabeleceu metas para o Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp). As metas são de curto e longo prazo para as escolas e, com base no pressuposto de que as condições sociais dos alunos tornam o desempenho cognitivo mais ou menos difícil, em 2013 uma medida de nível socioeconômico foi incorporada ao processo de monitoramento dessa meta. A medida possui neste contexto o papel de calibrar o avanço das escolas em relação ao Idesp indicando para cada escola o grau de influência das condições socioeconômicas sobre o desempenho, tendo em vista a meta proposta (Secretaria de Educação de São Paulo, 2014). O Idesp é um indicador que avalia a qualidade do ensino nos anos/séries iniciais e finais do Ensino Fundamental e do Ensino Médio em cada escola estadual paulista. E o “Plano de Metas Compromisso Todos pela Educação” foi instituído pelo governo federal por meio do decreto nº 6.094, de 24 de abril de 2007, enquanto uma política pública ou uma diretriz política na área educacional, e visa garantir uma Educação Básica de qualidade para todos os Estados brasileiros em 2021.

O Inep calcula bianualmente, para as escolas que fizeram a Anresc, o Índice de Desenvolvimento da Educação Básica – Ideb, e semelhante ao Idesp, tem o objetivo de acompanhar a qualidade da educação do país (Inep, [2007?]). O Ideb é obtido a partir de informações de desempenho dos alunos e rendimento escolar (aprovação) e diferentemente do Idesp não utiliza uma medida de nível socioeconômico. Desde 2014, o Inep calcula uma medida de nível socioeconômico (Inse) médio dos alunos por escola e a partir da experiência adquirida pela Secretaria de Educação de São Paulo, na incorporação de uma medida de nível socioeconômico ao Idesp, o Inep

deveria considerar incorporar também o Inse ao Ideb devido à grande importância do indicador neste âmbito.

Os indicadores de nível socioeconômico calculados pelo Inep e pela Secretaria de Educação de São Paulo são calculados utilizando a Teoria de Resposta ao Item - TRI. Essa teoria é a mesma utilizada nas avaliações para o cálculo das proficiências, porém, atualmente tem sido utilizada, também, para o cálculo de indicadores contextuais devido a sua alta aplicabilidade e relevância. Apesar de ser a mesma teoria, os modelos não são mais modelos relacionados ao acerto e erro (ou sim e não) e sim são utilizados modelos mais complexos relacionados à gradação da medida. Com isso, e devido a grande relevância dos indicadores contextuais no processo educacional, faz-se imprescindível levantar os procedimentos necessários para a criação de um indicador contextual a partir da TRI.

1.2 PROBLEMA DE PESQUISA

A criação de uma medida de contextualização, a partir da TRI, relacionada ao nível socioeconômico dos alunos permite aprofundar o debate educacional e levanta uma série de discussões e possibilidades, além do desenvolvimento de políticas públicas. Com isso, a medida permite contextualizar os resultados de desempenho tornando os resultados mais descritivos e possíveis de ser entendidos a partir do panorama de uma sociedade muito desigual, como a brasileira.

Devido a grande relevância dos indicadores contextuais no processo educacional, quais são os procedimentos mais adequados para criar um indicador válido e fidedigno utilizando a Teoria de Resposta ao Item, tomando como objeto de estudo o nível socioeconômico familiar dos alunos?

2 OBJETIVOS

2.1 OBJETIVO GERAL

Com o intuito de gerar debate e disseminação de conhecimentos específicos acerca do tema estudado, este trabalho propõe-se dissertar sobre a metodologia de cálculo de um indicador de nível socioeconômico, a partir das respostas aos questionários do Enem e Saeb (Aneb e Anresc), explicitando um conjunto de passos que devem ser seguidos na construção de um indicador que utiliza a TRI.

2.2 OBJETIVOS ESPECÍFICOS

Como desdobramento do objetivo geral faz-se necessário pontuar os objetivos específicos desta dissertação, os quais são apresentados a seguir:

- Verificar as características psicométricas dos instrumentos (questionários) Saeb e Enem de 2013 no que se refere ao tema em estudo;
- Analisar e comparar a adequação de alguns modelos da TRI em relação ao construto em questão;
- Analisar os pressupostos da unidimensionalidade, invariância e independência local das questões no modelo selecionado;
- Calcular o Nível Socioeconômico dos alunos a partir do modelo que melhor se adequou aos dados;
- Verificar a consistência do Indicador calculado;
- Propor metodologia de segmentação e descrição da escala.

Este trabalho está organizado em seis seções, incluindo introdução, objetivos, revisão da literatura, metodologia, resultados e considerações finais. Na próxima seção será apresentada a revisão de literatura das teorias que fundamentam a definição do construto nível socioeconômico e da TRI, utilizada como base para este trabalho.

3 REVISÃO DA LITERATURA

Nesta seção será apresentada a revisão de literatura no que se refere aos principais temas relacionados a dissertação, são eles: Nível Socioeconômico, o Nível Socioeconômico e a TRI, a TRI - pontos importantes sobre a modelagem e por último contribuições da TRI na construção de indicadores.

3.1 NÍVEL SOCIOECONÔMICO

O nível socioeconômico, segundo Alves e Soares (2009), é definido como uma medida observada indiretamente, um construto teórico, que é calculada operacionalmente pela agregação de informações sobre: a educação, a ocupação e a riqueza ou rendimento dos indivíduos. Os autores ressaltam que não há consenso na literatura sobre sua conceituação, bem como sobre a forma de medir os construto nas pesquisas empíricas e que a definição de uma medida de nível socioeconômico depende não só das opções teóricas do pesquisador, mas também da disponibilidade de dados adequados para esse tipo de análise.

De acordo com Buchmann (2002), o *status* socioeconômico familiar possui três componentes: educação dos pais, ocupação dos pais e renda familiar. A autora resalta que informações confiáveis sobre renda são difíceis de serem obtidas e, por essa razão, vários pesquisadores optam por outras medidas, tais como *proxies* para riqueza familiar, por exemplo, índices de bens domésticos ou características estruturais da casa.

A medida de nível socioeconômico está pautada em diversas áreas do conhecimento, como por exemplo, na área da saúde - a configuração espacial da dengue no contexto socioeconômico, na agricultura - perfil socioeconômico dos produtores de café orgânico do norte do estado, na economia - o cooperativismo e desenvolvimento socioeconômico de uma cooperativa de crédito. Em geral, indicadores de nível socioeconômico são utilizados para caracterização do público estudado ou como variável controle nos estudos de associação.

No Brasil, uma das medidas mais utilizadas para representar a condição socioeconômica, que segue uma linha semelhante ao desse estudo no que se refere aos dados coletados, é proposta pela Associação Brasileira de Empresas de Pesquisa (ABEP), o Critério de Classificação Econômica do Brasil. Esse critério tem como finalidade estimar o poder de compra das pessoas e famílias urbanas classificando as famílias em

classes econômicas. O instrumento é composto por questões sobre posse de bens e grau de instrução do chefe familiar. A escala é produzida a partir de uma amostra coletada em que é atribuída uma pontuação para cada categoria de resposta e por fim é somada a pontuação dos indivíduos pesquisados, situando os indivíduos em classes A1, A2, B1, B2, C1, C2, D e E (ABEP, 2014). Os questionários do Saeb e Enem apresentam algumas questões iguais às coletadas no instrumento da ABEP.

No âmbito da educação o construto nível socioeconômico tem grande importância. Buchmann (2002) relata vários estudos que apesar de usarem diferentes métodos, concluíram que o status socioeconômico determina fortemente os resultados educacionais e devido a isso, pesquisadores se interessam em uma medida do nível socioeconômico para isolar os fatores externos e assim verificar a importância nos resultados dos fatores escolares.

Na mesma linha de Buchmann (2002), Soares e Andrade (2006) destacam que no Brasil não se deve estudar a realidade educacional sem considerar o nível socioeconômico dos alunos. Os autores informam que importantes estudos empíricos realizados entre as décadas de 50 e 60 destacaram os fatores extraescolares como sendo os mais importantes para a explicação dos resultados escolares e que atualmente reconhece-se que os fatores que determinam o desempenho cognitivo do aluno pertencem a três grandes categorias: a estrutura escolar, a familiar e características do próprio aluno. Alves e Soares (2009) reforçam essa ideia, afirmando que diferentes hierarquias sociais estão associadas às oportunidades educacionais, às trajetórias ocupacionais, ao acesso aos bens e serviços, ao prestígio social, ao comportamento político e social, entre outras.

3.2 O NÍVEL SOCIOECONÔMICO E A TEORIA DE RESPOSTA AO ITEM

No Brasil, uma das medidas mais utilizadas para representar a condição socioeconômica é proposta pela ABEP, o Critério Brasil. Segundo Soares, Souza e Pereira (2004), uma medida calculada pela TRI se diferencia de um indicador como o Critério Brasil por discriminar melhor os indivíduos, pois consegue retirar mais informação dos dados, possibilitando, ainda, melhor conhecimento sobre cada item que compõe o critério, conduzindo então a uma interpretação mais adequada do índice produzido. Outra questão destacada pelos autores, é que quando existem dados faltantes não é possível traçar o estrato

socioeconômico do indivíduo pelo Critério Brasil, situação possível de ser calculada utilizando a TRI em caso de dados faltantes. Esse panorama de trabalhar com dados faltantes é um cenário comum quando se aplica questionário para estudantes dado que o preenchimento total ou parcial do questionário não é obrigatório.

Outra situação, também, que não seria possível a partir da metodologia adotada pelo Critério Brasil, seria a de utilizar questões que não estão presentes em um instrumento e presente em outro para a criação do indicador, por exemplo, situação possível utilizando a TRI. Esta contribuição, utilizar questões de um instrumento que não estão presentes em outro, no caso da produção de estatísticas oficiais é um grande avanço metodológico, pois, permite que uma escola tenha apenas um indicador de nível socioeconômico que represente o perfil familiar do seu alunado, por exemplo, e não um indicador de nível socioeconômico para cada avaliação/exame ou etapa avaliativa.

A utilização da TRI para cálculo de indicadores contextuais tem ganhado bastante espaço no Brasil, de modo que, o destaque não é mais apenas para a utilização do método nos cálculos de proficiência e não se limita mais apenas no ramo da educação. Soares (2005) em seu trabalho destaca a utilização da TRI em diversas áreas do conhecimento, como nas ciências sociais, ciências humanas e ciências da saúde, ressaltando que nestas áreas é comum o interesse em avaliar determinados construtos teóricos, que se caracterizam por serem variáveis latentes (não diretamente observadas), como por exemplo, a habilidade cognitiva de um aluno em língua portuguesa, a satisfação no trabalho, a ansiedade, a depressão e a condição socioeconômica.

No que se refere ao nível socioeconômico, nesta seção apresenta-se uma série de trabalhos que utilizam a TRI para o cálculo deste indicador e suas principais considerações partindo da definição de Alves e Soares (2009) de que nível socioeconômico é definido como uma medida observada indiretamente.

Com base nesse contexto, o trabalho de Soares (2005) propõe três indicadores de condições socioeconômica familiar dos alunos da 4ª e 8ª série (atual 5º e 9º ano) do ensino fundamental e da 3ª série do ensino médio das escolas estaduais (e, também, algumas municipais) do estado de Minas Gerais. Essas escolas participaram do SIMAVE (Sistema Mineiro de Avaliação Educacional) 2002 e para criação desse indicador, o autor utiliza modelos da TRI. O trabalho tem como objetivo, entre outros, comparar os resultados utilizando os modelos da TRI (modelo dicotômico de 2 parâmetros e o modelo de respostas graduais de Samejima (1969)). Estes modelos, segundo Andrade, Tavares e Valle

(2000), procuram representar a probabilidade de um indivíduo dar uma resposta a um item em função dos parâmetros do item e da característica que está sendo medida (traço latente). Soares (2005) ressalta a importância de selecionar apenas as variáveis realmente necessárias, pois com isso, reserva espaço no questionário para outras questões mais relevantes. O indicador é calculado inicialmente a partir das respostas dadas aos questionários sobre: se o aluno trabalha, posse de bens na residência (televisão em cores, geladeira, máquina de lavar roupa, aparelho de som, videocassete, freezer, telefone fixo, telefone celular, computador e automóvel), infraestrutura disponível na residência (água encanada, eletricidade, calçamento e coleta de lixo), presença ou quantidade de determinadas características no domicílio (cozinha, banheiro, sala e quarto) e escolaridade do pai ou responsável, totalizando 20 questões. Após a análise de unidimensionalidade, que é a forma de verificar se os itens utilizados no modelo medem apenas um traço latente, ou seja, nesse caso, o nível socioeconômico, optou-se por retirar itens sobre trabalho do aluno, infraestrutura disponível na residência e as sobre características do domicílio. Assim, o autor calcula os três índices utilizando os seguintes métodos: índice_1 - dicotomização direta das questões e modelos binários de dois parâmetros; o índice_2 - dicotomização das opções de resposta de cada item, selecionando-se as mais informativas, e modelos binários de dois parâmetros; e, o índice_3 - modelos para respostas graduadas. Finalmente, concluiu que do ponto de vista da informação do índice deve-se utilizar o índice_3, mas ressalta que esse argumento não invalida o emprego dos dois outros índices.

Soares e Andrade (2006) também propõe uma medida de nível socioeconômico (NSE) das escolas de ensino fundamental e médio de Belo Horizonte, Minas Gerais e contribuem para as discussões sobre o efeito da escola. Os autores destacam que no Brasil não se deve estudar a realidade educacional sem considerar o nível socioeconômico dos estudantes nem como os diferentes estabelecimentos tratam as diferenças entre grupos de alunos. O indicador é definido a partir das respostas dos alunos, que participaram do SIMAVE 2002 e dos vestibulares da UFMG em 2002, 2003 e 2004. As questões incluídas no estudo são sobre a posse de bens (televisão em cores, televisão a cabo, telefone fixo ou celular, geladeira duplex, freezer, videocassete, máquina de lavar roupa, microcomputador, aspirador de pó e automóvel), característica do domicílio (sala e banheiro na residência), contratação de serviços (empregada doméstica), ocupação (ocupação por grandes grupos dos pais), escolaridade (escolaridade dos pais ou

responsáveis) e renda familiar (em intervalos de salários mínimos), além de considerar também o turno e a rede (pública ou privada) dos alunos. Os autores utilizaram o modelo politômico de respostas graduais de Samejima (1969) para calcular o indicador e para analisar a qualidade dos itens verificaram a curva característica de cada item. Após a construção do indicador dos alunos, calculou-se o nível socioeconômico médio de cada escola da base de dados e em seguida utilizaram o método denominado *k-means*, fixando o número de grupos em cinco para posicionar as escolas em cinco grupos homogêneos. Segundo Cizek (2001), a utilização de métodos de agrupamentos para definição dos níveis é comumente utilizado, no *k-means* o número de grupos é definido previamente e os indivíduos são classificados nos *cluster* de modo que estes são mais semelhantes dentro dos grupos e diferentes entre os grupos. Esses *clusters* permitem ser ordenados dos que apresentam menores valores para os de maiores valores do traço latente, além disso, a média dentro dos grupos representa um perfil médio de cada cluster e a partir dessa média especialistas podem descrever cada grupo. Por fim, os autores calcularam medidas de equidade e qualidade a partir de um modelo de regressão hierárquica utilizando como nível 1 o NSE do aluno e nível 2 o NSE da escola, sendo a variável resposta a proficiência dos alunos.

Alves e Soares (2009) também utilizaram o modelo de Samejima (1969) como técnica de agregação de variáveis à uma variável latente. A medida de nível socioeconômico foi calculada com base nos dados, respostas de 4.636 alunos, da Pesquisa Geração Escolar (GERES) do Pólo Belo Horizonte. Foram utilizadas informações de educação (escolaridade da mãe), ocupação (ao responsável com maior status socioeconômico - principal ocupação e o que faz em sua principal ocupação) e renda (medida indiretamente a partir do local de residência do aluno - CEP). Os autores concluíram que o nível socioeconômico calculado, a partir dessas informações, discrimina muito bem as famílias ao longo da escala. Destacam que, a partir da metodologia adotada, educação, ocupação e renda perdem seu status teórico quando incorporadas em uma medida única. Entretanto, os autores indicam que para estudos que objetivam isolar o efeito de variáveis externas às escolas a medida tem sua utilidade e também como vantagem da técnica, a possibilidade de calcular a medida mesmo com a ausência de resposta que comumente ocorre na coleta de informações das variáveis isoladas. Outra contribuição, no que se refere a forte associação entre as variáveis, que poderia ser considerada prejudicial ao indicador por considerar mais de uma vez uma mesma característica, porém, os

autores destacam que a técnica usada para agregação dos indicadores considera naturalmente as associações entre os indicadores e mantém na medida final apenas a parte específica de cada um.

Soares e Alves (2013) contribuem com outra medida de nível socioeconômico utilizando dados da Prova Brasil nos anos 2005, 2007, 2009 e 2011. A medida é calculada a partir das respostas dos alunos ao questionário, utilizando as questões sobre nível de escolaridade dos pais, posse de bens de consumo duráveis e a contratação de serviços domésticos aplicando um modelo da TRI. No que se refere ao modelo, os autores não explicitam no artigo qual foi utilizado, e apenas fazem referência ao trabalho feito em 2009, quando utilizaram o modelo de Samejima (1969). O indicador de nível socioeconômico nesse trabalho foi utilizado como um dos componentes para realizar o estudo que mede o efeito de escolas e municípios na qualidade do ensino fundamental. Para realizar o estudo, eles compararam os efeitos das escolas e dos municípios avaliados com outros indicadores de qualidade educacional: Ideb (Índice de Desenvolvimento da Educação Básica); gastos municipais por aluno; e a eficiência das redes de ensino públicas dentro dos municípios. Os autores concluem, entre outras, que o uso do Ideb como única medida de qualidade do sistema educacional privilegia os municípios e escolas com alunos de melhor nível socioeconômico.

Já Costa (2010), propõe uma medida de nível socioeconômico (NSE) utilizando os dados dos alunos concluintes do Enem de 2006 a 2008. Propõe comparar as escolas que são similares entre si e estuda a relação do indicador de NSE médio das escolas com o desempenho médio dos alunos. Segundo o mesmo autor, a criação de um indicador de nível socioeconômico utilizando a TRI se motiva devido ao número de reclamações por parte das escolas que são comparadas com outras escolas não similares, dado que diversas ações são estabelecidas com base no desempenho dos alunos, com impacto inclusive financeiro. O autor construiu a medida em etapas, utilizando o modelo de Samejima (1969), primeiro calibrou os itens com os dados do Enem 2006, em seguida com os dados do Enem 2007 e por fim com a edição de 2008 e verificou que os parâmetros dos itens não oscilaram muito entre as edições. Com base nos resultados apresentados, o autor optou por estimar o NSE com as três edições juntas e por fim calculou o NSE médio das escolas pertencentes a estes bancos de dados. Foi adotada uma métrica com média 500 e desvio padrão 100, esta prática é bastante comum, uma vez que a métrica disponibilizada pelos softwares, por conveniência matemática, é com média 0 e desvio padrão 1. Nas métricas com média igual 0, alguns indivíduos apresentam NSE

negativo, situação que dificulta o entendimento das pessoas no que se refere a medida. Para construir a medida, o autor utilizou itens relacionados a: educação dos pais (escolaridade do pai e mãe), renda familiar (faixas de salário mínimo), ocupação dos pais (área de atuação) e itens de conforto doméstico (DVD, rádio, computador, automóvel, geladeira, telefone fixo, celular, internet, televisão, máquina de lavar e televisão a cabo). Inicialmente Costa (2010) propôs também considerar questões sobre cargo de ocupação dos pais e se o aluno trabalhou durante o ensino médio, entretanto, observou que essas questões não forneceram informações consistentes para a construção do nível socioeconômico. Após a verificação psicométrica dos itens e o cálculo da medida por aluno, procedeu-se calculando o NSE médio dos alunos de cada escola e classificando as escolas em 10 grupos com amplitude de $\frac{1}{2}$ desvio padrão. Por fim, o autor conclui que a metodologia proposta para comparar os desempenhos das escolas similares traz grandes contribuições, sobretudo em relação à simplicidade do procedimento.

Alves, Soares e Xavier (2014) propõem outra forma de medir o nível socioeconômico das escolas de educação básica do Brasil, utilizando o modelo de Samejima (1969), da TRI. Eles ressaltam que a construção do índice das escolas é uma contribuição à uma demanda que surge naturalmente, a partir da divulgação dos resultados educacionais por estabelecimentos de ensino. Neste estudo, foram utilizadas as respostas de 20.806.062 indivíduos sobre escolaridade, ocupação dos pais e a renda familiar. Eles consideraram os seguintes microdados: a Prova Brasil (Anresc) dos ciclos referentes aos anos de 2005, 2007, 2009 e 2011; a Aneb de 2003, 2005 e 2011; e o Enem dos anos de 2007, 2008, 2009 e 2011, e com isso destacaram que a utilização dessas bases de dados, no estudo, incluiu um amplo espectro de escolas públicas e privadas. Os autores, para verificar a adequação dos itens ao modelo, calcularam a matriz de correlação policórica dos itens, que mede o grau de associação das variáveis qualitativas ordinais. Verificaram, também, a unidimensionalidade a partir dos autovalores e os autovetores, um algoritmo que baseia-se na matriz de variância-covariância ou na matriz de correlação das variáveis, e em seguida analisaram as curvas características dos itens e as suas respectivas curvas de informação. Com os pressupostos necessários atendidos e os itens selecionados, calculou-se o NSE dos alunos e o NSE médio de cada escola pertencente ao banco de dados em estudo, totalizando 81.140 escolas. Com o objetivo de classificar as escolas em grupos, foram criados sete grupos (mais baixo, baixo, médio baixo, médio, médio alto, alto e mais alto) a partir do NSE médio das escolas, apenas daquelas com 15 ou mais

alunos com medida de NSE calculada. Essa classificação foi feita utilizando a técnica de análise de agrupamentos - método de agrupamento não hierárquico *k-means*. Por fim, para avaliar o NSE os pesquisadores correlacionaram os resultados encontrados com outras medidas de nível socioeconômico por escola, calculada em outras pesquisas, obtendo correlações maiores que 0,80. Com NSE agregado, relacionaram com o rendimento mensal per capita por município, com base no censo populacional 2010, obtendo uma correlação igual a 0,90. Além da validação quantitativa, avaliaram o resultado de forma qualitativa solicitando à educadores de várias partes do país o diagnóstico com base no conhecimento que cada um deles tinha da realidade local, apresentando um resultado segundo os autores bastante fiel. Por fim, os autores concluem que a medida de NSE das escolas possibilita estudos que investiguem as desigualdades educacionais entre alunos, escolas e municípios associadas ao nível socioeconômico e que o índice de NSE produzido reflete uma ordenação em grupos que descreve de forma fidedigna a realidade socioeconômica das escolas estudadas.

Machado (2014) propõe uma medida de nível socioeconômico para diferenciar os alunos do 5º ano do ensino fundamental, a partir do questionário do aluno do Saeb 2011. O autor modelou os dados utilizando o modelo de respostas graduais de Samejima (1969), considerando 11 questões relacionadas a posse de bens familiar (televisão em cores, rádio, videocassete ou DVD, geladeira, freezer junto a geladeira, máquina de lavar roupa, carro e computador), característica do domicílio (sala e quartos para dormir) e contratação de serviços (empregada doméstica). Após a criação da medida do aluno, a escala com média 100 e desvio-padrão 15 foi descrita a partir do posicionamento dos itens em quatro níveis, a saber: “menor que 85”, “de 85 a 115”, “de 115 a 130” e “maior que 130”. A partir das probabilidades do aluno responder uma alternativa e suas superiores, para cada item, foram identificados os itens âncoras. Com isso, a escala foi descrita a partir do posicionamento dos itens obtendo, por exemplo, no nível 2 a seguinte descrição: “As residências dos estudantes posicionados nesse nível possuem, além dos itens elencados no nível anterior, no máximo: rádio; duas tvs em cores; freezer junto à geladeira; máquina de lavar roupas; e computador”.

Outro estudo utilizando a TRI para calcular uma medida socioeconômica é feita por May (2006). O autor utilizou os dados do primeiro estudo de tendências internacionais em matemática e ciências que foi aplicado em 39 países e que coletava, entre outras, informações

sobre escolaridade dos pais e posse de bens em casa. Diferente das pesquisas encontradas no Brasil, no que se refere a posse de bens em casa, o questionário continha perguntas sobre calculadora, mesa de estudos e dicionário, por exemplo. Uma contribuição interessante desse estudo é que por envolver diversas nações o autor propõe o uso de itens internacionais que possuem comportamento semelhante em todas as nações e itens específicos de cada uma.

Observa-se que a proposta de criar um indicador de nível socioeconômico utilizando a TRI é frequente na literatura. Alves, Soares e Xavier (2014) destacam que uma vantagem da TRI, em relação aos métodos convencionais, é a possibilidade de estimar uma medida de nível socioeconômico mesmo com itens parcialmente comum entre os instrumentos, além de permitir trabalhar com dados faltantes.

A TRI, na área da educação, como técnica para estimar proficiência significou um grande avanço. Inicialmente era utilizado o método clássico chamado Teoria Clássica dos Testes (TCT), nesse método o aluno recebe um ponto quando acerta um item e zero quando erra. Pode-se fazer a analogia do que chamamos de TCT na área da educação para os métodos utilizados na estatística descritiva e que por muitas vezes, por serem mais simples e acessíveis, são utilizados para a criação de indicadores. Isso envolve não só técnicas mais simples, como as médias aritméticas simples e ponderadas, mas também outras técnicas estatísticas exploratórias como análise de regressão e análise fatorial exploratória.

Sartes e Souza-Formigoni (2013) citam Hambleton, Swaminathan e Rogers (1991) destacando seis vantagens da TRI em comparação a psicometria tradicional, são eles: 1) a estimação do traço latente do indivíduo independe da amostra de itens utilizados, 2) a estimação dos parâmetros dos itens independe da amostra de sujeitos utilizadas (significa que mesmo em amostras não representativas os parâmetros serão estimados de forma adequada pela TRI), 3) a TRI permite posicionar os indivíduos e os itens em uma mesma escala, conseguindo por exemplo identificar os itens que melhor descrevem cada nível do traço, 4) os erros de medida não são iguais para todos os indivíduos, 5) pessoas que responderam instrumentos diferentes podem ser comparados, desde que tenha itens comuns, sem ter a necessidade de fazer instrumentos “paralelos” como é o caso da TCT e 6) considera no modelo questões que “não se aplicam” à realidade de um determinado grupo. Em relação a primeira consideração as autoras ainda destacam que a teoria da TRI parte do pressuposto de que o traço latente existe, um traço latente “verdadeiro”, independente dos itens utilizados e que a

única exigência é de que os itens estejam realmente medindo o mesmo construto.

Com base nos trabalhos supracitados, pode-se destacar as potencialidades da TRI na criação de uma medida de nível socioeconômico, são eles:

- Técnica adequada para medir o que não é diretamente observado - traço latente;
- Permite trabalhar com dados faltantes;
- Permite estimar uma medida de nível socioeconômico mesmo com itens parcialmente comuns entre os instrumentos;
- Permite trabalhar inicialmente com variáveis fortemente correlacionadas;
- Permite conhecer características dos itens utilizados e escolher apenas aqueles que contribuíram de forma significativa com o modelo;
- Possibilidade de discriminar (diferenciar) bem os indivíduos/famílias ao longo da escala;
- Permite utilizar a medida como controle em estudos de fatores associados quando o objetivo é isolar o efeito de variáveis externas a escola;
- Possibilita estudos que investiguem as desigualdades educacionais entre alunos, escolas e municípios associadas ao nível socioeconômico;
- Permite calcular uma escala única para avaliações diferentes tendo a escola uma única medida de nível socioeconômico;
- Permite comparar escolas com outras escolas similares a ela de acordo com o indicador de nível socioeconômico e outras características demográficas como localização (urbana e rural), dependência administrativa (federal, estadual, municipal e privada) e até outros indicadores de contexto.

Na próxima seção será apresentada uma breve revisão sobre as principais características da TRI, destacando os modelos e procedimentos que serão empregados neste estudo.

3.3 A TEORIA DE RESPOSTA AO ITEM - MODELAGEM

A TRI surgiu a partir da necessidade de se trabalhar com medidas latentes na psicologia e na educação e atualmente vem sendo utilizada nas diversas áreas do conhecimento, inclusive para o cálculo de

indicadores contextuais. A técnica foi fundamentada a partir de dois postulados: (a) a performance do examinado no teste pode ser predita (ou explicada) por uma série de fatores chamados de traços latentes e (b) a relação entre o desempenho no item e o traço latente pode ser descrita por uma função monotônica crescente, chamada de curva característica do item - CCI, segundo Hambleton, Swaminathan e Rogers (1991). Com base nesses postulados, Andrade, Tavares e Valle (2000) definem a TRI, de forma geral, como um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma resposta a um item em função dos parâmetros do item e do traço latente dos respondentes. Desta forma, quanto maior for o valor do traço latente, maior é a probabilidade de um indivíduo responder positivamente a um item relacionado ao traço medido.

No caso de indicadores contextuais, a proposta não é medir proficiência em um teste, como usualmente é feito no ramo da educação, e sim, por exemplo, o nível socioeconômico dos indivíduos. Como o trabalho envolve a criação de um indicador contextual, a revisão de literatura será feita com base na metodologia que se adequa a esta situação.

O indicador de nível socioeconômico é entendido como um traço latente, isto é, uma medida que não é possível de ser obtida de forma direta, e essa medida será referenciada nas fórmulas com a letra grega θ (*theta*). Dessa forma, se espera que quanto maior for o nível socioeconômico (θ) dos indivíduos, maior é a probabilidade dele responder “sim” para uma questão dicotômica ou marcar uma alternativa de maior gradação no caso de questões politômicas, sendo essas questões associadas ao nível socioeconômico.

Existem vários modelos estatísticos propostos na literatura, segundo Andrade, Tavares e Valle (2000), e estes dependem de três fatores:

1. Natureza do item – que podem ser dicotômicos (duas opções de resposta) ou politômicos (mais de duas opções de respostas);
2. Número de populações envolvidas - apenas uma ou mais de uma; e
3. Quantidade de traços latentes mensurados - modelos unidimensional ou multidimensional.

O parâmetro *theta* (θ), neste trabalho, é calculado em função de dois parâmetros dos itens, são eles:

1. Parâmetro *b*: é o nível socioeconômico necessário ao indivíduo, relacionado a um modelo de probabilidade, de responder “sim”

para uma questão dicotômica ou marcar uma alternativa ou alternativas superiores no caso de questões politômicas.

2. Parâmetro a : é a discriminação, isto é representa a capacidade do item de diferenciar o nível socioeconômico entre os respondentes.

Alguns modelos da TRI pressupõem a unidimensionalidade dos dados ou a predominância de uma dimensão (traço latente). Uma das formas de verificar a unidimensionalidade é avaliando a relação entre as variáveis, por meio da técnica estatística análise fatorial, em que utiliza a matriz de correlações das variáveis.

Outro pressuposto da TRI, de acordo com Andrade, Tavares e Valle (2000), é a independência local ou independência condicional, o qual assume independência entre os itens de um instrumento. Quando o pressuposto da independência entre os itens não é atendido implica dizer, segundo Embreston e Reise (2000), que as respostas dos alunos dependem não apenas do nível socioeconômico, mas também da sua resposta a outras questões. Andrade, Tavares e Valle (2000) citam Hambleton, Swaminathan e Rogers (1991), destacando que a unidimensionalidade implica em independência local, assim ao medir a unidimensionalidade, a independência local também está sendo verificada. Entretanto, nas ciências sociais, de acordo com Schilling (2007), é praticamente impossível conseguir isolar um construto dos demais, com isso, verificando a predominância de um traço latente no instrumento é necessário verificar a independência local entre os itens do instrumento.

O pressuposto da invariância é outro ponto importante a ser avaliado quando se utiliza a TRI. Ele pode ser verificado analisando o comportamento diferencial dos itens (DIF), que significa verificar se indivíduos que estão no mesmo nível socioeconômico apresentam probabilidades diferentes de endossar determinada alternativa. Isto é, se uma mesma questão possui parâmetros diferentes entre grupos, o pressuposto da invariância não está sendo atendido e a questão apresenta DIF, Glas (2010).

Os modelos mais usuais utilizados quando se trabalha com variáveis qualitativas ordinais são os modelos dicotômico (em que as variáveis são dicotomizadas) e politômicos de dois parâmetros, tendo ambos os modelos parâmetros a e b 's dos itens. Esses modelos são especificados a seguir:

1. Modelo dicotômico de dois parâmetros - esse modelo foi proposto inicialmente por Lord (1952) e em seguida

- aperfeiçoado por Birnbaum (1968). Como o próprio nome diz, esse modelo serve para as questões dicotômicas, isto é, aquelas que são corrigidas como certo ou errado ou indicam a presença ou ausência de determinada característica, por exemplo.
2. Modelo politômico de respostas graduais de Samejima (1969) - segundo Linden e Hambleton (2013) foi o primeiro modelo proposto para respostas graduais. Ele é uma extensão ao modelo dicotômico de dois parâmetros e parte dos modelos que se ajustam as categorias ordinais politômicas, em que o foco não é mais um modelo dicotômico, tipo acerto e erro e sim a gradação das respostas. Hambleton, Linden e Wells (2010) destacam que o modelo de Samejima (1969) utiliza um processo em duas etapas para obter a probabilidade de o indivíduo responder uma determinada categoria, já que na primeira etapa se obtêm a probabilidade de o indivíduo responder uma categoria ou superior a ela para um determinado valor do traço latente.
 3. Modelo politômico de crédito parcial generalizado - esse modelo foi proposto por Muraki em 1992. De acordo com Hambleton, Linden e Wells (2010), o modelo é conhecido como um modelo direto, diferente do modelo de Samejima (1969) que é considerado um modelo indireto, isto é, mede em apenas uma etapa a probabilidade de um indivíduo responder a uma determinada categoria. Esse modelo também faz parte dos modelos que se ajustam a itens com categorias ordinais.
 4. Modelo politômico de respostas nominais - é definido como o modelo em que as alternativas não possuem a priori uma ordem determinada, foi proposto por Bock (1972). Thissen, Cai e Bock (2010) destacam três características em que esse modelo é indicado: 1) quando se deseja obter as características dos itens e uma medida a partir de itens puramente nominais; 2) validar empiricamente o ordenamento esperado das alternativas de uma questão; e 3) fornecer um modelo para as respostas aos *testlets* (conjunto de itens que possuem uma referência/estímulo comum, um texto, por exemplo).
 5. Modelo politômico de escala gradual - definido por Andrich (1978), esse modelo é semelhante ao proposto por Samejima (1969), pois também se ajusta às categorias ordinais politômicas. O modelo de escala gradual considera que os escores das categorias são igualmente espaçados, diferentemente do modelo de Samejima (1969). Segundo Andrade, Tavares e Valle (2000), se os itens que compõem o

instrumento tiverem suas próprias categorias de respostas, que podem diferir no número, então esse modelo não é adequado.

Os modelos da TRI que consideram um terceiro parâmetro, que é conhecido como o parâmetro de acerto casual, não são utilizados no contexto de indicadores contextuais, pois, parte-se do pressuposto de que não existe acerto casual para uma medida desse tipo e, por isso, não faria sentido considerar na modelagem tal parâmetro.

Após a escolha do modelo que será adotado para calcular o indicador deve-se optar pelo processo de estimação dos parâmetros dos itens, também conhecido como calibração. Os parâmetros podem ser estimados utilizando, além de outros, o método da máxima verossimilhança marginal ou método bayesiano. Andrade, Tavares e Valle (2000) apresentam um resumo considerando os pontos positivos e negativos dos dois métodos, os quais serão explicados a seguir:

- Máxima Verossimilhança Marginal (MVM): tem como pontos positivos a propriedade assintótica (as estimativas dos parâmetros são consistentes), que uma vez estimado os parâmetros dos itens pode-se estimar as habilidades a partir de métodos convencionais. Já os pontos negativos é que não é possível estimar parâmetro para acerto total ou erro total, pois a função não possui pontos de máximo nessas duas situações; é bastante trabalhoso computacionalmente; necessita-se estabelecer uma distribuição para o parâmetro θ ; apresenta problemas na estimação do parâmetro de acerto casual.
- Bayesiano: os pontos positivos são: definido para qualquer padrão de resposta, e uma vez estimado os parâmetros dos itens, pode-se estimar as habilidades a partir de métodos convencionais. Já os pontos negativos: é mais trabalhoso computacionalmente que o MVM, já que todos os parâmetros dos itens são considerados variáveis aleatórias e, para eles, é necessário estabelecer uma distribuição a priori.

Da mesma forma que existem processos de estimação para os parâmetros dos itens, existem processos de estimação para o traço latente que está sendo medido (θ). Os métodos mais comuns para estimar θ são os métodos da máxima verossimilhança, esperança a posteriori e moda a posteriori. Andrade, Tavares e Valle (2000) apresentam um resumo considerando os pontos positivos e negativos dos três métodos, a seguir:

- Máxima verossimilhança (MV): o ponto positivo, para testes com muitos itens produz estimadores não viciados. E negativo

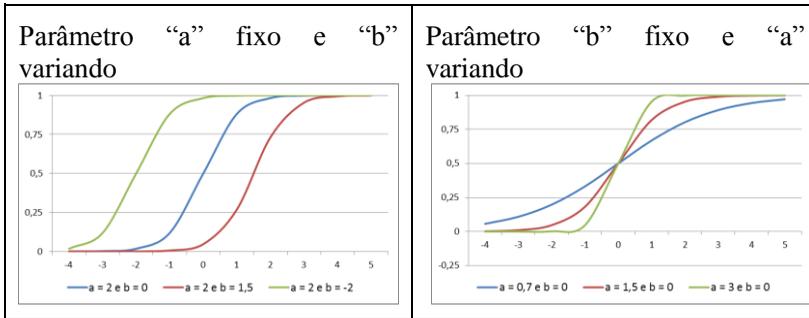
não está definido para alguns padrões de respostas, tais como todas as respostas negativas ou todas as respostas positivas.

- Bayesiano – EAP: a estimação é feita pela média da posteriori, os pontos positivos são: é possível para qualquer conjunto de resposta, possui o menor erro médio e não precisa utilizar métodos iterativos. Já os pontos negativos são: produz estimadores viciados, exige cálculos mais complexos do que o método MV e necessita de uma distribuição a priori para θ .
- Bayesiano – MAP: a estimação é feita pela moda da posteriori, o ponto positivo, é definido para qualquer padrão de resposta. Já os negativos: produz estimadores viciado, necessidade computacional para cálculos mais complexos do que o método MV e necessita de uma distribuição a priori para θ .

Em todo o processo de modelagem é comum ao final avaliar a qualidade do ajuste dos itens ao modelo e ainda quando utilizada a TRI verificar a contribuição de cada item na informação produzida. Isso pode ser feito verificando graficamente comportamento dos itens utilizados e a quantidade de informação fornecida por cada item, após a estimação dos parâmetros dos itens. Com base nisso, são elaborados gráficos que ilustram a características dos itens com base nos parâmetros calculados - Curva Característica do Item (CCI) e gráficos que ilustram a quantidade de informação que cada item fornece ao modelo e em que intervalo do traço latente medido - Curva de Informação do Item (CI).

A CCI graficamente apresenta no eixo horizontal o traço latente (θ) e no eixo vertical a probabilidade. Com isso, é apresentada a curva, a partir dos parâmetros dos itens, e pode-se observar que quanto maior for o parâmetro de discriminação maior é a inclinação da curva e quanto maior for o parâmetro b , maior é o deslocamento da curva a direita. Souza (2014) simula as curvas características fixando dois parâmetros e variando outro em um modelo dicotômico de três parâmetros. A seguir, na Ilustração 1, pode-se verificar o comportamento das curvas, semelhante ao simulado por Souza (2014), porém em um modelo dicotômico de dois parâmetros.

Ilustração 1. Comparação gráfica de curvas características dos itens

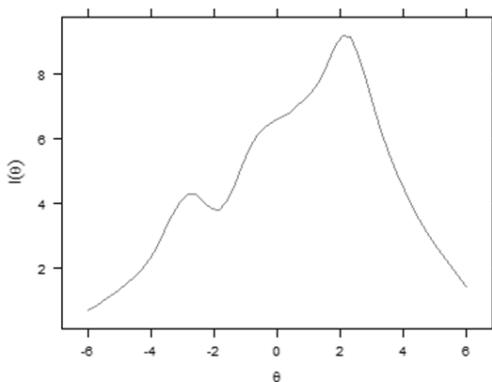


Fonte: Elaborado pela autora

A simulação da CCI para modelos politômicos apresentaria resultados semelhantes quando plotado as probabilidades acumuladas, já que hipoteticamente cada alternativa seria considerada como um item do modelo dicotômico.

Outra medida gráfica comumente utilizada é a Curva de Informação do item (CI) ou do teste. Segundo Andrade, Tavares e Valle (2000), ela permite analisar quanto um item ou teste possuem de informação para determinado intervalo do traço latente (θ) e quanto maior for a informação do item ou teste menor é o erro padrão associado a essa medida. A CI do teste é calculada pela soma das curvas de informação dos itens, a seguir no Gráfico 1 é apresentada a curva de informação de um teste.

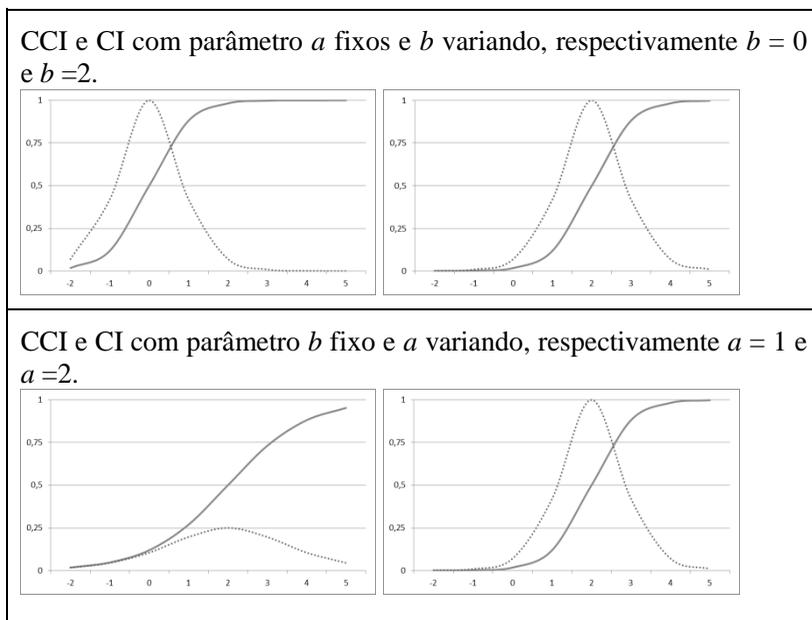
Gráfico 1. Curva de Informação



Fonte: Elaborado pela autora

Na Ilustração 2 é apresentada a curva de informação e a curva característica, respectivamente curva pontilhada e curva contínua, de dois itens. Na primeira parte, variou-se o parâmetro b e na segunda o parâmetro a . Observa-se que quanto maior o parâmetro b mais deslocada à direita encontra-se a CI, já que esta encontra seu ponto de máximo próximo ao parâmetro b . E quanto maior for a discriminação do item maior é o pico da curva em torno do parâmetro b . Segundo Embreston e Reise (2000) nos modelos politômicos os valores do parâmetro a não devem ser interpretados diretamente como discriminação do item, para avaliar diretamente a quantidade de discriminação que um item fornece é necessário avaliar a curva de informação do item.

Ilustração 2. Comparação gráfica das curvas de informação



Fonte: Elaborado pela autora

Como já destacado anteriormente, as análises gráficas das curvas características dos itens e das curvas de informação permitem avaliar a qualidade do ajuste dos dados ao modelo. Soares (2005), em seu trabalho, calcula três índices a partir do modelo dicotômico de dois parâmetros e do modelo politômico de respostas graduais. Para comparar os três índices, o autor compara a curva de informação dos itens e opta pelo modelo que possui maior informação. Entretanto,

ressalta que os três índices são válidos e que estão fortemente associados entre si, com correlações de Pearson maiores que 0,93.

3.4 TEORIA DE RESPOSTA AO ITEM - CONTRIBUIÇÕES

Após todos os procedimentos abordados na seção 3.3 e da definição de qual é o melhor modelo que será utilizado é possível calcular, pela TRI, a probabilidade que cada respondente tem de marcar uma alternativa de um item dado seu Inse (θ). É possível também, conhecer quais são as características dos itens, isto é, quais itens que discriminam mais os indivíduos ou os que são mais difíceis de o indivíduo apresentar aquela característica medida.

Apesar de muitos estudos apresentarem como objetivo final calcular a medida numérica a partir da TRI, para estudos de fatores associados, um dos grandes diferenciais de implementar esta técnica no cálculo de indicadores é que esta permite descrever de forma qualitativa, e não apenas quantitativa, o valor do traço latente atribuído ao indivíduo. É possível, por exemplo, dizer que um indivíduo com o Inse calculado igual a 500 está acima na hierarquia social de um indivíduo com valor 300 e que o primeiro indivíduo possui uma probabilidade maior de contratar empregada mensalista em sua residência e o indivíduo de valor 300, baixa probabilidade de contratar tal serviço.

A descrição de cada valor de Inse, apesar de ser possível, não é viável, dado o quantitativo alto de padrões de respostas, em geral, apresentados em um banco de dados. Com isso, opta-se por fazer cortes na escala criando grupos característicos, de forma que ocorra uma descrição para cada grupo, com base em um valor de referência. Segundo Cizek (2001), as definições para padronização (*standard settings*), ou pontos de corte, são tarefas em que se desejam obter níveis de desempenho, pelos quais as decisões serão tomadas ou pessoas serão classificadas como, por exemplo, em deficientes, proficientes e avançados. Definir a melhor forma de estratificar os dados visando a melhor descrição dos grupos não é uma tarefa elementar.

Desde os anos 50, surgiram várias propostas de como criar pontos de cortes, principalmente para certificações ou classificações de acordo com a proficiência, como os métodos de Nedelsky (1954) - estudante F-D, Angoff (1971) - mínimo aceitável, Ebel (1972) - o minimamente qualificado, o método do grupo limítrofe (1982) e o de contraste de grupos (1982). Segundo Cizek (2001), muitas dessas propostas foram criticadas na literatura por serem muito arbitrárias e foi fortemente recomendado que pontos de cortes fossem evitados de serem feitos,

entretanto a demanda continuava existindo e em algumas situações elas eram necessárias, inclusive por determinação da lei, pois certificavam médicos, advogados, arquitetos, e outros. Segundo o mesmo autor, na década de 70 e 80 a necessidade de definir pontos de corte se intensificou com os testes de competências aplicados nas escolas, que eram definidos também por lei, e com isso vários manuais de “como fazer” foram lançados para os profissionais dos departamentos de educação, que não eram necessariamente especialistas em medição nem familiarizados com a tarefa, definirem pontos de cortes em testes.

Esse processo de classificação de indivíduos, de fato, gera um estresse acadêmico, dado que está sendo feito um juízo de valor, isto é, as pessoas estão sendo classificadas como proficientes e não proficientes, como no caso das certificações, com base em critérios ditos como arbitrários. Desta forma, é necessário haver um elo entre a medida calculada e a teoria sobre o construto medido, para que, com isso, pontos de corte sejam definidos de forma mais científica.

A ideia geral em descrever qualitativamente uma escala pela TRI é baseada em torno de posicionar os questões/alternativas em níveis (faixas de θ) em que os indivíduos estão mais suscetíveis a dar uma determinada resposta em um nível em relação ao anterior.

A seguir será apresentada uma revisão teórica no que se refere a definição de pontos de corte e/ou descrição de escalas utilizando a TRI.

No Brasil, a escala de proficiência mais utilizada é a escala do Saeb. O Saeb utiliza o modelo logístico de três parâmetros da TRI, para respostas corrigidas como “certo” e “errado”, para calcular a proficiência dos alunos do 5º ano do Ensino Fundamental, 9º ano do Ensino Fundamental e 3ª série do Ensino Médio, em Língua Portuguesa e em Matemática, sendo portanto uma escala para cada uma das áreas. Segundo Klein (2009), descrever a escala significa escolher alguns pontos e descrever os conhecimentos e habilidades que os alunos demonstraram ter quando apresentam resultados próximos desses pontos. O mesmo indica que a metodologia das escalas inclui dois procedimentos principais: identificação de itens âncoras e a apresentação desses itens a um painel de especialistas. No artigo, o autor cita duas metodologias, uma utilizada no Saeb 1995 e 1997 e outra utilizada no Saeb 1999 e 2001. A primeira utilizou a metodologia descrita por Beaton e Allen (1992), em que são definidos itens âncoras que caracterizam os pontos de forma que a grande maioria dos alunos situados em um nível acerta o item, e os alunos do nível anterior menos da metade acertam. De forma prática três regras foram estabelecidas para um item ser âncora: que 65% ou mais dos respondentes em torno

do nível acertem o item, que menos de 50% dos alunos posicionados no nível anterior acertem o item e que a diferença entre os percentuais dos que acertaram seja maior que 30%, conceito definido por Beaton e Allen (1992). Klein (2009) destaca alguns problemas em utilizar essa metodologia: 1) para ter muitos itens âncoras é necessário pontos bem espaçados; 2) para ter mais itens para interpretar, é necessário usar o conceito de item quase-âncora; e 3) alguns itens mesmo apresentando boas características não foram utilizados pois não foram considerados âncoras. Já na segunda metodologia, o mesmo autor, destaca que em 1999 e 2001, os intervalos foram definidos considerando um intervalo de 25 pontos na escala do Saeb (média 250 e desvio-padrão 50). Foram utilizados os seguintes critérios para definir um item como âncora: 1) o número de alunos relacionado a cada ponto é maior que 50 alunos; 2) o percentual de acertos no ponto anterior é menor que 65%; 3) o percentual de acertos no ponto é maior que 65% e 4) o ajuste da curva é bom, com essa nova proposta todo item seria considerado âncora.

Soares, Souza e Pereira (2004) propõem uma medida calculada pela TRI e destacam que essa medida se diferencia de um indicador como o Critério Brasil de classificação econômica por discriminar melhor os indivíduos, pois consegue retirar mais informação dos dados, possibilitando, ainda, melhor conhecimento sobre cada item que compõe o critério, conduzindo então a uma melhor descrição do índice produzido. O modelo proposto foi o modelo de respostas graduais para dados politômicos e ao final do estudo eles descrevem a escala, que apresenta valores entre -2 e 2, produzida com o objetivo de, por exemplo, responder a pergunta “O que significa uma família apresentar um escore de 1,38 para o seu nível socioeconômico?”. Para a descrição das escalas foram definidos 10 faixas de *score* de nível socioeconômico, com amplitudes diferentes, sendo os seis intervalos extremos de amplitude 1 desvio-padrão e quatro intervalos centrais com amplitude $\frac{1}{2}$ desvio-padrão. Todos os itens foram utilizados e a sua descrição foi feita com base no percentual de respondentes em relação as alternativas, por exemplo, “possui mais de um item com chance maior que 50% e menor que 70%” e para os itens de escolaridade: “possui primário incompleto”, “possui até colegial incompleto”, “possui superior incompleto” e “possui superior completo”.

Alves e Soares (2009) utilizam os dados da Pesquisa Geração Escolar (GERES) do polo de Belo Horizonte para calcular um indicador de nível socioeconômico (NSE). Após o cálculo da medida do aluno, grupos de alunos foram definidos a partir dos quintis do NSE, isto é, o banco de dados foi dividido em cinco partes de igual tamanho. Em

seguida, descreveram os quintis obtidos a partir da maior frequência empírica de cada questão utilizada, concluindo, por exemplo: em relação ao total 12,7% das mães nunca estudaram ou não terminaram a 4ª série, já considerando apenas os dados do primeiro quintil há quase quatro vezes mais mães nessa categoria (47,3%), nos outros quintis a frequência de mães nessa categoria é muito pequena, quase nula ou ausente.

Os mesmos autores, em 2014, com os pressupostos necessários atendidos para aplicação do modelo, calcularam o NSE dos alunos e o NSE médio de cada escola pertencente ao banco de dados. Com o objetivo de classificar as escolas em grupos, foram criados 7 grupos (mais baixo, baixo, médio baixo, médio, médio alto, alto e mais alto) a partir do NSE médio das escolas, apenas daquelas com 15 ou mais alunos com medida de NSE calculada. Para o agrupamento, utilizaram a técnica de análise de conglomerados - método de agrupamento não hierárquico *k-means*. Nesse estudo os autores não descreveram a escala relacionando a com os itens, apenas determinaram uma ordenação de acordo com NSE médio das escolas.

Mafra (2010) propõe uma sistemática para a modelagem de risco de crédito sob uma abordagem da TRI. Ele utiliza dois modelos da TRI - modelo dicotômico logístico de 2 parâmetros e o modelo politômico de respostas graduais. O autor, calculou o modelo de 2 parâmetros dicotomizando os itens e em seguida selecionou apenas os itens com discriminação maior que 0,7 para modelar no modelo politômico. A escala foi descrita apenas para o modelo politômico e foi transformada para que tivesse média 50 e desvio-padrão 10. Os itens âncoras ou quase-âncoras (dois critérios atendidos) foram definidos a partir de 4 pontos: 30, 40, 50 e 60, utilizando o conceito de Beaton e Allen (1992) de itens âncora. Após a definição dos itens âncora, esses itens foram posicionados em níveis de forma que os pontos representassem o limite inferior de cada nível, por exemplo, o nível 40 é referente ao intervalo de 40 à 49,99. Com os níveis conhecidos, a escala foi descrita por nível como, por exemplo: “as empresas que se localizam no nível 40 caracterizam-se por atingir, além dos itens do nível anterior, a categoria ‘bom’ ou mais elevada dos itens ‘organização da empresa’ e ‘nível de transparência das informações’”. Por fim, Mafra (2010) descreve o percentual de empresas por nível.

Menegon (2013) calcula, duas medidas, uma escala de conforto e uma de desconforto em poltronas de aeronave pela TRI utilizando o modelo de respostas graduais e para descrever a escala, utiliza o conceito de Beaton e Allen (1992) de item âncora. Menegon (2013),

citando Mafra (2010), indica que estas condições são válidas para modelos dicotômicos da TRI, mas podem ser adaptadas aos modelos politômicos, sendo que nesses modelos não verifica-se se um item é âncora, e sim se uma categoria é âncora. Com isso Menegon, para definir os itens âncoras e quase âncoras (dois dos critérios de um item âncora), calcula as probabilidades acumuladas para os pontos 20, 30, 40, 50, 60, 70, 80 e 90 de uma escala com média 50 e desvio-padrão 10. A partir dos pontos, níveis são definidos e são posicionados os itens pertencentes a cada nível e suas categorias de resposta individualmente. Por fim, a autora mostra a distribuição de indivíduos em cada nível da escala obtida.

Em geral, observa-se que os pesquisadores optam em dividir os indivíduos em grupos, mas poucos sugerem uma interpretação qualitativa da medida trabalhada, isso porque boa parte dos trabalhos se propõe calcular os fatores associados ao desempenho e não a relação da medida com os itens. Tal situação se intensifica quando se trabalha com modelos politômicos dada a dificuldade de trabalhar com uma maior quantidade de parâmetros.

Com base no que foi exposto nessa seção, a definição de pontos de cortes e interpretação de escala são temas extensos na literatura, já que não existe apenas uma forma de se fazer, e que já se justificaria, talvez, como tema de uma dissertação de mestrado. Nesse trabalho, uma proposta será implementada na tentativa de diferenciação dos indivíduos estudados.

4. METODOLOGIA

Após a revisão de literatura no que se refere ao construto Nível Socioeconômico e a TRI, nesta seção será implementada uma proposta para a criação do Inse.

A seção metodológica será dividida em quatro etapas: 1) contexto de pesquisa; 2) instrumentos e questões; 3) seleção da amostra, 4) teoria de resposta ao item.

4.1 CONTEXTO DE PESQUISA

Utilizou-se para o cálculo do Inse os bancos de dados do Saeb e Enem, disponíveis no formato de microdados no site do Inep. Estes bancos abrangem de forma geral as respostas dos alunos de escolas públicas e privadas, rurais e urbanas matriculados no 5º ano e 9º ano do Ensino Fundamental e 3ª série do Ensino Médio. A seguir é apresentada de forma detalhada a descrição da população de referência desse estudo:

- Saeb: o questionário do aluno é composto pela Aneb e Anresc. A parte da Aneb é aplicada a uma amostra dos alunos das redes públicas com 10 a 19 alunos na escola e da rede privada com 10 ou mais alunos na escola, matriculados no 5º ano e no 9º ano do Ensino Fundamental e na 3ª série do Ensino Médio regular em alunos da rede pública e privada com 10 ou mais alunos na escola. Já a Anresc é uma avaliação censitária que abrange as escolas com no mínimo 20 alunos matriculados nos anos avaliados, envolvendo os alunos do 5º ano e do 9º ano do Ensino Fundamental regular das escolas públicas das redes municipais, estaduais e federal (fonte: Inep/Saeb). As duas avaliações são aplicadas no mesmo dia e um questionário é aplicado para cada ano. Além disso, na edição de 2013, foi coletado também informações da 3ª série do Ensino Médio Integrado.
- Enem: é um exame voluntário que engloba alunos de diversos perfis: concluintes do Ensino Médio em anos anteriores ao do exame, concluintes do Ensino Médio no ano do exame, concluinte do Ensino Médio após o ano do exame, indivíduos que não concluíram e não estão cursando o Ensino Médio. No Enem foi aplicado um questionário único, on-line, no momento da inscrição. Nos microdados são disponibilizados dois bancos de dados, o primeiro contém, além de outras informações, a declaração do aluno no que se refere a escola que está ou esteve

matriculado e sua situação de conclusão e um segundo banco de dados que contém informações gerais apenas dos alunos concluintes do ensino médio regular seriado, que participaram do exame e estavam matriculados em uma ou mais escolas, de acordo com o Censo Escolar da Educação Básica de 2013, que foram considerados no cálculo do Enem 2013 por escola (Inep, [2014?]). Optou-se por trabalhar apenas com os alunos deste segundo banco, dado que é uma informação originada do Censo Educacional, igual ao do Saeb, e também porque após o cruzamento desses dois bancos observou-se a existência de inconsistência na informação declarada pelo aluno comparada ao Censo Educacional de aproximadamente 5%, tanto para a informação da escola como para a informação da situação de conclusão.

Na Tabela 1, são apresentados os quantitativos de alunos presentes nos microdados e o quantitativo presente no Censo Educacional (Inep, [2014?]), de acordo com o público de referência definido. No total são mais de 6 milhões de registros e sua maior parte são alunos do 5º e 9º ano do Ensino Fundamental e em menor parte por alunos do 3ª série do Ensino Médio. Esses valores apresentados se fundamentam pela avaliação do Saeb ter uma parte censitária para o 5º e no 9º ano EF e a parte do 3ª série EM ser exclusivamente amostral. E no caso do Enem, devido ao recorte utilizado ser de apenas alunos concluintes do Ensino Médio, conforme com o que foi declarado pelas escolas no Censo Educacional de 2013 que fizeram parte da divulgação do Enem por Escola.

Tabela 1. Quantitativo de alunos nos Microdados e Censo Educacional 2013

		MICRODADOS	CENSO	% Cobertura Censo
Saeb	5º ano Ensino Fundamental	2.524.125	2.893.664	87%
	9º ano Ensino Fundamental	2.720.588	3.017.764	90%
	3ª série Ensino Médio	150.429	2.205.287	7%
Enem	3ª série Ensino Médio	878.853	2.278.156	39%
Total		6.273.995	-	-

Fonte: Elaborado pela autora

4.2 INSTRUMENTOS E QUESTÕES

Após a definição da população de referência que será utilizada para o cálculo do indicador, buscou-se conhecer os instrumentos do Saeb e Enem visando selecionar as questões que comporiam o indicador. É muito importante para a criação de um indicador de qualidade, que o processo estatístico esteja alinhado com a definição do que se deseja medir – o construto, isto porque a medida calculada pode não refletir aquilo que está sendo definido, por exemplo, o nível socioeconômico familiar.

Nos instrumentos avaliados, de forma geral, verificam-se dois eixos centrais relacionados às características socioeconômicas, um eixo está relacionado diretamente ao aluno que se expressa em questões como “Você exerce ou já exerceu atividade remunerada?”, “Você cursa ou já cursou a Educação de Jovens e Adultos – EJA?” e “Em que tipo de escola você estudou?” e o segundo eixo são características relacionadas à base familiar.

Serão consideradas para o cálculo do indicador as questões relacionadas às características familiares, entendendo-se que as questões relacionadas aos alunos se expressam decorrente de um suporte socioeconômico familiar. Com isso, selecionaram-se questões que abordam temas relacionados à escolaridade dos pais, posse de bens na residência, características do domicílio, contratação de serviços e renda familiar.

4.2.1 Mapeamento das questões

Como instrumentos diferentes estão sendo utilizados, é natural que as questões não sejam exatamente iguais em sua grafia, que a ordem de apresentação das alternativas não seja as mesmas ou que até mesmo exista uma questão em um instrumento e não em outro. Desta forma, é necessário unificar os bancos de dados e com isso algumas questões podem ter suas alternativas agrupadas ou até mesmo se agruparem com outra questão. Além dessas alterações, é necessário, para utilizar os modelos da TRI e o pacote estatístico escolhido, ordenar as alternativas em ordem crescente do que se está medindo no traço latente, recodificar as alternativas de letras para números e no caso de não resposta ou a resposta “não sei” recodificar para “NA”, isto é, “não se aplica”.

A seguir, no Quadro 1, são apresentadas as questões na forma como foram apresentadas nos questionários de origem e como elas foram tratadas para que pudessem ser modeladas pela TRI.

Quadro 1. Comparativo das questões para cálculo Inse, Saeb e Enem

Inse	SAEB	Enem
	5º ano, 9º ano e 3º série	3º série
Q01 - Televisão em cores em casa	Na sua casa tem televisão em cores?	Você tem em sua casa? TV em cores
1 - Não tem	A - Não tem	A - 1
2 - Sim, uma	B - Sim, uma	B - 2
3 - Sim, duas	C - Sim, duas	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q02 - TV por assinatura em casa		Você tem em sua casa? TV por assinatura
1 - Não tem		A - 1
2 - Sim, uma		B - 2
3 - Sim, duas		C - 3 ou mais
4 - Sim, três ou mais		D - Não tenho
Q03 - Videocassete e/ou DVD em casa	Na sua casa tem videocassete e/ou DVD?	Você tem em sua casa? Videocassete e/ou DVD
1 - Não tem	A - Não tem	A - 1
2 - Sim, um	B - Sim, um	B - 2
3 - Sim, dois	C - Sim, dois	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q04 - Computador em casa	Na sua casa tem computador?	Você tem em sua casa? Microcomputador
1 - Não tem	A - Não tem	A - 1
2 - Sim, um	B - Sim, um	B - 2
3 - Sim, dois	C - Sim, dois	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q05 - Internet em casa		Você tem em sua casa? Acesso à Internet
1 - Não tem		A - 1
2 - Sim, uma		B - 2
3 - Sim, duas		C - 3 ou mais
4 - Sim, três ou mais		D - Não tenho
Q06 - Rádio em casa	Na sua casa tem aparelho de rádio?	Você tem em sua casa? Rádio
1 - Não tem	A - Não tem	A - 1
2 - Sim, um	B - Sim, um	B - 2
3 - Sim, dois	C - Sim, dois	C - 3 ou mais

Inse	SAEB	Enem
	5º ano, 9º ano e 3ª série	3ª série
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q07 - Telefone fixo em casa		Você tem em sua casa? Telefone fixo
1 - Não tem		A - 1
2 - Sim, um		B - 2
3 - Sim, dois		C - 3 ou mais
4 - Sim, três ou mais		D - Não tenho
Q08 - Telefone celular em casa		Você tem em sua casa? Telefone celular
1 - Não tem		A - 1
2 - Sim, uma		B - 2
3 - Sim, duas		C - 3 ou mais
4 - Sim, três ou mais		D - Não tenho
Q09 - Aspirador de pó em casa		Você tem em sua casa? Aspirador de pó
1 - Não tem		A - 1
2 - Sim, um		B - 2
3 - Sim, dois		C - 3 ou mais
4 - Sim, três ou mais		D - Não tenho
Q10 - Geladeira em casa	Na sua casa tem geladeira?	Você tem em sua casa? Geladeira
1 - Não tem	A - Não tem	A - 1
2 - Sim, uma	B - Sim, uma	B - 2
3 - Sim, duas	C - Sim, duas	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q11 - Freezer (aparelho independente ou parte da geladeira) em casa	Na sua casa tem freezer (parte da geladeira duplex)? Na sua casa tem freezer separado da geladeira? (mesmas opções de reposta)	Você tem em sua casa? Freezer (aparelho independente ou parte da geladeira duplex)'
1 - Não tem	A - Não tem	A - 1
2 - Sim, um ou mais	B - Sim, um	B - 2
	C - Sim, dois	C - 3 ou mais
	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	

Inse	SAEB	Enem
	5º ano, 9º ano e 3º série	3º série
Q12 - Máquina de lavar roupa em casa	Na sua casa tem máquina de lavar roupa (O tanquinho NÃO deve ser considerado)?	Você tem em sua casa? Máquina de lavar roupa
1 - Não tem	A - Não tem	A - 1
2 - Sim, uma	B - Sim, uma	B - 2
3 - Sim, duas	C - Sim, duas	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q13 - Automóvel em casa	Na sua casa tem carro?	Você tem em sua casa? Automóvel
1 - Não tem	A - Não tem	A - 1
2 - Sim, uma	B - Sim, um	B - 2
3 - Sim, duas	C - Sim, dois	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q14 - Banheiro em casa	Na sua casa tem banheiro?	Você tem em sua casa? Banheiro
1 - Não tem	A - Não tem	A - 1
2 - Sim, um	B - Sim, um	B - 2
3 - Sim, dois	C - Sim, dois	C - 3 ou mais
4 - Sim, três ou mais	D - Sim, três	D - Não tenho
	E - Sim, quatro ou mais	
Q15 - Quartos para dormir em casa	Na sua casa tem quartos para dormir?	
1 - Não tem	A - Não tem	
2 - Sim, um	B - Sim, um	
3 - Sim, dois	C - Sim, dois	
4 - Sim, três	D - Sim, três	
5 - Sim, quatro ou mais	E - Sim, quatro ou mais	
Q16 - Situação residência		A residência de sua família é
1 - Outra situação (loteamento não regularizado, ocupação etc.)		A - Própria e quitada
2 - Cedida		B - Própria e em pagamento (financiada)
3 - Alugada		C - Alugada
4 - Própria e em		D - Cedida

Inse	SAEB	Enem
	5º ano, 9º ano e 3ª série	3ª série
pagamento (financiada)		
5 - Própria e quitada		E - Outra situação (loteamento não regularizado, ocupação etc.)
Q17 - Localização residência		A residência de sua família está localizada em
1 - Comunidade indígena ou quilombola		A - Zona rural
2 - Zona rural		B - Zona urbana
3 - Zona urbana		C - Comunidade indígena
		D - Comunidade quilombola
Q18 - Quantas pessoas moram na casa (incluindo o aluno)	Incluindo você, quantas pessoas vivem atualmente em sua casa?	Quantas pessoas moram em sua casa (incluindo você)?
1 - Uma, pois moro sozinho(a)	A - Uma, pois moro sozinho(a)	A - 1
2 - Duas	B - Duas	B - 2
3 - Três	C - Três	C - 3
4 - Quatro	D - Quatro	D - 4
5 - Cinco	E - Cinco	E - 5
6 - Seis pessoas ou mais	F - Seis pessoas ou mais	F - 6
		...
		S - 19
		T - 20 ou mais
Q19 - Contrata empregada doméstica em casa	Em sua casa trabalha empregado(a) doméstico(a) pelo menos cinco dias por semana?	Você tem em sua casa? Empregada mensalista
1 - Não	A - Não	A - 1
2 - Sim, um ou mais	B - Sim, um(a) empregado(a)	B - 2
	C - Sim, dois(duas) empregados(as)	C - 3 ou mais
	D - Sim, três empregados(as)	D - Não tenho
	E - Sim, quatro ou mais empregados(as)	
Q20 - Renda mensal da família (considerando a renda do aluno)		Qual é a renda mensal de sua família? (Some a sua renda com a dos seus

Inse	SAEB	Enem
	5º ano, 9º ano e 3ª série	3ª série
		familiares)
1 - Nenhuma renda		A - Nenhuma renda
2 - Até um s.m		B - Até um salário mínimo
3 - Mais de um até 1,5 s.m		C - Mais de um até 1,5
4 - Mais de 1,5 até 3 s.m		D - Mais de 1,5 e até 2
5 - Mais de 3 até 7 s.m		E - Mais de 2 e até 2,5
6 - Mais de 7 até 9 s.m		F - Mais de 2,5 e até 3
7 - Mais de 9 até 12 s.m		G - Mais de 3 e até 4
8 - Acima de 12 s.m		H - Mais de 4 e até 5
		I - Mais de 5 e até 6
		J - Mais de 6 e até 7
		K - Mais de 7 e até 8
		L - Mais de 8 e até 9
		M - Mais de 9 e até 10
		N - Mais de 10 e até 12
		O - Mais de 12 e até 15
		P - Mais de 15 e até 20
		Q - Acima 20 salários mínimos
Q21- Mãe, ou a mulher responsável por você, sabe ler e escrever	Sua mãe, ou a mulher responsável por você, sabe ler e escrever?	
1 - Não	A - Sim	
2 - Sim	B - Não	
Q22- Pai, ou homem responsável por você, sabe ler e escrever	Seu pai, ou homem responsável por você, sabe ler e escrever?	
1 - Não	A - Sim	
2 - Sim	B - Não	
Q23 - Até que série a mãe estudou	Até que série sua mãe, ou a mulher responsável por você, estudou?	Até quando sua mãe estudou?
1 - Nunca estudou.	A - Nunca estudou	A - Não estudou
2 - Ensino Fundamental cursando ou completo	B - Não completou a 4ª série/5º ano	B - Da 1ª à 4ª série do Ensino Fundamental (antigo primário)
3 - Ensino Médio completo	C - Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano	C - Da 5ª à 8ª série do Ensino Fundamental (antigo ginásio)

Inse	SAEB	Enem
	5º ano, 9º ano e 3ª série	3ª série
4 - Ensino Superior Completo	D - Completou a 8ª série/9º ano, mas não completou o Ensino Médio	D - Ensino Médio (antigo 2º grau) incompleto
	E - Completou o Ensino Médio, mas não completou a Faculdade	E - Ensino Médio (antigo 2º grau)
	F - Completou a Faculdade	F - Ensino Superior incompleto
	G - Não sei	G - Ensino Superior
		H - Pós-graduação
		I - Não sei
Q24 - Até que série o pai estudou	Até que série seu pai, ou o homem responsável por você, estudou?	Até quando seu pai estudou?
1 - Nunca estudou.	A - Nunca estudou	A - Não estudou
2 - Ensino Fundamental cursando ou completo	B - Não completou a 4ª série/5º ano	B - Da 1ª à 4ª série do Ensino Fundamental (antigo primário)
3 - Ensino Médio completo	C - Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano	C - Da 5ª à 8ª série do Ensino Fundamental (antigo ginásio)
4 - Ensino Superior Completo	D - Completou a 8ª série/9º ano, mas não completou o Ensino Médio	D - Ensino Médio (antigo 2º grau) incompleto
	E - Completou o Ensino Médio, mas não completou a Faculdade	E - Ensino Médio (antigo 2º grau)
	F - Completou a Faculdade	F - Ensino Superior incompleto
	G - Não sei	G - Ensino Superior
		H - Pós-graduação
		I - Não sei

Fonte: Elaborado pela autora

Na revisão de literatura Soares (2005) utiliza itens relacionados com a infraestrutura disponível na residência (água encanada, eletricidade, calçamento e coleta de lixo), com base nisso, optou-se por iniciar as análises considerando as questões “Q17 - localização

residência” e “Q16 - Situação residência” e verificar a qualidade do ajuste dessas questões no modelo.

É necessário, também, definir o número máximo de não respostas que será aceito no banco de dados para a calibração e para o cálculo do indicador. Acredita-se que se de 24 questões o aluno respondeu apenas uma questão a qualidade da medida proposta para esse aluno não será boa. Não se encontrou na literatura uma proposta de metodologia para definir o número máximo de não respostas, com isso optou-se por considerar aqueles que responderam mais de cinco questões.

Segundo Morettin e Bussab (2004), quando se estuda uma variável, o maior interesse do pesquisador é conhecer o comportamento dessa variável, analisando a ocorrência de suas possíveis realizações. Uma medida útil para avaliar o comportamento das variáveis em estudo é utilizar a porcentagem, assim possibilita a comparação das variáveis já que estas possuem a mesma magnitude.

No que se refere à manipulação dos bancos de dados e criação das tabelas, optou-se por utilizar o software estatístico *SAS Enterprise Guide* versão 4.2 e o Microsoft Excel 2010.

4.3 SELEÇÃO DA AMOSTRA

Os procedimentos que serão descritos para a verificação da correlação e a dimensionalidade dos itens serão feitos com os quatro bancos de dados utilizando toda a população, que totalizam mais de cinco milhões de respostas. Porém, trabalhar com esses quantitativos computacionalmente é bastante oneroso, por isso para os procedimentos relacionados a modelagem optou-se por selecionar uma amostra de 25.000 alunos de cada banco de dados.

A partir da técnica análise fatorial para dados ordinais é possível calcular os escores fatoriais de cada indivíduo, o qual tem o mesmo significado do escore calculado através da análise clássica, com isso o banco de dados foi dividido a partir dos escores fatoriais em quatro partes sendo a primeira referente a 33% dos escores iniciais, a segunda e terceira referente aos seguintes 17% e a última referente aos 33% escores mais altos. Optou-se por utilizar essa técnica para o cálculo dos escores, dado que a técnica permite com que seja respeitada a natureza das questões, que são qualitativas ordinais no cálculo dos escores fatoriais. Esse método permite que a estimação dos parâmetros seja feita com indivíduos que possuam nível socioeconômico diversos, do baixo ao alto. Com isso, uma amostra aleatória estratificada será retirada de acordo com os quantitativos especificado na Ilustração 3, abaixo:

Ilustração 3. Desenho amostral

0%	33%	50%	67%	100%
n = 8.333	n = 4.167	n = 4.167	n = 8.333	

Fonte: Elaborado pela autora

Para retirar a amostra foi utilizado o pacote *Sampling* versão 2.7(MATEI e TILLÉ, 2015) do *software* R (R *CORE TIME*, 2015).

4.4 TEORIA DE RESPOSTA AO ITEM

A TRI, segundo Hambleton, Linden e Wells (2010), faz parte de um grupo de teorias modernas que são caracterizadas por uma forte modelagem de dados, que modelam as respostas dos indivíduos à um instrumento.

Os modelos da TRI que serão utilizados nesse estudo possuem os seguintes pressupostos: unidimensionalidade dos dados, a independência local e, conseqüentemente, a invariância dos dados.

4.4.1 Pressuposto da Unidimensionalidade

Dizer que um instrumento é unidimensional significa dizer que existe apenas um único traço latente sendo medido, significa também dizer que as variáveis são correlacionadas entre si.

Nas ciências sociais é praticamente impossível conseguir isolar um construto dos demais, por isso os estudos de unidimensionalidade buscam a verificação da predominância de um traço latente e não a verificação de apenas um traço. Com isso, Schilling (2007) destaca a necessidade de diferenciar dois conceitos importantes – a unidimensionalidade estrita e a essencial. A unidimensionalidade estrita sugere que qualquer associação entre as questões avaliadas está relacionada a um único fator, já a unidimensionalidade essencial é mais flexível, pois apenas sugere a predominância de um fator.

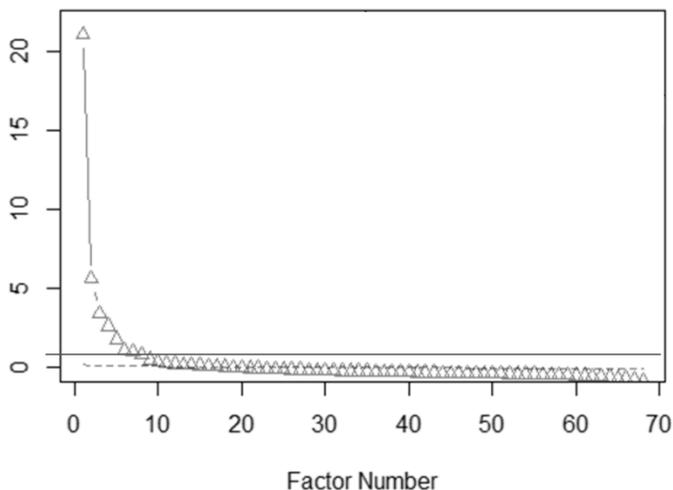
Na literatura existem várias propostas para verificar a unidimensionalidade dos dados e este é um dos pressupostos fundamentais para utilizar os modelos descritos nesse trabalho.

A verificação da unidimensionalidade de um instrumento pode ser feita por meio da técnica estatística análise fatorial, em que se utiliza

a matriz de correlações. Quando as variáveis são qualitativas ordinais, usa-se a matriz de correlação policórica e utiliza-se a análise fatorial para dados com respostas ordinais. Segundo Maroco (2010), utilizar a correlação de Pearson, medida de correlação mais popular entre os pesquisadores, em dados qualitativos ordinais é bastante controverso, pois, são calculadas médias e desvio de uma medida que deveria ser minimamente intervalar. O mesmo autor cita ainda estudos de Bollen (1989), que demonstram que utilizar medidas de associação como a correlação de Pearson em medidas ordinais resulta em valores, em geral, inferiores, quando comparado à utilização da técnica adequada. Bortolotti et al. (2012) em seu trabalho citam o critério de Reckase (1979), o qual sugere a predominância de uma dimensão quando o primeiro fator corresponde a pelo menos 20% da variância total.

Outra forma, também utilizada, é a análise do gráfico *screeplot* em que se avalia os autovalores (total da variância explicada por um fator) e se o primeiro fator se destaca em relação aos demais considera-se que existe uma dimensão predominante. É possível, também, determinar o número de dimensões pelo número de autovalores maiores que 1, isso pode ser verificado também no gráfico *screeplot*, Hair (2005). No Gráfico 2 é apresentado um exemplo de um *screeplot*, o eixo y indica os autovalores, o eixo x o número de fatores e a linha paralela ao eixo x é referente ao autovalor 1.

Gráfico 2. Screeplot



Fonte: Elaborado pela autora

Schilling (2007) destaca que a verificação da unidimensionalidade essencial pode ser feita avaliando os resíduos, após a modelagem dos dados. Calcula-se a raiz da média das correlações residuais quadráticas e quando esse valor é menor que 0,05 e as correlações forem, em geral, menores que 0,10, as questões utilizadas para a modelagem compõem um instrumento essencialmente unidimensional.

No que se refere ao cálculo das matrizes de correlações policóricas e a análise de componentes principais utilizou-se o pacote *psych* versão 1.5.1 (REVELLE, 2015) do *software* livre R (R CORE TIME, 2015) e para a análise dos resíduos o pacote *mirt* versão 1.9 (CHALMERS, 2015) que também foi utilizado para a modelagem dos dados.

4.4.2 Pressuposto da Independência local

A independência local das questões utilizadas na TRI é um dos pressupostos necessários para garantir a validade do modelo. A falta de independência local entre as questões indica que as respostas dos alunos dependem não apenas do nível socioeconômico, mas também da sua resposta a outras questões (Embreston e Reise 2000).

De acordo com Ayala (2009), a dependência local pode acontecer em diversas situações, cabe destacar, principalmente, no que se refere a questionários: questões que se referem ao mesmo tema ou são dependentes de um mesmo enunciado, tamanho do questionário, tempo para preenchimento não adequado e quantitativo baixo de variáveis relacionadas ao traço latente.

Embreston e Reise (2000) citam o estudo empírico de Yen (1993) que demonstra que questões dependentes entre si afetam a estimação da informação do teste e o parâmetro de discriminação das questões. Chen e Thissen (1997) também estudaram o efeito da dependência local na modelagem pela TRI e verificaram que a presença de questões dependentes entre si impactam nas estimação dos parâmetros, inclusive no parâmetro *b*. Os mesmos destacam que existir dependência local entre um par de questões implica em redundância, é como se o instrumento tivesse menos questões do que realmente possui.

Existem várias técnicas na literatura que objetivam verificar se existe independência local entre questões de um instrumento, a seguir será apresentada a estatística Q_3 , disponível no pacote utilizado para modelagem.

Embreston e Reise (2000) indicam a estatística Q_3 de Yen (1984) como um meio de identificar pares de questões que apresentam dependência local. Em consonância com esses pesquisadores, Ayala (2009) destaca que dentre vários índices disponíveis para verificar independência local nenhum pode ser considerado o mais ideal em termos de alta qualidade na identificação da dependência local versus baixo quantitativo de falsos positivos, porém ressalta que o índice Q_3 é bastante razoável. Outra vantagem destacada pelo autor é que este índice não é muito influenciado pelo tamanho da amostra como os demais índices e que é influenciado pela quantidade de questões (instrumentos pequenos – mais valores de Q_3 negativos do que instrumentos grandes).

O índice Q_3 é calculado após a modelagem dos dados e representa a correlação residual entre questões, isto é a correlação após extrair o traço latente medido, Yen (1993). Assim se a correlação é igual a 1 ou -1, as duas questões são totalmente dependentes e caso seja igual a zero não existe dependência linear entre as questões. Segundo o mesmo, espera-se para independência local valores de Q_3 próximos de $-1/(n-1)$, sendo n igual ao número de questões do instrumento e para instrumentos grandes espera-se obter valores de Q_3 próximos de zero. Valores positivos e altos de Q_3 indicam pares de itens que compartilham algum outro fator que pode ser uma causa de preocupação, Yen (1993). Ayala (2009) cita estudo de Yen (1993) que para instrumentos com pelo menos 17 questões um valor de corte razoável é $|Q_3| > 0,2$ indicando dependência local.

Por ser uma medida de correlação, outra forma de utilizar o índice Q_3 é elevando seu valor ao quadrado e, com isso, esse valor representa a quantidade de variância residual compartilhada pelo par de questões e valores acima de 5% podem indicar dependência entre as questões, Ayala (2009). O mesmo autor destaca que a dependência local pode ser explicada também quando existe mais de uma dimensão predominante, isto é a multidimensionalidade.

4.4.3 Pressuposto da Invariância - DIF

Analisar comportamento diferencial dos itens (DIF) significa verificar se indivíduos que possuem o mesmo Inse apresentam probabilidades diferentes de endossar determinada alternativa. Isto é, se uma mesma questão possui parâmetros diferentes entre grupos, o pressuposto da invariância não está sendo atendido e a questão apresenta DIF (Glas, 2010). Com isso, a análise de DIF é feita apenas nas questões

que são comuns entre os grupos estudados e é uma forma de verificar o ajuste do modelo.

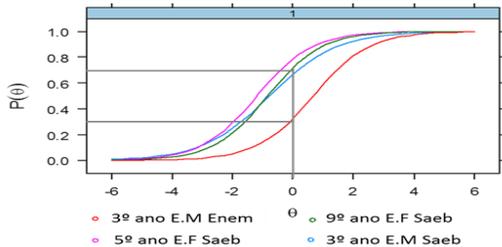
Segundo Andriola (2000), o DIF tem sido estudado principalmente nas investigações acerca das diferentes características sociodemográficas dos grupos, tais como o sexo, idade, classe social, região de moradia e outras. O mesmo autor cita Kim e Baker (1993), indicando os dois objetivos de se medir DIF: detecção do DIF que é quando o objetivo é identificar a possível diferença dos parâmetros dos itens de acordo com os grupos comparados; e verificar o impacto do DIF que significa investigar as causas (psicológicas, educativas, culturais, sociais, entre outras) que ocasionaram o comportamento diferente do item. Neste trabalho o foco será apenas na detecção do DIF entre os grupos: 3ª série Ensino Médio Enem, 5º ano Ensino Fundamental Saeb, 9º ano Ensino Fundamental Saeb e 3ª série Ensino Médio Saeb. A verificação do impacto não será o foco, pois abrange outras áreas de conhecimentos e vai além do escopo definido para este estudo.

É possível ao elaborar questões adotar determinados procedimentos para que se minimize a possibilidade de DIF. Ercikan (1998) destaca as situações que deverão ser evitadas ao se traduzir questões, estas situações podem ser generalizadas para o contexto de questionário de forma a evitar DIF, cabe destacar: termos cujo entendimento está relacionado ao pertencimento de um grupo específico, termos cujo entendimento é diferente dependendo do grupo ou contexto e questões muito longas (sentença). Andriola (2000) cita Muñiz (1997) destacando que não existem instrumentos que não apresentam DIF e que o que objetiva fazer nas análises de DIF é detectar a quantidade de DIF aceitável em uma questão.

Existem diversos métodos para se verificar DIF, Andriola (2000) cita Mellenbergh (1989) indicando que uma das formas de medir DIF é comparar o comportamento das curvas características das questões para o grupo de referência e o grupo focal. Assim, uma questão com comportamento diferencial entre grupos indica que o θ medido não está em função apenas dos parâmetros das questões e sim de características de um grupo e não outro. No Gráfico 3, mede-se a existência de freezer em casa e é ilustrado uma situação em que existe DIF entre os grupos estudados. Os itens foram calibrados na mesma escala considerando grupos múltiplos e com isso para cada grupo foram obtidos parâmetros distintos. Observa-se claramente que a probabilidade de um aluno responder “sim” para a pergunta se possui freezer em casa é diferente

dos alunos do grupo da 3ª série do E.M Enem em relação aos demais grupos.

Gráfico 3. Comportamento diferencial do item



Fonte: Elaborado pela autora

Assim, uma questão com comportamento diferencial entre grupos indica que o θ medido não está em função apenas dos parâmetros das questões e sim de características que favorecem um grupo e não outro.

Neste trabalho será utilizada uma das opções disponíveis no pacote *mirt* (CHALMERS, 2015), do *software* R, para a investigação das questões que possuem DIF. Serão comparadas graficamente as curvas características das questões por grupo quando calibrados por meio de grupos múltiplos, garantindo a comparabilidade por estarem na mesma escala. Com isso, para chegar nesse resultado primeiramente os parâmetros das questões serão estimados nos grupos e as questões que indicarem menos DIF serão fixadas e novamente as demais questões terão seus parâmetros calibrados livremente. Este procedimento é indicado no manual do pacote e é feito para obter melhores resultados.

4.4.4 Modelos da TRI

Nesta seção serão apresentados o modelo dicotômico e os modelos politômicos de dois parâmetros que serão utilizados neste trabalho. Os modelos da TRI que consideram um terceiro parâmetro, que é conhecido como o parâmetro de acerto casual, não serão utilizados pois parte-se do pressuposto de que não existe resposta casual de alunos com nível socioeconômico baixo e por isso não faria sentido considerar na modelagem tal parâmetro.

Para a modelagem dos dados será utilizado o pacote *mirt* (CHALMERS, 2015) do software livre R (R CORE TIME, 2015), e de forma geral manteve-se o *default* proposto pelo programa alterando

apenas a informação relacionada ao modelo utilizado. O pacote proposto utiliza para a estimação dos parâmetros dos itens o método de máxima verossimilhança marginal, e o *default* é o algoritmo EM - Expectativa de Maximização (BOCK AND AITKIN, 1981), que de acordo com Chalmers (2015) é uma opção razoável para modelos unidimensionais.

4.4.4.1 Modelo Dicotômico

O modelo dicotômico foi proposto inicialmente por Lord (1952) e em seguida aperfeiçoado por Birnbaum (1968). Como o próprio nome diz, esse modelo serve para as questões dicotômicas, isto é, aquelas que são corrigidas como certo ou errado ou indicam a presença ou ausência de determinada característica, por exemplo.

Neste estudo utilizou-se o modelo de dois parâmetros (Andrade, Tavares e Valle, 2000), o qual define a probabilidade de um indivíduo j marcar a categoria 1 (presença do atributo) da questão i como:

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

com questão $i = 1, 2, \dots, I$ e o indivíduo $j = 1, 2, \dots, n$ onde:

U_{ij} é uma variável dicotômica com valores 0 ou 1, sendo 1 quando o indivíduo j marcar para a questão i a alternativa que indica a presença do atributo e 0 a ausência.

θ_j é a medida de nível socioeconômico do indivíduo j .

$P(U_{ij} = 1 | \theta_j)$ é a probabilidade de um indivíduo j com um nível socioeconômico θ marcar a categoria 1 da questão i .

a_i é o parâmetro de discriminação da questão i .

b_i é o parâmetro de dificuldade da questão i .

A partir das definições do modelo pode-se dizer que indivíduos com valores mais altos do Inse possuem maior probabilidade de responder positivamente a um determinado item do que os indivíduos com valores mais baixos do Inse.

O modelo dicotômico é bastante utilizado em provas que medem conhecimento e também é uma opção para criação de indicadores contextuais. Quando o objetivo é criar um indicador de nível socioeconômico a partir de um questionário com questões que possuem mais de duas categorias, o pesquisador recodifica as questões de forma a indicar, por exemplo, a ausência ou presença do atributo medido. A

facilidade de trabalhar com modelos dicotômicos no lugar de modelos politômicos está no fato de trabalhar com menos parâmetros e bancos de dados mais simples, porém, de acordo com Ostini e Nering (2010), trabalhar com modelos politômicos é mais vantajoso por se obter mais informações psicométricas do construto que está sendo medido.

4.4.4.2 Modelos Politômicos

Nesta seção serão apresentados com mais detalhes os modelos politômicos que serão adotados nesse trabalho. Segundo Hambleton, Linden e Wells (2010), a contribuição dos modelos da TRI para dados politômicos está em descrever a probabilidade de um indivíduo responder a uma específica categoria dado o seu nível de habilidade (nível socioeconômico) e bem como as características das questões.

Serão utilizados o Modelo de Respostas Nominais, Modelo de Respostas Graduais de Samejima e Modelo de Crédito Parcial Generalizado. A proposta inicial deste trabalho considerava também a possibilidade de verificar o ajuste dos dados ao Modelo de Escala Gradual, por Andrich (1978). Este modelo, diferente do modelo de Samejima (1969), considera que as categorias são igualmente espaçadas. Segundo Andrade, Tavares e Valle (2000), se os itens que compõem o instrumento tiverem suas próprias categorias de respostas, que podem diferir no número, então esse modelo não é adequado. Com base nas considerações feitas por estes autores, verifica-se a impossibilidade de aplicar o modelo politômico de escala gradual aos dados que caracterizam o nível socioeconômico dos alunos. A seguir serão apresentados, com mais detalhes, os modelos que serão utilizados neste estudo.

4.4.4.2.1 Modelo de respostas nominais

O modelo de respostas nominais se ajusta às categorias politômicas e se caracteriza por não considerar a priori a ordem das alternativas, foi proposto por Bock em 1972. Três são os motivos para utilizá-lo: 1) quando se deseja obter as características dos itens e deseja-se uma medida a partir de itens puramente nominais; 2) validar empiricamente o ordenamento esperado das alternativas de uma questão; e 3) para fornecer um modelo para as respostas a um conjunto de itens, os *testlets* (THISSEN, CAI e BOCK 2010).

Andrade, Tavares e Valle (2010) citam o modelo de Bock definindo a probabilidade de um indivíduo j escolher uma particular categoria k do item i como:

$$P_{i,k}(\theta_j) = \frac{e^{a_{i,k}(\theta_j - b_{i,k})}}{\sum_{h=1}^{m_i} e^{a_{i,k}(\theta_j - b_{i,k})}},$$

com questão $i = 1, 2, \dots, I$; o indivíduo $j = 1, 2, \dots, n$; e a alternativa $k = 1, 2, \dots, m_i$ onde:

$P_{i,k}(\theta_j)$ é a probabilidade de um indivíduo j com um nível socioeconômico θ marcar a categoria k da questão i .

θ_j é a medida de nível socioeconômico do indivíduo j .

$a_{i,k}$ é o parâmetro de discriminação da questão i relacionado a alternativa k .

$b_{i,k}$ é o parâmetro de dificuldade da questão i relacionado a alternativa k .

Pinheiro, Costa e Cruz (2010) sugerem a utilização do modelo nominal como alternativa eficaz de análise da qualidade dos itens, visto que o modelo não utiliza a priori a ordenação das alternativas. Assim, é possível verificar a validação do modelo empírico com o proposto em um modelo politômico nominal.

Para identificar a ordem empírica das categorias deve-se inspecionar os valores de $a_{i,k}$ e se para cada questão o ordenamento das alternativas estiver em ordem crescente significa que coincide com a ordem teórica (disposto no banco de dados) e caso não, sugere uma outra ordenação (CHALMERS, 2015). Por exemplo, uma questão com quatro alternativas, esperava-se que $a_1 < a_2 < a_3 < a_4$ indicando que valores menores de a_k estão relacionados com menores valores de Inse e valores maiores a_k com valores maiores de Inse, caso contrário essas alternativas precisam ser recodificadas.

4.4.4.2.2. Modelo de respostas graduais – Samejima

O modelo de respostas graduais de Samejima faz parte dos modelos que se ajustam às categorias ordinais politômicas e o foco não é mais um modelo dicotômico de presença ou ausência de uma característica, mas sim a gradação das respostas. Segundo Hambleton, Linden e Wells (2010), foi o primeiro modelo proposto para respostas

graduais. O modelo permite calcular a probabilidade de um aluno com determinado nível socioeconômico responder determinada alternativa, a partir dos parâmetros a dos itens e dos b 's das alternativas.

Hambleton, Linden e Wells (2010) destacam que o modelo de Samejima (1969) utiliza um processo em duas etapas para obter a probabilidade de o aluno responder uma determinada alternativa, já que na primeira etapa se obtêm a probabilidade de o aluno responder uma alternativa e as suas superiores, dado um nível socioeconômico.

Assim, para o cálculo da primeira etapa, a probabilidade de um aluno j escolher uma alternativa e as suas superiores do item i é dada por Andrade, Tavares e Valle (2010) como:

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}}$$

com questão $i = 1, 2, \dots, I$; o indivíduo $j = 1, 2, \dots, n$; e a alternativa $k = 1, 2, \dots, m_i$ onde:

$P_{i,k}^+(\theta_j)$ é a probabilidade de um indivíduo j com um nível socioeconômico θ marcar a categoria k e suas posteriores da questão i .

θ_j é a medida de nível socioeconômico do indivíduo j .

a_i é o parâmetro de *discriminação* da questão i .

$b_{i,k}$ é a dificuldade da questão i e relacionado a alternativa k .

E para a segunda etapa, a probabilidade de um aluno j escolher uma alternativa k no item i é:

$$P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j)$$

Assim, no caso de um item com três alternativas, a probabilidade de um aluno marcar a primeira alternativa é apenas a probabilidade de não passar o primeiro limiar e ela diminui à medida que o nível socioeconômico aumenta. A probabilidade de marcar a terceira alternativa aumenta a medida que o nível socioeconômico aumenta, as duas alternativas dos extremos apresentam uma curva monotômica e a do meio curva em formato de sino. Já a probabilidade de o aluno marcar a segunda alternativa é a probabilidade de passar o primeiro, mas não o segundo, isto é marcar a alternativa 2 significa que o aluno possui um Inse entre quem marca a alternativa 1 e 3. Isso porque a probabilidade de uma resposta de uma determinada alternativa é dada pela diferença

de probabilidades de responder a um limiar inferior e superior, por isso, o modelo de resposta graduada é dito ser um modelo de diferença 1.

4.4.4.2.3 Modelo de crédito parcial e modelo de crédito parcial generalizado

O modelo de crédito parcial é outro modelo além dos já indicados, nesta seção, que se ajustam as categorias ordinais politômicas. O modelo de crédito parcial, segundo De São Paulo et al (2007), foi introduzido por Masters em 1982, no modelo o item é pontuado parcialmente e o respondente ganha mais créditos na medida em que sua resposta se aproxima da completa. Andrade, Tavares e Valle (2000) destacam que esse modelo pertence à família de Modelos de Rasch, indicando que é uma extensão do Modelo de Rasch para itens dicotômicos, em que o parâmetro de discriminação é igual a um para todos os itens.

De São Paulo et al (2007) citam o modelo de crédito parcial generalizado, que foi proposto por Muraki (1992), como generalização do primeiro e que considera o parâmetro de discriminação no modelo. Segundo os mesmos, o modelo explicita os pontos na escala onde uma alternativa se torna relativamente mais provável que a alternativa precedente dado que o respondente completou os passos anteriores. Isso corrobora com a afirmação de Hambleton, Linden e Wells (2010) que destacam que o modelo, diferentemente do proposto por Samejima (1969), é considerado um modelo direto, pois calcula a probabilidade de um aluno responder uma determinada alternativa já na primeira etapa.

A probabilidade de um aluno j responder uma determinada alternativa k do item i , é dada por Andrade, Tavares e Valle (2010) como:

$$P_{i,k}(\theta_j) = \frac{e^{[\sum_{u=0}^k a_i(\theta_j - b_{i,u})]}}{\sum_{u=0}^{m_i} e^{[\sum_{v=0}^u a_i(\theta_j - b_{i,v})]}}$$

com questão $i = 1, 2, \dots, I$; o indivíduo $j = 1, 2, \dots, n$; e a alternativa $k = 0, 1, 2, \dots, m_i$ onde:

$P_{i,k}(\theta_j)$ é a probabilidade de um indivíduo j com um nível socioeconômico θ marcar a categoria k da questão i .

θ_j é a medida de nível socioeconômico do indivíduo j .

a_i é o parâmetro de discriminação da questão i .

$b_{i,k}$ são os pontos na escala de j em que as curvas de $P_{i,k-1}(\theta_j)$ e $P_{i,k}(\theta_j)$ se interceptam.

Segundo Andrade, Tavares e Valle (2010) o modelo de crédito parcial generalizado, diferente do modelo de resposta gradual, permite que os valores de $b_{i,k}$ não necessariamente sejam ordenados.

Na próxima seção serão apresentadas as técnicas utilizadas para escolher o modelo que melhor se ajusta aos dados.

4.4.5 Escolha do modelo

A Curva de Informação (CI) é um recurso gráfico que permite analisar quanto uma questão ou um conjunto de questões (soma das curvas de informação) possuem de informação para determinado intervalo de nível socioeconômico, e quanto maior for a informação menor é o erro padrão associado a essa medida. Com isso, Soares (2005) compara a curva de informação a partir de três propostas de modelagem e opta pelo modelo que possui maior informação.

Neste trabalho, serão testados quatro modelos na tentativa de ajustá-los aos dados e a técnica que permitirá escolher o melhor modelo será a partir da análise gráfica da curva de informação, dado que quanto maior for a informação menor será o erro padrão associado à medida.

4.4.6 Ajuste dos modelos

O processo de criação de um indicador pela TRI é bastante trabalhoso dado que envolve uma série de etapas como, por exemplo, discussão do construto, seleção das questões, preparação dos bancos de dados e modelagem. Entretanto, todos os esforços e resultados encontrados perdem sua validade quando o modelo utilizado não se ajusta adequadamente aos dados. Glas (2010) destaca que nenhum modelo vai se ajustar aos dados perfeitamente e que o poder da estatística dos testes para a avaliação do ajuste do modelo é muito sensível ao tamanho da amostra. O mesmo autor destaca, também, que a avaliação do ajuste é um processo subjetivo.

Kang e Chen (2007) citam diversos autores (SHEPARD, CAMILLI & WILLIAMS, 1984; BOLT, 2002; RUPP & ZUMBO, 2004) destacando a importância de um bom ajuste dos dados ao modelo, já que o desajuste compromete o pressuposto da invariância dos dados.

Na literatura existem diversas formas de verificar ajuste do modelo aos dados. Nesta seção serão apresentados alguns dos procedimentos disponíveis no pacote *mirt* do *software* R.

4.4.6.1 RMSEA – Raiz do Erro Quadrático Médio de Aproximação

Uma forma de medir ajuste dos dados ao modelo, segundo Hair (2005), é calculando o RMSEA - a raiz do erro quadrático médio de aproximação, uma medida que tenta corrigir a fragilidade da estatística qui-quadrado em rejeitar a hipótese nula quando a amostra é suficientemente grande. Essa medida representa a qualidade do ajuste que seria esperado se o modelo fosse estimado na população e não somente na amostra utilizada. O autor ainda destaca que para definir se ocorreu um bom ajuste valores dessa medida entre 0,05 e 0,08 são considerados aceitáveis.

4.4.6.2 Análise dos parâmetros dos itens

Conhecer os produtos gerados na modelagem é um processo fundamental para a produção de um indicador de qualidade. Esse procedimento é feito analisando os valores dos parâmetros encontrados e a partir de gráficos.

Sartes e Souza-Formigoni (2013) destacam que quando uma questão apresenta pouca ou excessiva discriminação no modelo de respostas graduais um procedimento possível é reagrupar suas alternativas. Também destacam que é possível retirar questões ou reagrupar alternativas quando, em uma escala com média zero e desvio-padrão um, encontrar parâmetro a menor que 0,70 ou maior que 4,0 e/ou parâmetro b maior que 4 ou -4 e/ou estimativas com erros padrão muito grande.

A análise gráfica dos itens e do teste corresponde a uma etapa do processo de construção do indicador, a partir dos gráficos é possível observar características importantes dos itens e do teste. Segundo Alves e Soares (2009), a análise da qualidade de cada item como componente da medida característica latente pode ser feita analisando a curva característica e a curva de informação.

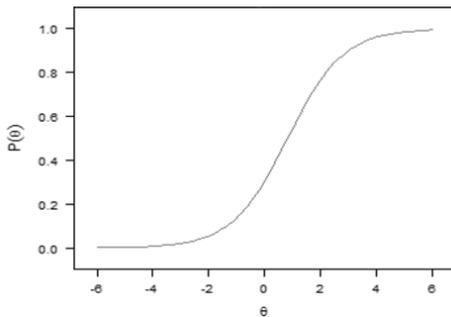
A análise gráfica permite ao final da modelagem que seja verificado a qualidade do ajuste dos itens ao modelo e a contribuição de cada item na informação produzida.

4.4.6.3 Análise gráfica – Curva Característica do Item (CCI)

A partir do modelo de probabilidade utilizado é possível construir a Curva Característica do Item (CCI), esse gráfico mostra no eixo vertical a probabilidade de resposta de determinada categoria e no eixo horizontal o nível socioeconômico dos alunos. Como a curva característica do item é calculada a partir dos parâmetros dos itens quanto maior for o parâmetro a maior é a inclinação da curva e quanto maior for o parâmetro b maior é o deslocamento da curva a direita.

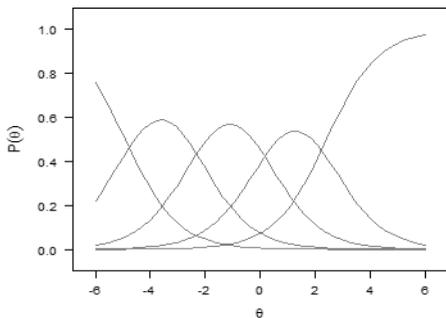
Nos Gráficos 4 e 5, a seguir, são apresentados um exemplo de curva característica para questão dicotômica e um para questão politômica. No caso de questões com duas categorias a curva é monótona e com mais de duas categorias a curva é no formato de sino para as probabilidades das categorias do meio e monótona para os extremos.

Gráfico 4. Curva característica questão dicotômica



Fonte: Elaborado pela autora

Gráfico 5. Curva característica questão politômica



Fonte: Elaborado pela autora

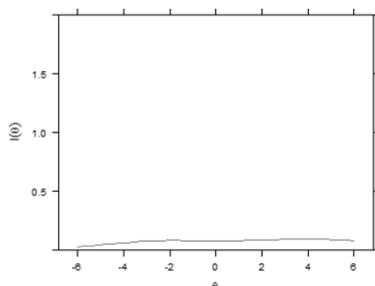
A CCI também é usada para verificar se o pressuposto da invariância está sendo atendido, pois analisa se o item está apresentando DIF entre grupos e já foi citada em seções anteriores.

4.4.6.4 Curva de Informação (CI)

A Curva de Informação do item ou de todos os itens também é outra medida gráfica utilizada, ela permite analisar quanto um item ou um conjunto de itens (soma das curvas de informação dos itens) possuem de informação para determinado intervalo de nível socioeconômico, e quanto maior for a informação menor é o erro padrão associado a essa medida.

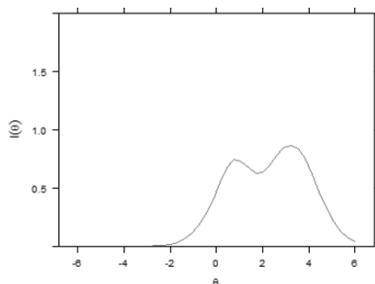
Este gráfico é dado em função das características da questão, por isso quanto maior for o parâmetro b mais deslocada à direita encontra-se a curva e quanto maior for o parâmetro a maior é o pico da curva. A seguir, são apresentados os Gráficos 6 e 7 que contém exemplos de CI de itens com baixa informação e com alta informação para um teste.

Gráfico 6. Curva de Informação da questão com baixa informação



Fonte: Elaborado pela autora

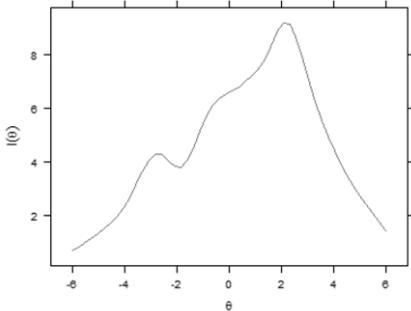
Gráfico 7. Curva de Informação da questão com alta informação



Fonte: Elaborado pela autora

A curva de informação do teste é calculada pela soma das curvas de informação das questões, a seguir no Gráfico 8 é apresentada a curva de informação de um teste.

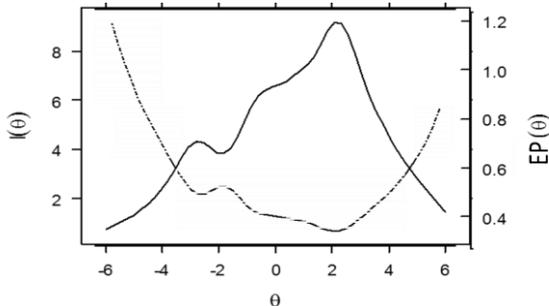
Gráfico 8. Curva de Informação do Teste



Fonte: Elaborado pela autora

É importante analisar a curva de informação, pois quanto maior for à informação do item ou teste menor é o erro padrão associado a essa medida. A seguir, no Gráfico 9, são apresentadas a curva de informação de um teste hipotético e do erro padrão, respectivamente a linha contínua e linha pontilhada.

Gráfico 9. Curva de informação do teste e erro padrão



Fonte: Elaborado pela autora

Segundo Embreston e Reise (2000) nos modelos politômicos os valores não devem ser interpretados diretamente como discriminação do item, para avaliar diretamente a quantidade de discriminação que um item fornece é necessário avaliar a curva de informação do item.

4.4.7 Estimação do nível socioeconômico

Após o processo de calibração é possível para cada aluno do banco de dados estimar o nível socioeconômico. Da mesma forma que existem processos de estimação para os parâmetros dos itens, existem processos de estimação para o traço latente que está sendo medido (θ), optou-se nesse trabalho estimar o θ pelo método bayesiano EAP - esperança *a posteriori*, que é um dos métodos mais utilizados na literatura.

De forma geral, as técnicas utilizadas para criação de indicadores geram medidas quantitativas e estas podem ser intervalares ou de razão. Segundo Hair (2005), as duas escalas se diferenciam unicamente pela escala de razão ter um zero absoluto. Ele cita um exemplo de uma escala intervalar a temperatura Fahrenheit e Celsius, elas possuem um ponto zero arbitrário diferente e nenhuma indica uma quantia nula ou ausência de temperatura, já que é possível registrar temperatura abaixo de zero em ambas. Ainda relacionado ao exemplo, Hair (2005), não se pode considerar que um dia de 80°F corresponda a um dia duas vezes mais quente de 40°F, dado que 80°F corresponde 26,7°C e 40°F não é igual a 13,35°C.

Desta forma, não é possível dizer que qualquer valor em uma escala intervalar é múltiplo de algum outro ponto da escala, e a escala permite que seus valores sejam ordenados, somados, subtraídos, porém não multiplicados e nem divididos entre si.

A escala obtida pela TRI é uma escala de natureza intervalar, dado que possui um zero arbitrário, seu valor pode variar de menos infinito a mais infinito e quando um indivíduo possui um Inse zero não significa ausência de nível socioeconômico. Conhecer a natureza da escala de um indicador é muito importante, pois evita o mal uso da medida, desta forma pelo fato de o Inse ser uma escala intervalar é um interpretação errada dizer por exemplo, que um indivíduo com Inse 80 tem duas vezes mais Inse do que um indivíduo com Inse 40.

Após a etapa da calibração dos parâmetros e da estimação da proficiência na amostra selecionada, para obter o valor do indicador de nível socioeconômico de todos os alunos é necessário fazer algumas transformações, isso porque:

- Na estimação dos parâmetros das questões uma escala “teórica” é arbitrada para que o grupo de referência tenha média 0 e desvio padrão 1, porém em termos práticos quando se calcula a média e o desvio padrão do grupo de

referência ele não possui exatamente média 0 e nem desvio 1, mas os valores são próximos.

- Como não é necessário calibrar as questões com toda a população o processo de estimação foi feito com uma amostra. Em geral, opta-se quando um indicador é criado pela TRI que o grupo de referência tenha a média e o desvio “teórico”, é como se fosse um marco zero da escala. Essa opção permite e facilita a comparação de uma série histórica, por exemplo.
- Além disso, pelo fato da média ser igual a 0 existirão valores negativos na escala, isso não é conveniente e até é mais confuso para quem tiver contato com a medida entender valores de nível socioeconômico negativo. Por isso, é possível transformar a escala para que tenha média 500 e desvio padrão 100, por exemplo, sem interferir nas características da medida.

Devido às considerações acima, é necessário para que a escala tenha média 500 e desvio 100, valor arbitrário escolhido para o indicador, fazer a seguinte transformação linear no θ :

$$\theta_{(500,100)} = 100 * \left(\frac{\theta_{estimado} - média_{grupo\ ref.}}{desvio_{grupo\ ref.}} \right) + 500$$

Além disso, é possível também colocar os parâmetros na nova escala:

$$a_{(500,100)} = \frac{a_{estimado}}{100 / desvio_{grupo\ ref.}}$$

$$b_{(500,100)} = 100 * \left(\frac{b_{estimado} - média_{grupo\ ref.}}{desvio_{grupo\ ref.}} \right) + 500$$

4.4.8 Segmentação e interpretação da escala de Inse

Um dos grandes diferenciais da TRI é poder relacionar a medida de Inse calculada com as questões utilizadas na sua criação. Ou seja, é possível calcular a probabilidade de um indivíduo responder determinada alternativa, dado o seu nível socioeconômico familiar.

Se selecionarmos três alunos como exemplo, A, B e C com os seguintes valores de nível socioeconômico respectivamente 250, 500 e 800, o que poderia ser dito em relação a essa medida calculada para eles?

Primeiramente pode-se dizer:

- o aluno B possui maior Inse do que o aluno A e
- o aluno C possui maior Inse do que o aluno B e o A.

Além disso, pela medida ser de natureza intervalar não é possível dizer, como descrito na seção 4.4.7, que o indivíduo B possui o dobro de Inse do aluno A. É possível dizer, também, por exemplo, que o aluno A possui alta probabilidade de não possuir automóvel, o aluno B alta probabilidade de não possuir ou de possuir um automóvel e o aluno C alta probabilidade de possuir dois ou mais automóveis. Fala-se em probabilidade porque os modelos da TRI são modelos probabilístico. Com isso, se selecionarmos alguns alunos semelhantes aos alunos A, B e C será encontrado mais situações parecidas ao que foi descrito do que o contrário.

Descrever a probabilidade que o aluno teria de responder as alternativas do questionário para cada θ existente seria uma atividade muito trabalhosa e, em geral, o objetivo é classificar os indivíduos em faixas de θ , isto é, busca-se dividir os indivíduos em grupos e cada grupo é interpretado.

Por ser uma abordagem nova, o uso de modelos politômicos, quando comparado com os dicotômicos, e por não se encontrar na literatura opções claras de como dividir e interpretar escalas, nesse trabalho decidiu-se fazer uma proposta diferente do que foi encontrado na literatura, no que se refere à interpretação de escala, isso porque, considera-se que as propostas estudadas possuem alguns pontos negativos quando aplicados nos modelos politômicos para posicionar cada alternativa de um item:

- critérios de item âncora utilizados para interpretação de escala dos modelos dicotômicos aplicados para interpretar alternativas nos modelos politômicos - talvez essa não

seja a melhor opção, dado que as curvas de probabilidade das categorias/alternativas do meio, nos modelos politômicos, são em formato de sino, isto é, crescem, alcançam o seu pico e depois decrescem e as curvas do modelo dicotômico só crescem a medida que o θ aumenta.

- na divisão dos grupos - quando se divide a escala em tamanhos iguais por um múltiplo do desvio padrão não se busca um ponto ótimo em relação às características das questões, podendo, por exemplo, posicionar uma alternativa em um nível e o seu ponto máximo estar em outro.
- em relação a interpretação da escala - posicionar as alternativas pela probabilidade de o indivíduo responder ela e suas superiores e na hora de interpretar a escala o pesquisador interpretar apenas a alternativa específica.
- dividir os alunos em grupos - a partir das características da população estudada utilizando procedimentos em que os grupos são divididos em quartis ou quintis e com isso a escala é interpretada em função do banco de dados sendo o primeiro grupo o de menor nível socioeconômico e o último o de maior nível socioeconômico. Tal procedimento é totalmente dependente do banco de dados disponível e com isso em um momento um mesmo valor de Inse pode estar classificado como nível socioeconômico muito baixo e em outro estar classificado como baixo pelo fato do banco de dados conter mais indivíduos de Inse baixo.

Com base no que foi exposto, optou-se por aplicar uma técnica comumente utilizada para segmentação, a análise de agrupamentos hierárquicos, a partir do método de Ward. Essa técnica, segundo Hair (2005), é utilizada para classificar indivíduos/elementos em grupos de modo que cada indivíduo é muito semelhante aos indivíduos do grupo que pertence e diferente dos demais grupos. Segundo o mesmo autor, essa análise é utilizada em diversas áreas como psicologia, biologia, sociologia, economia, engenharia e administração e é chamada também de análise Q, construção de tipologia, análise de classificação e taxonomia numérica.

Para aplicar a análise de agrupamentos hierárquicos, serão utilizadas como variáveis de entrada os valores de Inse e as

probabilidades de resposta por questões/alternativas do modelo. Os valores utilizados de Inse variam de 100 a 900, com intervalos de 10 unidades, totalizando 81 valores, de uma escala que será transformada para ter média 500 e desvio-padrão 100. Optou-se por utilizar intervalos pequenos para conseguir separar de forma mais eficiente os grupos. Com isso, utilizando apenas as informações que realmente são importantes para segmentação, que são os valores de Inse e as probabilidades de resposta por questões/alternativas do modelo, a segmentação não dependerá mais do banco de dados utilizado, ficando relacionado apenas às variáveis utilizadas na modelagem.

Com o intuito de definir o número de clusters será considerado o coeficiente de aglomeração e o gráfico de dendograma. O coeficiente de aglomeração é utilizado para ajudar a identificar grandes aumentos relativos na homogeneidade dos agrupamentos e o gráfico de dendograma o nível de similaridade entre os valores de Inse, (HAIR, 2005).

A ideia é que essa metodologia forneça direcionamentos que facilite a formação de grupos, porém a referência será sempre na relação entre as questões e a medida. O procedimento será realizado no *software Ibm Spss Statistics*, versão 23.

Após a definição das faixas de Inse por grupo, a partir da análise de agrupamentos, serão calculadas as médias das probabilidades do modelo referentes a cada alternativa/questão, por grupo, e a partir dessa média, a alternativa será descrita. Optou-se por essa metodologia por considerar que dentro do grupo as probabilidades seriam bastante parecidas e evitaria também utilizar apenas um ponto para poder representar o todo.

Por fim, com base nas probabilidades médias, os grupos serão descritos e para a descrição da escala serão utilizados três critérios, a saber:

1. valores maiores que 0,3 quando alguma probabilidade da mesma coluna e questão estiver entre 0,5 e 0,6;
2. valores maiores ou iguais a 0,5 e
3. quando nenhuma das situações anteriores ocorrerem e verificar-se uma concentração de probabilidades que somando as alternativas seja maior que 0,6.

5. RESULTADOS

A seção de resultados foi dividida em quatro subseções, primeiramente foi feita uma análise descritiva dos dados e em seguida, aplicados os procedimentos relacionados a TRI. Por fim, foi estimado o nível socioeconômico dos alunos e esses alunos foram divididos em grupos. Os grupos foram descritos de acordo com o nível socioeconômico calculado e as características das questões.

5.1 ANÁLISE DESCRITIVA

Analisar de forma descritiva um conjunto de dados é um procedimento que pode ser feito de várias formas, uma delas é analisar o percentual de respostas por alternativa de cada questão. Esta etapa é importante para o processo de produção de indicadores, isto porque situações em que a questão apresenta, por exemplo, grande parte das respostas concentradas em uma alternativa ou quando existe um grande número de alternativas pode não ser interessante. Esse procedimento de análise é feito sempre que é necessário conhecer melhor uma questão durante o processo de criação do indicador.

Para o cálculo do indicador utilizou-se inicialmente 24 questões e antes de calcular o percentual de resposta por questão, optou-se por retirar do banco de dados todos os alunos que responderam cinco ou menos questões, conforme indicado na seção 4.2. A partir do procedimento proposto, obtiveram-se os seguintes quantitativos por avaliação/exame, Tabela 2:

Tabela 2. Quantitativo de alunos por avaliação/exame

MICRODADOS		Qtd.
Saeb	5º ano Ensino Fundamental	2.013.019
	9º ano Ensino Fundamental	2.019.851
	3ª série Ensino Médio	90.446
Enem	3ª série Ensino Médio	878.853
Total		5.002.169

Fonte: Elaborado pela autora

A seguir, na Tabela 3, são apresentados os percentuais de resposta por alternativa de cada questão, por ano/série, após as questões terem sido recodificadas para a composição de um banco único. Em

geral, observa-se que o percentual de resposta por alternativa segue distribuição semelhante nas questões quando comparado entre as séries/anos, o Saeb e o Enem. A única questão que apresentou comportamento diferente do padrão foi a questão “Q11 - Freezer (aparelho independente ou parte da geladeira) em casa” em que o padrão apresentado para o Enem é o inverso do Saeb para todos as séries/anos avaliados. Esse tipo de situação pode acontecer, pois as questões não são perguntadas da mesma forma nos dois instrumentos e pode estar ocorrendo entendimentos diferentes. Caso essa questão seja entendida de forma diferente pelos grupos, provavelmente apresentará DIF e, conseqüentemente, problemas de ajuste no modelo.

Tabela 3. Porcentagens de respostas por questão e avaliação/exame

Questões/Alternativas	Saeb			Enem
	5º ano	9º ano	3ª série	3ª série
Q01 - Televisão em cores em casa				
1 - Não tem	6%	2%	2%	1%
2 - Sim, uma	36%	34%	38%	48%
3 - Sim, duas	31%	36%	32%	32%
4 - Sim, três ou mais	21%	24%	24%	19%
NA	6%	4%	4%	0%
Total	100%	100%	100%	100%
Q02 - TV por assinatura em casa				
1 - Não tem	-	-	-	66%
2 - Sim, uma	-	-	-	29%
3 - Sim, duas	-	-	-	3%
4 - Sim, três ou mais	-	-	-	2%
NA	-	-	-	0%
Total	-	-	-	100%
Q03 - Videocassete e/ou DVD em casa				
1 - Não tem	14%	12%	15%	16%
2 - Sim, um	59%	61%	61%	68%
3 - Sim, dois	17%	20%	17%	13%
4 - Sim, três ou mais	6%	5%	5%	3%
NA	5%	2%	3%	0%
Total	100%	100%	100%	100%
Q04 - Computador em casa				
1 - Não tem	37%	31%	24%	26%
2 - Sim, um	41%	46%	46%	54%
3 - Sim, dois	12%	15%	18%	13%
4 - Sim, três ou mais	5%	7%	11%	7%
NA	5%	1%	1%	0%
Total	100%	100%	100%	100%
Q05 - Internet em casa				
1 - Não tem	-	-	-	26%

Questões/Alternativas	Saeb			Enem
	5º ano	9º ano	3ª série	3ª série
2 - Sim, uma	-	-	-	65%
3 - Sim, duas	-	-	-	5%
4 - Sim, três ou mais	-	-	-	4%
NA	-	-	-	0%
Total	-	-	-	100%
Q06 - Rádio em casa				
1 - Não tem	20%	16%	23%	25%
2 - Sim, um	52%	55%	54%	61%
3 - Sim, dois	16%	19%	14%	11%
4 - Sim, três ou mais	7%	6%	5%	3%
NA	6%	3%	4%	0%
Total	100%	100%	100%	100%
Q07 - Telefone fixo em casa				
1 - Não tem	-	-	-	45%
2 - Sim, um	-	-	-	49%
3 - Sim, dois	-	-	-	4%
4 - Sim, três ou mais	-	-	-	2%
NA	-	-	-	0%
Total	-	-	-	100%
Q08 - Telefone celular em casa				
1 - Não tem	-	-	-	2%
2 - Sim, uma	-	-	-	26%
3 - Sim, duas	-	-	-	22%
4 - Sim, três ou mais	-	-	-	50%
NA	-	-	-	0%
Total	-	-	-	100%
Q09 - Aspirador de pó em casa				
1 - Não tem	-	-	-	73%
2 - Sim, um	-	-	-	26%
3 - Sim, dois	-	-	-	1%
4 - Sim, três ou mais	-	-	-	0%
NA	-	-	-	0%
Total	-	-	-	100%
Q10 - Geladeira em casa				
1 - Não tem	3%	2%	2%	2%
2 - Sim, uma	79%	83%	84%	92%
3 - Sim, duas	12%	12%	10%	6%
4 - Sim, três ou mais	2%	2%	1%	0%
NA	4%	2%	3%	0%
Total	100%	100%	100%	100%
Q11 - Freezer (aparelho independente ou parte da geladeira) em casa				
1 - Não tem	26%	30%	35%	61%
2 - Sim, um ou mais	66%	65%	60%	39%
NA	8%	5%	5%	0%

Questões/Alternativas	Saeb			Enem
	5º ano	9º ano	3ª série	3ª série
Total	100%	100%	100%	100%
Q12 - Máquina de lavar roupa em casa				
1 - Não tem	20%	24%	28%	23%
2 - Sim, uma	64%	66%	67%	74%
3 - Sim, duas	11%	8%	4%	3%
4 - Sim, três ou mais	2%	1%	0%	0%
NA	3%	1%	0%	0%
Total	100%	100%	100%	100%
Q13 - Automóvel em casa				
1 - Não tem	49%	49%	45%	40%
2 - Sim, uma	35%	37%	38%	44%
3 - Sim, duas	9%	10%	12%	13%
4 - Sim, três ou mais	4%	3%	4%	3%
NA	3%	1%	0%	0%
Total	100%	100%	100%	100%
Q14 - Banheiro em casa				
1 - Não tem	2%	1%	1%	1%
2 - Sim, uma	66%	65%	54%	60%
3 - Sim, duas	21%	25%	28%	26%
4 - Sim, três ou mais	6%	7%	15%	13%
NA	5%	2%	3%	0%
Total	100%	100%	100%	100%
Q15 - Quartos para dormir em casa				
1 - Não tem	1%	1%	1%	-
2 - Sim, um	13%	8%	6%	-
3 - Sim, dois	38%	39%	34%	-
4 - Sim, três	31%	37%	42%	-
5 - Sim, quatro ou mais	12%	12%	13%	-
NA	5%	3%	4%	-
Total	100%	100%	100%	-
Q16 - Situação residência				
1 - Outra situação (loteamento não regularizado, ocupação etc.)	-	-	-	1%
2 - Cedida	-	-	-	7%
3 - Alugada	-	-	-	15%
4 - Própria e em pagamento (financiada)	-	-	-	8%
5 - Própria e quitada	-	-	-	68%
NA	-	-	-	0%
Total	-	-	-	100%
Q17 - Localização residência				
1 - Comunidade indígena ou quilombola	-	-	-	0%
2 - Zona rural	-	-	-	12%
3 - Zona urbana	-	-	-	88%
NA	-	-	-	0%
Total	-	-	-	100%

Questões/Alternativas	Saeb			Enem
	5º ano	9º ano	3ª série	3ª série
Q18 - Quantas pessoas moram na casa (incluindo o aluno)				
1 - Uma, pois moro sozinho(a)	1%	0%	1%	0%
2 - Duas	6%	7%	9%	7%
3 - Três	17%	19%	23%	23%
4 - Quatro	28%	31%	33%	38%
5 - Cinco	22%	22%	19%	20%
6 - Seis pessoas ou mais	22%	18%	13%	12%
NA	5%	2%	3%	0%
Total	100%	100%	100%	100%
Q19 - Contrata empregada doméstica em casa				
1 - Não	84%	84%	84%	91%
2 - Sim, um ou mais	12%	12%	13%	9%
NA	4%	3%	4%	0%
Total	100%	100%	100%	100%
Q20 - Renda mensal da família (considerando a renda do aluno)				
1 - Nenhuma renda	-	-	-	1%
2 - Até um s.m	-	-	-	25%
3 - Mais de um até 1,5 s.m	-	-	-	18%
4 - Mais de 1,5 até 3 s.m	-	-	-	26%
5 - Mais de 3 até 7 s.m	-	-	-	18%
6 - Mais de 7 até 9 s.m	-	-	-	4%
7 - Mais de 9 até 12 s.m	-	-	-	3%
8 - Acima de 12 s.m	-	-	-	5%
NA	-	-	-	0%
Total	-	-	-	100%
Q21- Mãe, ou a mulher responsável por você, sabe ler e escrever				
1 - Não	5%	6%	6%	-
2 - Sim	92%	94%	94%	-
NA	3%	1%	1%	-
Total	100%	100%	100%	-
Q22- Pai, ou homem responsável por você, sabe ler e escrever				
1 - Não	9%	10%	9%	-
2 - Sim	86%	87%	88%	-
NA	5%	4%	3%	-
Total	100%	100%	100%	-
Q23 - Até que série a mãe estudou				
1 - Nunca estudou.	3%	3%	3%	2%
2 - Ensino Fundamental cursando ou completo	36%	46%	38%	40%
3 - Ensino Médio completo	13%	25%	31%	33%
4 - Ensino Superior Completo	11%	9%	21%	22%
NA	37%	17%	7%	3%

Questões/Alternativas	Saeb			Enem
	5º ano	9º ano	3ª série	3ª série
Total	100%	100%	100%	100%
Q24 - Até que série o pai estudou				
1 - Nunca estudou.	4%	4%	4%	3%
2 - Ensino Fundamental cursando ou completo	28%	41%	39%	44%
3 - Ensino Médio completo	10%	19%	27%	28%
4 - Ensino Superior Completo	10%	7%	15%	16%
NA	48%	30%	15%	8%
Total	100%	100%	100%	100%

Fonte: Elaborado pela autora

Após avaliação das frequências, a partir dos resultados da Tabela 3, optou-se por não agrupar nenhuma alternativa e verificar como será o comportamento de cada questão no próprio modelo. A opção de não agrupar as alternativas permite que o modelo seja mais informativo e, com isso, seja possível diferenciar melhor os indivíduos, porém, se necessário, as categorias poderão ser agrupadas posteriormente, caso melhore o ajuste ao modelo.

Nas próximas seções serão analisados os comportamentos das questões como um conjunto de questões que compõem um construto, isto é, será analisada a forma como se relacionam e se o conjunto delas expressam uma medida que está sendo denominada de nível socioeconômico.

5.2 TEORIA DE RESPOSTA AO ITEM

Nesta seção serão apresentados todos os procedimentos utilizados para verificação dos pressupostos necessários para a utilização da TRI, como a modelagem dos dados utilizando os quatro modelos descritos na metodologia, escolha do modelo mais adequado, bem como todos os procedimentos para validação do modelo escolhido e ajuste dos dados.

5.2.1 Análise de correlação e dimensionalidade

A análise das correlações e dimensionalidade do conjunto de questões se faz necessário para verificar o pressuposto da unidimensionalidade dos dados e, conseqüentemente, sua independência local. Optou-se calcular a matriz de correlação policórica para cada avaliação/exame e, em seguida, a análise fatorial calculando também o escore fatorial de cada aluno.

Segundo Hair (2005), se a inspeção visual da matriz de correlações não apresentar um grande quantitativo de correlações maiores que 0,3 a análise fatorial é inapropriada. De forma geral as questões apresentaram correlações positivas e maiores ou próximas de 0,3. Nas Tabelas 4, 5, 6 e 7 pode-se verificar as correlações entre os itens e destacado em azul as correlações maiores que 0,3 nos quatro grupos estudados.

Tabela 4. Correlação policórica entre questões, 5º ano do EF Saeb

	Q01	Q03	Q04	Q06	Q10	Q11	Q12	Q13	Q14	Q15	Q18	Q19	Q21	Q22	Q23	Q24
Q01	1,00															
Q03	0,39	1,00														
Q04	0,41	0,32	1,00													
Q06	0,28	0,29	0,22	1,00												
Q10	0,31	0,28	0,27	0,25	1,00											
Q11	0,29	0,25	0,40	0,21	0,26	1,00										
Q12	0,30	0,27	0,42	0,22	0,33	0,34	1,00									
Q13	0,34	0,25	0,52	0,21	0,26	0,37	0,38	1,00								
Q14	0,35	0,27	0,43	0,20	0,35	0,29	0,31	0,40	1,00							
Q15	0,28	0,20	0,26	0,19	0,27	0,16	0,21	0,28	0,47	1,00						
Q18	0,01	0,01	-0,06	0,01	0,10	-0,06	0,01	-0,03	0,05	0,22	1,00					
Q19	0,14	0,17	0,25	0,11	0,19	0,19	0,23	0,25	0,24	0,11	-0,05	1,00				
Q21	0,20	0,14	0,34	0,05	0,10	0,24	0,23	0,27	0,20	0,12	-0,06	-0,02	1,00			
Q22	0,20	0,14	0,33	0,07	0,10	0,24	0,23	0,31	0,20	0,11	-0,05	-0,01	0,54	1,00		
Q23	0,25	0,19	0,41	0,08	0,13	0,24	0,21	0,30	0,30	0,18	-0,07	0,12	0,61	0,35	1,00	
Q24	0,25	0,19	0,40	0,08	0,13	0,24	0,21	0,29	0,28	0,16	-0,05	0,11	0,37	0,63	0,65	1,00

Fonte: Elaborado pela autora

Tabela 5. Correlação policórica entre questões, 9º ano do EF Saeb

	Q01	Q03	Q04	Q06	Q10	Q11	Q12	Q13	Q14	Q15	Q18	Q19	Q21	Q22	Q23	Q24
Q01	1,00															
Q03	0,50	1,00														
Q04	0,48	0,36	1,00													
Q06	0,34	0,35	0,25	1,00												
Q10	0,40	0,33	0,31	0,29	1,00											
Q11	0,34	0,29	0,42	0,24	0,29	1,00										
Q12	0,38	0,30	0,46	0,25	0,35	0,38	1,00									
Q13	0,39	0,28	0,53	0,24	0,31	0,40	0,43	1,00								
Q14	0,41	0,30	0,44	0,22	0,41	0,33	0,34	0,43	1,00							
Q15	0,26	0,19	0,22	0,20	0,31	0,17	0,19	0,29	0,50	1,00						
Q18	0,01	0,02	-0,08	0,00	0,09	-0,06	0,00	-0,01	0,04	0,25	1,00					
Q19	0,12	0,12	0,13	0,07	0,17	0,08	0,16	0,15	0,17	0,09	0,04	1,00				
Q21	0,28	0,21	0,45	0,10	0,17	0,30	0,33	0,37	0,28	0,09	-0,09	0,02	1,00			
Q22	0,28	0,20	0,43	0,12	0,16	0,30	0,32	0,40	0,27	0,08	-0,08	0,02	0,55	1,00		
Q23	0,26	0,19	0,44	0,06	0,14	0,25	0,24	0,31	0,31	0,12	-0,13	0,09	0,79	0,39	1,00	
Q24	0,29	0,21	0,45	0,09	0,16	0,27	0,27	0,32	0,31	0,10	-0,11	0,09	0,46	0,77	0,59	1,00

Fonte: Elaborado pela autora

Tabela 6. Correlação policórica entre questões, 3º série do EM Saeb

	Q01	Q03	Q04	Q06	Q10	Q11	Q12	Q13	Q14	Q15	Q18	Q19	Q21	Q22	Q23	Q24
Q01	1,00															
Q03	0,57	1,00														
Q04	0,56	0,43	1,00													
Q06	0,37	0,39	0,28	1,00												
Q10	0,46	0,38	0,38	0,30	1,00											
Q11	0,42	0,35	0,49	0,27	0,32	1,00										
Q12	0,44	0,34	0,50	0,24	0,40	0,44	1,00									
Q13	0,51	0,36	0,62	0,24	0,40	0,51	0,51	1,00								
Q14	0,55	0,39	0,58	0,25	0,46	0,46	0,43	0,61	1,00							
Q15	0,36	0,27	0,35	0,23	0,36	0,31	0,28	0,41	0,58	1,00						
Q18	0,04	0,04	-0,04	0,02	0,08	-0,03	0,02	0,03	0,04	0,30	1,00					
Q19	0,28	0,24	0,26	0,11	0,29	0,21	0,22	0,37	0,39	0,25	0,08	1,00				
Q21	0,35	0,25	0,50	0,10	0,21	0,34	0,36	0,47	0,40	0,18	-0,06	0,08	1,00			
Q22	0,35	0,24	0,49	0,12	0,22	0,36	0,38	0,49	0,39	0,18	-0,07	0,09	0,63	1,00		
Q23	0,37	0,26	0,53	0,09	0,21	0,34	0,33	0,46	0,47	0,23	-0,10	0,26	0,82	0,51	1,00	
Q24	0,42	0,30	0,55	0,13	0,22	0,37	0,36	0,47	0,49	0,22	-0,09	0,25	0,57	0,82	0,66	1,00

Fonte: Elaborado pela autora

Tabela 7. Correlação policórica entre questões, 3ª série do EM Enem

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q16	Q17	Q18	Q19	Q20	Q23	Q24
Q01	1,00																				
Q02	0,60	1,00																			
Q03	0,60	0,44	1,00																		
Q04	0,57	0,58	0,45	1,00																	
Q05	0,52	0,62	0,39	0,73	1,00																
Q06	0,38	0,27	0,41	0,32	0,27	1,00															
Q07	0,49	0,59	0,35	0,54	0,57	0,27	1,00														
Q08	0,52	0,39	0,38	0,48	0,44	0,26	0,22	1,00													
Q09	0,50	0,58	0,44	0,57	0,54	0,36	0,56	0,33	1,00												
Q10	0,45	0,39	0,39	0,40	0,37	0,33	0,34	0,35	0,40	1,00											
Q11	0,43	0,47	0,34	0,47	0,44	0,31	0,44	0,29	0,53	0,29	1,00										
Q12	0,45	0,48	0,34	0,51	0,50	0,27	0,45	0,36	0,57	0,42	0,50	1,00									
Q13	0,50	0,51	0,38	0,56	0,51	0,28	0,48	0,42	0,59	0,42	0,49	0,52	1,00								
Q14	0,58	0,53	0,41	0,55	0,51	0,28	0,50	0,47	0,50	0,45	0,45	0,43	0,59	1,00							
Q16	0,10	0,03	0,10	0,04	0,02	0,13	0,07	0,03	0,08	0,13	0,10	0,03	0,16	0,17	1,00						
Q17	0,36	0,43	0,20	0,42	0,47	0,00	0,46	0,30	0,30	0,13	0,08	0,30	0,18	0,33	-0,11	1,00					
Q18	0,00	-0,03	0,02	-0,07	-0,09	-0,01	-0,05	0,07	-0,08	0,05	-0,04	-0,03	0,05	0,02	0,07	-0,14	1,00				
Q19	0,49	0,62	0,44	0,54	0,52	0,27	0,53	0,30	0,53	0,44	0,52	0,41	0,62	0,60	0,18	0,30	0,00	1,00			
Q20	0,54	0,63	0,39	0,64	0,61	0,26	0,57	0,44	0,61	0,38	0,50	0,51	0,64	0,61	0,03	0,42	-0,03	0,66	1,00		
Q23	0,34	0,45	0,25	0,47	0,44	0,10	0,43	0,25	0,40	0,23	0,30	0,31	0,42	0,45	0,01	0,40	-0,12	0,55	0,57	1,00	
Q24	0,39	0,51	0,29	0,50	0,47	0,15	0,48	0,28	0,47	0,26	0,34	0,35	0,43	0,48	0,00	0,43	-0,10	0,53	0,61	0,66	1,00

Fonte: Elaborado pela autora

A partir das tabelas apresentadas acima, observa-se que as questões que apresentaram maiores correlações com as demais questões foram a “Q04 – Possui computador em casa” e a “Q20 - Renda mensal da família (considerando a renda do aluno)”. E as que apresentaram menores correlações foram a “Q18 - Quantas pessoas moram na casa

(incluindo o aluno)” e a “Q16 - Situação residência”, com valores próximos de zero e até negativo para algumas questões.

Quando a correlação entre duas questões é negativa, sugere-se a revisão do ordenamento das alternativas, pois o ordenamento pode estar em ordem decrescente em uma questão e em outra crescente, porém no caso da questão Q16 e Q18 rever o ordenamento da alternativa não é a solução, pois a questão se relaciona positivamente com algumas questões e negativamente com outras.

A unidimensionalidade dos dados foi verificada a partir da análise fatorial e observando o gráfico *screeplot*. Na análise fatorial, verificou-se nos quatro grupos que o primeiro fator corresponde à 25%, 29%, 36% e 41% da variância total para o 5º ano do EF, 9º ano EF, 3ª série EM do Saeb e 3ª série EM Enem, respectivamente. Apesar dos valores apresentados indicarem a predominância de uma dimensão, as questões Q16 e Q18 apresentaram cargas fatoriais (correlação entre a variável e o fator, Hair (2005)) menores que 0,10, enquanto as demais questões valores mais altos, como indicadas na Tabela 8.

Tabela 8. Cargas fatoriais do 1º fator por grupo

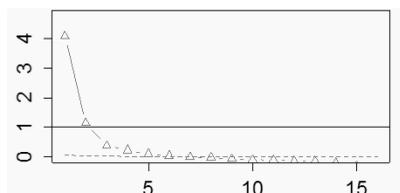
Questões	5º ano EF Saeb	9º ano EF Saeb	3ª série EM Saeb	3ª série EM Enem
Q01	0,55	0,60	0,68	0,72
Q02	-	-	-	0,77
Q03	0,44	0,48	0,52	0,56
Q04	0,72	0,74	0,78	0,79
Q05	-	-	-	0,76
Q06	0,32	0,34	0,33	0,39
Q07	-	-	-	0,70
Q08	-	-	-	0,53
Q09	-	-	-	0,74
Q10	0,42	0,46	0,50	0,53
Q11	0,52	0,54	0,60	0,62
Q12	0,53	0,58	0,61	0,64
Q13	0,63	0,65	0,76	0,73
Q14	0,59	0,59	0,74	0,74
Q15	0,41	0,34	0,48	-
Q16	-	-	-	0,10
Q17	-	-	-	0,48
Q18	-0,02	-0,05	0,00	-0,05
Q19	0,29	0,18	0,37	0,76
Q20	-	-	-	0,82

Questões	5º ano EF Saeb	9º ano EF Saeb	3ª série EM Saeb	3ª série EM Enem
Q21	0,51	0,64	0,66	-
Q22	0,51	0,63	0,67	-
Q23	0,59	0,61	0,69	0,60
Q24	0,59	0,63	0,72	0,65

Fonte: Elaborado pela autora

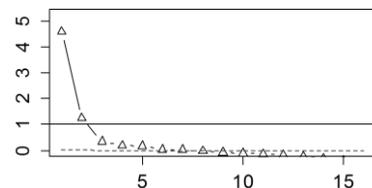
Nos Gráficos 10 a 13 verifica-se a existência de um ou dois fatores, porém, observa-se a predominância de um fator dado o distanciamento desse fator em relação aos demais. O gráfico *screepplot* indica o total da variância explicada pelo número de dimensões, sendo no eixo y os autovalores (variância explicada) e eixo x o número de fatores.

Gráfico 10. Screeplot – Questões do 5º ano do EF Saeb



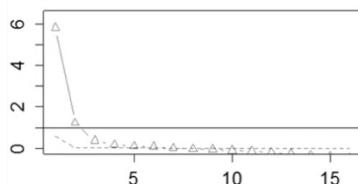
Fonte: Elaborado pela autora

Gráfico 11. Screeplot – Questões do 9º ano do EF Saeb



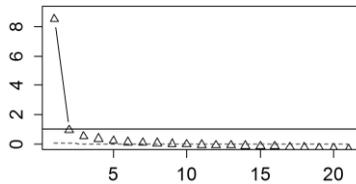
Fonte: Elaborado pela autora

Gráfico 12. Screeplot – Questões da 3ª série do EM Saeb



Fonte: Elaborado pela autora

Gráfico 13. Screeplot – Questões da 3ª série do EM Enem



Fonte: Elaborado pela autora

Os procedimentos relatados nessa seção são ambos exploratórios e foram feitos utilizando toda a população dos quatro bancos de dados. Inicialmente as questões que apresentaram baixa correlação ou valores negativos não foram retiradas das análises, porém serão avaliadas com mais cautela nas próximas seções.

Para dar seguimento as análises, utilizou-se uma amostra dos quatro bancos de dados, conforme já descrito na seção 4.3.

5.2.2 Modelagem dos dados

Busca-se, nesta seção, aplicar os modelos propostos e selecionar aquele que apresentar melhor ajuste aos dados. Optou-se por apresentar inicialmente a modelagem pelos modelos politômicos, dado que será utilizado o modelo de respostas nominais para validar o ordenamento das categorias. Por fim, será apresentada a modelagem ao modelo dicotômico de dois parâmetros.

É importante destacar que referente ao parâmetro b , ele não é calculado diretamente no processo de calibração, no pacote *mirt*, calcula-se o parâmetro d e, em seguida, o parâmetro b é calculado pela fórmula $b = -d/a$. Com base nisso, o pacote não fornece o erro-padrão do parâmetro b e sim do parâmetro d , por isso, nas tabelas desta seção serão apresentados apenas os parâmetro a e b e no apêndice serão apresentados os parâmetro a e b com o erro-padrão do parâmetro d . O erro-padrão do parâmetro a é calculado de forma direta na calibração.

5.2.2.1 Modelos Politômicos

A seguir, serão apresentados os resultados da modelagem dos dados a partir do modelo de respostas nominais, modelo de respostas graduais e o modelo de créditos parciais generalizados.

5.2.2.1.1 Modelo de respostas nominais

O modelo de respostas nominais foi adotado com o intuito de validar empiricamente o ordenamento das alternativas, isto é, verificar se a forma como as alternativas estão ordenadas no banco de dados seguem a mesma ordem sugerida pelo modelo.

Com base no exposto, o modelo nominal foi ajustado aos dados da amostra e verificou-se o ordenamento dos valores de a_k . Na primeira calibração observou-se que as questões exclusivas do Enem Q16, Q17 e Q18 apresentaram valores de a_k sugerindo ordenamento diferente do proposto. Na Tabela 9, são apresentados os valores encontrados para estas questões. As alternativas que apresentaram problemas de ordenamento apresentaram erro-padrão mais alto do que as demais, conforme apresentado no Apêndice A.

Tabela 9. Parâmetros primeira calibração modelo nominal

Q16 - Situação residência		
1 - Outra situação (loteamento não regularizado, ocupação etc.)	a_0	0,00
2 - Cedida	a_1	0,33
3 - Alugada	a_2	2,90
4 - Própria e em pagamento (financiada)	a_3	10,05
5 - Própria e quitada	a_4	4,00
Q17 - Localização residência		
1 - Comunidade indígena ou quilombola	a_0	0,00
2 - Zona rural	a_1	7,02
3 - Zona urbana	a_2	2,00
Q18 - Quantas pessoas moram na casa (incluindo o aluno)		
1 - Uma, pois moro sozinho(a)	a_0	0,00
2 - Duas	a_1	6,07
3 - Três	a_2	-1,34
4 - Quatro	a_3	-4,67
5 - Cinco	a_4	-0,72
6 - Seis pessoas ou mais	a_5	5,00

Fonte: Elaborado pela autora

Com base no exposto optou-se por fazer uma segunda calibração com as seguintes alterações: Q16 - agrupar as alternativas 4 e 5, Q17 - desconsiderar a questão da análise dado que teoricamente não faz sentido considerar que em geral os indivíduos que residem na área rural possuem maior nível socioeconômico dos que residem na área urbana e

Q18 – retirar a questão da análise por também não fazer sentido o ordenamento proposto.

Os resultados da segunda calibração indicaram que a questão Q16 continua com ordenamento empírico diferente do proposto pelo modelo e por isso optou-se por retirar a questão do estudo. A partir dos resultados apresentados, acredita-se que essa medida não possui uma relação direta com o construto nível socioeconômico. Talvez isso se deva ao fato de, por exemplo, existirem famílias residindo em residências pequenas e quitadas e outras em grandes mansões em loteamentos não regularizados, sendo assim, a questão não mede diretamente a dimensão socioeconômica. Na Tabela 10, são apresentados os valores encontrados para esta questão.

Tabela 10. Parâmetros segunda calibração modelo nominal

Q16 - Situação residência		
1 - Outra situação (loteamento não regularizado, ocupação etc.)	a_0	0,00
2 – Cedida	a_1	-0,11
3 – Alugada	a_2	1,45
4 - Própria e em pagamento (financiada) ou quitada	a_3	3,00

Fonte: Elaborado pela autora

Por fim, uma terceira calibração foi feita desconsiderando das análises as questões Q16, Q17 e Q18 e observou-se que os ordenamentos propostos empiricamente são os mesmos propostos pelo modelo nominal. Na Tabela 11, são apresentados os valores de a_k encontrados para as questões.

Tabela 11. Parâmetros e erro padrão terceira calibração modelo nominal

	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
Q01	0,00	0,28	1,59	3,00				
Q02	0,00	1,13	2,11	3,00				
Q03	0,00	0,70	2,04	3,00				
Q04	0,00	1,16	2,20	3,00				
Q05	0,00	1,63	2,36	3,00				
Q06	0,00	0,52	2,08	3,00				
Q07	0,00	1,33	2,56	3,00				
Q08	0,00	1,24	1,73	3,00				
Q09	0,00	1,50	2,88	3,00				
Q10	0,00	1,32	2,36	3,00				
Q11	0,00	1,00						
Q12	0,00	1,68	2,55	3,00				

	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
Q13	0,00	1,26	2,45	3,00				
Q14	0,00	0,91	1,87	3,00				
Q15	0,00	0,80	2,01	2,98	4,00			
Q19	0,00	1,00						
Q20	0,00	0,52	1,83	2,90	4,45	5,53	6,03	7,00
Q21	0,00	1,00						
Q22	0,00	1,00						
Q23	0,00	1,38	2,20	3,00				
Q24	0,00	1,32	2,17	3,00				

Fonte: Elaborado pela autora

5.2.2.1.2 Modelo de respostas graduais – Samejima

Nesta seção serão apresentados os resultados da modelagem ao modelo de respostas graduais. A partir dos resultados encontrados na modelagem pelo modelo nominal, optou-se por seguir com as análises já excluindo as questões Q16, Q17 e Q18. A seguir, a Tabela 12 apresenta os parâmetros das questões da primeira calibração.

Tabela 12. Parâmetros primeira calibração modelo respostas graduais

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q01	1,37	-3,02	-0,04	1,54				
Q02	1,72	0,79	2,87	3,66				
Q03	0,97	-1,99	1,87	3,88				
Q04	1,88	-0,55	1,35	2,38				
Q05	2,13	-0,75	2,11	2,68				
Q06	0,61	-2,06	2,64	5,32				
Q07	1,52	-0,06	2,72	3,72				
Q08	0,97	-4,22	-1,01	0,20				
Q09	1,57	1,32	4,26	5,66				
Q10	1,06	-4,12	2,64	4,86				
Q11	0,89	-0,30						
Q12	1,12	-1,16	3,00	5,09				
Q13	1,44	0,07	2,03	3,37				
Q14	1,51	-3,44	0,81	2,34				
Q15	0,80	-6,02	-2,74	0,26	3,03			
Q19	0,73	3,44						
Q20	2,39	-2,72	-0,70	-0,07	0,85	1,80	2,12	2,49
Q21	1,31	-2,64						
Q22	1,16	-2,31						

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q23	1,02	-3,76	0,30	2,08				
Q24	1,12	-2,90	0,59	2,25				

Fonte: Elaborado pela autora

A partir dos resultados apresentados acima, observa-se que os parâmetros das questões estão dentro do esperado e com isso não foi necessário optar por outra calibração.

As questões variaram de acordo com o parâmetro a de 0,61 a 2,39 e do parâmetro b de -6,02 a 5,66. A grande variação nos valores encontrados para o parâmetro b indica que existem questões/alternativas que estão medindo o construto em seus diversos níveis.

5.2.2.1.3 Modelo de crédito parcial generalizado

Da mesma forma como foi feito no modelo de respostas graduais, optou-se por nesta seção retirar as questões Q16, Q17 e Q18. Com isso, para as demais questões foi aplicado o modelo de crédito parcial generalizado. A seguir, na Tabela 13, são apresentados os parâmetros encontrados na primeira calibração.

Tabela 13. Parâmetros primeira calibração modelo de crédito parcial generalizado

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q01	1,07	-3,21	0,23	1,27				
Q02	1,54	0,88	2,87	2,37				
Q03	0,79	-2,11	2,28	2,79				
Q04	1,53	-0,51	1,51	1,95				
Q05	1,86	-0,80	2,38	1,86				
Q06	0,46	-2,08	3,35	3,22				
Q07	1,30	-0,04	2,96	2,33				
Q08	0,59	-4,90	0,01	-0,95				
Q09	1,54	1,34	3,85	3,61				
Q10	0,97	-4,35	2,88	3,45				
Q11	0,93	-0,26						
Q12	1,04	-1,18	3,09	3,63				
Q13	1,15	0,22	2,08	2,62				
Q14	1,27	-3,75	1,05	1,98				
Q15	0,55	-4,97	-2,80	0,33	2,77			
Q19	0,78	3,25						

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q20	1,35	-3,46	-0,32	-0,27	0,91	2,27	1,63	1,57
Q21	1,31	-2,59						
Q22	1,17	-2,25						
Q23	0,74	-4,29	0,76	1,55				
Q24	0,83	-3,27	0,99	1,75				

Fonte: Elaborado pela autora

A partir dos resultados apresentados na Tabela 13, observa-se que os parâmetros das questões estão dentro do esperado e com isso não foi necessário optar por outra calibração.

5.2.2.2 Modelo dicotômico

Nesta seção, o modelo que será ajustado aos dados será o modelo dicotômico de dois parâmetros e como o próprio nome diz serve para questões dicotômicas em que indicam a presença ou ausência de determinada característica. Com base nas informações adquiridas na seção referente aos modelos politômicos optou-se já na primeira calibração excluir as questões Q16, Q17 e Q18. E para prosseguir com a modelagem foi necessário dicotomizar as questões para que o banco de dados indique apenas a presença ou ausência de determinadas características. Desta forma as questões foram reagrupadas da seguinte forma:

Quadro 2. Reagrupamento das questões para ajuste modelo dicotômico

Reagrupamento		
Q01 - Televisão em cores em casa	1 - Não tem	2 - Sim, uma ou mais
Q02 - TV por assinatura em casa	1 - Não tem	2 - Sim, uma ou mais
Q03 - Videocassete e/ou DVD em casa	1 - Não tem	2 - Sim, um ou mais
Q04 - Computador em casa	1 - Não tem	2 - Sim, um ou mais
Q05 - Internet em casa	1 - Não tem	2 - Sim, um ou mais
Q06 - Rádio em casa	1 - Não tem	2 - Sim, um ou mais
Q07 - Telefone fixo em casa	1 - Não tem	2 - Sim, um ou mais
Q08 - Telefone celular em casa	1 - Não tem	2 - Sim, um ou mais
Q09 - Aspirador de pó em casa	1 - Não tem	2 - Sim, um ou mais
Q10 - Geladeira em casa	1 - Não tem	2 - Sim, uma ou mais
Q11 - Freezer em casa	1 - Não tem	2 - Sim, um ou mais

Reagrupamento		
Q12 - Máquina de lavar roupa em casa	1 - Não tem	2 - Sim, uma ou mais
Q13 - Automóvel em casa	1 - Não tem	2 - Sim, um ou mais
Q14 - Banheiro em casa	1 - Não tem	2 - Sim, um ou mais
Q15 - Quartos para dormir em casa	1 - Não tem	2 - Sim, um ou mais
Q19 - Contrata empregada doméstica em casa	1 - Não	2 - Sim, um ou mais
Q20 - Renda mensal da família	1 - Nenhuma renda até 7 s.m	Acima de 7 s.m
Q21 - Mãe, ou a mulher responsável por você, sabe ler e escrever	1 - Não	2 - Sim
Q22 - Pai, ou homem responsável por você, sabe ler e escrever	1 - Não	2 - Sim
Q23 - Até que série a mãe estudou	1 - Até Ens. Fund. Completo	2 - Ens. Médio Completo ou mais
Q24 - Até que série o pai estudou	1 - Até Ens. Fund. Completo	2 - Ens. Médio Completo ou mais

Fonte: Elaborado pela autora

Após a dicotomização das questões, os dados dos quatro bancos foram ajustados ao modelo dicotômico de dois parâmetros. A seguir na Tabela 14 são apresentados os resultados da primeira calibração.

Tabela 14. Parâmetros primeira calibração modelo dicotômico

Questões	a	b_1
Q01	0,87	-4,36
Q02	1,61	0,84
Q03	0,62	-3,03
Q04	2,24	-0,61
Q05	3,34	-0,65
Q06	0,31	-4,11
Q07	1,64	-0,05
Q08	1,29	-3,42
Q09	1,54	1,34
Q10	1,39	-3,49
Q11	0,76	-0,49
Q12	1,31	-1,17
Q13	1,33	-0,05
Q14	1,48	-3,61
Q15	1,30	-4,22
Q19	0,54	4,24
Q20	2,97	1,68
Q21	1,54	-2,53

Questões	a	b_1
Q22	1,40	-2,19
Q23	1,26	0,18
Q24	1,36	0,44

Fonte: Elaborado pela autora

A partir dos resultados apresentados acima, observa-se que as questões apresentaram parâmetros dentro do esperado para a modelagem.

Na próxima seção será escolhido o melhor modelo, dentre os utilizados nesta seção, para a criação do indicador de nível socioeconômico.

5.2.3 Escolha do modelo

Nas seções anteriores ocorreram várias tentativas de ajustar modelos da TRI aos dados. Algumas questões foram eliminadas a medida que apresentavam resultados que indicavam algum grau de desajuste com modelo e com isso, de forma interativa, as questões eram recodificadas ou eliminadas e novas calibrações eram feitas até encontrar estimativas que indicavam uma maior adequabilidade aos modelos.

A seguir será apresentado o procedimento que nortearam a escolha do melhor modelo.

5.2.3.1 Curva de Informação (CI)

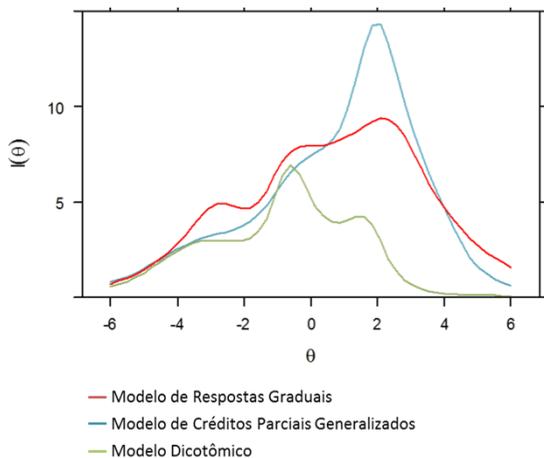
A seguir, são apresentadas as CI's do instrumento dos três modelos e para facilitar a comparação os gráficos foram sobrepostos, porém a análise desse gráfico precisa ser feita com cautela, pois as medidas de Inse (θ) para cada modelo não estão na mesma escala.

A partir do Gráfico 14, verifica-se que nas três curvas o modelo dicotômico é o que apresenta menor informação, com valor máximo próximo de 6, e a sua curva é mais assimétrica à direita com pouquíssima informação para alunos com nível socioeconômico mais alto. Esse tipo de comportamento é natural quando comparado com os demais modelos, dado que agrupar as alternativas implica em perder informação.

Já em relação ao modelo de créditos parciais generalizado, a curva apresenta máxima informação em valores de Inse próximos de 14 e a informação é maior com valores maiores que zero.

Por fim, o modelo de respostas graduais apresenta valor máximo de informação próximo de 9 e a sua CI está mais distribuída para os diversos valores de nível socioeconômico.

Gráfico 14. Curva de Informação do Instrumento dos três modelos



Fonte: Elaborado pela autora

A partir dos resultados apresentados, optou-se por escolher o modelo de respostas graduais para cálculo do Inse, dado que o objetivo é estimar de forma melhor possível o Inse para todas as faixas de Inse e optar pelo modelo de créditos parciais generalizados apesar de ser mais eficiente em algumas faixas em outras a informação seria inferior ao modelo de respostas graduais. Caso o objetivo desse indicador fosse diferenciar melhor os indivíduos de nível socioeconômico mais alto e não os de Inse mais baixo, utilizar o modelo de créditos parciais generalizados seria uma boa opção.

5.2.4 Ajuste do modelo

Pode-se verificar que é bastante trabalhoso o processo de criação de um indicador pela TRI, entretanto todos os esforços feitos nas seções anteriores e os resultados encontrados perdem sua validade quando o modelo utilizado não se ajusta adequadamente aos dados. Entretanto, Glas (2010) destaca que nenhum modelo vai se ajustar aos dados perfeitamente e que o ajuste ao modelo cabe tanto aos itens como aos respondentes.

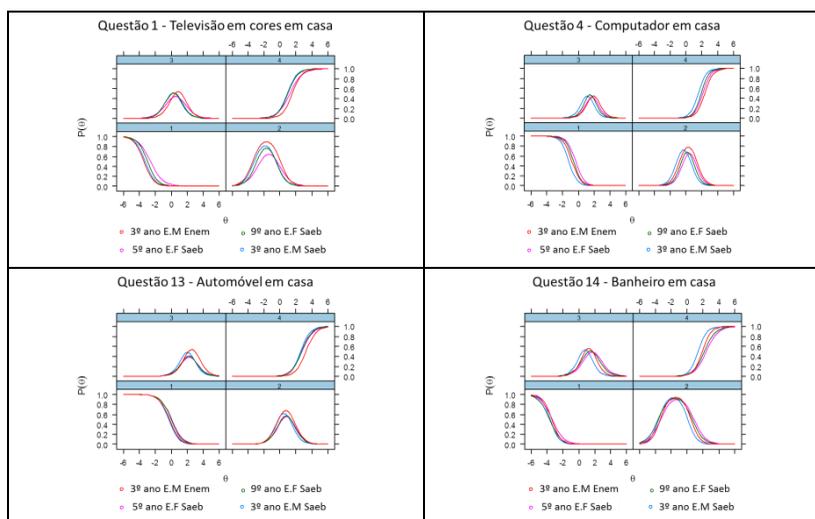
Existem diversos testes para verificar ajuste dos dados, porém esses testes são bastante sensíveis ao tamanho da amostra, e quando a amostra é muito grande, tende-se a rejeitar a hipótese nula. Por isso optou-se por escolher procedimentos que são poucos ou não são influenciado pelo tamanho da amostra. Outra questão relevante é que muito desses procedimentos não trabalham com dados faltantes, tendo como opção imputar dados ou retirar aqueles que não responderam todas as questões selecionadas. Neste trabalho optou-se, quando necessário, pela segunda opção.

5.2.4.1 Análise do Comportamento Diferencial dos Itens – DIF

Para a investigação das questões que possuem DIF primeiramente os parâmetros das questões foram estimados livremente por meio de grupos múltiplos e as questões que indicaram menos evidência de DIF foram fixadas e, novamente, as demais questões foram estimadas livremente por meio de grupos múltiplos. Este procedimento é indicado no manual do pacote e é feito para obter melhores resultados.

A seguir são apresentadas as curvas características dos grupos das quatro questões selecionadas para serem fixadas. Nos gráficos são apresentados em cada quadrante as alternativas separadamente para facilitar a comparação entre os grupos.

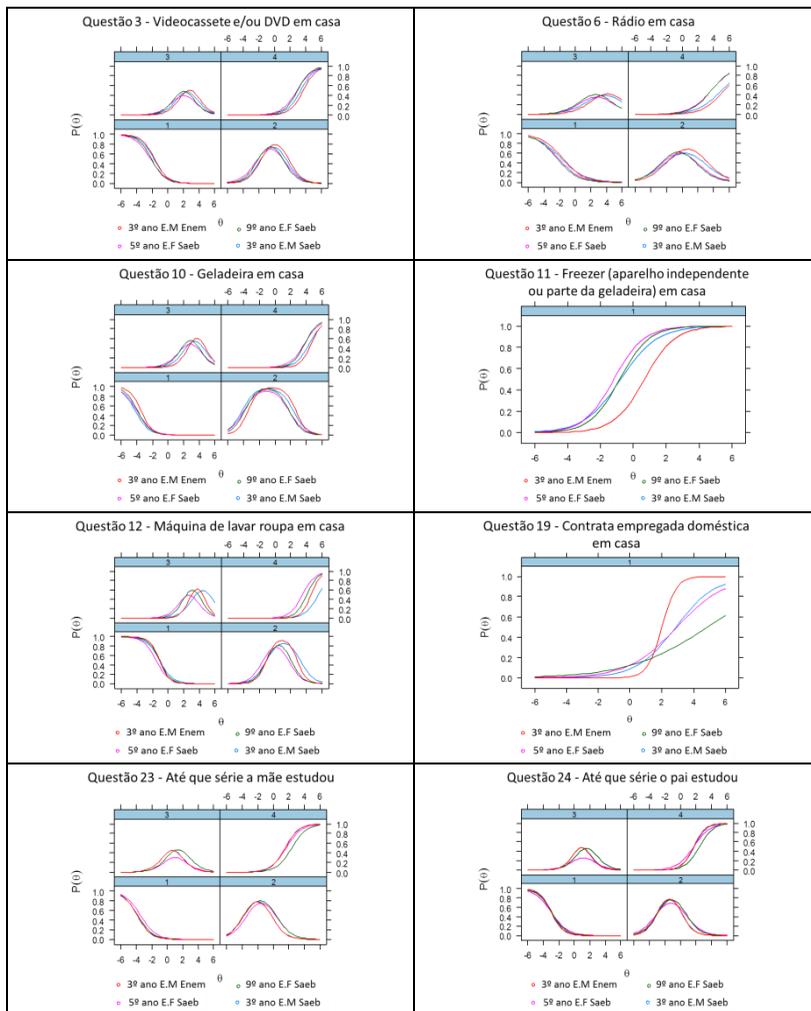
Ilustração 4. CCI questões relacionadas para serem fixadas



Fonte: Elaborado pela autora

Após o procedimento descrito na Ilustração 4, as curvas características foram calculadas novamente para as questões em que os parâmetros foram estimados livremente.

Ilustração 5. CCI questões com parâmetros estimados livremente



Fonte: Elaborado pela autora

A partir dos gráficos apresentados na ilustração acima observa-se que existe comportamento diferencial nas questões: “Questão 11 -

Freezer (aparelho independente ou parte da geladeira) em casa”, “Questão 12 - Máquina de lavar roupa em casa” e “Questão 19 - Contrata empregada doméstica em casa”.

A questão 11 apresentou DIF principalmente entre os grupos do Saeb e do Enem isso talvez se deva ao fato de a pergunta, apesar de aparentemente significar a mesma coisa, estar sendo feita de forma diferente, e talvez o entendimento de quem responde possa ocorrer de outra forma refletindo em “questões diferentes”. Este tema foi abordado no Saeb por meio de duas questões já no Enem a questão já englobava as duas questões do Saeb. A seguir, o Quadro 3 indica a forma como a questão foi perguntada nos grupos.

Quadro 3. Comparativo questão freezer Saeb e Enem

SAEB	Na sua casa tem freezer (parte da geladeira duplex)?
	Na sua casa tem freezer separado da geladeira?
Enem	Você tem em sua casa? Freezer (aparelho independente ou parte da geladeira duplex)
Recodificação Inse	Freezer (aparelho independente ou parte da geladeira) em casa

Fonte: Elaborado pela autora

Como ao final da modelagem deseja-se propor uma escala interpretada para o Inse, não é adequado considerar essas questões como novas, pois a interpretação seria para cada grupo nas questões que apresentarem DIF, tornando o entendimento da escala mais confusa. Com isso, optou-se por considerar essas três questões (Q11, Q12 e Q19) apenas no grupo de referência.

5.2.4.2 Independência local e unidimensionalidade

A independência local das questões e unidimensionalidade do instrumento utilizado na TRI é um dos pressupostos necessários para garantir a validade do modelo.

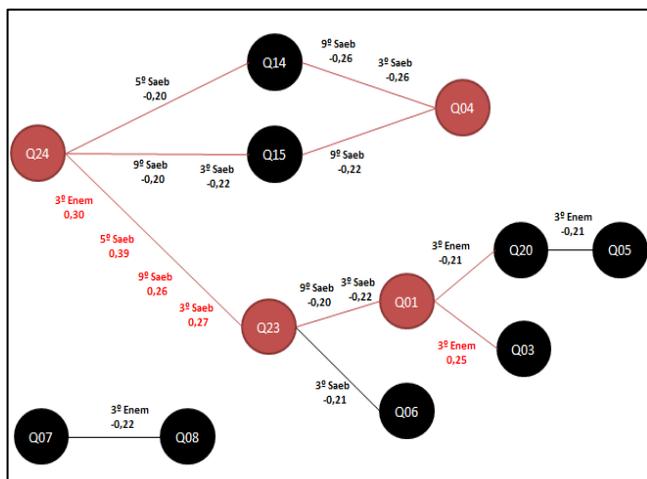
5.2.4.2.1 Independência local - Índice Q₃

É muito importante verificar o pressuposto da independência local, pois a violação dele impacta na estimação dos parâmetros das questões e conseqüentemente na informação do modelo. Desta forma, a

análise de independência local foi feita por grupo e apenas considerando os indivíduos que responderam todas as questões selecionadas, dado que o pacote utilizado não permite trabalhar com dados faltantes. Com isso o banco de dados da amostra apresentou inicialmente para 5º ano EF, 9º ano EF e 3ª série EM Saeb e 3ª série EM Enem respectivamente 9.090, 14.984, 18.742 e 22.591 respondentes.

O Esquema 1, a seguir, ilustra as relações entre pares de questões apenas para os valores do índice Q_3 maiores que $|0,2|$. Destacou-se em vermelho as correlações com valores positivos e as questões que tiveram mais incidências de correlações com as outras questões.

Esquema 1. Relações entre pares de questões Índice Q3 primeira calibração



Fonte: Elaborado pela autora

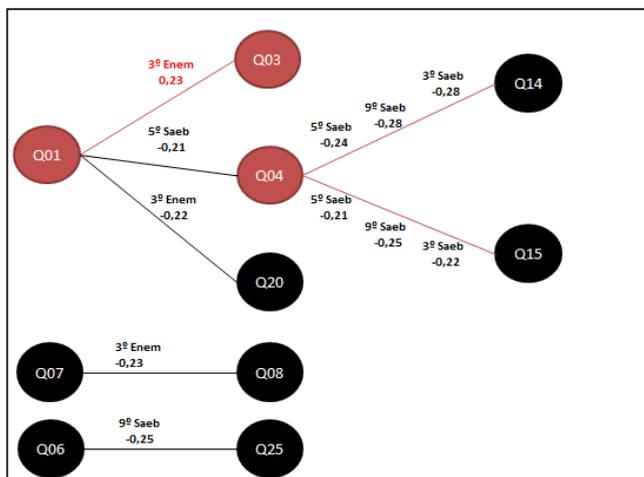
Observa-se nos resultados apresentados, Esquema 1, que as questões “Q23 - Até que série a mãe estudou” e “Q24 - Até que série o pai estudou” apresentaram valores positivos do índice Q_3 para os quatro grupos, indicando dependência local entre as questões e que possivelmente compartilham algum outro fator. A questão “Q01 - Televisão em cores em casa” apresentou correlação positiva também com a “Q03 - Videocassete e/ou DVD em casa”.

Inicialmente como as questões Q23 e Q24 apresentaram valores não desejados para o modelo no índice Q_3 nos quatro grupos optou-se por agrupar as duas questões em uma nova questão Q25, que irá representar a maior escolaridade dos pais ou responsáveis e fazer uma

nova tentativa de modelagem. Outra opção seria trabalhar apenas com a “Q23 - Até que série a mãe estudou”, porém optou-se por criar uma nova questão, pois agregaria mais informação ao indicador e, no caso de *missing* em uma das duas questões, diminuiria a taxa de não resposta.

O Esquema 2 ilustra as relações entre as questões que apresentaram valores do índice Q_3 maiores que $|0,2|$ para a nova modelagem.

Esquema 2. Relações entre pares de questões Índice Q_3 segunda calibração



Fonte: Elaborado pela autora

Observa-se que o par “Q01 - Televisão em cores em casa” e “Q03 - Videocassete e/ou DVD em casa” apresentou para o grupo do Enem valor positivo indicando dependência local entre as questões e que possivelmente compartilham algum outro fator. Além disso, continuaram apresentando indicativos de dependência local nessa segunda calibração a questão “Q04 - Computador em casa” com as “Q14 - Banheiro em casa” e “Q15 - Quartos para dormir em casa” e a questão “Q07 - Telefone fixo em casa” e “Q08 - Telefone celular em casa”. Desta forma, optou-se fazer uma terceira modelagem excluindo as questões Q01, Q04 e Q07.

A seguir nas Tabelas 15, 16, 17 e 18, são apresentadas as tabelas com o índice Q_3 para os quatro grupos com os valores obtidos na terceira calibração, em cinza as células com valores maiores que $|0,2|$. Optou-se por não fazer outras tentativas de calibração com os pares de questões que indicaram dependência local, isto porque o índice Q_3 é

sensível ao número de questões do instrumento e com isso os resultados apresentados por não terem aparecidos nas calibrações anteriores podem indicar “falsos positivos” devido ao baixo quantitativo de questões para o grupo do Saeb. No grupo do Enem nenhum par de questões apresentou índice Q_3 maior que $|0,2|$.

Tabela 15. Índice Q_3 , 5º ano do Ensino Fundamental Saeb

	Q03	Q06	Q10	Q13	Q14	Q15	Q21	Q22	Q25
Q03									
Q06	0,12								
Q10	0,04	0,08							
Q13	-0,10	-0,05	-0,09						
Q14	-0,16	-0,13	-0,08	-0,27					
Q15	-0,10	-0,03	-0,05	-0,17	-0,04				
Q21	-0,06	-0,07	-0,08	-0,05	-0,10	-0,12			
Q22	-0,06	-0,07	-0,09	-0,05	-0,13	-0,14	0,10		
Q25	-0,09	-0,14	-0,15	-0,11	-0,20	-0,18	0,06	0,07	

Fonte: Elaborado pela autora

Tabela 16. Índice Q_3 , 9º ano do Ensino Fundamental Saeb

	Q03	Q06	Q10	Q13	Q14	Q15	Q21	Q22	Q25
Q03									
Q06	0,16								
Q10	0,03	0,07							
Q13	-0,10	-0,05	-0,07						
Q14	-0,17	-0,13	-0,09	-0,28					
Q15	-0,11	-0,03	-0,02	-0,18	0,02				
Q21	-0,05	-0,06	-0,05	-0,04	-0,09	-0,15			
Q22	-0,05	-0,06	-0,05	0,00	-0,11	-0,17	0,13		
Q25	-0,08	-0,15	-0,15	-0,11	-0,17	-0,25	0,12	0,09	

Fonte: Elaborado pela autora

Tabela 17. Índice Q_3 , 3º série do Ensino Médio Saeb

	Q03	Q06	Q10	Q13	Q14	Q15	Q21	Q22	Q25
Q03									
Q06	0,20								
Q10	0,02	0,05							
Q13	-0,12	-0,09	-0,11						
Q14	-0,19	-0,15	-0,14	-0,22					
Q15	-0,10	-0,04	-0,04	-0,18	-0,01				
Q21	-0,04	-0,06	-0,05	-0,03	-0,07	-0,11			
Q22	-0,05	-0,06	-0,06	-0,02	-0,09	-0,13	0,18		
Q25	-0,10	-0,18	-0,16	-0,09	-0,12	-0,25	0,11	0,09	

Fonte: Elaborado pela autora

Tabela 18. Índice Q₃, 3º série do Ensino Médio Enem

	Q02	Q03	Q05	Q06	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q19	Q20	Q25
Q02														
Q03	-0,01													
Q05	0,04	-0,04												
Q06	-0,05	0,19	-0,03											
Q08	-0,06	0,08	-0,03	0,04										
Q09	-0,07	-0,01	-0,07	0,05	-0,11									
Q10	-0,06	0,03	-0,07	0,06	0,03	-0,02								
Q11	-0,05	-0,01	-0,06	0,06	-0,09	0,00	-0,04							
Q12	-0,06	-0,05	-0,08	0,00	-0,06	0,03	0,02	0,04						
Q13	-0,16	-0,06	-0,13	-0,04	-0,04	0,00	-0,03	-0,02	-0,01					
Q14	-0,15	-0,01	-0,13	-0,03	0,02	-0,16	0,00	-0,07	-0,09	-0,03				
Q19	-0,03	-0,05	-0,09	-0,05	-0,08	-0,06	-0,05	0,01	-0,05	-0,02	-0,04			
Q20	-0,14	-0,17	-0,19	-0,15	-0,13	-0,15	-0,17	-0,14	-0,19	-0,11	-0,19	-0,08		
Q25	-0,05	-0,11	-0,01	-0,15	-0,11	-0,10	-0,10	-0,13	-0,12	-0,08	-0,03	0,00	0,10	

Fonte: Elaborado pela autora

Nesta seção verificou-se a independência local entre as questões que serão utilizadas para criar o indicador de nível socioeconômico, decidiu-se por retirar as questões Q01, Q04 e Q07 e agrupar as questões Q23 e Q24 devido o indicativo de dependência local. Esse procedimento é de fundamental importância para o processo de criação do indicador, pois, manter questões dependentes entre si pode impactar na estimação dos parâmetros sem contar que implica em redundância. Retirar questões dependentes implica também em uma coleta com mais qualidade dado que reflete em questionários menores.

5.2.4.2.2 Questões e parâmetros das questões

Conhecer as características das questões utilizados na modelagem é um processo fundamental para a produção de um indicador de qualidade. Esse procedimento é feito analisando os valores dos parâmetros encontrados e também a partir de gráficos.

O modelo que melhor se ajustou aos dados foi o modelo politômico de resposta graduais e as questões selecionadas foram:

- Q02 - TV por assinatura em casa;
- Q03 - Videocassete e/ou DVD em casa;
- Q05 - Internet em casa;
- Q06 - Rádio em casa;
- Q08 - Telefone celular em casa;
- Q09 - Aspirador em pó em casa;
- Q10 - Geladeira em casa;

- Q11 - Freezer (aparelho independente ou parte da geladeira) em casa;
- Q12 - Máquina de lavar roupa em casa;
- Q13 - Automóvel em casa;
- Q14 - Banheiro em casa;
- Q15 - Quartos para dormir em casa;
- Q19 - Contrata empregada doméstica em casa;
- Q20 - Renda mensal da família (considerando a renda do aluno);
- Q21 - Mãe, ou a mulher responsável por você, sabe ler e escrever;
- Q22 - Pai, ou homem responsável por você, sabe ler e escrever e
- Q25 - Maior escolaridade pais ou responsáveis.

Além disso, as questões Q03, Q06, Q10, Q13, Q14 e Q25 são comuns entre os quatro grupos; as Q15, Q21, Q22 são presente nos grupos do Saeb; e as questões Q02, Q05, Q08, Q09, Q11, Q12, Q19, Q20 pertencentes apenas ao grupo do Enem.

Na Tabela 19, a seguir, são apresentados os parâmetros das questões da última calibração.

Tabela 19. Parâmetros questões

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q02	1,638	0,809	2,917	3,717				
Q03	0,844	-2,341	1,922	4,151				
Q05	1,850	-0,811	2,222	2,834				
Q06	0,559	-2,343	2,702	5,595				
Q08	0,966	-4,233	-1,015	0,189				
Q09	1,618	1,289	4,123	5,486				
Q10	1,060	-4,214	2,508	4,716				
Q11	1,009	0,825						
Q12	1,579	-1,062	3,032	4,852				
Q13	1,408	-0,044	1,923	3,274				
Q14	1,840	-3,173	0,644	2,028				
Q15	1,039	-4,962	-2,367	0,108	2,410			
Q19	2,366	2,109						
Q20	2,507	-2,685	-0,696	-0,078	0,836	1,757	2,069	2,416
Q21	1,271	-2,838						
Q22	1,060	-2,619						
Q25	0,981	-4,372	-0,212	1,617				

Fonte: Elaborado pela autora

Observa-se que o parâmetro a variou de 0,56 a 2,51, e o parâmetro b , variou de -4,96 a 5,60.

A questão “Q06 - Rádio em casa” foi a que obteve a menor discriminação e a única questão com valor inferior a 0,7. Algumas tentativas de reagrupamento das categorias foram feitas objetivando o aumento da discriminação, porém não foram bem sucedidas. Devido a isso e pelo fato das estimativas de erro padrão serem baixas, (Apêndice E) optou-se por não alterar esta questão.

A questão “Q09 - Aspirador em pó em casa” apresentou medida de erro padrão para os parâmetros “ d_2 ” e “ d_3 ” superiores aos das demais questões bem como a questão “Q12 - Máquina de lavar roupa em casa” para o parâmetro “ d_3 ”. Com base nesses resultados optou-se por agrupar as alternativas 2, 3 e 4 da Q09 e 3 e 4 da Q12. Na Tabela 20, a seguir, são apresentados os resultados dessa nova calibração.

Tabela 20. Parâmetros das questões – Calibração Final

Questões	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7
Q02	1,638	0,808	2,917	3,717				
Q03	0,843	-2,345	1,922	4,153				
Q05	1,850	-0,813	2,222	2,833				
Q06	0,558	-2,348	2,704	5,601				
Q08	0,967	-4,232	-1,016	0,188				
Q09	1,575	1,310						
Q10	1,059	-4,220	2,508	4,716				
Q11	1,008	0,825						
Q12	1,571	-1,067	3,041					
Q13	1,407	-0,046	1,922	3,273				
Q14	1,841	-3,175	0,642	2,026				
Q15	1,039	-4,966	-2,369	0,106	2,409			
Q19	2,372	2,105						
Q20	2,510	-2,687	-0,697	-0,079	0,834	1,754	2,066	2,412
Q21	1,271	-2,840						
Q22	1,060	-2,622						
Q25	0,981	-4,374	-0,213	1,615				

Fonte: Elaborado pela autora

Com base na Tabela 20, observa-se que o parâmetro a variou de 0,56 a 2,51, as questões que apresentaram maiores valores foram a “Q20 - Renda mensal da família (considerando a renda do aluno)”, “Q19 - Contrata empregada doméstica em casa” e “Q05 - Internet em casa”. E

as que apresentaram menores valores foram “Q06 - Rádio em casa” e “Q03 - Videocassete e/ou DVD em casa”.

Já no que se refere ao parâmetro b , variou de -4,97 a 5,60 e as questões “Q15 - Quartos para dormir em casa” e “Q25 - Maior escolaridade dos pais ou responsáveis” apresentaram menores valores e as “Q06 - Rádio em casa” e “Q10 - Geladeira em casa” maiores valores.

Ao analisar as medidas de erro padrão, Apêndice E, verifica-se que para o parâmetro a os valores variaram de 0,01 a 0,05 e para parâmetro d variaram de 0,00 a 0,09.

A partir dos resultados encontrados acima e nas seções anteriores optou-se por não fazer mais nenhuma outra calibração e prosseguir com as análises.

5.2.4.2.3 Unidimensionalidade

Na seção 5.2.1, a unidimensionalidade foi avaliada considerando 24 questões, estas questões haviam sido selecionadas inicialmente para a composição do Inse e após todos os procedimentos descritos nas seções anteriores optou-se por um modelo com 17 questões.

No decorrer do trabalho foram feitos vários procedimentos que alteraram a estrutura do banco de dados inicial, indicados a seguir:

1. As questões Q16, Q17 e Q18 foram retiradas do estudo, pois o ordenamento das alternativas do banco de dados não coincidiu com o proposto pelo modelo nominal, sugerindo que não existia uma relação direta entre as questões com o construto nível socioeconômico e a proposta do modelo nominal não fazia sentido teórico.
2. Após a escolha do melhor modelo, por meio da curva de informação, verificou-se o comportamento diferencial dos itens e constatou-se que as questões Q11, Q12 e Q19 apresentavam DIF. Com isso, optou-se por considerar essas questões apenas no grupo de referência, a 3ª série EM Enem.
3. Foi analisada a existência de dependência entre pares de questões. As questões Q01, Q04 e Q07 foram retiradas do modelo por apresentarem dependência com outras questões e as questões Q23 e Q24, por também apresentarem dependência, foram agrupadas de modo a selecionar a maior resposta desse par compondo a questão Q25, que corresponde ao maior grau de escolaridade dos pais ou responsáveis.

4. Por fim, verificou-se as características de cada questão no que se refere aos parâmetros e seu erro padrão. A medida de erro padrão para os parâmetros “ b_2 ” e “ b_3 ” da questão Q09 apresentaram valores superiores aos das demais questões bem como os da questão Q12 para o parâmetro “ b_3 ”. Com base nesses resultados optou-se por agrupar as alternativas 2, 3 e 4 da questão Q09 e 3 e 4 da Q12.

Com base no exposto, faz-se necessário avaliar a unidimensionalidade novamente com base apenas nas questões que foram selecionadas, isto é, no novo banco de dados, para compor o indicador de nível socioeconômico.

De acordo com o critério de Reckase (1979), o qual sugere unidimensionalidade quando o primeiro fator da análise fatorial corresponde a pelo menos 20% da variância total, os instrumentos estudados apresentaram para o 5º ano EF, 9º ano EF e 3ª série EM do Saeb e 3ª série EM do Enem, respectivamente 25%, 29%, 33% e 39% da variância total. Esses resultados, de acordo com o critério apresentado, sugerem predominância de uma dimensão nos dados. A seguir, na Tabela 21, são apresentadas as cargas fatoriais do 1º fator por grupo, verifica-se que em geral as correlações são altas apresentando valor em torno de 0,5.

Tabela 21. Cargas Fatorias do 1º fator por grupo no modelo final

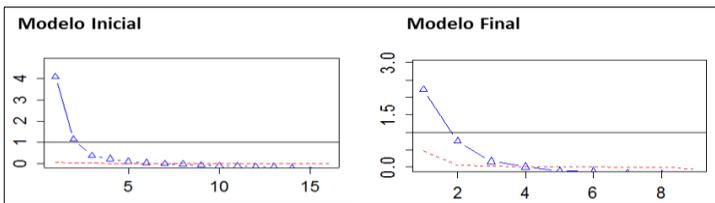
Questões	5º ano EF Saeb	9º ano EF Saeb	3º série EM Saeb	3º série EM Enem
Q02	-	-	-	0,68
Q03	0,43	0,43	0,40	0,49
Q05	-	-	-	0,69
Q06	0,32	0,30	0,23	0,34
Q08	-	-	-	0,45
Q09	-	-	-	0,68
Q10	0,41	0,44	0,40	0,52
Q11	-	-	-	0,53
Q12	-	-	-	0,61
Q13	0,60	0,62	0,65	0,71
Q14	0,63	0,60	0,66	0,67
Q15	0,48	0,38	0,43	-
Q19	-	-	-	0,81
Q20	-	-	-	0,79

Questões	5º ano EF Saeb	9º ano EF Saeb	3º série EM Saeb	3º série EM Enem
Q21	0,49	0,68	0,75	-
Q22	0,47	0,62	0,68	-
Q25	0,55	0,66	0,71	0,57

Fonte: Elaborado pela autora

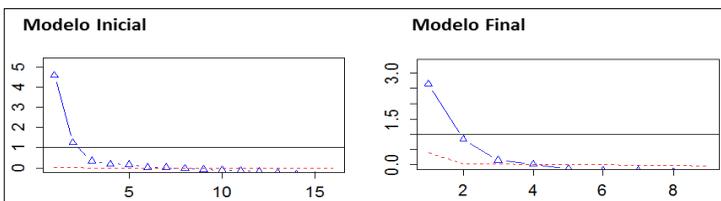
As Ilustrações de 6 a 9, apresentam o *screepplot* do total da variância explicada pelo número de dimensões considerando o modelo inicial, com todas as 24 questões, e o modelo final para os quatro grupos. Nos gráficos o eixo y indica os autovalores, o eixo x o número de fatores e a linha paralela ao eixo x é referente ao autovalor 1. Pela análise dos gráficos verificam-se dois fatores, porém observa-se que no modelo final, utilizando o critério do autovalor maior que 1, existe a predominância de uma dimensão pela análise fatorial.

Ilustração 6. *Screepplot* comparativo 5º ano do Ensino Fundamental Saeb

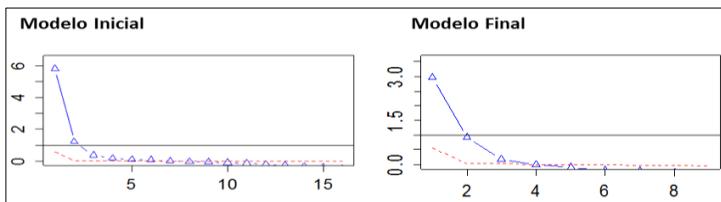


Fonte: Elaborado pela autora

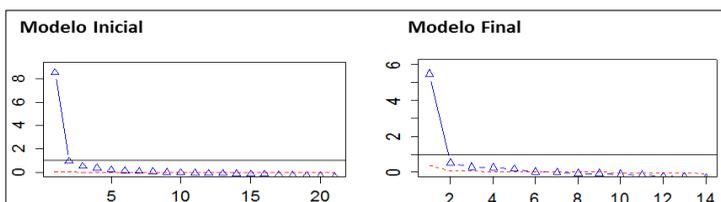
Ilustração 7. *Screepplot* comparativo 9º ano do Ensino Fundamental Saeb



Fonte: Elaborado pela autora

Ilustração 8. *Screeplot* comparativo 3ª série do Ensino Médio Saeb

Fonte: Elaborado pela autora

Ilustração 9. *Screeplot* comparativo 3ª série do Ensino Médio Enem

Fonte: Elaborado pela autora

Verifica-se também a partir dos gráficos apresentados e da Tabela 22, em que apresenta-se os cinco primeiros autovalores para cada grupo, que em relação ao critério dos autovalores existe a predominância de um fator no modelo final, resultado diferente em relação ao modelo inicial.

Tabela 22. Autovalores modelo inicial e final

Autovalores	Modelo Inicial				Modelo Final			
	5º ano EF Saeb	9º ano EF Saeb	3ª série EM Saeb	3ª série EM Enem	5º ano EF Saeb	9º ano EF Saeb	3ª série EM Saeb	3ª série EM Enem
1	4,07	4,59	5,82	8,53	2,21	2,64	2,96	5,44
2	1,14	1,25	1,23	0,94	0,74	0,83	0,92	0,50
3	0,36	0,34	0,38	0,50	0,16	0,16	0,18	0,28
4	0,21	0,19	0,17	0,36	0,01	0,02	-0,01	0,25
5	0,09	0,18	0,11	0,22	-0,11	-0,14	-0,09	0,16
6

Fonte: Elaborado pela autora

Outra verificação feita foi a proposta por Schilling (2007), em que a verificação da unidimensionalidade essencial é feita avaliando os resíduos. Com isso calculou-se a raiz da média das correlações residuais quadráticas obtendo para o 5º ano EF, 9º ano EF e 3ª série EM do Saeb e

3ª série EM Enem respectivamente 0,11; 0,11; 0,11 e 0,08. Apesar desses valores não serem menores que 0,05, observa-se nas Tabelas 15 a 18 que, de forma geral, as correlações residuais são menores que $|0,10|$.

5.2.4.2.4 RMSEA – Raiz do Erro Quadrático Médio de Aproximação

Além das medidas já propostas, foi calculado também a raiz do erro quadrático médio de aproximação – RMSEA – que representa a qualidade do ajuste que seria esperado se o modelo fosse estimado na população e não somente na amostra utilizada.

No que se refere a qualidade do ajuste dos dados, observou-se valores aceitáveis do RMSEA, para o 5º ano EF, 9º ano EF e 3ª série EM do Saeb e 3ª série EM Enem, respectivamente 0,06; 0,08; 0,08 e 0,06.

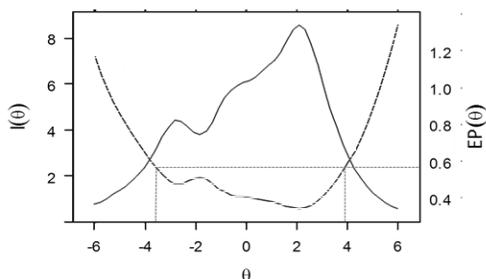
5.2.5 Modelo Proposto

Após todas as análises apresentadas nas seções anteriores em que verificou-se resultados razoáveis no que se refere ao ajuste do modelo escolhido, atendendo aos pressupostos de unidimensionalidade, independência dos dados e invariância entre os grupos chegou-se a uma proposta final. O modelo explica 38% da variância total dos dados, é construído a partir de 17 questões e a modelagem se dá a partir de grupos múltiplos, sendo o grupo de referência os alunos da 3ª série EM Enem e os demais grupos os alunos do 5º ano EF, 9º ano EF e 3ª série EM do Saeb.

A seguir, são apresentados gráficos que caracterizam a relação entre a medida de Inse proposta e as questões. Essa é uma das grandes contribuições da TRI quando comparada com os métodos convencionais de criação de indicadores, ela permite conhecer onde cada questão é mais informativa para determinada faixa de Inse e também onde o indicador está melhor estimado, por haver mais informação e conseqüentemente menos erro.

A partir do Gráfico 15, observa-se maior quantidade de informação e conseqüentemente menor erro padrão nos valores próximos de -4 e 4.

Gráfico 15. Curva de Informação e erro padrão do conjunto de questões



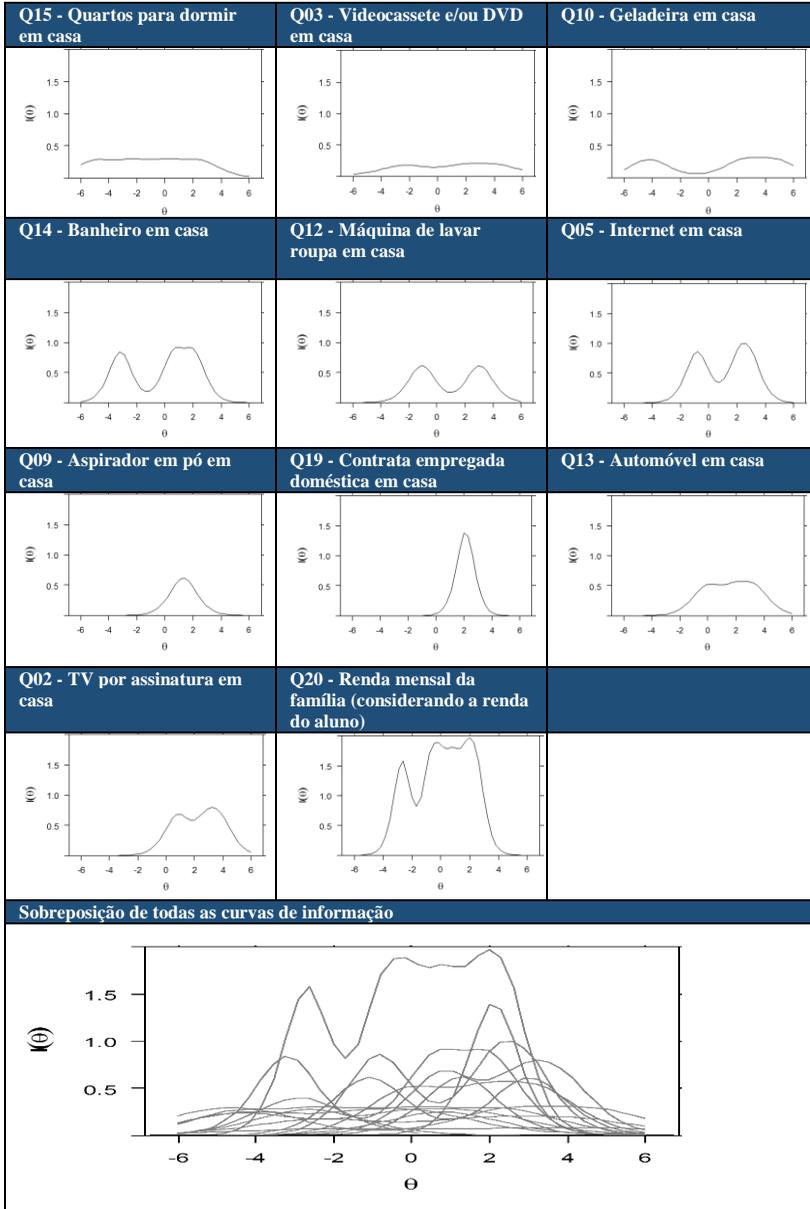
Fonte: Elaborado pela autora

A partir da curva de informação de cada questão é possível avaliar quantidade de informação que cada questão possui e em que região esta questão contribui na estimação da medida do Inse.

Na Ilustração 10 são apresentados os gráficos com as CI's de cada questão e um último gráfico que reúne todas as curvas sobrepostas. Tentou-se apresentar as curvas de informação dispostas em ordem crescente da escala, isto é, no início aquelas curvas que contribuem mais para a estimação das medidas de Inse mais baixa e ao final, aquelas que contribuem para as faixas de Inse mais alta.

Ilustração 10. Curva de Informação por questão

Q21 - Mãe, ou a mulher responsável por você, sabe ler e escrever	Q22 - Pai, ou homem responsável por você, sabe ler e escrever	Q08 - Telefone celular em casa
Q11 - Freezer (aparelho independente ou parte da geladeira) em casa	Q06 - Rádio em casa	Q25 - Maior escolaridade pais ou responsáveis



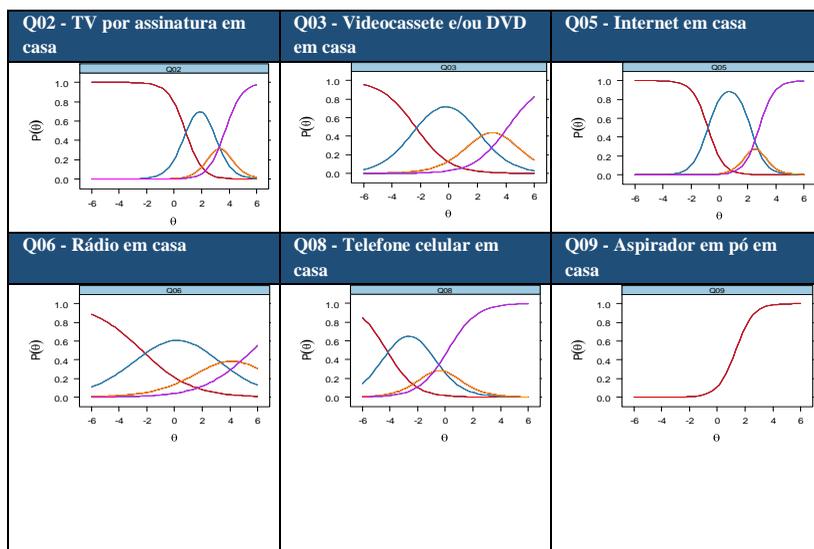
Fonte: Elaborado pela autora

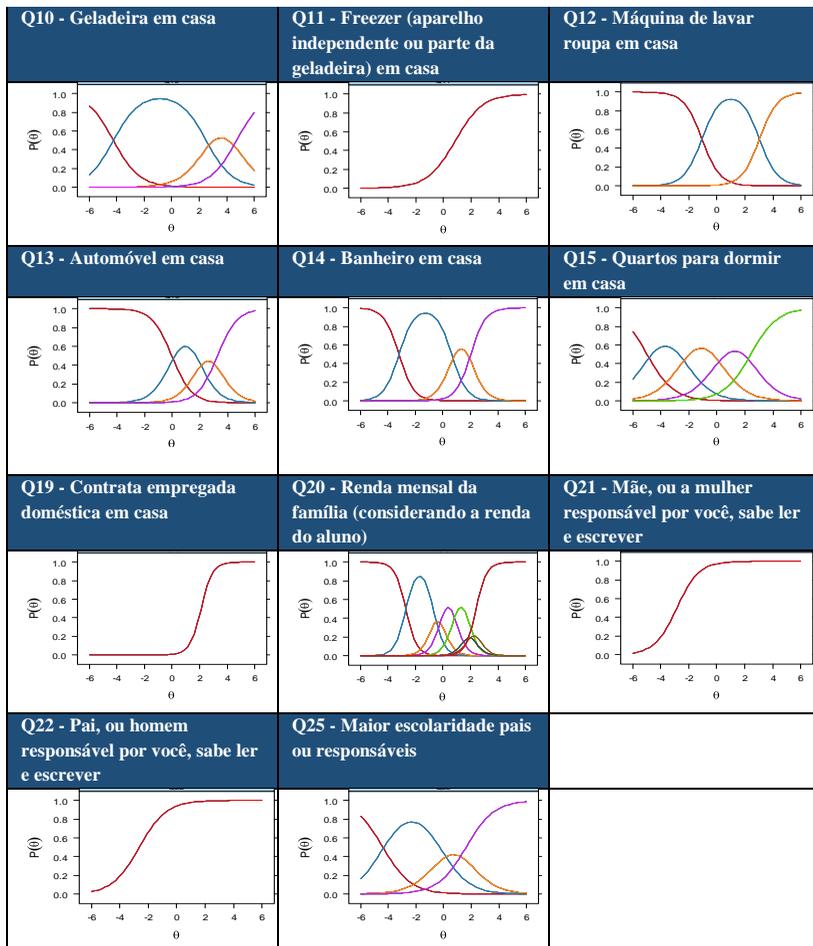
A partir dos gráficos apresentados, na ilustração acima, observa-se que as questões “Q21 - Mãe, ou a mulher responsável por você, sabe ler e escrever”, “Q22 - Pai, ou homem responsável por você, sabe ler e escrever”, “Q08 - Telefone celular em casa” são mais informativas para faixas de Inse mais baixas. Já as questões “Q02 - TV por assinatura em casa”, “Q13 - Automóvel em casa” e “Q19 - Contrata empregada doméstica em casa” são mais informativas para as faixas de Inse mais altas. Além disso, existem questões que são informativas para praticamente quase toda a escala como “Q20 - Renda mensal da família (considerando a renda do aluno)”, “Q14 - Banheiro em casa” e “Q05 - Internet em casa”.

A partir do último gráfico da Ilustração 10, em que as curvas estão sobrepostas umas nas outras, observa-se a questão “Q20 - Renda mensal da família (considerando a renda do aluno)” é a que possui a maior informação bem como engloba várias faixas de Inse.

Por fim, na Ilustração 11, são apresentadas as curvas características de cada questão. Observa-se que de forma geral as alternativas alcançam seu ponto de máximo em diversos valores de Inse e, em geral, em lugares distintos das demais alternativas da mesma questão. Tal situação, é bastante interessante já que se deseja descrever a escala de forma a diferenciar os alunos pelo nível socioeconômico familiar.

Ilustração 11. Curva característica por questão





Fonte: Elaborado pela autora

5.3 ESTIMAÇÃO DO NÍVEL SOCIOECONÔMICO DOS ALUNOS

Após todo o processo de calibração e definição do melhor modelo que se ajusta aos dados, é possível para cada aluno do banco de dados estimar o nível socioeconômico deles. Como o quantitativo de questões não é o mesmo do quantitativo inicial, quando selecionou-se o banco de dados, faz-se necessário selecionar novamente apenas os alunos que responderam mais de cinco questões. Na Tabela 23, a seguir,

apresenta-se o quantitativo de alunos por avaliação/exame que tiveram o seu Inse estimado.

Tabela 23. Quantitativo de alunos selecionados para cálculo do Inse

		Qtd.
Saeb	5º ano Ensino Fundamental	1.985.587
	9º ano Ensino Fundamental	2.013.080
	3ª série Ensino Médio	90.255
Enem	3ª série Ensino Médio	878.853
Total		4.967.775

Fonte: Elaborado pela autora

Para obter a medida do nível socioeconômico de todos os alunos da população foi necessário fazer algumas transformações com base no que foi indicado na seção 4.4.7. A seguir, são apresentadas as fórmulas para que sejam feitas a transformação da escala estimada pelo software para a escala que tenha média 500 e desvio padrão 100.

Transformação linear do θ :

$$\theta_{(500,100)} = 100 * \left(\frac{\theta_{estimado} - 0,17280396}{1,04577759} \right) + 500$$

Transformação parâmetros:

$$a_{(500,100)} = \frac{a_{estimado}}{100/1,04577759}$$

$$b_{(500,100)} = 100 * \left(\frac{b_{estimado} - 0,17280396}{1,04577759} \right) + 500$$

A seguir, na Tabela 24, são apresentados os resultados das estatísticas descritivas do Inse por avaliação/exame e da população como um todo. O grupo dos alunos estudados da 3ª série do Ensino Médio do Enem possui média 500 e desvio padrão 100, pois é o grupo de referência. Salienta-se que essa média 500 não tem nenhuma relação com o desempenho nas provas do aluno no exame ENEM.

Tabela 24. Estatísticas Descritivas Inse por Avaliação/Exame

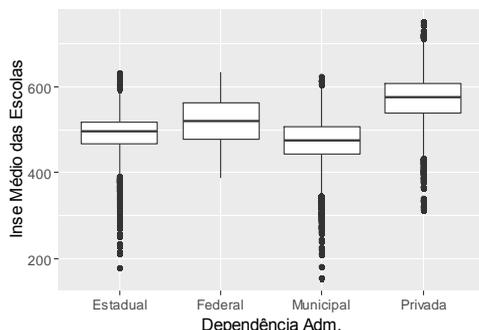
	Saeb			Enem	Total
	5º ano EF	9º ano EF	3ª série EM	3ª série EM	
Média	490,9	500,5	522,2	500,0	497,0
Desvio Padrão	87,5	84,5	97,5	100,0	89,0
Mínimo	103,3	115,2	123,6	142,9	103,3
Q ₁	434,4	443,0	452,2	427,5	437,1
Mediana	488,2	498,2	517,7	496,6	494,2
Q ₃	546,1	555,2	588,6	567,5	554,2
Máximo	824,7	821,0	837,5	855,6	855,6
Quantitativo escolas	41.026	33.119	2.832	14.715	69.758
Quantitativo alunos	1.985.587	2.013.080	90.255	878.853	4.967.775
Qtd. alunos escola pública e (%)	1.947.491 (98%)	1.970.150 (98%)	63.647 (71%)	618.942 (70%)	4.600.230 (93%)
Qtd. alunos escola particular e (%)	38.096 (2%)	42.930 (2%)	26.608 (29%)	259.911 (30%)	367.545 (7%)

Fonte: Elaborado pela autora

A partir dos resultados apresentados na Tabela 24 observa-se coerência entre os resultados apresentados e o que se esperava com base nas características do Saeb e Enem. Verifica-se que o grupo dos alunos do 5º ano e 9º ano do Ensino Fundamental possuem Inse médio mais baixo em relação ao dos alunos da 3ª série do Ensino Médio do Saeb e também menor variabilidade, isso se deve ao fato de os dois primeiros grupos possuírem em quase que em sua totalidade alunos de escolas públicas.

Após a criação do indicador, é possível também realizar algumas validações que são bastante intuitivas a partir do cruzamento do indicador com outras variáveis que foram coletadas na pesquisa. O Gráfico 16 apresenta o Inse médio das escolas por dependência administrativa, no qual observam-se resultados dentro do esperado, em que as escolas privadas apresentam Inse maior do que as públicas e entre as públicas, as escolas federais se destacam com Inse maior e as escolas estaduais e municipais apresentam distribuição mais semelhante.

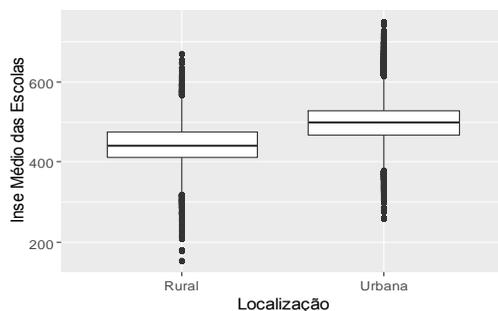
Gráfico 16. Inse por Dependência Administrativa



Fonte: Elaborado pela autora

Já em relação à localização das escolas, apresentadas no Gráfico 17, os resultados também estão bastante consistente dado que, observa-se que o Inse das escolas urbanas é maior, em geral, do que o das escolas rurais.

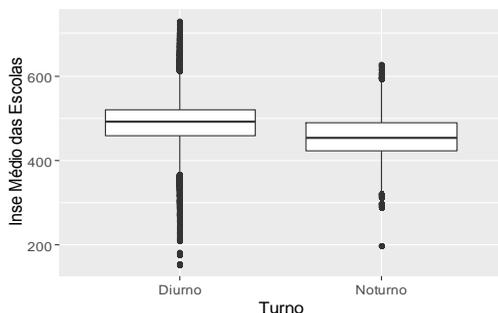
Gráfico 17. Inse por Localização



Fonte: Elaborado pela autora

Apenas em relação aos alunos do Saeb, pois, no microdados do Enem essa variável não é disponibilizada, no que se refere ao turno em que estudam, verifica-se também uma coerência no indicador, dado que geralmente as escolas recebem no turno noturno alunos com nível socioeconômico menor, Gráfico 18.

Gráfico 18. Inse por Turno Saeb



Fonte: Elaborado pela autora

Outra forma de verificar a consistência do indicador é relacioná-lo com outras variáveis existentes que medem o mesmo construto ou que se relacionam a ele. A seguir serão apresentadas algumas medidas de correlação entre o Inse, calculado neste trabalho, e outras medidas de Inse ou relacionadas a este.

O Inep possui uma medida de nível socioeconômico familiar dos alunos e disponibiliza em seu site o Inse médio, média aritmética simples, por escola, ele foi calculado com base nas edições do Saeb e Enem de 2011 e 2013. Foi calculada a correlação linear de Pearson entre o Inse médio, média aritmética simples, da Escola proposto neste trabalho e o calculado pelo Inep, e encontrou-se uma correlação de 0,93.

O Índice de Desenvolvimento Humano Municipal – IDHM é uma medida composta por três indicadores o IDHM Longevidade, IDHM Educação e IDHM Renda. O IDHM Longevidade é calculado a partir da informação sobre expectativa de vida ao nascer, o IDHM Educação é calculado com base na informação sobre escolaridade da população adulta e fluxo escolar da população jovem e o IDHM Renda é a renda per capita. Essas medidas são calculadas a cada dez anos, pois, utiliza o Censo demográfico Populacional, desta forma a última versão disponível é a de 2010 (Atlas do Desenvolvimento Humano no Brasil, 2013). A correlação de Pearson entre o Inse médio municipal e o IDHM foi igual a 0,87 e quando se considera os indicadores separadamente observa-se uma correlação de 0,89 com o IDHM Renda e 0,76 com o IDHM Educação.

O IDHM também é disponibilizado considerando o resultado agregado por Unidade da Federação, desta forma para correlacionar os dados do Inse com essa medida por UF foi calculado o Inse médio dos

alunos por UF. Encontrou-se uma correlação de 0,92 com o IDHM, 0,93 com o IDHM Renda e IDHM Educação 0,85.

No Atlas do Desenvolvimento Humano do Brasil outros indicadores também são disponibilizados, além do IDHM, e alguns estão relacionados com construto proposto neste trabalho. Desta forma na Tabela 25, a seguir, apresenta-se a correlação de Pearson ente o Inse e algumas dessas medidas. É importante considerar que o Inse é calculado com base nas informações coletadas em 2013 e as informações disponibilizadas pelo Atlas de Desenvolvimento Humano são de 2010.

Tabela 25. Correlação Inse e medidas do Atlas de Desenvolvimento Humano

Indicadores	Correlação
RDPC1 - Renda per capita média do 1º quinto mais pobre	0,86
RDPC2 - Renda per capita média do 2º quinto mais pobre	0,88
RDPC3 - Renda per capita média do 3º quinto mais pobre	0,89
RDPC4 - Renda per capita média do 4º quinto mais pobre	0,87
RDPC - Renda per capita	0,84
RPOB - Renda per capita média dos vulneráveis à pobreza	0,83
RENOCUP - Rendimento médio dos ocupados - 18 anos ou mais	0,78
T_BANAGUA - % da população em domicílios com banheiro e água encanada	0,77
T_FUND18A24 - % de 18 a 24 anos com Ensino Fundamental completo	0,75
T_MED18A24 - % de 18 a 24 anos com Ensino Médio completo	0,73
T_AGUA - % da população em domicílios com água encanada	0,58
T_LIXO - % da população em domicílios com coleta de lixo	0,46
T_LUZ - % da população em domicílios com energia elétrica	0,39
PPOB - % de vulneráveis à pobreza	-0,91
T_ANALF15M - Taxa de analfabetismo - 15 anos ou mais	-0,87
T_FUNDIN_TODOS_MMEIO - % de pessoas em domicílios vulneráveis à pobreza e em que ninguém tem fundamental completo.	-0,85
RENI - % dos ocupados com rendimento de até 1 s.m. - 18 anos ou mais	-0,81

Fonte: Elaborado pela autora

A partir dos resultados apresentados verifica-se que o Inse apresenta correlações fortes com outras medidas que estão relacionadas com o construto, bem como, os temas que o definem como a renda, ocupação (mesmo não entrando questões no cálculo) e escolaridade.

5.4 O INSE E O PARÂMETRO DAS QUESTÕES

Antes de definir a forma que os alunos serão divididos em grupos, buscou-se conhecer melhor a relação entre a escala e as questões. Foram definidos valores de Inse de 100 a 900 com diferença de 10 unidades entre um valor e outro, totalizando 81 valores. Após isso, calculou-se a probabilidade, a partir do modelo, para esses valores de Inse, para cada alternativa/questão e buscou-se o valor onde a probabilidade era mais alta dentre os valores calculados.

Na Tabela 26, a seguir, são apresentadas as probabilidades calculadas, a partir do modelo, para alguns valores de θ e destacado os valores onde as probabilidades alcançam o ponto de máximo por alternativa na questão Q20, por exemplo.

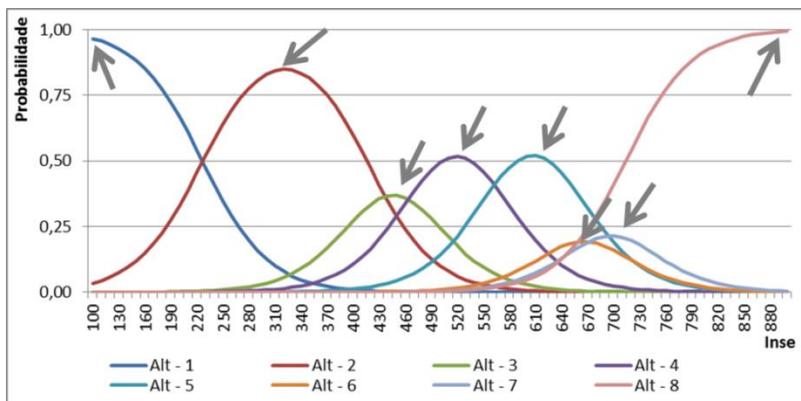
Tabela 26. Probabilidade por alternativa da questão Q20

Q20	100	320	450	520	610	670	700	900
Alt - 1	0,97	0,08	0,00	0,00	0,00	0,00	0,00	0,00
Alt - 2	0,03	0,85	0,29	0,06	0,01	0,00	0,00	0,00
Alt - 3	0,00	0,06	0,37	0,18	0,02	0,00	0,00	0,00
Alt - 4	0,00	0,01	0,29	0,52	0,20	0,05	0,02	0,00
Alt - 5	0,00	0,00	0,04	0,21	0,52	0,32	0,19	0,00
Alt - 6	0,00	0,00	0,00	0,02	0,12	0,19	0,16	0,00
Alt - 7	0,00	0,00	0,00	0,01	0,07	0,19	0,21	0,00
Alt - 8	0,00	0,00	0,00	0,01	0,06	0,24	0,41	0,99

Fonte: Elaborado pela autora

Tal procedimento descrito seria o equivalente a olhar os pontos de máximo das curvas no Gráfico 19. O resultado é equivalente a CCI apresentada na metodologia, porém já na escala com média 500 e desvio 100.

Gráfico 19. CCI Questão Q20 destacado os pontos de máximo

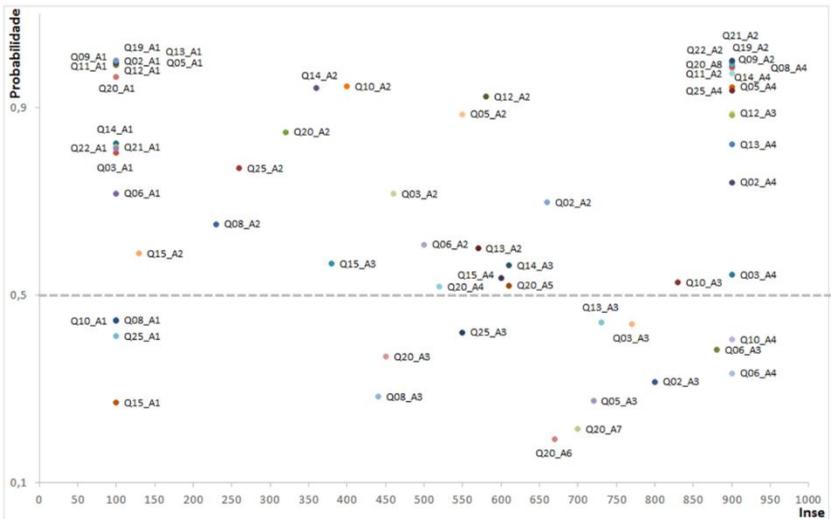


Fonte: Elaborado pela autora

Observa-se tanto no Gráfico 19 como na Tabela 26 que em alguns valores de Inse algumas alternativas se destacam com probabilidades bastante superiores, como as alternativas 1, 2 e 8 para um Inse de 100, 320 e 900 respectivamente, com probabilidades superiores a 0,80. Outras se destacam com probabilidades maior que 0,5, isto é, a maioria dos indivíduos, como as alternativas. 4 e 5 nos valores de 520 e 610 respectivamente para o Inse. Já as alternativas 3, 6 e 7 sozinhas não seriam possíveis de serem interpretadas para esses valores, pois as probabilidades não são suficientes para caracterizar um grupo, porém é possível descreve-las agrupando com outras alternativas.

O Gráfico 20, a seguir, posiciona as questões e suas alternativas na escala e onde elas apresentam probabilidade máxima de acordo com os valores de Inse apresentados. Ao todo são 62 alternativas/questões que foram utilizadas para construir o Inse, destas 46 alternativas/questões encontram um ponto de máximo com probabilidade maior que 0,5; 30 alternativas/questões com probabilidade maior que 0,8; e apenas 16 alternativas/questões com probabilidade menor que 0,5.

Gráfico 20. Ponto de máxima probabilidade das questões e suas alternativas, por Inse



Fonte: Elaborado pela autora

Além de buscar avaliar os pontos de máximo buscou-se verificar para cada alternativa/questão a sua cobertura, isto é, até que faixa de Inse ela se destaca em relação às demais.

A Tabela 27 indica a probabilidade de resposta de cada alternativa para cada valor de Inse indicado. As células em verde indicam os valores cuja probabilidade é maior ou igual a 0,5; em vermelho destacados o valor do Inse onde a probabilidade alcançou o ponto de máximo para cada alternativa; e as células em cor cinza são aquelas com valores de probabilidade menores que 0,3.

Observa-se pela Tabela 27:

- existem alternativas para descrever os diversos níveis de Inse.
- algumas alternativas sozinhas poderiam ser classificadas como âncoras, isto é, possíveis de descrever aquele valor de Inse por apresentarem probabilidade maiores que 0,5 e menores que 0,3, como por exemplo, a questão Q14 nas suas quatro alternativas.
- algumas questões, quando a escala for elaborada, deverão ser descritas tendo algumas de suas alternativas agrupadas para poder caracterizar o grupo de respondentes, isto

porque a probabilidade de um indivíduo responder aquela alternativa sozinha é muita baixa e não representaria a maioria ou maior parte dos respondentes.

- as questões dicotômicas precisam ser consideradas com cautela quando forem criados os grupos. Por exemplo, se for criado um grupo que inclui indivíduos com Inse entre 100 e 260 esse grupo irá incluir as duas únicas opções de resposta das questões Q21 e Q22: pais que sabem ler e pais que não sabem ler dentro do mesmo grupo. Com isso, essas duas questões apesar de contribuírem na construção da escala não estariam sendo consideradas na construção dos grupos, dado que não haverá em um grupo apenas pais que não sabem ler e escrever em sua maioria.
- no caso das questões politômicas é possível que duas ou mais questões se agrupem em um mesmo grupo.

Tabela 27. Probabilidade por alternativa das questões para diversos valores de Inse

Questão/ alternativa	Nível																							
	100	130	230	260	360	380	440	450	460	500	520	570	580	600	610	660	670	700	730	800	820	880	900	
Q02	Alt1	1,0	1,0	1,0	1,0	1,0	1,0	0,9	0,9	0,9	0,7	0,7	0,6	0,5	0,4	0,3	0,3	0,2	0,1	0,1	0,1	0,0	0,0	0,0
	Alt2	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,3	0,3	0,4	0,5	0,5	0,6	0,6	0,7	0,7	0,6	0,3	0,3	0,1	0,1
	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,3	0,3
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1	0,3	0,4
Q03	Alt1	0,8	0,8	0,6	0,5	0,3	0,3	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alt2	0,2	0,2	0,4	0,5	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,7	0,6	0,6	0,6	0,6	0,5	0,5	0,4	0,4	0,2	0,2	0,1
	Alt3	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,2	0,3	0,3	0,3	0,4	0,4	0,4	0,4	0,4	0,4	0,3
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,3	0,4	0,5
Q05	Alt1	1,0	1,0	1,0	0,9	0,7	0,6	0,3	0,3	0,3	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alt2	0,0	0,0	0,0	0,1	0,3	0,4	0,7	0,7	0,7	0,8	0,9	0,9	0,9	0,8	0,8	0,7	0,6	0,5	0,3	0,1	0,1	0,0	0,0
	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,2	0,2	0,3	0,3	0,2	0,1	0,1	0,0
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,2	0,3	0,4	0,7	0,8	0,9	0,9
Q06	Alt1	0,7	0,7	0,5	0,5	0,4	0,3	0,3	0,3	0,2	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
	Alt2	0,3	0,3	0,4	0,4	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,6	0,5	0,5	0,5	0,5	0,4	0,4	0,3	0,3
	Alt3	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,3	0,3	0,4	0,4	0,4	0,4
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,3	0,3
Q08	Alt1	0,5	0,4	0,2	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alt2	0,5	0,6	0,7	0,6	0,5	0,5	0,3	0,3	0,3	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0
	Alt3	0,0	0,1	0,1	0,1	0,2	0,3	0,3	0,3	0,3	0,3	0,2	0,2	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
	Alt4	0,0	0,0	0,1	0,1	0,2	0,2	0,4	0,4	0,4	0,5	0,6	0,6	0,7	0,7	0,7	0,8	0,8	0,9	0,9	1,0	1,0	1,0	1,0
Q09	Alt1	1,0	1,0	1,0	1,0	1,0	1,0	0,9	0,9	0,9	0,9	0,8	0,7	0,7	0,6	0,5	0,5	0,3	0,3	0,2	0,1	0,0	0,0	0,0
	Alt2	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,3	0,4	0,4	0,5	0,5	0,7	0,7	0,8	0,9	1,0	1,0	1,0
Q10	Alt1	0,4	0,4	0,2	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alt2	0,6	0,6	0,8	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,8	0,8	0,8	0,8	0,7	0,6	0,6	0,5	0,3	0,3	0,2
	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,3	0,3	0,4	0,4	0,5	0,5	0,5
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,2	0,4	0,4
Q11	Alt1	1,0	1,0	1,0	1,0	0,9	0,9	0,8	0,8	0,8	0,7	0,6	0,5	0,5	0,5	0,4	0,4	0,3	0,2	0,2	0,2	0,1	0,1	0,0
	Alt2	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,2	0,3	0,3	0,4	0,5	0,5	0,6	0,6	0,6	0,7	0,8	0,8	0,9	0,9	1,0	1,0
Q12	Alt1	1,0	1,0	0,9	0,9	0,6	0,5	0,3	0,2	0,2	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alt2	0,0	0,0	0,1	0,1	0,4	0,5	0,7	0,8	0,8	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,8	0,8	0,7	0,4	0,3	0,2	0,1

Questão/ alternativa	Nível																								
	100	130	230	260	360	380	440	450	460	500	520	550	570	580	600	610	660	670	700	730	800	820	880	900	
Q13	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,2	0,2	0,3	0,6	0,7	0,9	0,9	
	Alt1	1,0	1,0	1,0	1,0	0,9	0,8	0,6	0,6	0,6	0,4	0,4	0,3	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,0	0,0	0,0	0,0	0,1	0,2	0,3	0,4	0,4	0,5	0,5	0,6	0,6	0,6	0,6	0,5	0,4	0,3	0,3	0,1	0,1	0,0	0,0	
	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,2	0,2	0,2	0,2	0,4	0,4	0,4	0,4	0,3	0,2	0,2	
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,3	0,5	0,6	0,8	0,8
Q14	Alt1	0,8	0,7	0,3	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,2	0,3	0,7	0,8	0,9	0,9	0,9	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,3	0,2	0,1	0,1	0,1	0,0	0,0	0,0	0,0	
	Alt3	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,3	0,3	0,5	0,5	0,5	0,6	0,6	0,5	0,5	0,3	0,2	0,1	0,1	0,0	
	Alt4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,2	0,4	0,5	0,6	0,7	0,9	0,9	1,0	1,0
Q15	Alt1	0,3	0,2	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,6	0,6	0,5	0,4	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt3	0,1	0,2	0,4	0,4	0,6	0,6	0,5	0,5	0,5	0,4	0,4	0,3	0,3	0,3	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	
	Alt4	0,0	0,0	0,1	0,1	0,2	0,2	0,3	0,3	0,4	0,4	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,4	0,4	0,3	0,2	0,1	0,1
	Alt5	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,2	0,4	0,4	0,5	0,5	0,7	0,8	0,9	0,9
Q19	Alt1	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,9	0,9	0,9	0,7	0,6	0,4	0,3	0,1	0,0	0,0	0,0	
	Alt2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,4	0,4	0,6	0,8	1,0	1,0	1,0	
Q20	Alt1	1,0	0,9	0,5	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,0	0,1	0,5	0,7	0,8	0,7	0,4	0,3	0,2	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt3	0,0	0,0	0,0	0,0	0,1	0,2	0,4	0,4	0,3	0,2	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt4	0,0	0,0	0,0	0,0	0,0	0,1	0,2	0,3	0,3	0,5	0,5	0,5	0,4	0,3	0,2	0,2	0,1	0,1	0,0	0,0	0,0	0,0	0,0	
	Alt5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,4	0,4	0,5	0,5	0,5	0,4	0,3	0,2	0,1	0,0	0,0	0,0	
	Alt6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,1	0,0	0,0	0,0	
	Alt7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,2	0,2	0,2	0,1	0,0	0,0	
	Alt8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	0,2	0,4	0,6	0,9	0,9	1,0	1,0
Q21	Alt1	0,8	0,8	0,4	0,4	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,2	0,3	0,6	0,7	0,9	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	
Q22	Alt1	0,8	0,8	0,5	0,4	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,2	0,2	0,5	0,6	0,8	0,8	0,9	0,9	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	
Q25	Alt1	0,4	0,3	0,2	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
	Alt2	0,6	0,6	0,8	0,8	0,7	0,7	0,5	0,5	0,5	0,4	0,4	0,3	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	
	Alt3	0,0	0,0	0,1	0,1	0,2	0,2	0,3	0,3	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,3	0,3	0,3	0,2	0,1	0,1	0,1	
	Alt4	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,3	0,3	0,4	0,4	0,4	0,6	0,6	0,7	0,7	0,8	0,9	0,9	0,9

5.5 O INSE E A INTERPRETAÇÃO DA ESCALA

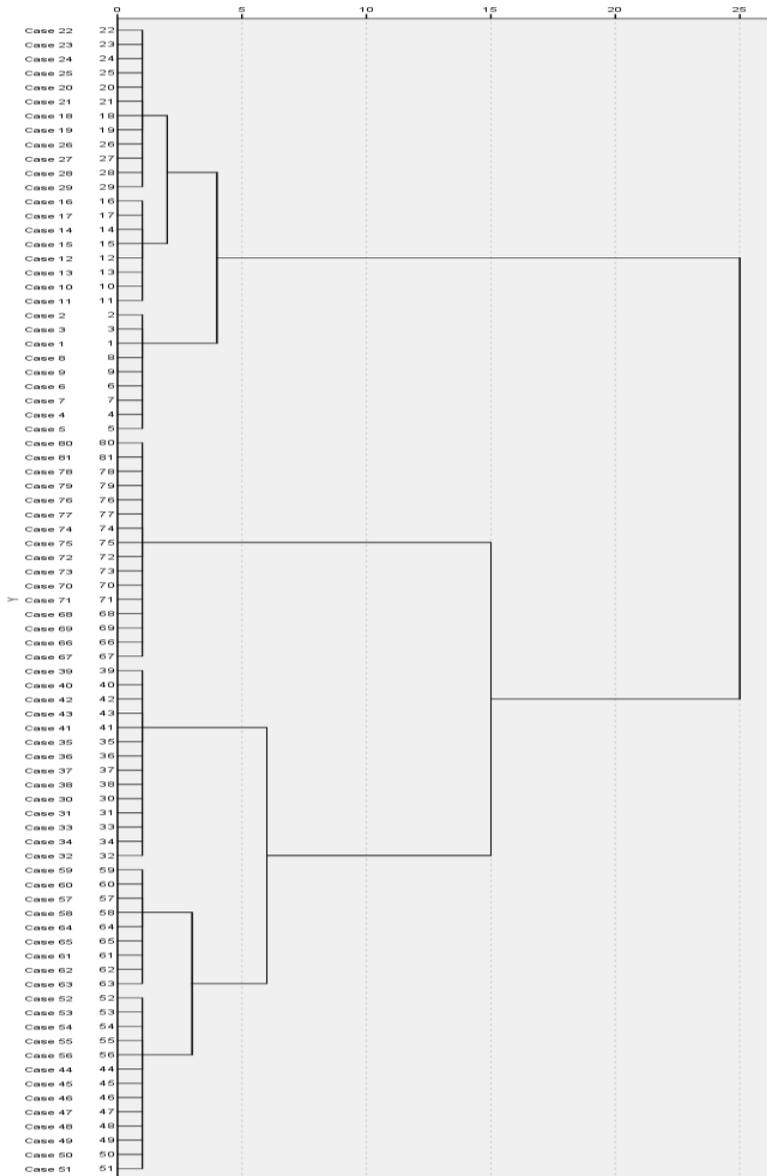
Após conhecer a forma como as questões se relacionam com a medida criada de nível socioeconômico, buscou-se uma forma de dividir os alunos em grupos de forma que eles sejam mais parecidos dentro dos grupos e diferentes entre os grupos. Não existe um método específico na literatura que direcione a forma de segmentar grupos e interpretar escalas. E no caso dos modelos politômicos, a interpretação da escala é uma proposta mais complexa ainda, devido ao número de parâmetros envolvidos.

O procedimento utilizado, como já mencionado, foi a análise de agrupamentos hierárquico a partir do método de Ward e foi selecionada a opção de padronizar as variáveis, dado que esse procedimento é influenciado quando as variáveis possuem dimensões diferentes. Como variáveis de entrada utilizou-se os 81 valores de Inse, definidos na seção 5.4, e as probabilidades de resposta por questões/alternativas do modelo. Essa estrutura é semelhante a indicada na Tabela 27, porém, para aplicar a técnica alterou-se o layout desta tabela disponibilizando as questões/alternativas nas colunas e os valores de Inse nas linhas.

A partir dos resultados dessa análise utilizou-se como critério para selecionar o número de grupos a análise do coeficiente de aglomeração, vide Hair (2005), que mostrou grandes aumentos quando se passa de quatro para três agrupamentos e seis para cinco. Com isso, as duas soluções de quatro e seis agrupamentos serão consideradas.

Outra análise considerada é pelo dendograma, ele representa o procedimento de agrupamento no qual cada caso (valores de Inse) é colocado no eixo vertical e no eixo horizontal o número de agrupamentos, Gráfico 21, em que o primeiro caso corresponde ao valor do Inse 100 e o último caso ao valor do Inse 900. Com isso, será considerada também a opção de 7 agrupamentos que é referente aos menores nichos.

Gráfico 21. Dendograma



Fonte: Elaborado pela autora

Com base no exposto as opções de quatro, seis e sete agrupamentos serão consideradas e são apresentadas no Quadro 4.

Quadro 4. Opções de agrupamentos por meio de métodos hierárquicos

Opção: 4 Grupos	Opção: 6 Grupos	Opção: 7 Grupos
1 - De 100 a 380	1 - De 100 a 180	1 - De 100 a 180
	2 - De 190 a 380	2 - De 190 a 260
		3 - De 270 a 380
2 - De 390 a 520	3 - De 390 a 520	4 - De 390 a 520
3 - De 530 a 740	4 - De 530 a 650	5 - De 530 a 650
	5 - De 660 a 740	6 - De 660 a 740
4 - De 750 a 900	6 - De 750 a 900	7 - De 750 a 900

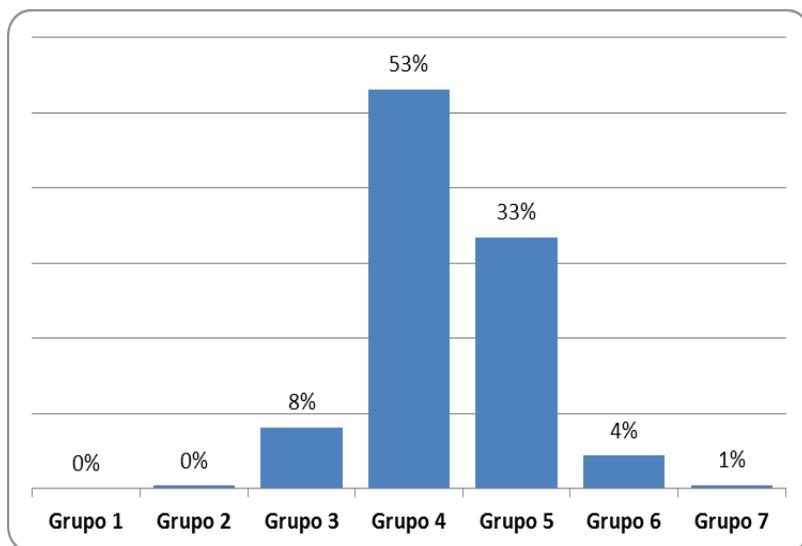
Fonte: Elaborado pela autora

A partir dos resultados apresentados, observa-se que existem grupos que possuem a mesma faixa de Inse e outros correspondem a exatamente o agrupamento de uma ou mais faixas. Com base nisso, optou-se por estudar primeiro a proposta mais abrangente que seria a de sete grupos e caso as questões/alternativas não tivessem diferenciando muito os grupos eles seriam reagrupados com base nessas soluções menores.

Foram calculados para cada grupo a média das probabilidades do modelo referentes a cada alternativa/questão e a partir dessa média que a alternativa foi posicionada no grupo. Optou-se por essa metodologia por considerar que dentro do grupo as probabilidades seriam bastante parecidas e evitaria também utilizar apenas um ponto para poder representar o todo.

A opção de segmentação em sete grupos apresentou-se bastante interessante no que se refere ao posicionamento das questões/alternativas. Entretanto, observa-se que o grupo 4, que varia de 390 a 520, possui uma amplitude muito grande e desta forma, por incluir a média da distribuição irá concentrar um grande quantitativo de indivíduos. Desta forma, seria interessante conseguir diferenciar melhor esses alunos ao invés de um grande número de alunos estarem concentrados em um único grupo. A seguir, no Gráfico 22, é apresentado como seriam classificados os alunos do banco de dados estudado caso fossem divididos com base nos sete grupos.

Gráfico 22. Distribuição dos Alunos em 7 grupos de Inse



Fonte: Elaborado pela autora

Com base no resultado apresentado optou-se por fazer o mesmo procedimento relatado de agrupamentos hierárquicos considerando apenas os valores de Inse referentes ao grupo 4 e grupo 5 separadamente, buscando uma proposta de divisão para esses dois grupos.

A seguir, na Tabela 28, são apresentadas as probabilidades médias de respostas por alternativa/questão por grupo criado. Foram destacados os seguintes critérios:

1. em verde as células com valores maiores ou iguais a 0,5;
2. em azul as células com valores maiores que 0,3 quando alguma da mesma coluna e questão estava entre 0,5 e 0,6;
3. em laranja quando nenhuma das situações anteriores ocorreu e verificou-se uma concentração de probabilidades que somando as alternativas ficasse maior que 0,6; e
4. em cinza as células restantes.

Observa-se que os dois grupos que foram segmentados novamente possuem algumas questões/alternativas que permitiram a diferenciação dos alunos, porém se fosse possível sugerir melhorias para a escala seria interessante a inclusão de mais questões que pudessem

diferenciar os indivíduos no que se refere às faixas de nível socioeconômico entre 390 e 650, grupos 4 e 5, para assim poder caracterizar melhor a diferença entre esses grupos.

Tabela 28. Probabilidade média por questões/alternativas nos grupos

	Grupo 1	Grupo 2	Grupo 3	Grupo 4_1	Grupo 4_2	Grupo 5_1	Grupo 5_2	Grupo 6	Grupo 7
Q02 - TV por assinatura em casa									
1 - Não tem	1,00	1,00	0,98	0,93	0,79	0,48	0,24	0,09	0,01
2 - Sim, uma	0,00	0,00	0,02	0,07	0,20	0,48	0,66	0,65	0,27
3 - Sim, duas	0,00	0,00	0,00	0,00	0,01	0,03	0,07	0,17	0,27
4 - Sim, três ou mais	0,00	0,00	0,00	0,00	0,00	0,01	0,03	0,09	0,45
Q03 - Videocassete e/ou DVD em casa									
1 - Não tem	0,74	0,58	0,36	0,21	0,13	0,07	0,04	0,02	0,01
2 - Sim, um	0,25	0,41	0,59	0,70	0,71	0,65	0,54	0,41	0,20
3 - Sim, dois	0,01	0,02	0,04	0,08	0,13	0,23	0,32	0,40	0,41
4 - Sim, três ou mais	0,00	0,00	0,01	0,02	0,03	0,06	0,10	0,17	0,38
Q05 - Internet em casa									
1 - Não tem	0,99	0,97	0,81	0,48	0,20	0,05	0,01	0,00	0,00
2 - Sim, uma	0,01	0,03	0,19	0,52	0,78	0,87	0,77	0,48	0,10
3 - Sim, duas	0,00	0,00	0,00	0,00	0,01	0,05	0,14	0,25	0,14
4 - Sim, três ou mais	0,00	0,00	0,00	0,00	0,00	0,03	0,09	0,27	0,76
Q06 - Rádio em casa									
1 - Não tem	0,67	0,55	0,41	0,29	0,22	0,14	0,10	0,07	0,04
2 - Sim, um	0,30	0,40	0,51	0,58	0,61	0,59	0,55	0,49	0,35
3 - Sim, dois	0,02	0,04	0,07	0,10	0,14	0,20	0,25	0,30	0,37
4 - Sim, três ou mais	0,01	0,01	0,02	0,03	0,04	0,07	0,09	0,14	0,25
Q08 - Telefone celular em casa									
1 - Não tem	0,35	0,19	0,08	0,03	0,02	0,01	0,00	0,00	0,00
2 - Sim, uma	0,57	0,65	0,57	0,41	0,26	0,14	0,07	0,04	0,01
3 - Sim, duas	0,05	0,11	0,21	0,27	0,27	0,21	0,14	0,08	0,03
4 - Sim, três ou mais	0,03	0,06	0,15	0,28	0,45	0,65	0,79	0,88	0,96
Q09 - Aspirador de pó em casa									
1 - Não tem	1,00	1,00	0,99	0,96	0,89	0,67	0,41	0,19	0,04
2 - Sim, um ou mais	0,00	0,00	0,01	0,04	0,11	0,33	0,59	0,81	0,96
Q10 - Geladeira em casa									
1 - Não tem	0,34	0,17	0,07	0,03	0,01	0,00	0,00	0,00	0,00
2 - Sim, uma	0,66	0,83	0,92	0,94	0,92	0,85	0,74	0,56	0,26
3 - Sim, duas	0,00	0,00	0,01	0,03	0,06	0,13	0,23	0,37	0,50
4 - Sim, três ou mais	0,00	0,00	0,00	0,00	0,01	0,02	0,03	0,07	0,24
Q11 - Freezer (aparelho independente ou parte da geladeira) em casa									
1 - Não tem	0,99	0,97	0,92	0,83	0,70	0,49	0,33	0,19	0,07
2 - Sim, um ou mais	0,01	0,03	0,08	0,17	0,30	0,51	0,67	0,81	0,94
Q12 - Máquina de lavar roupa em casa									
1 - Não tem	0,98	0,92	0,70	0,38	0,17	0,05	0,02	0,01	0,00
2 - Sim, uma	0,02	0,08	0,30	0,61	0,82	0,92	0,90	0,76	0,32
3 - Sim, duas ou mais	0,00	0,00	0,00	0,00	0,01	0,03	0,09	0,24	0,68
Q13 - Automóvel em casa									
1 - Não tem	0,99	0,98	0,90	0,73	0,50	0,22	0,10	0,04	0,01
2 - Sim, uma	0,01	0,02	0,09	0,25	0,44	0,59	0,53	0,34	0,10
3 - Sim, duas	0,00	0,00	0,01	0,02	0,05	0,16	0,29	0,41	0,30
4 - Sim, três ou mais	0,00	0,00	0,00	0,00	0,01	0,03	0,08	0,20	0,59
Q14 - Banheiro em casa									
1 - Não tem	0,67	0,31	0,07	0,01	0,00	0,00	0,00	0,00	0,00
2 - Sim, uma	0,33	0,69	0,91	0,92	0,76	0,41	0,17	0,05	0,01
3 - Sim, duas	0,00	0,00	0,02	0,06	0,21	0,48	0,54	0,34	0,07
4 - Sim, três ou mais	0,00	0,00	0,00	0,01	0,03	0,11	0,29	0,60	0,93
Q15 - Quartos para dormir em casa									
1 - Não tem	0,20	0,09	0,03	0,01	0,01	0,00	0,00	0,00	0,00
2 - Sim, um	0,58	0,50	0,29	0,15	0,08	0,03	0,02	0,01	0,00
3 - Sim, dois	0,20	0,36	0,53	0,55	0,45	0,28	0,17	0,09	0,03
4 - Sim, três	0,02	0,05	0,13	0,25	0,39	0,51	0,53	0,44	0,21
5 - Sim, quatro ou mais	0,00	0,01	0,02	0,04	0,08	0,17	0,29	0,46	0,76

	Grupo 1	Grupo 2	Grupo 3	Grupo 4_1	Grupo 4_2	Grupo 5_1	Grupo 5_2	Grupo 6	Grupo 7
Q19 - Contrata empregada doméstica em casa									
1 - Não	1,00	1,00	1,00	1,00	0,99	0,95	0,79	0,42	0,05
2 - Sim, um ou mais	0,00	0,00	0,00	0,00	0,01	0,06	0,21	0,58	0,95
Q20 - Renda mensal da família (considerando a renda do aluno)									
1 - Nenhuma renda	0,89	0,51	0,09	0,01	0,00	0,00	0,00	0,00	0,00
2 - Até um s.m	0,11	0,48	0,80	0,54	0,18	0,02	0,00	0,00	0,00
3 - Mais de um até 1,5 s.m	0,00	0,01	0,08	0,30	0,30	0,08	0,01	0,00	0,00
4 - Mais de 1,5 até 3 s.m	0,00	0,00	0,02	0,14	0,41	0,39	0,13	0,03	0,00
5 - Mais de 3 até 7 s.m	0,00	0,00	0,00	0,02	0,10	0,40	0,48	0,21	0,02
6 - Mais de 7 até 9 s.m	0,00	0,00	0,00	0,00	0,01	0,05	0,16	0,15	0,02
7 - Mais de 9 até 12 s.m	0,00	0,00	0,00	0,00	0,00	0,03	0,11	0,19	0,05
8 - Acima de 12 s.m	0,00	0,00	0,00	0,00	0,00	0,02	0,11	0,42	0,91
Q21- Mãe, ou a mulher responsável por você, sabe ler e escrever									
1 - Não	0,72	0,46	0,19	0,07	0,03	0,01	0,00	0,00	0,00
2 - Sim	0,28	0,54	0,81	0,93	0,97	0,99	1,00	1,00	1,00
Q22- Pai, ou homem responsável por você, sabe ler e escrever									
1 - Não	0,73	0,52	0,27	0,12	0,06	0,03	0,01	0,01	0,00
2 - Sim	0,27	0,48	0,73	0,88	0,94	0,97	0,99	0,99	1,00
Q25 - Até que série a mãe, pai ou responsável estudou									
1 - Nunca estudou.	0,32	0,17	0,07	0,03	0,01	0,01	0,00	0,00	0,00
2 - Ensino Fund. cursando ou completo	0,64	0,76	0,73	0,60	0,44	0,26	0,15	0,08	0,03
3 - Ensino Médio Completo	0,03	0,07	0,16	0,28	0,37	0,42	0,37	0,27	0,11
4 - Ensino Superior Completo	0,01	0,01	0,04	0,09	0,17	0,32	0,48	0,65	0,86

Fonte: Elaborado pela autora

Com base nessa tabela de probabilidades médias, os grupos tiveram sua escala descrita e para a descrição da escala foram utilizadas as questões/alternativas pintadas de azul, verde e laranja. Ao escrever a escala algumas alternativas foram omitidas para simplificar a escrita sem comprometer a escala. Por exemplo, o primeiro grupo de menor nível socioeconômico as famílias em geral não possuem renda e por isso, apesar de apresentarem probabilidades mais altas, é desnecessário dizer que não contratam serviço de empregada mensalista e que não possuem automóvel em casa. Da mesma forma, no último grupo que é caracterizado pelos alunos de nível socioeconômico mais alto é desnecessário dizer que os pais ou responsáveis sabem ler e escrever.

Outra alteração feita ao descrever a escala, foi no nome dos “grupos” para “níveis”, isto porque, a palavra grupo não fornece a ideia de hierarquia e ordenamento e a palavra nível se mostra mais adequada. Com isso, para facilitar o entendimento utilizou-se as seguintes correspondências:

Quadro 5. Correspondência nomes grupos

Grupo 1	Grupo 2	Grupo 3	Grupo 4_1	Grupo 4_2	Grupo 5_1	Grupo 5_2	Grupo 6	Grupo 7
Nível 1	Nível 2	Nível 3	Nível 4	Nível 5	Nível 6	Nível 7	Nível 8	Nível 9
Inse de: 100 a 180	Inse de: 180 a 260	Inse de: 260 a 380	Inse de: 380 a 430	Inse de: 430 a 520	Inse de: 520 a 600	Inse de: 600 a 650	Inse de: 650 a 740	Inse de: 740 a 900

Fonte: Elaborado pela autora

Importante destacar que esses nomes apenas correspondem a uma representação ordinal, indicando que o Nível 1 como o de mais baixo Inse e o Nível 9 como o de mais alto Inse. Não é possível devido ao nome atribuído fazer analogias em relação a outras escalas, isto porque cada uma parte de um referencial teórico e metodologias de cálculo e agrupamentos diferentes. Para conhecer os níveis deve-se fazer referência apenas a descrição da escala, apresentada no Quadro 6.

Quadro 6. Descrição dos Níveis

<p>Nível 1 – Inse de 100 a 180</p> <p>Esse grupo é composto pelos alunos com os menores valores de nível socioeconômico familiar. Em geral, os alunos deste nível indicaram que a família não possui renda familiar. No que se refere à escolaridade, os alunos informaram que os pais ou responsáveis não sabem ler e escrever e estes possuem Ensino Fundamental completo ou incompleto.</p> <p>A casa que moram contém em geral um quarto para dormir, não possui banheiro e algumas famílias possuem um telefone celular em casa e uma geladeira.</p>
<p>Nível 2 – Inse de 180 a 260</p> <p>Nesse grupo algumas famílias não possuem renda e outras recebem até 1 salário mínimo. Os pais ou responsáveis alguns sabem ler e escrever e possuem o Ensino Fundamental completo ou incompleto.</p> <p>A casa que moram contém em geral um ou dois quartos para dormir e possui um banheiro. A família possui um telefone celular, uma geladeira em casa e algumas possuem um rádio e um videocassete e/ou DVD.</p>
<p>Nível 3 – Inse de 260 a 380</p> <p>A família possui renda familiar, em geral, até um salário mínimo. Os pais ou responsáveis sabem ler e escrever e possuem o Ensino Fundamental completo ou incompleto.</p> <p>A casa que moram contém em geral dois quartos para dormir e possui um banheiro. A família possui um telefone celular, uma geladeira em casa e algumas possuem um rádio e um videocassete e/ou DVD.</p>

Nível 4 – Inse de 380 a 430

A família possui renda familiar, em geral, até 1,5 salário mínimo. Os pais ou responsáveis sabem ler e escrever e possuem o Ensino Fundamental completo ou incompleto.

A casa contém em geral dois quartos para dormir e possui um banheiro. A família possui em casa um rádio, um videocassete e/ou DVD, internet, uma máquina de lavar roupa, uma geladeira e mais de um telefone celular. Além disso, algumas famílias possuem internet em casa.

Nível 5 – Inse de 430 a 520

A renda mensal da família é de 1 até 3 salários mínimos. Os pais ou responsáveis sabem ler e escrever e possuem o Ensino Fundamental incompleto ou já concluíram o Ensino Fundamental e/ou o Ensino Médio.

A casa contém em geral dois ou três quartos para dormir e possui um banheiro. A família possui em casa um rádio, um videocassete e/ou DVD, internet, uma máquina de lavar roupa, uma geladeira e mais de um telefone celular. Além disso, algumas famílias possuem um automóvel.

Nível 6 – Inse de 520 a 600

A renda mensal da família é de 1,5 até 7 salários mínimos. Os pais possuem Ensino Médio ou Superior Completo.

A casa contém em geral três quartos para dormir e possui um ou dois banheiros. Possui também três ou mais telefones celulares, um rádio, um videocassete e/ou DVD, internet, um automóvel, uma máquina de lavar roupa e uma geladeira. Algumas famílias possuem alguns bens e outras não, como: freezer (aparelho independente ou parte da geladeira) e TV por assinatura.

Nível 7 – Inse de 600 a 650

A renda mensal da família é em maior parte de 3 até 7 salários mínimos e em menor parte de 7 até 9 salários mínimos. Os pais possuem Ensino Médio ou Superior Completo.

A casa contém em geral três quartos para dormir e possui dois banheiros. Possui também três ou mais telefones celulares, um rádio, um ou dois videocassetes e/ou DVD, TV por assinatura, internet, um automóvel, uma máquina de lavar roupa e uma geladeira, freezer (aparelho independente ou parte da geladeira). Algumas famílias possuem aspirador de pó e outras não.

Nível 8 – Inse de 650 a 740

A renda mensal da família é bastante variada existem famílias com renda de 3 a 7 salários mínimos, de 7 a 12 salários mínimos e acima de 12. Os pais ou responsáveis possuem Ensino Superior Completo.

A casa contém em geral três ou mais quartos para dormir e possui três ou

mais banheiros. Possui também em casa três ou mais telefones celulares, um ou dois rádios, um ou dois videocassetes e/ou DVD, TV por assinatura, internet, uma máquina de lavar roupa, uma ou duas geladeiras, freezer (aparelho independente ou parte da geladeira), aspirador de pó e um ou dois automóveis. Algumas famílias contratam o serviço de empregada doméstica e outras não.

Nível 9 – Inse de 740 a 900

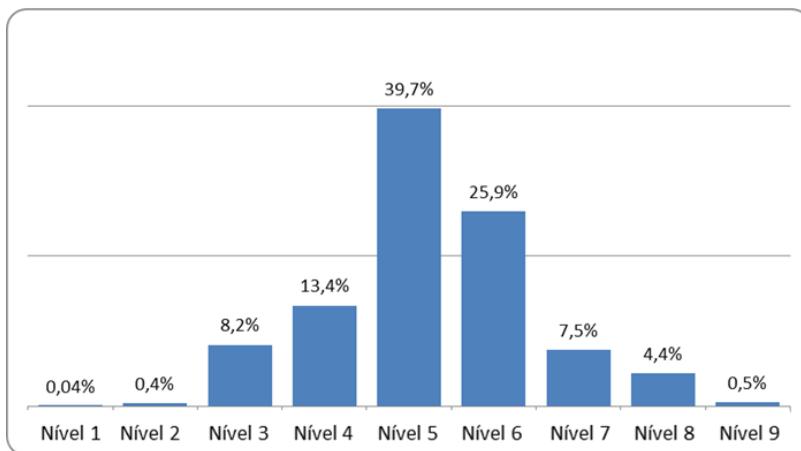
A renda mensal da família é maior que 12 salários mínimos. Os pais ou responsáveis possuem Ensino Superior Completo.

A casa contém em geral quatro ou mais quartos para dormir e possui três ou mais banheiros. Possui também em casa três ou mais telefones celulares, mais de um rádio, dois ou mais videocassetes e/ou DVD, TV por assinatura, internet, duas ou mais máquinas de lavar roupa, duas geladeiras, freezer (aparelho independente ou parte da geladeira), aspirador de pó e dois ou mais automóveis. E as famílias contratam o serviço de empregada doméstica.

Fonte: Elaborado pela autora

A seguir, no Gráfico 23, é apresentado como seriam classificados os alunos do banco de dados estudado a partir da definição dos nove grupos.

Gráfico 23. Distribuição dos alunos em 9 grupos do Inse



Fonte: Elaborado pela autora

Com base no Gráfico 23 e no Quadro 6, verifica-se que o indicador é sensível para captar camadas da sociedade que possuem um nível socioeconômico muito baixo, isto é aqueles que declararam que a

família não possui renda familiar, os pais ou responsáveis não sabem ler e escrever e possuem Ensino Fundamental completo ou incompleto. O indicador também é sensível a camadas da população com Inse mais alto já que consegue diferenciar os indivíduos com renda mensal familiar maior que 12 salários mínimos, cujos pais ou responsáveis possuem Ensino Superior Completo dos demais, por exemplo.

Esses resultados corroboram na vantagem em fazer a segmentação dos dados utilizando os parâmetros do modelo e não a partir da base de dados existente por meio de quintis, por exemplo.

6. CONSIDERAÇÕES FINAIS

Esse trabalho dissertou sobre os procedimentos necessários para a construção de indicadores contextuais utilizando a Teoria de Resposta ao Item – TRI e para tal, construiu-se um indicador cujo tema tem grande relevância na área educacional, o Indicador de Nível Socioeconômico familiar dos alunos – Inse. Um indicador de nível socioeconômico está cada vez mais presente nas discussões no âmbito da educação, pois permite uma visão mais contextualizada da realidade em que o aluno está inserido e subsidia, assim, o melhor desenvolvimento de políticas educacionais. Na revisão de literatura foram apresentados aspectos teóricos sobre o construto medido e a técnica estudada.

Para a construção de um indicador é necessário definir a população estudada e o instrumento que irá fornecer as informações para o cálculo. Estudou-se as respostas dos alunos do 5º e 9º ano do Ensino Fundamental e 3ª série do Ensino Médio, que responderam os questionários contextuais em 2013 do exame Enem e avaliação Saeb (Aneb e Anresc), realizados pelo Inep.

Após a seleção de uma amostra que representasse alunos diversos para a modelagem e da seleção das questões que poderiam compor o indicador, buscou-se estudar as características psicométricas dos instrumentos e a sua adequação aos modelos estudados da TRI. Verificou-se o ordenamento das alternativas pelo modelo nominal e algumas questões foram retiradas por não se adequarem ao ordenamento proposto, procedimento importante para validar as relações entre as alternativas das questões e construto medido. Além disso, foram aplicados três outros modelos e foi escolhido àquele que fornecia mais informação para os diversos valores da escala, o modelo de respostas graduais.

A TRI, como toda técnica estatística, possui pressupostos que precisam ser respeitados para que seu uso seja adequado. Desta forma, verificou-se a unidimensionalidade dos dados, a independência local e a invariância, avaliando a existência de um comportamento diferencial das questões entre os grupos avaliados. Tendo como referência esses pressupostos e outras medidas de ajuste realizados, finalizou-se o trabalho com uma proposta de modelo considerando 17 questões.

Após a definição do melhor modelo, buscou-se conhecer a consistência da medida criada fazendo algumas validações com variáveis como localização e dependência administrativa e verificou-se que os resultados apresentam bastante coerência. Buscou-se também

estudar como as questões se relacionam com a medida criada, estimou-se o Inse de cada aluno, e a partir das relações entre os parâmetros das questões e nível socioeconômico estudou-se a melhor forma de agrupar os alunos em faixas de nível socioeconômico de modo que eles fossem semelhantes dentro das faixas e diferentes entre as faixas e com isso nove grupos de nível socioeconômico distintos foram criados.

Por fim, cada grupo foi descrito a partir das questões/alternativas que melhor descreviam aquele grupo. A possibilidade de relacionar a medida criada com as questões que definem o indicador é um dos grandes diferenciais da TRI quando comparada com outros modelos.

O estudo demonstrou a relevância de se levantar os procedimentos necessários para a criação de indicadores a partir da técnica TRI, de modo a criar um indicador com qualidade e que se ajuste bem aos dados. Verificou-se também adequação dos dados estudados para a criação de um indicador de nível socioeconômico familiar dos alunos.

Este trabalho limitou-se a apresentar uma forma de segmentação de grupos e descrição de escala, acredita-se que trabalhos futuros possam ser desenvolvidos buscando outras técnicas de segmentação e descrição de escala para modelos politômicos. Além disso, outros indicadores também podem ser estudados a partir dos microdados do Saeb e Enem buscando melhor caracterização do contexto educacional que os alunos estão inseridos.

REFERÊNCIAS

- ABEP, Associação Brasileira de Empresas de Pesquisa. Critério de classificação econômica Brasil CCEB 2014 . Disponível em: <<http://www.abep.org/criterio-brasil>>. Acesso em: 29 jun. 2015.
- ANDRIOLA, W. B. Funcionamento diferencial dos itens (DIF): estudo com analogias para medir o raciocínio verbal. *Psicol. Reflex. Crit.*, Porto Alegre , v. 13, n. 3, p. 475-483, 2000. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79722000000300015&lng=en&nrm=iso>. Acesso em 11 jun. 2015.
- _____. Descrição dos Principais Métodos para Detectar o Funcionamento Diferencial dos Itens (DIF). *Psicologia: Reflexão e Crítica*, 2001, 14(3), pp. 643-652. Disponível em: <<http://www.scielo.br/pdf/prc/v14n3/7850.pdf>>. Acesso em 05 jan. 2016.
- ALVES, M. T. G., SOARES, J. F. Medidas de nível socioeconômico em pesquisas sociais: uma aplicação aos dados de uma pesquisa educacional. *Opinião Pública*, vol. 15, núm.1, p.1-30. Campinas, junho de 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-62762009000100001>. Acesso em 22 mar. 2015.
- _____. O nível socioeconômico das escolas de educação básica brasileiras. Belo Horizonte: Grupo de Avaliação e Medidas Educacionais (GAME); São Paulo: Instituto Unibanco, 2012. Disponível em: <<http://www.scielo.br/pdf/ensaio/v22n84/a05v22n84.pdf>>. Acesso em 22 mar. 2015.
- ALVES, M.T. G.; SOARES, J. F.; XAVIER, F. P. Índice socioeconômico das escolas de educação básica brasileiras. *Ensaio: aval.pol.públ.Educ.*, Rio de Janeiro , v. 22, n. 84, p. 671-703, Sept. 2014 . Disponível em: <<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362014000300005&lng=en&nrm=iso>>. Acesso em 20 jun. 2015.
- ANDRADE D.F.; TAVARES, H.R.; VALLE, R.C. (2000). Teoria da Resposta ao Item: Conceitos e Aplicações. São Paulo: Associação Brasileira de Estatística. Disponível em: <<http://www.ufpa.br/heliton/arquivos/LivroTRI.pdf>>. Acesso em 04 de jun. 2015.

AYALA, R.J. *The Theory and Practice of Item Response Theory*. New York, NY: The Guilford Press, 2009.

ATLAS DO DESENVOLVIMENTO HUMANO NO BRASIL. Rio de Janeiro, PNUD, IPEA, Fundação João Pinheiro, 2013. Disponível em: <<http://www.atlasbrasil.org.br/2013/>> Acesso em 29 Mar. 2016.

BEATON, A. E. & ALLEN, N. L. Interpreting Scales Through Scale Anchoring. *Journal of Educational Statistics*, 17, 191-204, (1992).

BOCK, R. D., & AITKIN, M. Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443-459 (1981). Disponível em: <<http://link.springer.com/article/10.1007%2FBF02293801>> Acesso em 29 Mar. 2016.

BORTOLOTTI, S.; VINCENZI L. ET AL. . Avaliação do nível de satisfação de alunos de uma instituição de ensino superior: uma aplicação da teoria da resposta ao item. *Gest. Prod.*, São Carlos, v. 19, n. 2, p. 287-302, 2012. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2012000200005&lng=en&nrm=iso>. Acesso em 05 Jun. 2015.

BUCHMANN, CLAUDIA. Measuring family background in international studies of education: Conceptual issues and methodological challenges. *Methodological advances in cross-national surveys of educational achievement* (2002): 150-197.

CHALMERS, P. (2015). Package mirt version 1.9. Disponível em: <<http://cran.r-project.org/web/packages/mirt/mirt.pdf>>. Acesso em: 1 abr. 2015.

CHEN, Wen-Hung; THISSEN, David. Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, v. 22, n. 3, p. 265-289, 1997.

CIZEK, J. GREGORY, (Ed.). *Setting Performance Standards: Theory and Applications*. Routledge, 2001.

COSTA, F. M. O Modelo de Samejima no estudo da relação NSE x Desempenho via Teoria da Resposta ao Item. *Dissertação de Mestrado*. Instituto de Matemática e Estatística Universidade Federal do Pará. Belém, 2010.

DE SÃO PAULO, E., SILVESTRI, M. B., GONÇALVES, M. N. J., RODRIGUES P. L. A. Aplicação do modelo de crédito parcial generalizado na avaliação do projeto SESI - por um brasil alfabetizado. REICE - Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación [online] 2007, 5 () : ISSN. Disponível em: <<http://www.redalyc.org/articulo.oa?id=55150204>> Acesso em: 15 jun. 2015.

EMBRETSON, S. E., REISE, S. Item response theory for psychologists. Mahwah, NJ: Erlbaum Publishers, 2000.

ERCIKAN, K. Translation effects in international assessments. International Journal of Educational Research, 29, 543-553, (1998). Disponível em: <<http://carme-educ.sites.olt.ubc.ca/files/2015/11/ercikan-1998.pdf>> . Acesso em: 05 jan. 2016.

FRANCO, C. Quais as contribuições da avaliação para as políticas educacionais. In: BONAMINO, A; BESSA, N; FRANCO, C. Avaliação da educação básica. Rio de Janeiro: Editora PUC-Rio; São Paulo: Loyola, 2004.

GLAS, C. A. W. Testing Fit to IRT Models for Polytomously Scored Items. In: M. L. Nering and R. Ostini (Eds.). Handbook of Polytomous Item Response Models (pp. 185-208). New York, NJ: Taylor and Francis Group, 2010.

HAMBLETON, R. K., H. SWAMINATHAN, e H. J. ROGERS. Fundamentals of item response theory. Vol. 2. Sage publications, 1991.

HAMBLETON, R. K., LINDEN VAN DER, W. J. and WELLS, C. S.. IRT Models for the Analysis of Polytomously Scored Data. In: M. L. Nering and R. Ostini (Eds.). Handbook of Polytomous Item Response Models (pp. 21-42). New York, NJ: Taylor and Francis Group, 2010.

HAIR JR., J. F.; TATHAM, R. L.; ANDERSON, R. E.; BLACK, W. Análise multivariada de dados. 5. ed. Porto Alegre: Bookman, 2005.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Microdados Enem. Microdados da Aneb e da Anresc 2013. Brasília: Inep, 2015 Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 3 abr. 2015.

_____. Microdados Saeb. Banco de dados, 2013. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 3 abr. 2015.

_____. Microdados Censo Escolar. Banco de dados 2013, [2014?]. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 3 abr. 2015.

_____. Nota Técnica Ideb. Disponível em: <http://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf>. Acesso em: 12 out. 2015.

_____. Nível Socioeconômico das escolas. Disponível em: <<http://portal.inep.gov.br/indicadores-educacionais>>. Acesso em: 22 abr. 2016.

KLEIN, Ruben. 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. ISSN 2175-2753. Disponível em: <<http://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/38>>. Acesso em: 06 Sep. 2015.

MACHADO, D. C. O. Análise de fatores associados ao desempenho escolar de alunos do quinto ano do ensino fundamental com base na construção de indicadores. Dissertação – Universidade Federal de Santa Catarina, Centro de Ciências Físicas e Matemáticas. Programa de Pós-Graduação em Métodos e Gestão em Avaliação. Florianópolis, SC, 2014. 227p.

MAY, H. A multilevel Bayesian item response theory method for scaling socioeconomic status in international studies of education. Journal of Educational and Behavioral Statistics, 31(1), 63-79. Disponível em : <http://repository.upenn.edu/cgi/viewcontent.cgi?article=1122&context=gse_pubs>. Acesso em: 06 jul. 2015.

MAFRA, P. M. R. Proposta de uma sistemática para a modelagem de risco de crédito sob a perspectiva da teoria da criação do conhecimento: uma abordagem da teoria da resposta ao item. 2010. 257f. Tese (Doutorado em Engenharia e Gestão do Conhecimento) - Universidade Federal de Santa Catarina, Florianópolis, 2010. Disponível em:

<<https://repositorio.ufsc.br/handle/123456789/95148>>. Acesso em: 09 set. 2015.

MAROCO, J.. Integração do R nos menus do PASW Statistics: Um exemplo de aplicação com o package 'polycor' do R. Boletim da Sociedade Portuguesa de Estatística (2010), 71-80.

MENEGON, L. D. S. Mensuração de conforto e desconforto em poltrona de aeronave pela teoria da resposta ao item. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 2013. Disponível em: <http://btd.egc.ufsc.br/wp-content/uploads/2011/04/Priscilla_Martins_-Ramos_Mafra.pdf> Acesso em: 09 set. 2015.

MICROSOFT. Office Excel. Versão 2010 [S.L.]: Microsoft Corporation, 2010. Software.

MORETTIN, P.A.; BUSSAB, W.O. Estatística básica, São Paulo: Saraiva. 5.ed., 2004. 526p.

MURAKI, E. A Generalized partial credit model: application of an EM algorithm. Applied Psychological Measurement, n.16, p. 159-176, 1992.

OSTINI, R. NERING, M. L. (2010). New perspectives and applications. In M. Nering & R. Ostini (Eds.), Handbook of polytomous item response theory models (Pp. 3-20). New York, NY: Routledge, 2010.

PINHEIRO, I. R.; COSTA, F. R.; CRUZ, R. M. Modelo nominal da teoria de resposta ao item: uma alternativa. Aval. psicol., Porto Alegre, v. 9, n. 3, dez. 2010. Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712010000300010&lng=pt&nrm=iso>. Acesso em 08 jun. 2015.

RECKASE, M. D. Unifactor latent trait models applied to multifactor tests: Results and implications. Journal of Educational Statistics, v. 4, p. 207-230, 1979.

REVELLE, W. (2015). Package 'psych' version 1.5.1. Disponível em: <<http://cran.r-project.org/web/packages/psych/psych.pdf>>. Acesso em: 3 abr. 2015.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em : <www.R-project.org>. Acesso em: 3 abr. 2015.

SAMEJIMA FA. Estimation of latent ability using a response pattern of graded scores. Psychometric Monograph.1969.

SARTES, L. M. A.; SOUZA-FORMIGONI, M. L. O. d. Avanços na psicometria: da Teoria Clássica dos Testes à Teoria de Resposta ao Item. Psicol. Reflex. Crit., Porto Alegre , v. 26, n. 2, p. 241-250, 2013. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79722013000200004>. Acesso em: 05 abr. 2016.

SAS software, versão 4.2. SAS Institute Inc., Cary, NC, USA.

SECRETARIA DE ESTADO DA EDUCAÇÃO DE SÃO PAULO. Nota Técnica - Programa de Qualidade da Escola. Março de 2014. Disponível em: <<http://idesp.edunet.sp.gov.br/Arquivos/NotaTecnica2013.pdf>> Acesso em: 17 ago. 2015.

SCHILLING, S. G. The role of psychometric modeling in test validation: An application of multidimensional item response theory. Measurement, v. 5, n. 2-3, p. 93-106, 2007.

SIEDENBERG, D. R. Indicadores de Desenvolvimento Socioeconômico - uma Síntese. Desenvolvimento em Questão, n. 1, p. 45-71, jan./jun. 2003.

SOARES, J. F.; ALVES, M. T. G. Efeitos de escolas e municípios na qualidade do ensino fundamental. Cadernos de Pesquisa (Fundação Carlos Chagas.), v. 43, p. 492-517, 2013a. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-15742013000200007&script=sci_abstract&tlng=pt>. Acesso em: 3 abr. 2015.

SOARES, T. M. Utilização da Teoria da Resposta ao Item na Produção de Indicadores Sócio-Econômico. Pesquisa Operacional, v.25, n.1, p.83-112, 2005. Disponível em: <<http://www.scielo.br/pdf/pope/v25n1/24252>>. Acesso em: 20 abr. 2015.

SOARES, J. F.; ANDRADE, R. D. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. Ensaio: aval. pol. públ. Educ., Rio de Janeiro, v.14, n.50, p. 107-126, jan./mar. 2006. Disponível em: <<http://www.scielo.br/pdf/ensaio/v14n50/30410.pdf>>. Acesso em: 20 mar. 2015.

SOARES, T. M.; SOUZA, R. C.; PEREIRA, V. R. Métodos Alternativos ao Critério Brasil para Construção de Indicadores Sócio-

Econômicos: Teoria da Resposta ao Item. In: XXXVI Simpósio Brasileiro de Pesquisa Operacional, 2004, São João Del Rei - MG. Anais, Rio de Janeiro, SOBRAPO, 2004. Disponível em <<http://www.din.uem.br/sbpo/sbpo2004/pdf/arq0089.pdf>>. Acesso em: 23 ago. 2015.

SOUZA, P. R. Avaliação Educacional utilizando a Teoria da Resposta ao Item. – TRI. Dissertação de Mestrado 2014, Universidade Federal do Acre - UFAC. Disponível em: <http://bit.proformat-sbm.org.br/xmlui/bitstream/handle/123456789/1229/2012_01013_PAULO_ROBERTO_DE_SOUZA.pdf?sequence=1>. Acesso em: 29 jun. 2015.

TILLÉ, Y.; MATEI, A. Sampling: Survey Sampling. R package version 2.7. Online: <<https://cran.r-project.org/web/packages/sampling/index.html>>

THISSEN, D., CAI, L., BOCK, R.D.. The nominal categories item response model. In M. Nering & R. Ostini (Eds.), Handbook of polytomous item response theory models (Pp. 43-75). New York, NY: Routledge, 2010.

VAN DER LINDEN, W. J., & HAMBLETON, R. K. (Eds.). (2013). Handbook of modern item response theory. Springer Science & Business Media.

YEN, W. M. Scaling performance assessments: Strategies for managing local item dependence. Journal of educational measurement, v. 30, n. 3, p. 187-213, 1993. Disponível em: <<http://www.education.umd.edu/MARC/mdarch/pdf/1000001.pdf>>. Acesso em: 07 jan. 2016.

KANG, T. & CHEN, T. An investigation of the performance of the generalized S-X2 item-fit index for polytomous IRT models. ACT (2007). Disponível em: <http://www.act.org/research/researchers/reports/pdf/%20ACT_RR2007-1.pdf>. Acesso em: 07 jan. 2016.

APÊNDICE A – PARÂMETROS E ERRO PADRÃO MODELO NOMINAL – ESCOLHA MODELO

Parâmetros e erro padrão primeira calibração modelo nominal								
Q16 - Situação residência								
1 - Outra situação (loteamento não regularizado, ocupação etc.)	a_0	0,00						
2 – Cedida	a_1 (erro-padrão)	0,33 (0,86)						
3 – Alugada	a_2 (erro-padrão)	2,90 (0,36)						
4 - Própria e em pagamento (financiada)	a_3 (erro-padrão)	10,05 (1,35)						
5 - Própria e quitada	a_4	4,00						
Q17 - Localização residência								
1 - Comunidade indígena ou quilombola	a_0	0,00						
2 - Zona rural	a_1 (erro-padrão)	7,02 (1,69)						
3 - Zona urbana	a_2	2,00						
Q18 - Quantas pessoas moram na casa (incluindo o aluno)								
1 - Uma, pois moro sozinho(a)	a_0	0,00						
2 – Duas	a_1 (erro-padrão)	6,07 (0,55)						
3 – Três	a_2 (erro-padrão)	-1,34 (1,48)						
4 – Quatro	a_3 (erro-padrão)	-4,67 (2,25)						
5 – Cinco	a_4 (erro-padrão)	-0,72 (1,33)						
6 - Seis pessoas ou mais	a_5	5,00						
Parâmetros e erro padrão segunda calibração modelo nominal								
Q16 - Situação residência								
1 - Outra situação (loteamento não regularizado, ocupação etc.)	a_0	0,00						
2 – Cedida	a_1 (erro-padrão)	-0,11 (0,76)						
3 – Alugada	a_2 (erro-padrão)	1,45 (0,40)						
4 - Própria e em pagamento (financiada) ou quitada	a_3	3,00						
Parâmetros e erro padrão terceira calibração modelo nominal								
	\hat{a}_0 (erro-padrão)	\hat{a}_1 (erro-padrão)	\hat{a}_2 (erro-padrão)	\hat{a}_3 (erro-padrão)	\hat{a}_4 (erro-padrão)	\hat{a}_5 (erro-padrão)	\hat{a}_6 (erro-padrão)	\hat{a}_7 (erro-padrão)
Q01	0,00	0,28 (0,02)	1,59 (0,01)	3,00				
Q02	0,00	1,13 (0,04)	2,11 (0,07)	3,00				

Q03	0,00	0,70 (0,01)	2,04 (0,02)	3,00				
Q04	0,00	1,16 (0,01)	2,20 (0,01)	3,00				
Q05	0,00	1,63 (0,02)	2,36 (0,04)	3,00				
Q06	0,00	0,52 (0,01)	2,08 (0,03)	3,00				
Q07	0,00	1,33 (0,02)	2,56 (0,05)	3,00				
Q08	0,00	1,24 (0,04)	1,73 (0,04)	3,00				
Q09	0,00	1,50 (0,16)	2,88 (0,32)	3,00				
Q10	0,00	1,32 (0,02)	2,36 (0,02)	3,00				
Q11	0,00	1,00						
Q12	0,00	1,68 (0,02)	2,55 (0,04)	3,00				
Q13	0,00	1,26 (0,01)	2,45 (0,02)	3,00				
Q14	0,00	0,91 (0,01)	1,87 (0,01)	3,00				
Q15	0,00	0,80 (0,06)	2,01 (0,04)	2,98 (0,02)	4,00			
Q19	0,00	1,00						
Q20	0,00	0,52 (0,05)	1,83 (0,05)	2,90 (0,06)	4,45 (0,06)	5,53 (0,08)	6,03 (0,08)	7,00
Q21	0,00	1,00						
Q22	0,00	1,00						
Q23	0,00	1,38 (0,01)	2,20 (0,01)	3,00				
Q24	0,00	1,32 (0,01)	2,17 (0,01)	3,00				

Fonte: desenvolvido pelo autor

APÊNDICE B – PARÂMETROS E ERRO PADRÃO MODELO DE RESPOSTAS GRADUAIS – ESCOLHA MODELO

Parâmetros e erro padrão primeira calibração modelo de respostas graduais								
Questões	a <small>(erro-padrão)</small>	b_1 <small>(erro-padrão d)</small>	b_2 <small>(erro-padrão d)</small>	b_3 <small>(erro-padrão d)</small>	b_4 <small>(erro-padrão d)</small>	b_5 <small>(erro-padrão d)</small>	b_6 <small>(erro-padrão d)</small>	b_7 <small>(erro-padrão d)</small>
Q01	1,37 (0,01)	-3,02 (0,02)	-0,04 (0,00)	1,54 (0,01)				
Q02	1,72 (0,02)	0,79 (0,02)	2,87 (0,06)	3,66 (0,09)				
Q03	0,97 (0,00)	-1,99 (0,01)	1,87 (0,01)	3,88 (0,02)				
Q04	1,88 (0,01)	-0,55 (0,01)	1,35 (0,01)	2,38 (0,02)				
Q05	2,13 (0,02)	-0,75 (0,02)	2,11 (0,04)	2,68 (0,06)				
Q06	0,61 (0,00)	-2,06 (0,00)	2,64 (0,01)	5,32 (0,01)				
Q07	1,52 (0,02)	-0,06 (0,01)	2,72 (0,04)	3,72 (0,07)				
Q08	0,97 (0,01)	-4,22 (0,04)	-1,01 (0,01)	0,20 (0,01)				
Q09	1,57 (0,03)	1,32 (0,02)	4,26 (0,11)	5,66 (0,30)				
Q10	1,06 (0,01)	-4,12 (0,02)	2,64 (0,01)	4,86 (0,03)				
Q11	0,89 (0,01)	-0,30 (0,00)						
Q12	1,12 (0,01)	-1,16 (0,01)	3,00 (0,01)	5,09 (0,03)				
Q13	1,44 (0,01)	0,07 (0,00)	2,03 (0,01)	3,37 (0,02)				
Q14	1,51 (0,01)	-3,44 (0,03)	0,81 (0,01)	2,34 (0,02)				
Q15	0,80 (0,00)	-6,02 (0,04)	-2,74 (0,01)	0,26 (0,00)	3,03 (0,01)			
Q19	0,73 (0,01)	3,44 (0,01)						
Q20	2,39 (0,02)	-2,72 (0,06)	-0,70 (0,02)	-0,07 (0,02)	0,85 (0,02)	1,80 (0,04)	2,12 (0,05)	2,49 (0,06)
Q21	1,31 (0,02)	-2,64 (0,02)						
Q22	1,16 (0,01)	-2,31 (0,01)						
Q23	1,02 (0,00)	-3,76 (0,02)	0,30 (0,00)	2,08 (0,01)				
Q24	1,12 (0,01)	-2,90 (0,02)	0,59 (0,01)	2,25 (0,01)				

Fonte: desenvolvido pelo autor

APÊNDICE C – PARÂMETROS E ERRO PADRÃO MODELO DE CRÉDITO PARCIAL GENERALIZADO – ESCOLHA MODELO

Parâmetros e erro padrão primeira calibração modelo crédito parcial generalizado								
Questões	a <small>(erro-padrão)</small>	b_1 <small>(erro-padrão d)</small>	b_2 <small>(erro-padrão d)</small>	b_3 <small>(erro-padrão d)</small>	b_4 <small>(erro-padrão d)</small>	b_5 <small>(erro-padrão d)</small>	b_6 <small>(erro-padrão d)</small>	b_7 <small>(erro-padrão d)</small>
Q01	1,07 (0,00)	-3,21 (0,02)	0,23 (0,02)	1,27 (0,02)				
Q02	1,54 (0,02)	0,88 (0,02)	2,87 (0,07)	2,37 (0,14)				
Q03	0,79 (0,00)	-2,11 (0,01)	2,28 (0,01)	2,79 (0,02)				
Q04	1,53 (0,01)	-0,51 (0,01)	1,51 (0,01)	1,95 (0,03)				
Q05	1,86 (0,02)	-0,80 (0,02)	2,38 (0,04)	1,86 (0,08)				
Q06	0,46 (0,00)	-2,08 (0,00)	3,35 (0,01)	3,22 (0,02)				
Q07	1,30 (0,01)	-0,04 (0,01)	2,96 (0,04)	2,33 (0,08)				
Q08	0,59 (0,01)	-4,90 (0,04)	0,01 (0,04)	-0,95 (0,04)				
Q09	1,54 (0,03)	1,34 (0,02)	3,85 (0,13)	3,61 (0,36)				
Q10	0,97 (0,01)	-4,35 (0,02)	2,88 (0,02)	3,45 (0,04)				
Q11	0,93 (0,04)	-0,26 (0,00)						
Q12	1,04 (0,00)	-1,18 (0,01)	3,09 (0,01)	3,63 (0,04)				
Q13	1,15 (0,00)	0,22 (0,00)	2,08 (0,01)	2,62 (0,03)				
Q14	1,27 (0,01)	-3,75 (0,03)	1,05 (0,03)	1,98 (0,03)				
Q15	0,55 (0,00)	-4,97 (0,04)	-2,80 (0,04)	0,33 (0,04)	2,77 (0,04)			
Q19	0,78 (0,01)	3,25 (0,01)						
Q20	1,35 (0,01)	-3,46 (0,06)	-0,32 (0,06)	-0,27 (0,06)	0,91 (0,06)	2,27 (0,07)	1,63 (0,08)	1,57 (0,11)
Q21	1,31 (0,02)	-2,59 (0,02)						
Q22	1,17 (0,01)	-2,25 (0,01)						
Q23	0,74 (0,00)	-4,29 (0,02)	0,76 (0,02)	1,55 (0,02)				
Q24	0,83 (0,00)	-3,27 (0,02)	0,99 (0,02)	1,75 (0,02)				

Fonte: desenvolvido pelo autor

**APÊNDICE D – PARÂMETROS E ERRO PADRÃO MODELO
DICOTÔMICO DE DOIS PARÂMETROS – ESCOLHA
MODELO**

Parâmetros e erro padrão primeira calibração modelo dicotômico		
Questões	a <i>(erro-padrão)</i>	b_1 <i>(erro-padrão d)</i>
Q01	0,87 (0,02)	-4,36 (0,02)
Q02	1,61 (0,03)	0,84 (0,02)
Q03	0,62 (0,01)	-3,03 (0,01)
Q04	2,24 (0,02)	-0,61 (0,01)
Q05	3,34 (0,07)	-0,65 (0,04)
Q06	0,31 (0,00)	-4,11 (0,00)
Q07	1,64 (0,02)	-0,05 (0,01)
Q08	1,29 (0,04)	-3,42 (0,06)
Q09	1,54 (0,03)	1,34 (0,02)
Q10	1,39 (0,02)	-3,49 (0,04)
Q11	0,76 (0,01)	-0,49 (0,00)
Q12	1,31 (0,01)	-1,17 (0,01)
Q13	1,33 (0,01)	-0,05 (0,00)
Q14	1,48 (0,03)	-3,61 (0,05)
Q15	1,30 (0,04)	-4,22 (0,07)
Q19	0,54 (0,01)	4,24 (0,01)
Q20	2,97 (0,08)	1,68 (0,10)
Q21	1,54 (0,03)	-2,53 (0,03)
Q22	1,40 (0,02)	-2,19 (0,02)
Q23	1,26 (0,01)	0,18 (0,01)
Q24	1,36 (0,01)	0,44 (0,01)

Fonte: desenvolvido pelo autor

APÊNDICE E – PARÂMETROS E ERRO PADRÃO MODELO DE RESPOSTAS GRADUAIS – MODELO FINAL

Parâmetros e erro padrão modelo final								
Questões	a (erro-padrão)	b ₁ (erro-padrão d)	b ₂ (erro-padrão d)	b ₃ (erro-padrão d)	b ₄ (erro-padrão d)	b ₅ (erro-padrão d)	b ₆ (erro-padrão d)	b ₇ (erro-padrão d)
Q02	1,638 (0,028)	0,809 (0,022)	2,917 (0,061)	3,717 (0,093)				
Q03	0,844 (0,008)	-2,341 (0,010)	1,922 (0,010)	4,151 (0,018)				
Q05	1,850 (0,02)	-0,811 (0,02)	2,222 (0,04)	2,834 (0,06)				
Q06	0,559 (0,00)	-2,343 (0,00)	2,702 (0,00)	5,595 (0,01)				
Q08	0,966 (0,01)	-4,233 (0,04)	-1,015 (0,01)	0,189 (0,01)				
Q09	1,618 (0,03)	1,289 (0,02)	4,123 (0,11)	5,486 (0,31)				
Q10	1,060 (0,01)	-4,214 (0,02)	2,508 (0,01)	4,716 (0,03)				
Q11	1,009 (0,02)	0,825 (0,01)						
Q12	1,579 (0,02)	-1,062 (0,02)	3,032 (0,04)	4,852 (0,16)				
Q13	1,408 (0,01)	-0,044 (0,00)	1,923 (0,01)	3,274 (0,02)				
Q14	1,840 (0,01)	-3,173 (0,03)	0,644 (0,01)	2,028 (0,02)				
Q15	1,039 (0,01)	-4,962 (0,04)	-2,367 (0,01)	0,108 (0,00)	2,410 (0,01)			
Q19	2,366 (0,05)	2,109 (0,07)						
Q20	2,507 (0,03)	-2,685 (0,06)	-0,696 (0,02)	-0,078 (0,02)	0,836 (0,02)	1,757 (0,04)	2,069 (0,05)	2,416 (0,07)
Q21	1,271 (0,02)	-2,838 (0,03)						
Q22	1,060 (0,01)	-2,619 (0,02)						
Q25	0,981 (0,00)	-4,372 (0,02)	-0,212 (0,00)	1,617 (0,01)				

Parâmetros e erro padrão ajuste modelo final

Questões	a (erro-padrão)	b_1 (erro-padrão d)	b_2 (erro-padrão d)	b_3 (erro-padrão d)	b_4 (erro-padrão d)	b_5 (erro-padrão d)	b_6 (erro-padrão d)	b_7 (erro-padrão d)
Q02	1,638 (0,02)	0,808 (0,02)	2,917 (0,06)	3,717 (0,09)				
Q03	0,843 (0,00)	-2,345 (0,01)	1,922 (0,01)	4,153 (0,01)				
Q05	1,850 (0,02)	-0,813 (0,02)	2,222 (0,04)	2,833 (0,06)				
Q06	0,558 (0,00)	-2,348 (0,00)	2,704 (0,00)	5,601 (0,01)				
Q08	0,967 (0,01)	-4,232 (0,04)	-1,016 (0,01)	0,188 (0,01)				
Q09	1,575 (0,03)	1,310 (0,02)						
Q10	1,059 (0,01)	-4,220 (0,02)	2,508 (0,01)	4,716 (0,03)				
Q11	1,008 (0,02)	0,825 (0,01)						
Q12	1,571 (0,02)	-1,067 (0,02)	3,041 (0,04)					
Q13	1,407 (0,01)	-0,046 (0,00)	1,922 (0,01)	3,273 (0,02)				
Q14	1,841 (0,01)	-3,175 (0,03)	0,642 (0,01)	2,026 (0,02)				
Q15	1,039 (0,01)	-4,966 (0,04)	-2,369 (0,01)	0,106 (0,00)	2,409 (0,01)			
Q19	2,372 (0,05)	2,105 (0,07)						
Q20	2,510 (0,03)	-2,687 (0,06)	-0,697 (0,02)	-0,079 (0,02)	0,834 (0,02)	1,754 (0,04)	2,066 (0,05)	2,412 (0,07)
Q21	1,271 (0,02)	-2,840 (0,03)						
Q22	1,060 (0,01)	-2,622 (0,02)						
Q25	0,981 (0,00)	-4,374 (0,02)	-0,213 (0,00)	1,615 (0,01)				

Fonte: desenvolvido pelo autor