Universidade de Brasília
IE - Departamento de Estatística
Trabalho de Conclusão de Curso 2

Avaliação da validade e da fidedignidade do questionário para a mensuração da conduta antissocial

Rodrigo Marques dos Santos

Relatório Final

Orientador: Prof. Dr. Luís Gustavo do Amaral Vinha

Brasília

11 de dezembro de 2017

Rodrigo Marques dos Santos

Avaliação da validade e da fidedignidade do questionário para a mensuração da conduta antissocial

Orientador: Prof. Dr. Luís Gustavo do Amaral Vinha

Projeto apresentado para obtenção do título de Bacharel em Estatística.

Brasília

11 de dezembro de 2017

RESUMO

O presente estudo tem por objetivo validar o instrumento de coleta de dados acerca do comportamento antissocial de jovens utilizado no estudo de Murta (2017). Para o alcance dos objetivos propostos, foi utilizada a Teoria de Resposta ao Item. Os itens foram analisados a partir do cálculo das correlações de Pearson entre eles e também do Alfa de Cronbach. Observou-se que, no geral, todos os itens tem boa correlação entre si e que, apesar de alguns itens não terem informações altas, a teoria não exige a retirada de nenhum dos itens do questionário. Mesmo assim, 3 itens foram selecionados para futuras melhorias. O questionário se mostrou bom para medir níveis de comportamento antissocial médios e altos, pecando ao medir os níveis baixos.

Palavras Chave: Comportamento Antissocial; Teoria de Resposta ao Item; Validade; Fidedignidade.

Sumário

1 INTRODUÇÃO **8**

2 METODOLOGIA **9**

2.1 Estimação da Fidedignidade 9

2.1.1 Tipo Teste-Reteste 10

2.1.2 Formas Paralelas 10

2.1.3 Inter-Avaliadores 10

2.1.4 Tipo Consistência Interna 11

2.1.5 O Coeficiente Alfa de Cronbach 11

2.2 Tipos da Validade 12

2.2.1 Validade de Conteúdo 12

2.2.2 Validade de Critério 12

2.2.3 Validade de Construto 13

2.3 Teoria de Resposta ao Item 13

2.3.1 Modelos para Itens Dicotômicos ou Dicotomizados 14

2.3.2 Modelo para Itens Não Dicotômicos 15

2.3.3 Curva de Informação do Item 16

2.3.4 TRI para a Análise da Validade 17

2.3.5 Comparação entre dois modelos 18

2.4 Instrumento de Coleta 18

3 RESULTADOS **20**

3.1 Análise Descritiva 20

3.2 Fidedignidade 22

3.3 Teoria de Resposta ao Item 23

3.3.1 Curva Característica do Item 23

3.3.2 Curvas de informação dos Itens 23

3.3.3 Curva de Informação do Teste 25

3.3.4 Curvas Características dos Itens Modificados 26

3.3.5 Curvas de Informação dos Itens 26

3.3.6 Curva de Informação do Teste Modificado 29

4 CONCLUSÃO **29**

Referências **31**

5 APÊNDICE **33**
5.1 Informação dos Itens Separadamente 33
5.2 Informação dos Itens modificados Separadamente 34
5.3 Curvas Operacionais dos Itens Separadamente 35
5.4 Código em R 35

Lista de Figuras

1 Exemplos de curvas de probabilidade de acerto segundo habilidade ou traço latente 15
2 Exemplo de um item bem ajustado ao Modelo de Resposta Gradual . 17
3 Questionário sobre comportamento antissocial 19
4 Distribuição amostral do sexo dos respondentes 20
5 Distribuição amostral da idade dos respondentes 21
6 Distribuição amostral da escolaridade dos responsáveis dos respondentes 21
7 Correlação entre os itens do questionário 22
8 Curvas Características para os 15 itens 24
9 Curvas de informação para os 15 itens 25
10 Curvas de informação para o teste 25
13 Curvas de informação para os 15 itens modificados 26
11 Frequências por categorias em cada item 27
12 Curvas Características para os 15 itens modificados 28
14 Curvas de informação para o teste modificado 29

1 INTRODUÇÃO

Um estudo psicométrico busca analisar e estudar o comportamento humano com base nas técnicas quantitativas existentes. Para isso, este tipo de pesquisa analisa uma série de respostas dadas por um respondente a várias tarefas diferentes, tipicamente chamadas de "itens"(PASQUALI, 2009). O principal objetivo dos estudos psicométricos é transformar características psicológicas e pessoais em números, para permitir interpretações mais fáceis. Por isso, a estatística se torna fundamental, com a introdução de métodos quantitativos os quais ajudam não só na interpretação, mas também na coleta de dados, desde a criação do instrumento até o modo como será aplicado.

Esses instrumentos, em sua maioria, são questionários. Isso se dá pois estes são maneiras interessantes de agrupar variáveis diferentes de um mesmo indivíduo e, assim, utilizá-las em um mesmo estudo. Além disso, ao utilizarem-se estruturas de correlação em agrupamento de perguntas de um mesmo construto, a confiança desse instrumento aumenta (DANCEY; REIDY, 2013). Porém, sem a devida atenção à fidedignidade e à validade do questionário, a coleta de dados feita por esse instrumento pode ser não merecedora de crédito algum. Segundo Martins (2006), a fidedignidade de um instrumento é a consistência dos dados, ou seja, a capacidade dos dados serem reproduzidos por outros pesquisadores usando as mesmas técnicas. Já a validade, o autor define como a capacidade de um instrumento de realmente medir um construto.

O termo construto é definido por Freitas (1994) como um conceito abstrato criado ou adotado intencionalmente com o objetivo de se conseguir definições e especificações que permitam sua observação e mensuração, por meio de pesquisas científicas. São exemplos de construtos muito utilizados na área de Psicologia: inteligência, atitude, amor, medo, dor, entre outros.

No presente estudo, o construto a ser estudado é o comportamento antissocial. Em uma pesquisa sobre o uso de drogas por jovens adolescentes participantes do programa Fortalecendo Famílias, Murta et al. (2017) aplicaram um questionário que coletava dados relacionados a vários construtos, além do comportamento antissocial. Alguns deles são: autoeficácia, vida escolar, relacionamento com os pais, uso de drogas, entre outros. De agora em diante, usaremos a definição de comportamento antissocial proposta por Bringas (2003), que a define como "aquele comportamento que não está de acordo com as normas sociais ou morais."

Apesar do comportamento antissocial já ter sido abordado em vários estudos, poucos correlacionam diretamente esse construto e o uso de drogas na ado-

lescência. Uma pesquisadora que cita essa relação é Boldin (2000). Ela coloca que, frequentemente, como foi observado em algumas pesquisas estrangeiras, é comum o comportamento antissocial de um indivíduo levá-lo a quebrar regras e aumentar a probabilidade de um comportamento de risco, um exemplo disso é o uso de drogas. Um outro estudo que cita esses construtos é o de Pacheco (2004), o qual descreve que os jovens que iniciam um comportamento antissocial precocemente têm maior probabilidade de desenvolverem comportamentos tipicamente definidos como delinquentes, entre eles o alcoolismo e uso de drogas.

Assim, o objetivo principal deste estudo é analisar o questionário utilizado por Murta et al. (2017), tendo como foco os itens relacionados com a avaliação do comportamento antissocial. Para isso, o estudo deverá alcançar, por meio da Teoria de Resposta ao Item, os seguintes objetivos específicos: avaliar e estimação da fidedignidade do questionário; avaliar a validade dele; e analisar a característica dos itens individualmente.

Desta maneira, poderiam ser reproduzidos estudos mais aprofundados com estes itens, ou ainda se estudar melhorias para estes que os tornem ainda mais relevantes. Posteriormente, poderia ser criado um banco de dados que permita múltiplas aplicações diferentes em populações diferentes ou na mesma população tempos depois e, com as qualidades devidas do teste e dos itens nele presentes, permitisse a comparação dos resultados dessas aplicações para possíveis estudos posteriores.

2 METODOLOGIA

Para uma correta aplicação de um instrumento de coleta, é preciso se atentar a duas características essenciais: a fidedignidade e a validade. Essas duas características e seus respectivos métodos de estimação serão apresentados a seguir.

2.1 Estimação da Fidedignidade

Ao aplicar um questionário para uma amostra ou população, tem-se por objetivo coletar dados referentes a um ou mais construtos dos indivíduos respondentes. Assim, observa-se o escore de cada respondente para aquele conjunto de itens. Mas, como geralmente lidamos com seres humanos e traços pouco precisos, esse escore está sujeito a um erro aleatório, que não necessariamente é causado pela falta de qualidade do questionário em si, mas sim por algo não controlável (SILVA NETO, 2008).

Para ajudar a entender e controlar este erro, é calculada a fidedignidade deste questionário. A fidedignidade, como é definida por Anastasi e Urbina (1997), é a consistência dos escores obtidos pelas mesmas pessoas quando elas são submetidas ao mesmo teste ou a um teste com itens equivalentes àquele.

Existem diferentes formas de se obter uma medida que expresse o nível de fidedignidade de um questionário, listadas e explicadas a seguir.

2.1.1 Tipo Teste-Reteste

É a aplicação do mesmo teste para o mesmo grupo de respondentes em momentos diferentes. Assim, o coeficiente de correlação entre os escores em tempos diferentes é usado como medida de fidedignidade do teste.

O erro existente neste tipo de estimação de fidedignidade está associado à situação de testagem, ou seja, a todas as diferenças nos momentos das duas aplicações do questionário. Além disso, a correlação calculada tende a ser mais alta quanto menor for o tempo entre as aplicações

2.1.2 Formas Paralelas

É a aplicação de dois testes diferentes, porém equivalentes, aos mesmos respondentes em uma mesma situação. A partir dessas aplicações, calcula-se a correlação entre os escores dos dois testes.

Para este tipo, o erro associado vem de uma possível flutuação de desempenho ao se utilizar um teste ou outro. Além disso, é importante salientar que, se houver um intervalo de tempo entre as aplicações, esse método levará em conta também o erro do tipo Teste-Reteste, já que será uma combinação dos dois procedimentos.

2.1.3 Inter-Avaliadores

Consiste na aplicação de um só teste e a correção deste feita por dois ou mais avaliadores diferentes. Assim, é calculada a correlação dos escores baseados nas diferentes correções do teste.

Aqui, deseja-se observar se estão de acordo os métodos de pontuação para as mesmas questões, se elas possuem regras claras não só para o ato de respondê-la, mas também para avaliá-la, mesmo que seja por diferentes avaliadores.

2.1.4 Tipo Consistência Interna

Na busca de corrigir os erros existentes nos tipos Teste-Reteste e Formas Paralelas, a consistência interna é medida com base em um só questionário aplicado em um só momento aos respondentes. Após essa aplicação, avalia-se as correlações existentes entre os itens daquele teste.

Para isso, várias fórmulas foram criadas e podem ser utilizadas, como analisar a correlação entre duas metades do teste ou separar os itens em pares e impares para obter a mesma medida. A medida mais comum é o Alfa de Cronbach, o qual será utilizado neste trabalho.

2.1.5 O Coeficiente Alfa de Cronbach

Este estudo terá foco no cálculo do Alfa de Cronbach (α), pois temos a aplicação de apenas um questionário, não sendo possível a aplicação de nenhuma das outras técnicas acima citadas.

O Coeficiente α é um estimador de fidedignidade do tipo de consistência interna muito indicado para questionários cujos itens tem múltiplas escolhas. De uma maneira simplificada, o coeficiente mede a correlação entre o desvio dos escores do item e os escores totais dos respondentes (HOGAN, 2006).

A fórmula do α de Cronbach é representada pela expressão

$$\alpha = \left(\frac{K}{K-1}\right) \cdot \left(1 - \frac{\sum S_i^2}{S_x^2}\right), \tag{1}$$

onde K é o número de itens no teste, S_x é o desvio padrão dos escores totais dos respondentes, e S_i é o desvio padrão dos escores de cada item.

Segundo Maroco e Garcia-Marques (2006), o fato de esta fórmula não assumir homogeneidade de variâncias, ou seja, não assumir que as variâncias dos itens são iguais, faz com que o α subestime a verdadeira consistência em muitos casos. Uma forma de evitar esse problema seria padronizar os dados, transformando a equação 1 em

$$\alpha = \frac{K(\bar{r_{ij}})}{1 + (K-1)\bar{r_{ij}}}, \tag{2}$$

onde $\bar{r_{ij}}$ é a correlação média entre as questões i e j, e K é o número de questões.

Porém, apesar de ser amplamente utilizado, muitas vezes sem o devido entendimento, este coeficiente pode não ser apropriado para alguns casos, como os teste que procuram determinar velocidade de reação, por exemplo. Além disso, o α tende a crescer à medida que a quantidade de itens cresce, tornando a estimativa

imprecisa para um número alto de K.

Este coeficiente varia de 0 a 1, com 0 significando nenhuma semelhança dos itens e 1 uma identidade dos itens. Por isso, segundo vários autores, como Hogan (2006), Maroco e Garcia-Marques (2006) e Almeida et al. (2010), o coeficiente precisa estar, no mínimo, entre 0,6 e 0,7 para considerarmos um instrumento fidedigno. Por outro lado, o valor não pode ultrapassar 0,9. Isso se dá pois, para valores acima deste, há indícios de que muitos dos itens podem estar medindo exatamente a mesma coisa, tornando o teste redundante e podendo aumentar significativamente a variância (ALMEIDA et al., 2010).

2.2 Tipos da Validade

A validade é o mais importante dos aspectos para a avaliação de um questionário e de sua qualidade. Isso porque esta mede a capacidade do método de coleta de realmente medir o que se propõe medir (HOGAN, 2006). É importante salientar que a validade não pode existir sem a fidedignidade.

Assim como na fidedignidade, existem vários tipos de validade, os quais serão listados e resumidos a seguir.

2.2.1 Validade de Conteúdo

A validade de conteúdo se refere à intensidade da associação que existe entre o questionário criado e o conteúdo referente ao traço latente que se deseja medir (HOGAN, 2006). O exemplo clássico é a prova do Exame Nacional do Ensino Médio (ENEM), que procura abranger todo o conhecimento existente no currículo do Ensino Médio, um conteúdo finito e definido pelo Ministério da Educação.

Apesar de parecer fácil de se obter esse tipo de validade, raramente consegue-se obter uma definição do conteúdo finita e bem definida. Além disso, existem vários tipos de itens diferentes para se medir o mesmo elemento específico do conteúdo, tornando difícil saber se aquele teste mede o construto da melhor forma possível (HOGAN, 2006).

2.2.2 Validade de Critério

A validade de critério, de acordo com Pasquali (2009), "consiste no grau de eficácia que ele tem em predizer um desempenho específico de um sujeito". É importante dizer que a mensuração do construto deve ser independente do teste que se está querendo validar. Dito isso, o coeficiente de validade será a correlação entre

o escore dos respondentes no teste aplicado e a medida referente ao desempenho desejado.

Alguns exemplos deste tipo de validade são os exames que medem um traço psicológico, como o nível de depressão, a ser comparado com a avaliação que um psicólogo faz desse mesmo indivíduo, ou ainda um inventário de personalidade que pretende predizer a probabilidade do sujeito de se suicidar no futuro (HOGAN, 2006).

2.2.3 Validade de Construto

Essa validade é considerada por muitos a forma mais fundamental de validade e se refere a quanto os itens que se propõem a medir um mesmo construto se correlacionam entre si (PASQUALI, 2009).

O meio mais utilizado para medir este tipo de validade é a Análise Fatorial, a qual separa diferentes itens em um ou mais fatores, dependendo do teste. Porém, outro método para mensurar a validade é a Teoria de Resposta ao Item (TRI) e, neste estudo, o foco será dado para este método.

2.3 Teoria de Resposta ao Item

A Teoria Clássica dos Testes (TCT) procura analisar os escores totais dos respondentes, ou seja, o somatório dos acertos em todos os itens presentes no teste. Uma adaptação desta teoria seria adicionar pesos a cada um dos itens, transformando o somatório total no somatório dos acertos multiplicados pelos seus respectivos pesos. Essa técnica tenta priorizar itens considerados mais importantes em relação aos outros.

Porém, alguns problemas surgem deste tipo de abordagem, como a escolha dos pesos de cada questão (SABOIA, 2013), ou ainda a incapacidade de comparação do traço latente entre duas pessoas que não fizeram a mesma prova. Isto se dá pois a estimativa deste traço depende do conjunto de todas as questões, tornando cada prova completamente diferente da outra (ANDRADE et al., 2000).

A TRI é uma alternativa para o método dos pesos e tem algumas vantagens com relação a ele. A TRI, ao invés de olhar para o teste todo, procura explicar cada item e analisar a capacidade dele sozinho de medir o traço latente desejado, permitindo que sejam determinados pesos para a estimação deste.

Atualmente, no Brasil, vem crescendo muito a utilização da TRI, que se propõe a criar modelos diferentes para os mais variados tipos de itens existentes (ANDRADE et al., 2000). Vamos analisar os principais a seguir.

2.3.1 Modelos para Itens Dicotômicos ou Dicotomizados

Itens dicotômicos são aqueles que são considerados itens que tem duas opções de escolha. São considerados exemplos de itens dicotômicos ou dicotomizados:

- itens de múltipla escolha com apenas uma resposta certa.
- itens de resposta livre que são corrigidos de maneira dicotômica, ou seja, como certo ou errado, sem a possibilidade de um meio certo.
- itens de resposta dicotômica sem uma resposta certa, ou seja, itens distintos que descrevem uma proficiência distinta.

Seguindo este raciocínio, podemos considerar a probabilidade de acerto de um item dicotomizado como sendo uma Bernoulli de parâmetro p_{ij}, sendo p_{ij} a probabilidade do j-ésimo indivíduo de acertar a i-ésima questão (SABOIA, 2013).

Assim, os três modelos mais utilizados são os modelos de regressão logística com um, dois ou três parâmetros. São eles:

- Apenas o grau de dificuldade;
- O grau de dificuldade e o grau de discriminação; e
- O grau de dificuldade, o grau de discriminação e a probabilidade de acerto ao acaso.

Considerando que o modelo com 3 parâmetros é o mais completo e os outros dois podem ser facilmente obtidos a partir dele, vamos abordar apenas este modelo no presente estudo.

Assim o modelo de 3 parâmetros, segundo Andrade et al. (2000), se dá por

$$P(U_{ij} = 1|\theta j) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta j - b_i)}}, \tag{3}$$

onde θj é o traço latente do indivíduo j, $P(U_{ij} = 1|\theta j)$ é a probabilidade de um indivíduo j acertar a questão i dado que tem habilidade θj, a_i é o grau de discriminação da questão i, b_i é o grau de dificuldade da questão i, c_i é a probabilidade de acerto ao acaso da questão i, se esta existir, e D é um fator de escala, constante, igual a 1,7, para que a função logística obtida tenha resultados parecidos com a função ogiva normal.

A probabilidade de acerto, a partir do cálculo dos parâmetros, pode ser representada por uma curva característica. Alguns exemplos de curvas características

podem ser observados na Figura 1. Essa imagem também é útil para entender a interpretação dos parâmetros a_i, b_i e c_i.

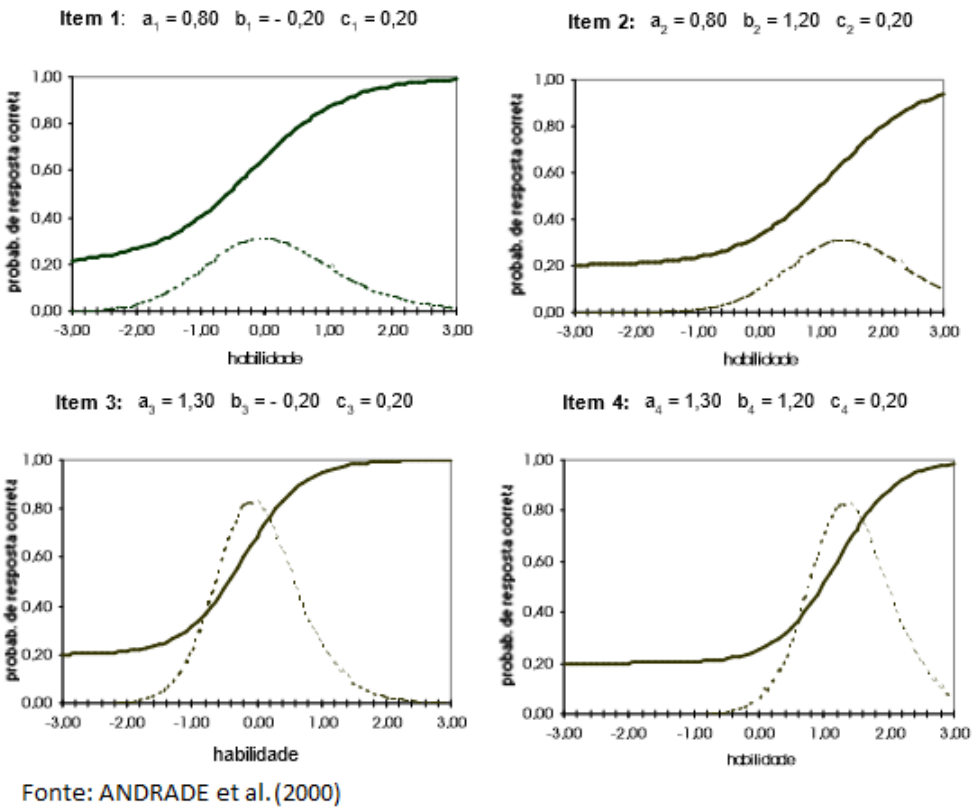


Figura 1: Exemplos de curvas de probabilidade de acerto segundo habilidade ou traço latente

O parâmetro a_i é o grau de discriminação, ou seja, é a capacidade do item de discriminar pequenas diferenças no traço latente estudado. Esse parâmetro é proporcional à derivada da curva no ponto de inflexão (ANDRADE et al., 2000). Além disso, ele determina a inclinação da curva característica (SABOIA, 2013). O parâmetro b_i é o grau de dificuldade, ou seja, quanto mais alto, maior deverá ser o traço latente para se acertar a questão. Na curva característica, ele pode ser encontrado como o valor de θ que proporciona uma probabilidade de acerto igual a (1+c)/2 (ANDRADE et al., 2000). Já o parâmetro c_i é a probabilidade de acerto ao acaso. No gráfico, ele pode ser observado como o limite inferior da curva característica (ANDRADE et al., 2000).

2.3.2 Modelo para Itens Não Dicotômicos

Aqui, são considerados modelos para itens abertos (de resposta livre), ou para itens de múltipla escolha na qual essas escolhas seguem uma graduação

de intensidade (ANDRADE et al., 2000). Neste momento, será analisado apenas o modelo para os itens politômicos chamado "Modelo de Resposta Gradual", já que os itens do questionário a serem analisados no presente estudo seguem essa ideia.

Desta forma, a análise será focada no Modelo de Resposta Gradual de Samejima (1969). Este modelo assume que as respostas estejam ordenadas de forma crescente segundo algum critério pré-definido, e define as probabilidades de um respondente selecionar uma particular categoria ou outra mais alta segundo a seguinte equação

$$P^{+}_{i,k}(\theta j) = \frac{1}{1 + e^{-Da_i(\theta j - b_{i,k})}}, \tag{4}$$

onde $b_{i,k}$ é o parâmetro de dificuldade da k-ésima categoria do item i.

Assim, como esta equação trata a probabilidade como uma função acumulada, a probabilidade de um respondente selecionar exatamente a categoria k é dada por

$$P_{i,k}(\theta j) = P^{+}_{i,k}(\theta j) - P^{+}_{i,k+1}(\theta j). \tag{5}$$

Por definição de Samejima (1969), temos que a probabilidade do primeiro k é 1 e a probabilidade do último k é 0.

A Figura 2 representa como deveria ser um modelo bem ajustado pelo Modelo de Resposta Gradual. Este gráfico mostra que as diferentes curvas estão bem espaçadas, representando que cada categoria k está estimando um traço latente diferente, seguindo a ordenação da relação pré-definida das categorias com o traço latente desejado (SABOIA, 2013).

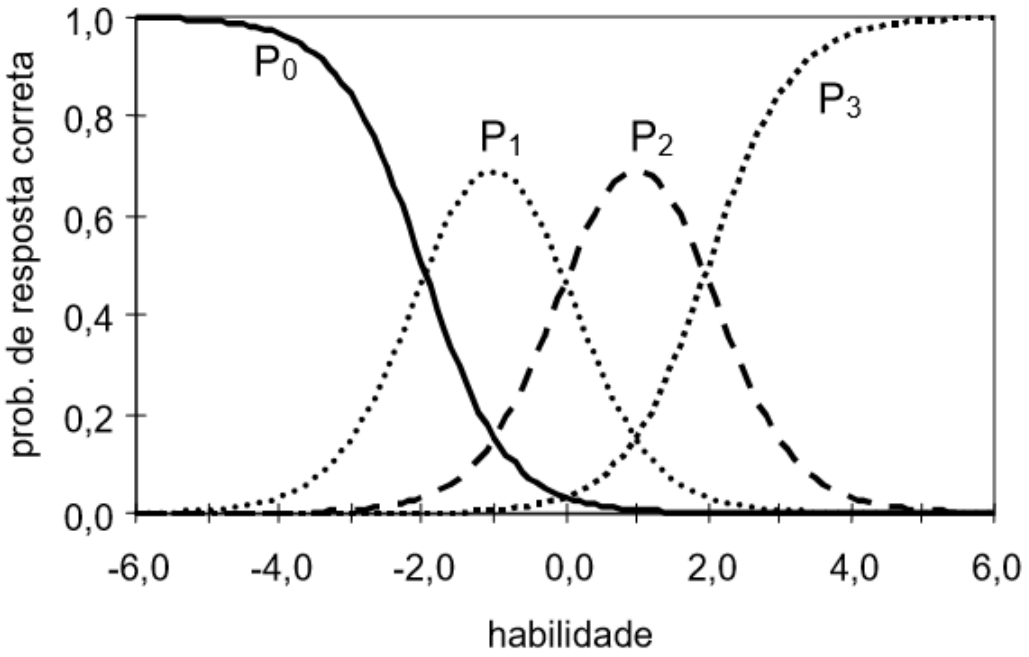
2.3.3 Curva de Informação do Item

A curva de informação do item é utilizada para se estudar os valores do traço latente os quais aquele item consegue estimar com precisão, ou seja, provém uma informação alta, e quais valores o item não consegue estimar bem, ou seja, provém uma informação baixa (SABOIA, 2013).

Assim, a função de informação do item i relativa ao traço latente θ é dada por:

$$I_i(\theta) = \sum_{x=1}^{k_i} \frac{\left[P'_{i,x}(\theta)\right]^2}{P_{i,x}(\theta)Q_{i,x}(\theta)} \tag{6}$$

onde $P'_{i,x}(\theta)$ é a primeira derivada da equação 5, k_i é o número de categorias do item i e $Q_{i,x}(\theta)$ é $1 - P_{i,x}(\theta)$



Fonte: ANDRADE et al. (2000)

Figura 2: Exemplo de um item bem ajustado ao Modelo de Resposta Gradual

Além disso, a informação fornecida por um teste nada mais é do que a soma de todas as informações dos itens que compõem este teste (ANDRADE et al., 2000).

2.3.4 TRI para a Análise da Validade

A TRI faz uma análise completa, item por item, de todo o instrumento de coleta. Desta forma, ela consegue indicar quais itens devem ser excluídos por não se encaixarem com o resto do questionário ou ainda por indicarem viés.

Assim, itens que podem ser marcados como aptos para exclusão são os itens que tem baixa ou negativa correlação com o escore total. Além disso, um número alto de não respondentes pode ser um indicativo de que o item não é compreensível aos respondentes, ou ainda um possível indicativo de um problema no item seria se 25% de respostas fossem da categoria mais extrema do construto (JONES, 2009).

Além disso, itens com correlação maior que 0,7 com outros itens podem ser redundantes. Outro tipo de indicativo de exclusão seria uma correlação muito alta com alguma variável externa, como sexo ou idade, indicando que a resposta daquele item não é dada apenas pelo traço latente, mas sim, em boa parte, pela

resposta àquela outra variável externa também (JONES, 2009).

Desta forma, com base na TRI, pode-se certificar que todos os itens do questionário estão realmente medindo o construto que este se propõe a medir, ou seja, que todos os itens estão sendo respondidos, em grande parte, de acordo com o valor do traço latente que aquele respondente possui.

Por isso, neste estudo, a análise do instrumento de coleta será feita por esta teoria, na intenção de classificar os itens e na busca de mais informações sobre estes.

2.3.5 Comparação entre dois modelos

Segundo Emiliano (2010), ao selecionarmos modelos é preciso ter em mente que não existem modelos verdadeiros. Há apenas modelos aproximados da realidade que causam perda de informações. Deste modo, é necessário fazer a seleção do melhor modelo possível dentre aqueles que foram ajustados para explicar o fenômeno sob estudo.

Desta forma, ao utilizar-se da TRI, um modelo será construído, e seus parâmetros serão estimados. Porém, é preciso que consigamos alguma forma de comparar dois modelos para que se tenha ideia do impacto que alguma mudança na modelagem terá no modelo resultante.

Assim, este estudo utilizará os critérios de informação de Akaike (AIC) e Bayesiano (BIC). O AIC estima que o viés causado pelas estimações dos parâmetros e da verossimilhança é o próprio número de parâmetros (EMILIANO, 2010). Assim, a fórmula fica como segue

$$AIC = -2LogL(\hat{\theta}) + 2(p), \tag{7}$$

onde L é a função de verossimilhança do modelo e p é o número de parâmetros.

Já o BIC estima que o viés seja $\frac{pLogn}{2}$ (EMILIANO, 2010), resultando na seguinte fórmula

$$BIC = -2LogL(\hat{\theta}) + pLogn, \tag{8}$$

onde L é a função de verossimilhança do modelo e n é o tamanho da amostra.

2.4 Instrumento de Coleta

O questionário a ser estudado se propõe a medir o comportamento antissocial dos respondentes em uma pesquisa de avaliação sobre o Programa Famílias

Fortes[1] .

Este questionário possui 15 itens a serem respondidos em uma escala que varia de “Nunca” a “Muitas vezes”, indicando, de forma ordenada, a frequência com que o respondente pratica ações consideradas antissociais.

44. NOS ÚLTIMOS SEIS MESES, você já praticou alguma das ações abaixo?

	Nunca	Pouquíssimas vezes	Algumas vezes	Várias vezes	Muitas vezes
a) Mentiu para tirar vantagem	☐	☐	☐	☐	☐
b) Xingou pessoas	☐	☐	☐	☐	☐
c) Ameaçou pessoas	☐	☐	☐	☐	☐
d) Entrou em lugares proibidos	☐	☐	☐	☐	☐
e) Jogou objetos em carros na rua	☐	☐	☐	☐	☐
f) Carregou algum objeto que serve como arma para brigar com alguém	☐	☐	☐	☐	☐
g) Pegou sem permissão coisas que não são suas	☐	☐	☐	☐	☐
h) Respondeu mal aos mais velhos	☐	☐	☐	☐	☐
i) Provocou briga entre outras pessoas	☐	☐	☐	☐	☐
j) Cuspiu em alguém	☐	☐	☐	☐	☐
k) Pegou comida ou bebida a mais mesmo sabendo que outra pessoa podia ficar sem	☐	☐	☐	☐	☐
l) Culpou outra pessoa por algo que você fez	☐	☐	☐	☐	☐
m) Pichou lugares não permitidos	☐	☐	☐	☐	☐
n) Estragou coisas na rua (ex, placa, banco, parada de ônibus, janelas, vitrines)	☐	☐	☐	☐	☐
o) Pegou coisas dos outros quando não estavam olhando	☐	☐	☐	☐	☐

Figura 3: Questionário sobre comportamento antissocial

[1]Para mais informações a cerca do programa Famílias Fortes, acesse o link ⟨http://prevencaoepesquisa.icict.fiocruz.br/⟩

3 RESULTADOS

Primeiramente, serão apresentados os resultados de uma análise exploratória e descritiva dos dados obtidos pelos 699 adolescentes e crianças que responderam ao questionário. Este é constituído por 15 questões que pretendem medir o grau de comportamento antissocial dos respondentes.

3.1 Análise Descritiva

Foi calculado que apenas 2,4% dos dados foram faltantes para os itens sobre sexo, idade e escolaridade dos pais. Isso pode se dar por uma série de motivos, como itens mal preenchidos ou deixados em branco, alguma dificuldade de entendimento, entre outros tipos de inconsistência.

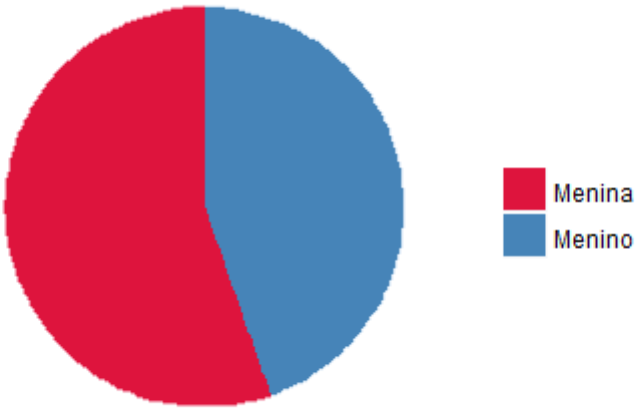


Figura 4: Distribuição amostral do sexo dos respondentes

Pela Figura 4, é possível observar uma maioria feminina dos respondentes com as mulheres representando 55% deles contra 45% são do sexo masculino. Já com relação a idade, destaca-se, na Figura 5, que apenas a idade de 11 anos foge do padrão, com a porcentagem consideravelmente abaixo de 25%. As outras idades tiveram os respondentes bem distribuídos, com uma predominância ligeiramente maior para a média, ou seja, entre 12 e 13 anos.

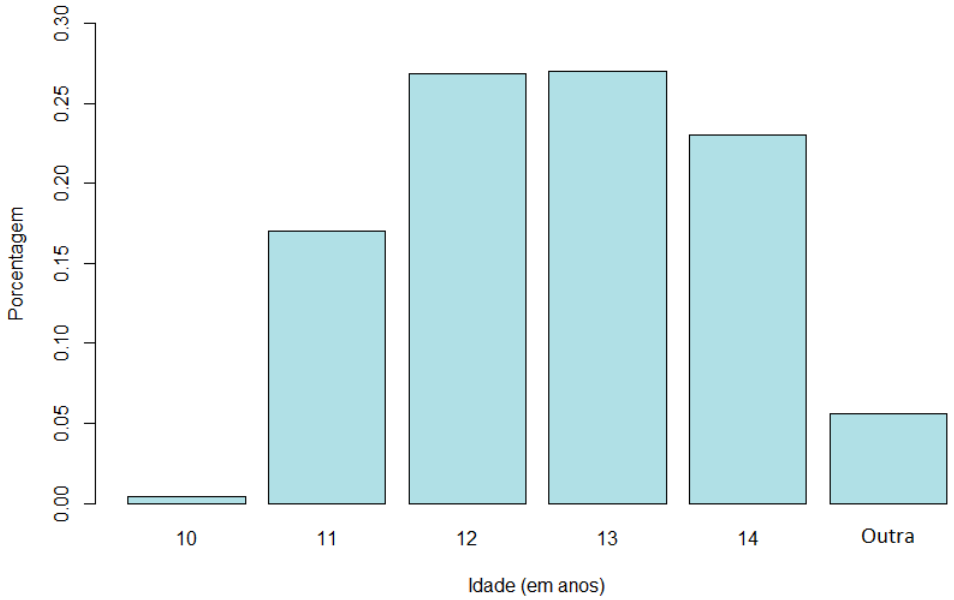


Figura 5: Distribuição amostral da idade dos respondentes

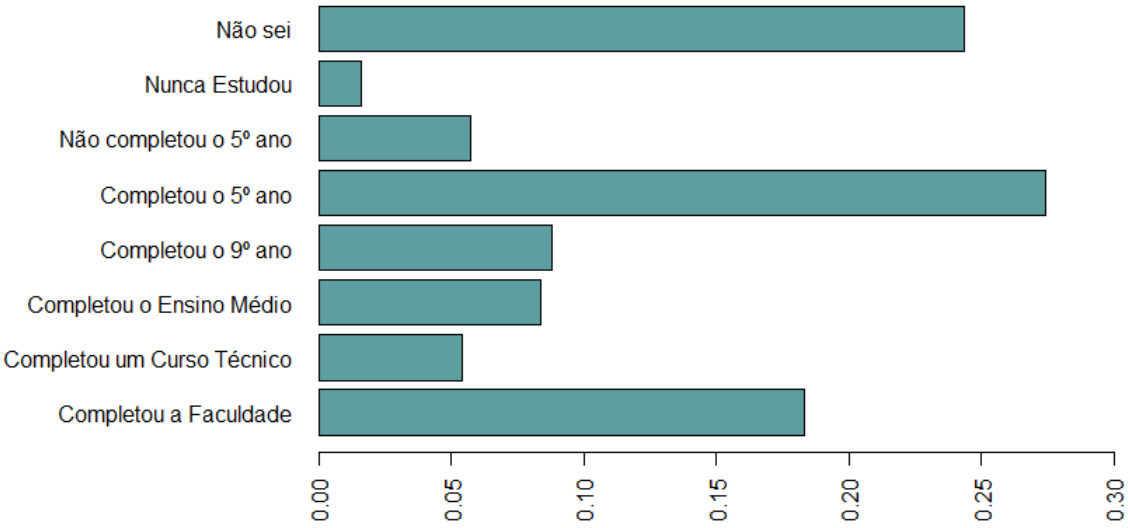


Figura 6: Distribuição amostral da escolaridade dos responsáveis dos respondentes

Por fim, temos, na Figura 6, a distribuição da escolaridade dos pais ou responsáveis. Pela figura, verfica-se uma quantidade relevante de respondentes que não sabiam essa informação ao responder o questionário. Mesmo assim, dentre os

que sabiam, a maioria tem ensino fundamental completo, com proporção consideravelmente maior aos que tem ensino superior completo.

3.2 Fidedignidade

Em uma etapa inicial de análise do questionário sobre o comportamento antissocial, foram calculadas os coeficientes de correlação τ de Kendall item a item do teste, que se resumem na Figura 7. Observa-se que algumas correlações são baixas, com valores abaixo de $0,2$, mas, no geral, as correlações são médias, com valores entre $0,2$ e $0,6$.

	Correlações (τ de Kendall)														
Itens	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
A	1,00														
B	0,45	1,00													
C	0,38	0,43	1,00												
D	0,33	0,35	0,44	1,00											
E	0,27	0,21	0,35	0,36	1,00										
F	0,18	0,19	0,39	0,36	0,38	1,00									
G	0,35	0,33	0,33	0,31	0,22	0,23	1,00								
H	0,35	0,41	0,40	0,28	0,21	0,19	0,38	1,00							
I	0,33	0,35	0,50	0,38	0,41	0,34	0,34	0,40	1,00						
J	0,16	0,17	0,30	0,27	0,36	0,41	0,23	0,24	0,34	1,00					
K	0,33	0,32	0,35	0,31	0,25	0,18	0,38	0,40	0,36	0,30	1,00				
L	0,44	0,37	0,40	0,33	0,23	0,21	0,44	0,42	0,45	0,29	0,49	1,00			
M	0,14	0,15	0,32	0,29	0,42	0,48	0,14	0,21	0,35	0,35	0,18	0,20	1,00		
N	0,18	0,19	0,37	0,34	0,40	0,34	0,18	0,23	0,36	0,43	0,21	0,27	0,57	1,00	
O	0,34	0,30	0,35	0,34	0,29	0,32	0,54	0,36	0,36	0,30	0,39	0,43	0,31	0,34	1,00

Figura 7: Correlação entre os itens do questionário

Para incrementar os resultados, foi utilizado um método de fidedignidade do tipo consistência interna, nesse caso o cálculo do Alfa de Cronbach, para os dados coletados. Como os valores das correlações são altas, espera-se um valor alto para o α.

Com o resultado de $\alpha = 0{,}89$, podemos inferir que o questionário tem um bom nível de fidedignidade. Portanto, o valor de fidedignidade foi considerado ótimo, uma vez que se encontra entre os valores 0,7 e 0,9.

Porém, é importante ressaltar que o α de Cronbach utiliza a correlação de Pearson para o seu cálculo. Como os itens seguem uma escala ordinal, temos que esse estimador de correlação não seria o mais indicado, podendo tornar o estimador do α pouco preciso.

3.3 Teoria de Resposta ao Item

3.3.1 Curva Característica do Item

Utilizou-se, para o ajuste do Modelo de Resposta Gradual, a função "*grm*"encontrada no pacote "*ltm*" (RIZODOPOULOS, 2006), encontrado no *software* R. Com auxílio deste pacote, foram geradas estimativas para os parâmetros do modelo e gráficos que descrevem a informação dos itens e do teste em geral.

Vale ressaltar que, nesse estudo, a "Habilidade" expressada nos gráficos se refere ao nível de comportamento antissocial apresentado pelo respondente. Dessa forma, a curva referente a qualquer categoria se refere à probabilidade de um respondente com aquele nível de comportamento antissocial marcar aquela opção. As categorias variam de 1 a 5, sendo 1 "Nunca" e 5 "Muitas Vezes".

Assim, observamos, na Figura 8 que nenhum dos itens segue a curva característica de um item perfeitamente discriminante (Figura 2), com as 5 categorias tendo, para algum θ, a maior probabilidade de ser marcada. Além disso, muitos itens parecem ser dicotômicos, com probabilidades altas de marcarem 1 ou 5, enquanto as probabilidades de marcar 2, 3 ou 4 são muito baixas.

Ademais, muitas categorias ficaram com probabilidades muito próximas de 0, como nos casos da categoria 4 (curva azul escura) para os itens A e H, por exemplo. Os itens mais próximos de uma alta discriminação são os itens A e B, os quais possuem probabilidades elevadas para as categorias 2 e 4 em intervalos grandes de habilidade, como, por exemplo, um teta entre -1 e 4,5 para o item A.

3.3.2 Curvas de informação dos Itens

Observa-se, pela Figura 9, que a informação dos itens, em geral, é baixa. Com exceção do item 2, que corresponde à curva vermelha maior. Além disso, é fácil observar que vários itens tem sua informação menor que 1. Para melhor visualização da informação dos itens, os gráficos sobre a informação dos itens individuais estão no apêndice deste trabalho.

Este resultado, juntamente com os obtidos nas CCI's, pode significar que uma combinação de categorias deveria ser feita nos itens.

Essa combinação poderia agrupar duas ou mais categorias de resposta em uma só, deixando o item com menos categorias de escolha. Isso diminui a quantidade de parâmetros estimados pela TRI, o que maximiza a precisão, melhorando os resultados e a informação provida pelo teste.

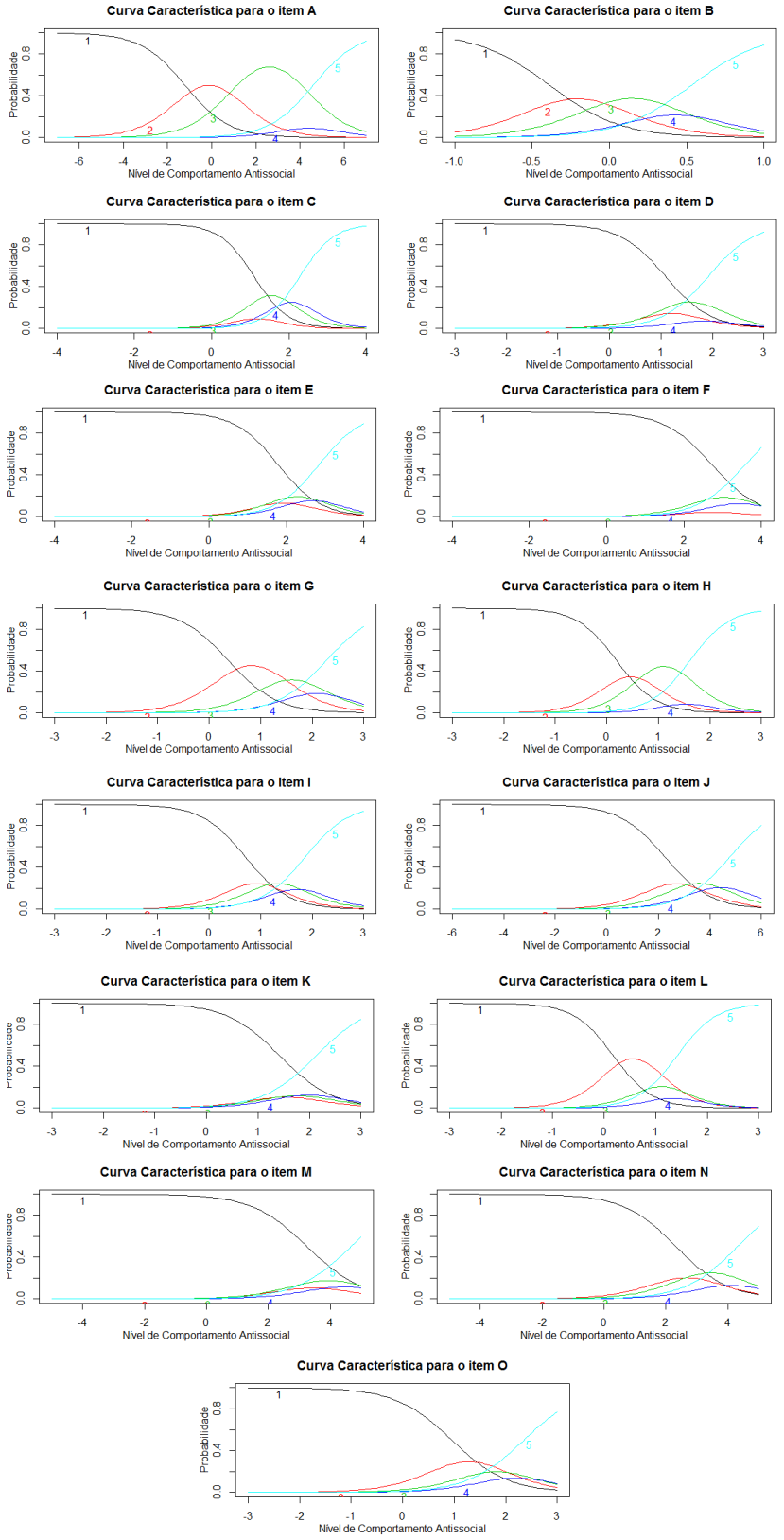


Figura 8: Curvas Características para os 15 itens

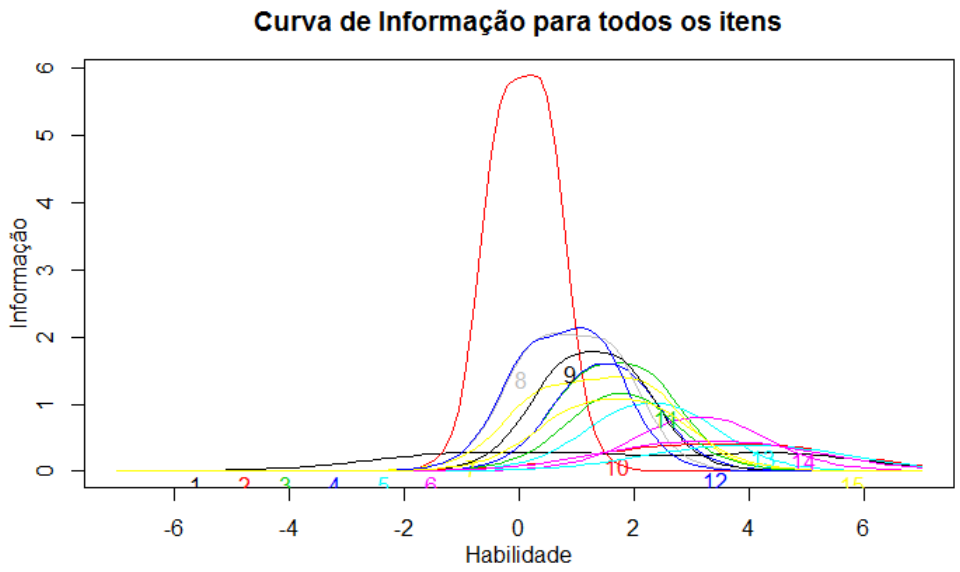


Figura 9: Curvas de informação para os 15 itens

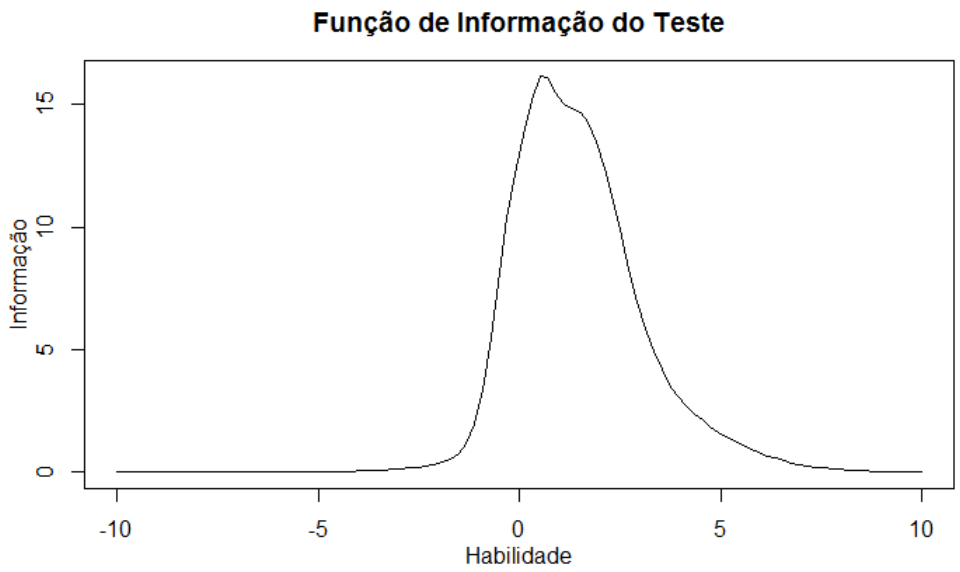


Figura 10: Curvas de informação para o teste

3.3.3 Curva de Informação do Teste

Como pode-se observar pela Figura 10, a curva de informação do teste respeita as curvas de informação dos itens como um todo. Desta forma, a informação do teste só é alta para os valores de habilidade entre -1 e 5. Fora deste intervalo, o teste tem dificuldade de medir tal habilidade.

Além disso, é possível observar que a curva não se resume a uma normal

por uma leve assimetria a esquerda e uma espécie de bico bimodal. Este bico pode ser explicado pelas Funções de Informações dos Itens. Existe uma grande concentração de itens com informação moderada um pouco a direita do item 2.

Desta forma foi feita uma combinação de categorias, de acordo com a Figura 8 e a Figura 11, na qual se encontram todas as frequências de resposta para cada item.Além disso, também foram utilizadas as Curvas Características Operacionais, as quais estão presentes no apêndice deste trabalho. Cada item X foi levado no item X' pela combinação. Por exemplo, o item A foi levado no item A', o item B foi levado no B' e assim por diante.

3.3.4 Curvas Características dos Itens Modificados

A Figura 12 mostra uma adequação melhor das curvas características de cada item. Mesmo em alguns casos, nos quais alguma categoria não tem, para alguma habilidade, a maior probabilidade de ser marcada, as probabilidades agora se encontram bem maiores do que antes da modificação, beirando quase sempre o 0,4.

3.3.5 Curvas de Informação dos Itens

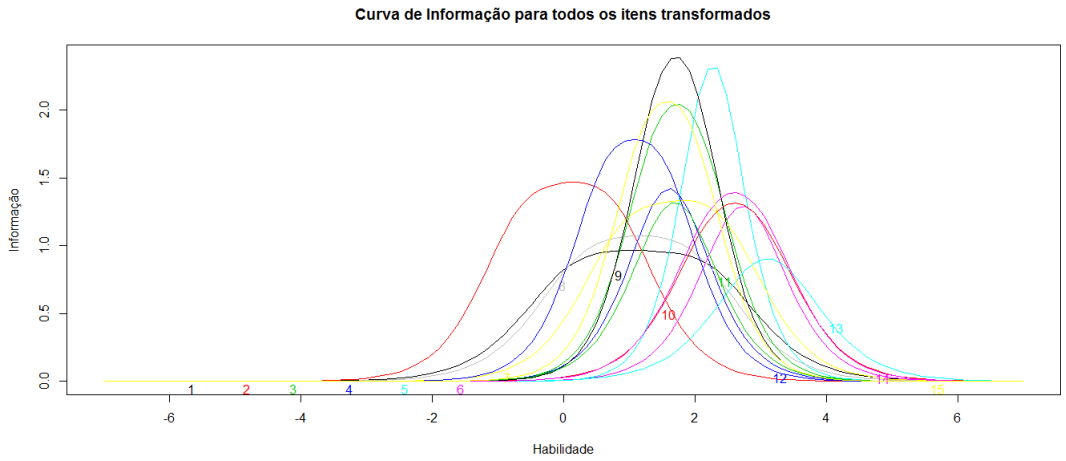


Figura 13: Curvas de informação para os 15 itens modificados

Fica claro, na Figura 13, que os itens agora possuem informação mais bem distribuída. Apenas 3 itens continuam com sua informação próxima a 1, os itens 1, 8 e 13. Além disso, fica claro que o item 2 deixou de concentrar a grande maioria da informação do teste, que agora está mais bem distribuído pelo teste como um todo.

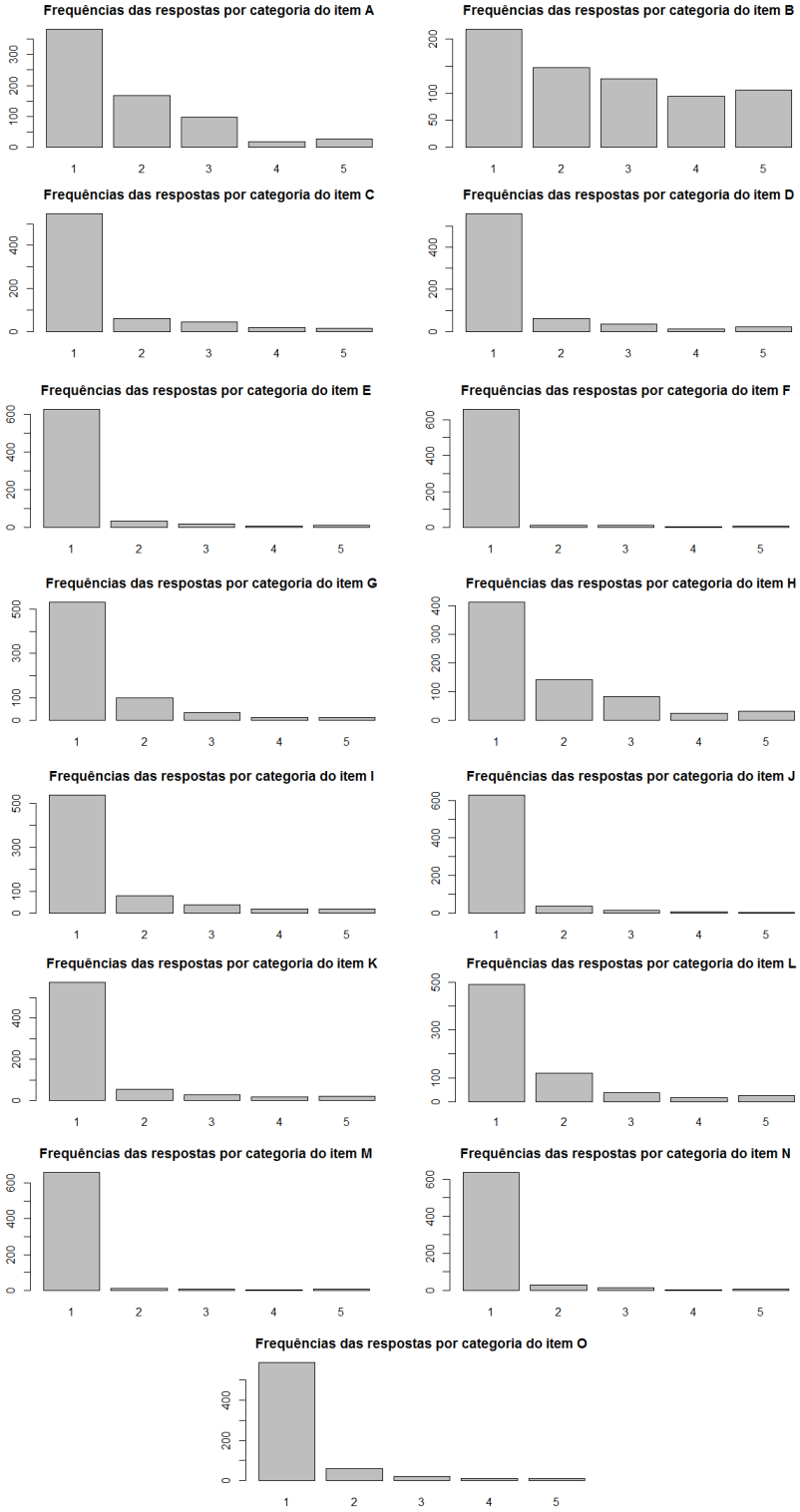


Figura 11: Frequências por categorias em cada item

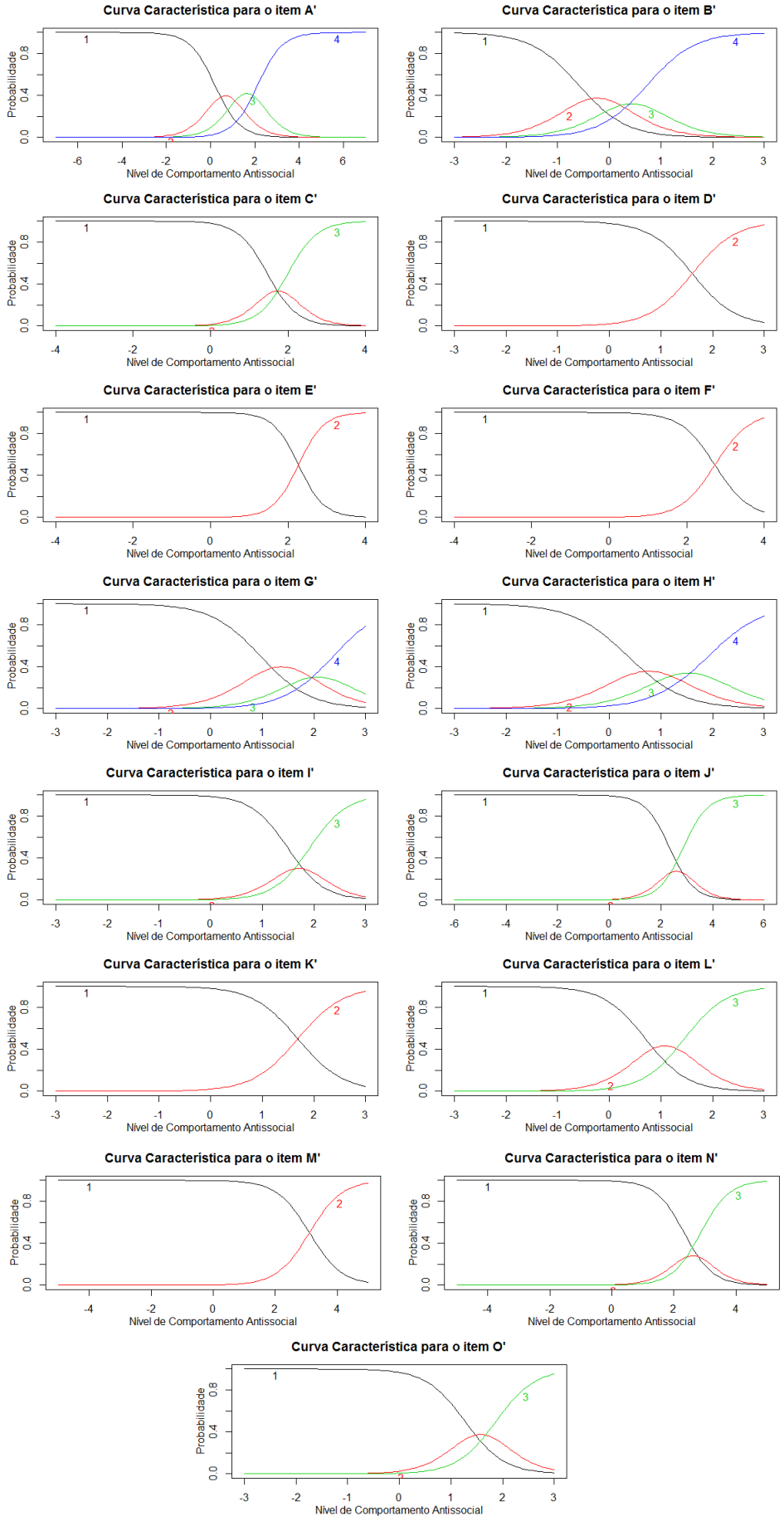


Figura 12: Curvas Características para os 15 itens modificados

Novamente, estão disponíveis, no apêndice, todos os gráficos referentes à informação dos itens separadamente.

3.3.6 Curva de Informação do Teste Modificado

A Figura 14 mostra que a informação do teste cresceu em aproximadamente 3 unidades. A curva ficou mais suave em toda a sua extensão, sem bicos, e a assimetria diminuiu, apesar de ainda existir assimetria à direita. A amplitude de habilidade a qual o teste consegue medir se manteve a mesma, como esperado.

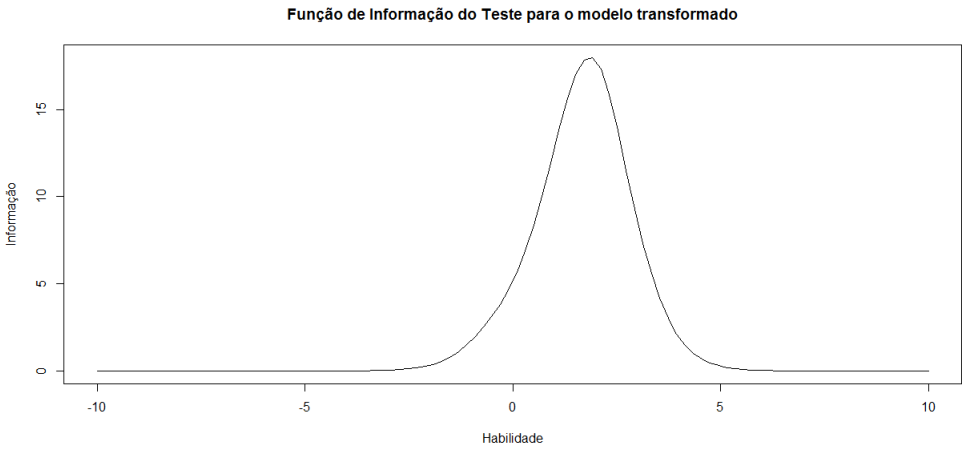


Figura 14: Curvas de informação para o teste modificado

Além disso, ao fazer-se uma comparação dos dois modelos, foram calculados o BIC e o AIC para permitir a comparação entre os modelos modificado e não modificado.

Os dois estimadores mostraram significativa melhora na combinação de categorias com relação à adequação do modelo, tendo o BIC passado de $14181,39$ para 9166.23 e o AIC passado de 13840.17 para 8966.048.

Desta forma, pode-se concluir que o modelo modificado é consideravelmente mais bem ajustado do que o modelo original, demonstrando uma certa razão na combinação de categorias aplicada.

4 CONCLUSÃO

O objetivo principal do estudo foi validar o questionário utilizado por Murta (2017) sobre comportamento antissocial e avaliar os itens deste questionário para, possivelmente, propor mudanças e melhorias. A análise de TRI apresentada

conseguiu êxito ao mostrar a correlação alta entre os itens e o papel que cada um deles teve na avaliação do questionário como um todo.

Os resultados aqui apresentados mostram que a combinação de categorias (página 25 em diante) se mostrou satisfatória e, desta, percebeu-se que o teste como um todo tem dificuldade em estimar o nível de comportamento antissocial para os indivíduos que possuem este nível baixo. Isso se deve à grande quantidade de respostas da categoria "Nunca" em todos os itens, dificultando, assim, a discriminação de quem tem um nível de comportamento antissocial baixo ou muito baixo.

Além disso, mostrou-se que os itens com menor informação, ou seja, os itens 1, 6 e 13 são os maiores candidatos a saírem do questionário. Mesmo assim, nenhum item obteve coeficientes insatisfatórios, ou seja, uma discriminação muito baixa (menor que 1), na estimação de parâmetros da TRI, o que diz não ser estritamente necessária a retirada de nenhum item.

Ademais, o teste se mostrou muito eficiente em estimar o nível de comportamento antissocial no intervalo de 1 a 3 do traço latente. Isso significa que, se o indivíduo tiver um comportamento antissocial dentro deste intervalo, o teste se mostrou muito bom em identificar e estimar este construto.

Por fim, vale a pena ressaltar que, ao fazer a combinação das categorias dos itens, a informação do teste aumentou. Porém, na contramão disso, a informação dos itens se aproximou de uma média. Desta forma, itens como o 2 ou o 4 tiveram sua informação significativamente diminuída.

O presente estudo apresentou, portanto, um estudo detalhado da qualidade dos itens presentes no questionário estudado. Os dados apresentados são relevantes para a certeza de boa coleta de informações relacionadas ao comportamento antissocial entre os jovens. Ainda assim, se faz necessário aprimorar o questionário e reaplicá-lo, para comparar se as melhorias propostas aqui de fato melhorariam os resultados obtidos em campo.

Referências

aaaa.

ALMEIDA, D.; REIS DOS SANTOS, M. A.; COSTA, A. F. B. Aplicação do coeficiente alfa de Cronbach nos resultados de um questionário para avaliação de desempenho da saúde pública. **Encontro Nacional de Engenharia de Produção**, São Carlos, p. 1-12, 2010.

ANASTASI, A.; URBINA, S. **Psychological Testing**. 7. ed. Michigan: Prentice Hall, p. 721, 1997.

ANDRADE, D. F., TAVARES, H. R. e VALLE, R. C. **Teoria da resposta ao item**: conceitos e aplicações. Associação Brasileira de Estatística, Caxambu, 2000.

BORDIN, I. A.; OFFORD, D. R. Transtorno da Conduta e Comportamento Antisocial, **Revista Brasileira de Psiquiatria**, vol.22, s.2, São Paulo: Scielo, Dezembro de 2000.

BRINGAS, C. et al. La conducta antisocial en adolescentes no conflictivos: Adaptación del Inventario de Conductas Antisociales (ICA). **Revista Electrónica de Metodología Aplicada**, 11(2), 1-10. Recuperado en 10 diciembro 2009, de http://www.psico.uniovi.es/rema/v11n2/antisoc.pdf, 2006.

COUTO, G.; PRIMI, R. **Teoria de resposta ao item (TRI)**: Conceitos elementares dos modelos para itens dicotômicos. Bol. psicol, São Paulo , v. 61, n. 134, p. 1-15, jun. 2011 . Disponível em ⟨⟨http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0006-59432011000100002&lng=pt&nrm=iso⟩⟩. acessos em 01 nov. 2017.

DANCEY, C.P.; REIDY, J. **Estatística Sem Matemática Para Psicologia** 5. ed. Porto Alegre: Penso, 2013.

EMILIANO, P. A. et al. "Critérios de Informação de Akaike Versus Bayesiano: Análise Comparativa."**Simpósio Nacional de Probabilidade e Estatística** - 19º SINAPE. São Pedro-SP, 2010.

FREITAS, E.L. **Alguns Aspectos da Linguagem Científica Sitientibus**, *Feira de Santana*, n. 12, p.101-112, 1994.

HOGAN, T. P. **Introdução À Prática de Testes Psicológicos**. Rio de Janeiro: LTC, 540 p. 2006.

JONES, P. W. et al. ***European Respiratory Journal*** *Development and first validation of the COPD Assessment Test*, [S. l.] Vol. 34, p. 648–654, 2009.

MAROCO, J.; GARCIA-MARQUES, T. Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas? **Laboratório de Psicologia**, Braga, v. 4, n. 1, p. 65-90, 2006.

MARTINS, G.A. **Sobre Confiabilidade e Validade RBGN**, São Paulo, Vol. 8,

n. 20, p.1-12, jan./abr. 2006.

MURTA, S. G. et al. **Avaliação da efetividade, qualidade de implementação e validade social do programa fortalecendo famílias**. Relatório não publicado.

PACHECO, J. T. B.**A construção do comportamento anti-social em adolescentes autores de atos infracionais** : uma análise a partir das práticas educativas e dos estilos parentais. Tese de Doutorado pela UFRS. 2004.

PASQUALI, L. **Psicometria** *Rev Esc Enferm* USP, Vol. 43, n. esp, p. 992-9, 2009.

RIZODOPOULOS, D. ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses, **Journal of Statistical Software**, 17 (5), 1-25. URL http://www.jstatsoft.org/v17/i05/, 2006.

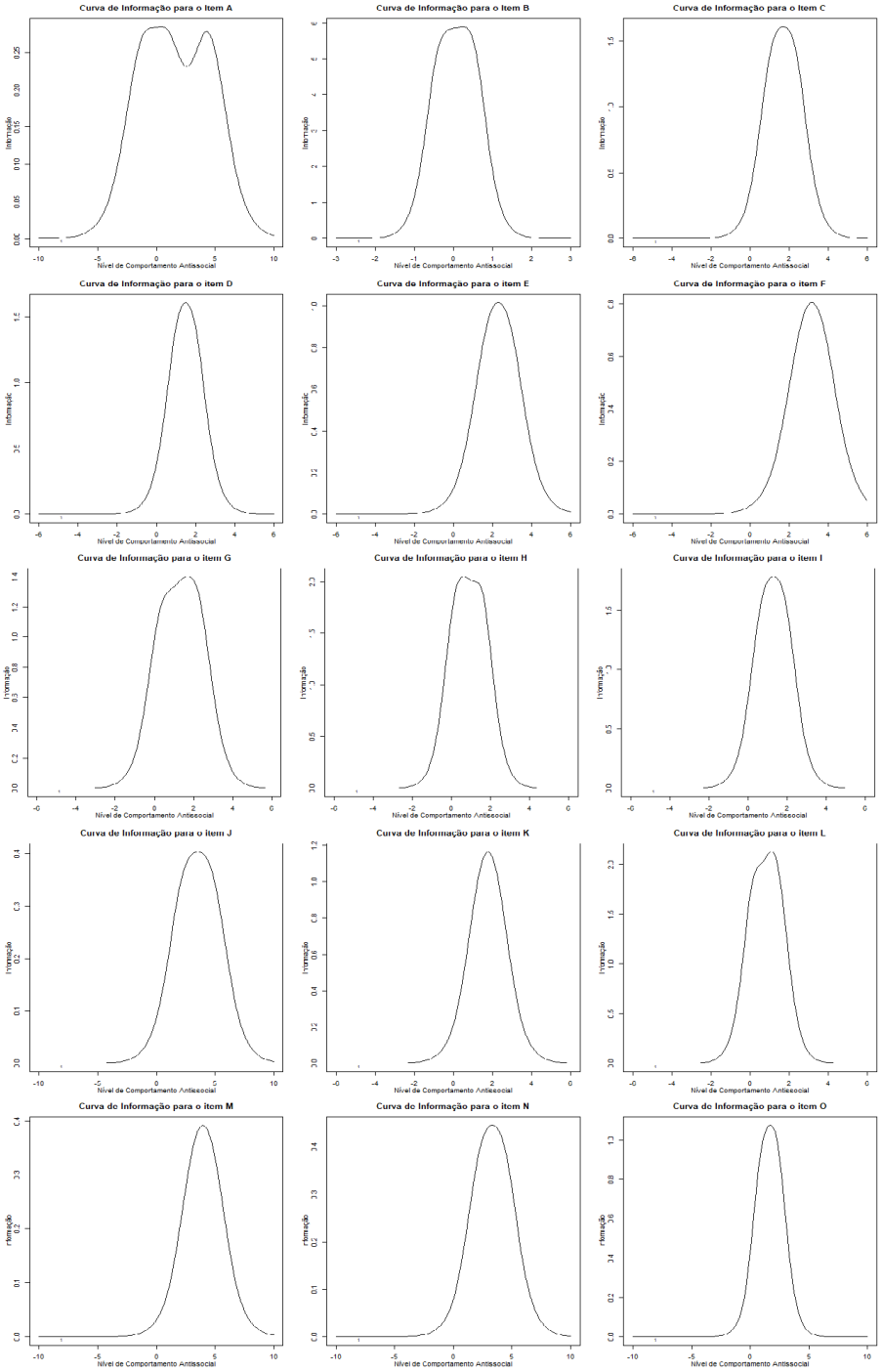
SABOIA, R. T. N. **Uso da teoria de resposta ao item para compreender diferenças sociais de gênero**. 2013. 50 f., il. Monografia (Bacharelado em Estatística)—Universidade de Brasília, Brasília, 2013.

SAMEJIMA, F. A. . *Estimation of latent ability using a response pattern of graded scores. Psychometric Monograph*, p.17, 1969.

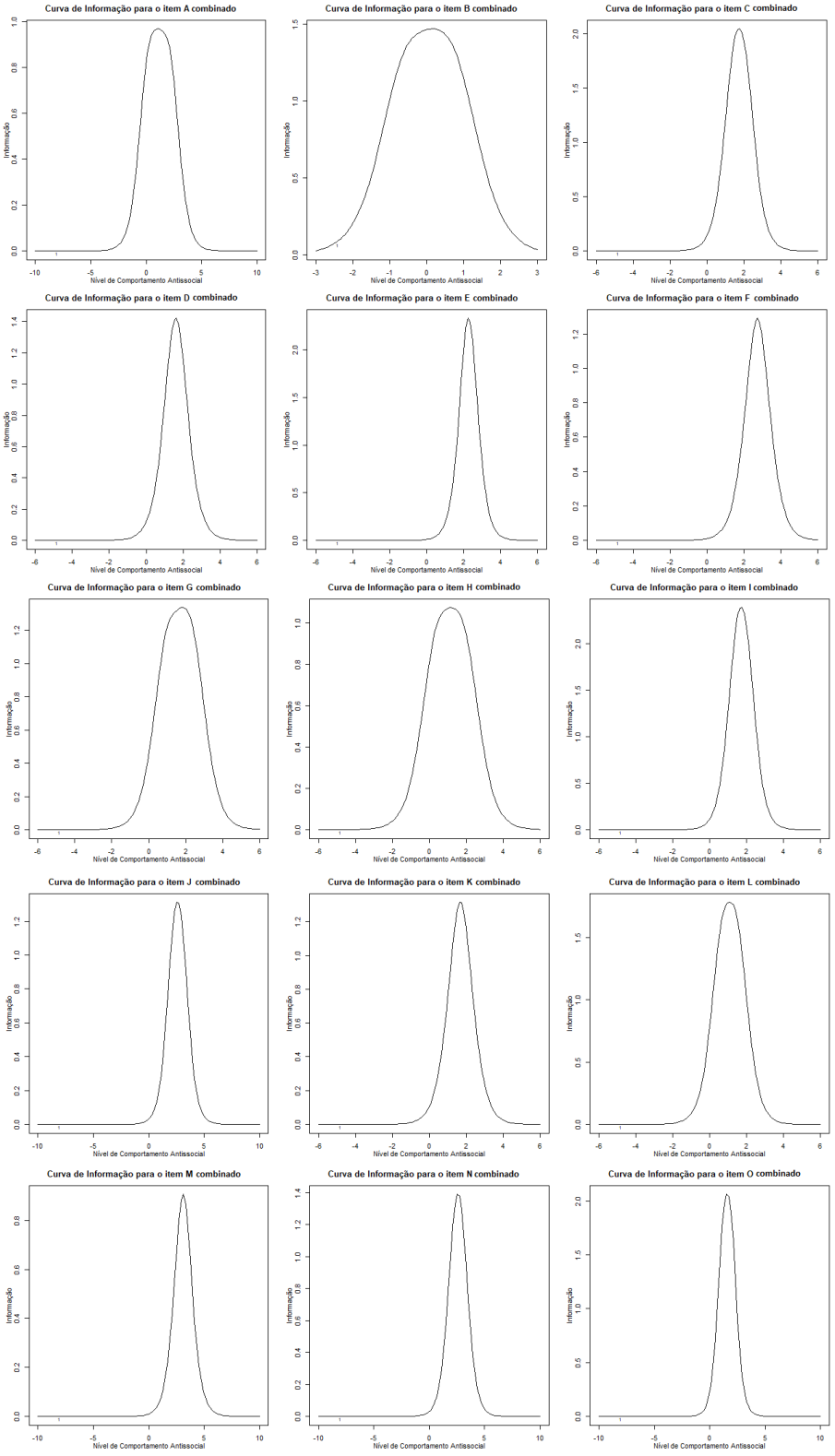
SILVA NETO, A. C. P. **Fidedignidade do sistema compreensivo do Rorschach**: revisão e estudo da estabilidade temporal em adultos da cidade de São Paulo. 2008. Tese (Doutorado em Psicologia Escolar e do Desenvolvimento Humano) - Instituto de Psicologia, Universidade de São Paulo, São Paulo, doi:10.11606/T.47.2008.tde-30112009-091310. Acesso em: 2017-06-16, 2008.

5 APÊNDICE

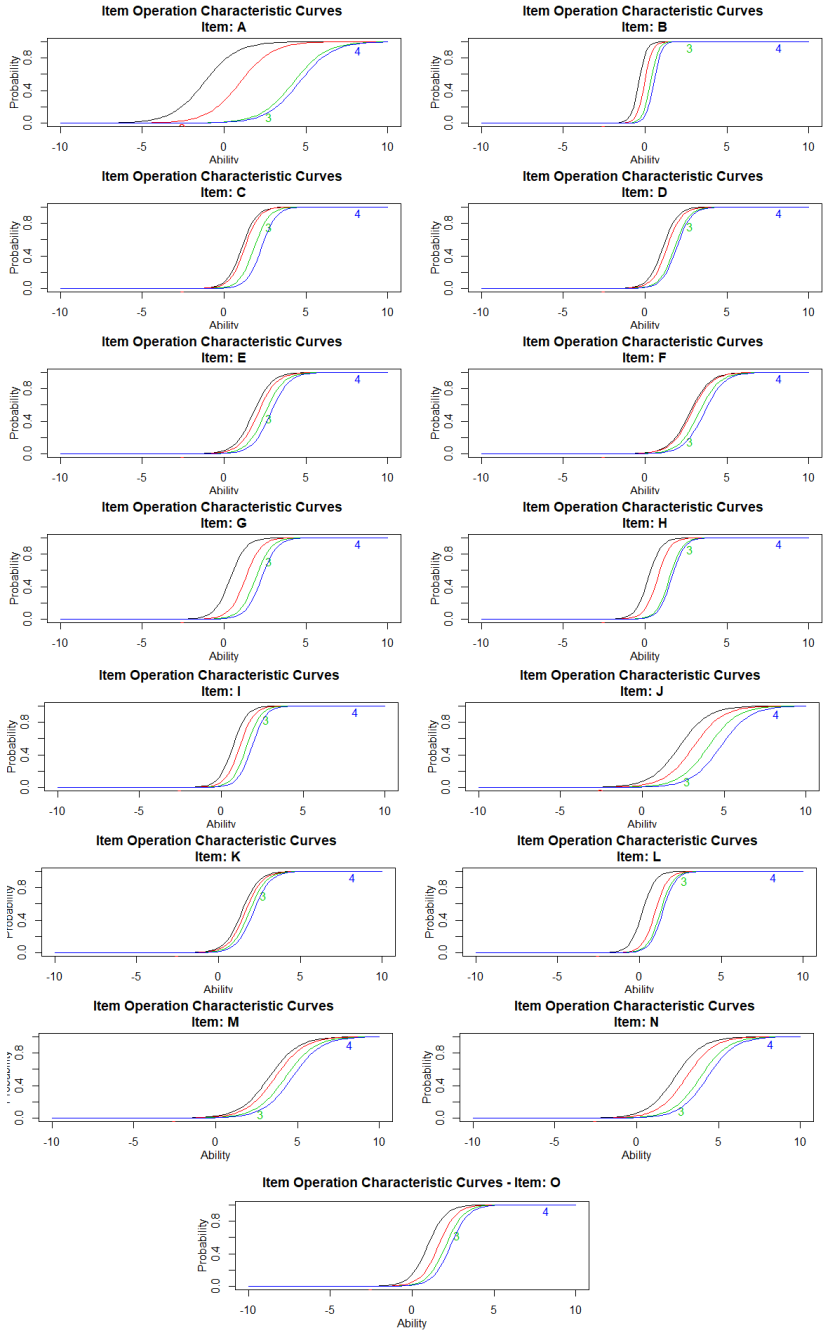
5.1 Informação dos Itens Separadamente



5.2 Informação dos Itens modificados Separadamente



5.3 Curvas Operacionais dos Itens Separadamente



5.4 Código em R

```
#Funções úteis para o código
asnumeric.factor = function(x){
  x = gsub(",", ".", x)
  x = as.numeric(as.character(gsub(",", ".", x)))
```

```
  x
}

re_levels <-
  function(dfanti) {
    if (is.factor(dfanti))  levels(dfanti) <- c(levels(
       dfanti), "0")
    dfanti
  }

#Instalando pacotes

if(!require(ggplot2)) {
  install.packages("ggplot2");
  require(ggplot2)
}
if(!require(foreign)) {
  install.packages("foreign");
  require(foreign)
}
if(!require(plotrix)) {
  install.packages("plotrix");
  require(plotrix)
}

if(!require(data.table)) {
  install.packages("data.table");
  library(data.table)
}

if(!require(ltm)) {
  install.packages("ltm"); require(ltm)}


#Importar os dados
setwd("C:/Users/rodms/OneDrive/Documentos/tcc")
dados = read.spss("dadosSpss.sav",
```

```
                    use.value.labels = TRUE,
                    to.data.frame = T,
                    max.value.labels = Inf,
                    trim.factor.names = FALSE,
                    trim_values = TRUE,
                    reencode = NA)

##### Análise Descritiva

#Sexo
descritivo = dados[,c(222:237)]
sexo = as.numeric(descritivo[,1])
s <- ggplot(descritivo[!is.na(descritivo$Q53),], aes(x="Sexo
    ", fill=factor(Q53)))
bpSexo <- s +
  geom_bar(position = position_dodge(width = 1.1), width =
      0.7)+
  scale_fill_discrete(name = "Sexo")
pieSexo <- s + geom_bar(position = 'stack', width = 1) +
    coord_polar(theta = "y", start = 0)+theme(title=
    element_blank(),axis.text.x = element_blank()) +
    scale_fill_manual(values=c("#8B0000", "#0000FF"))

#Idade
idade = descritivo[,4]
summary(idade)
sd(idade)
tabela1 = table(as.factor(idade))
tabela1.prop = prop.table(tabela1)

#Escolaridade
escolaridade = as.character(descritivo$Q57)
escolaridade[escolaridade == "Não sei"] = "Outros"
tabela2 = table(escolaridade)
tabela2.prop = prop.table(tabela2)
op = par(mar=c(3,12,4,3))
labels = c("Não sei", "Nunca Estudou", "Não completou o 5
```

```
    ano", "Completou o 5 ano", "Completou o 9 ano", "
    Completou o Ensino Médio", "Completou um Curso Técnico",
    "Completou a Faculdade")
barplot(tabela2.prop,col = "#5F9EA0",names.arg = labels,las
    =2, horiz = T)
rm(op)


##### TRI

#Separando os dados
dfanti = dados[,c(183:197)]

#Ajustando os 0's e NA's
dfanti <- sapply(dfanti,re_levels)
dfanti[is.na(dfanti)] <-  0
dfanti = data.frame(dfanti)

#transformando as questões para numérico de 1 a 5
source("vetoresQuestoes.r", encoding = "UTF8")

#Colocando as questões numéricas em um dataFrame
df = data.table(A=Q44a, B=Q44b, C=Q44c, D=Q44d, E=Q44e, F=
    Q44f,
                G=Q44g, H=Q44h, I=Q44i, J=Q44j, K=Q44k, L=
                    Q44l,
                M=Q44m, N=Q44n, O=Q44o)

#Retirando os NA'S e Missing Values que foram combinados
    para numérico
semMaior = df[, lapply(.SD, function(x) ifelse(as.character(
    x) == "99", NA, x))]
semMaior = semMaior[, lapply(.SD, function(x) ifelse(as.
    character(x) == "98", NA, x))]

#Calculando a TRI
tri = grm(semMaior)
```

```
tri$coefficients #Ver os coeficientes da TRI

par(mfrow=c(1,1))
#FUNÇÃO DE INFORMAÇÃO DO TESTE
plot(tri, type = "IIC",items = 0,zrange = c(-10,10), main =
    "Função de Informação do Teste", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")

#Ajustes de gráficos do R
par(mfrow=c(1,3))
title(line = 1.2)
par(mar=c(4,3,3,2))
par(mgp = c(2, 1, 0))

#Curvas Características das Categorias dos Itens
plot(tri, type = "ICC",items = 1,zrange = c(-7,7), main = "
    Curva Característica para o item A",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 2,zrange = c(-1,1), main = "
    Curva Característica para o item B",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 3,zrange = c(-4,4), main = "
    Curva Característica para o item C",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 4,zrange = c(-3,3), main = "
    Curva Característica para o item D",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 5,zrange = c(-4,4), main = "
    Curva Característica para o item E",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 6,zrange = c(-4,4), main = "
    Curva Característica para o item F",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 7,zrange = c(-3,3), main = "
    Curva Característica para o item G",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 8,zrange = c(-3,3), main = "
```

```
    Curva Característica para o item H",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 9,zrange = c(-3,3), main = "
    Curva Característica para o item I",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 10,zrange = c(-6,6), main = "
    Curva Característica para o item J",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 11,zrange = c(-3,3), main = "
    Curva Característica para o item K",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 12,zrange = c(-3,3), main = "
    Curva Característica para o item L",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 13,zrange = c(-5,5), main = "
    Curva Característica para o item M",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 14,zrange = c(-5,5), main = "
    Curva Característica para o item N",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri, type = "ICC",items = 15,zrange = c(-3,3), main = "
    Curva Característica para o item O",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")


#CURVAS DE INFORMAÇÃO DOS ITENS JUNTOS
plot(tri, type = "IIC",zrange = c(-7,7), main = "Curva de
    Informação para todos os itens", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")

#CURVAS DE INFORMAÇÃO DOS ITENS SEPARADOS
plot(tri, type = "IIC",items = 1,zrange = c(-10,10), main =
    "Curva de Informação para o item A", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 2,zrange = c(-3,3), main = "
    Curva de Informação para o item B", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
```

```
plot(tri, type = "IIC",items = 3,zrange = c(-6,6), main = "
    Curva de Informação para o item C", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 4,zrange = c(-6,6), main = "
    Curva de Informação para o item D", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 5,zrange = c(-6,6), main = "
    Curva de Informação para o item E", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 6,zrange = c(-6,6), main = "
    Curva de Informação para o item F", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 7,zrange = c(-6,6), main = "
    Curva de Informação para o item G", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 8,zrange = c(-6,6), main = "
    Curva de Informação para o item H", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 9,zrange = c(-6,6), main = "
    Curva de Informação para o item I", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 10,zrange = c(-10,10), main =
     "Curva de Informação para o item J", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 11,zrange = c(-6,6), main = "
    Curva de Informação para o item K", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 12,zrange = c(-6,6), main = "
    Curva de Informação para o item L", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 13,zrange = c(-10,10), main =
     "Curva de Informação para o item M", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 14,zrange = c(-10,10), main =
     "Curva de Informação para o item N", xlab = "Nível de
    Comportamento Antissocial", ylab = "Informação")
plot(tri, type = "IIC",items = 15,zrange = c(-10,10), main =
```

```
    "Curva de Informação para o item O", xlab = "Nível de
   Comportamento Antissocial", ylab = "Informação")


# annot é usado para tirar ou colocar os números das
   categorias.
# legend = TRUE, cx = "left",ncol = 2 é usado para colocar
   legendas, em alternativa ao annot.

#Curvas características operacionais dos itens. u para upper
    (crescente), l para lower (decrescente)
plot(tri, type = 'OCCu',items = 1, zrange = c(-10,10))
plot(tri, type = 'OCCl')

#information() calcula exatamente a informação abaixo da
   curva em um intervalo para cada item.

#Calculando os estimadores de fidedignidade
require(Lambda4)
guttman(semMaior, missing = "complete", standardize = FALSE)
require(psych)
alpha(semMaior,na.rm=T)

#Calculando os estimadores de comparadores de modelos (AIC E
    BIC)
AIC(tri2)
BIC(tri2)


#Análise exploratória descritiva dos itens individualmente.

barplot(table(semMaior$A), main = "Frequências das respostas
    por categoria do item A")
plot(tri, type = 'OCCu',items = 1, zrange = c(-10,10))

barplot(table(semMaior$B), main = "Frequências das respostas
    por categoria do item B")
```

```
plot(tri, type = 'OCCu',items = 2, zrange = c(-10,10))

barplot(table(semMaior$C), main = "Frequências das respostas
    por categoria do item C")
barplot(table(semMaior$Q44c))
plot(tri, type = 'OCCu',items = 3, zrange = c(-10,10))

barplot(table(semMaior$D), main = "Frequências das respostas
    por categoria do item D")
plot(tri, type = 'OCCu',items = 4, zrange = c(-10,10))

barplot(table(semMaior$E), main = "Frequências das respostas
    por categoria do item E")
plot(tri, type = 'OCCu',items = 5, zrange = c(-10,10))

barplot(table(semMaior$F), main = "Frequências das respostas
    por categoria do item F")
plot(tri, type = 'OCCu',items = 6, zrange = c(-10,10))

barplot(table(semMaior$G), main = "Frequências das respostas
    por categoria do item G")
plot(tri, type = 'OCCu',items = 7, zrange = c(-10,10))

barplot(table(semMaior$H), main = "Frequências das respostas
    por categoria do item H")
plot(tri, type = 'OCCu',items = 8, zrange = c(-10,10))

barplot(table(semMaior$I), main = "Frequências das respostas
    por categoria do item I")
plot(tri, type = 'OCCu',items = 9, zrange = c(-10,10))

barplot(table(semMaior$J), main = "Frequências das respostas
    por categoria do item J")
plot(tri, type = 'OCCu',items = 10, zrange = c(-10,10))

barplot(table(semMaior$K), main = "Frequências das respostas
    por categoria do item K")
```

```
plot(tri, type = 'OCCu',items = 11, zrange = c(-10,10))

barplot(table(semMaior$L), main = "Frequências das respostas
    por categoria do item L", ylim = c(0,500))
plot(tri, type = 'OCCu',items = 12, zrange = c(-10,10))

barplot(table(semMaior$M), main = "Frequências das respostas
    por categoria do item M")
plot(tri, type = 'OCCu',items = 13, zrange = c(-10,10))

barplot(table(semMaior$N), main = "Frequências das respostas
    por categoria do item N")
plot(tri, type = 'OCCu',items = 14, zrange = c(-10,10))

barplot(table(semMaior$O), main = "Frequências das respostas
    por categoria do item O")
plot(tri, type = 'OCCu',items = 15, zrange = c(-10,10))

#Buscar a melhor configuração de categorias para cada item.

semMaior2 = semMaior

semMaior2$A[semMaior2$A == 5] = 4

semMaior2$B[semMaior2$B == 5] = 4

semMaior2$C[semMaior2$C == 2] = 1
semMaior2$C[semMaior2$C == 3] = 2
semMaior2$C[semMaior2$C == 4] = 3
semMaior2$C[semMaior2$C == 5] = 3

semMaior2$D[semMaior2$D == 2] = 1
semMaior2$D[semMaior2$D == 3] = 2
semMaior2$D[semMaior2$D == 4] = 2
semMaior2$D[semMaior2$D == 5] = 2

semMaior2$E[semMaior2$E == 2] = 1
```

semMaior2$E [semMaior2$E == 3] = 1
semMaior2$E [semMaior2$E == 4] = 2
semMaior2$E [semMaior2$E == 5] = 2

semMaior2$F [semMaior2$F == 2] = 1
semMaior2$F [semMaior2$F == 3] = 1
semMaior2$F [semMaior2$F == 4] = 2
semMaior2$F [semMaior2$F == 5] = 2

semMaior2$G [semMaior2$G == 5] = 4

semMaior2$H [semMaior2$H == 5] = 4

semMaior2$I [semMaior2$I == 2] = 1
semMaior2$I [semMaior2$I == 3] = 2
semMaior2$I [semMaior2$I == 4] = 3
semMaior2$I [semMaior2$I == 5] = 3

semMaior2$J [semMaior2$J == 2] = 1
semMaior2$J [semMaior2$J == 3] = 2
semMaior2$J [semMaior2$J == 4] = 3
semMaior2$J [semMaior2$J == 5] = 3

semMaior2$K [semMaior2$K == 2] = 1
semMaior2$K [semMaior2$K == 3] = 2
semMaior2$K [semMaior2$K == 4] = 2
semMaior2$K [semMaior2$K == 5] = 2

semMaior2$L [semMaior2$L == 4] = 3
semMaior2$L [semMaior2$L == 5] = 3

semMaior2$M [semMaior2$M == 2] = 1
semMaior2$M [semMaior2$M == 3] = 1
semMaior2$M [semMaior2$M == 4] = 2
semMaior2$M [semMaior2$M == 5] = 2

semMaior2$N [semMaior2$N == 2] = 1

```
semMaior2$N[semMaior2$N == 3] = 2
semMaior2$N[semMaior2$N == 4] = 3
semMaior2$N[semMaior2$N == 5] = 3

semMaior2$O[semMaior2$O == 4] = 3
semMaior2$O[semMaior2$O == 5] = 3

#Refazendo a TRI
tri2 = grm(semMaior2)
tri2$coefficients #Ver os coeficientes da nova TRI

#FUNÇÃO DE INFORMAÇÃO DO TESTE
plot(tri2, type = "IIC",items = 0,zrange = c(-10,10), main =
    "Função de Informação do Teste para o modelo combinado",
    xlab = "Nível de Comportamento Antissocial", ylab = "
   Informação")

#Ajustes de gráficos do R
par(mfrow=c(1,1))
title(line = 1.2)
par(mar=c(4,3,3,2))
par(mgp = c(2, 1, 0))

#Curvas Características das Categorias dos Itens
plot(tri2, type = "ICC",items = 1,zrange = c(-7,7), main = "
   Curva Característica para o item A'",xlab = "Nível de
   Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 2,zrange = c(-3,3), main = "
   Curva Característica para o item B'",xlab = "Nível de
   Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 3,zrange = c(-4,4), main = "
   Curva Característica para o item C'",xlab = "Nível de
   Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 4,zrange = c(-3,3), main = "
   Curva Característica para o item D'",xlab = "Nível de
   Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 5,zrange = c(-4,4), main = "
```

```
    Curva Característica para o item E'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 6,zrange = c(-4,4), main = "
    Curva Característica para o item F'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 7,zrange = c(-3,3), main = "
    Curva Característica para o item G'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 8,zrange = c(-3,3), main = "
    Curva Característica para o item H'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 9,zrange = c(-3,3), main = "
    Curva Característica para o item I'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 10,zrange = c(-6,6), main =
    "Curva Característica para o item J'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 11,zrange = c(-3,3), main =
    "Curva Característica para o item K'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 12,zrange = c(-3,3), main =
    "Curva Característica para o item L'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 13,zrange = c(-5,5), main =
    "Curva Característica para o item M'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 14,zrange = c(-5,5), main =
    "Curva Característica para o item N'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")
plot(tri2, type = "ICC",items = 15,zrange = c(-3,3), main =
    "Curva Característica para o item O'",xlab = "Nível de
    Comportamento Antissocial", ylab = "Probabilidade")


#CURVAS DE INFORMAÇÃO DOS ITENS JUNTOS
plot(tri2, type = "IIC",zrange = c(-7,7), main = "Curva de
    Informação para todos os itens combinados", xlab = "Nível
```

```
    de Comportamento Antissocial", ylab = "Informação")

#CURVAS DE INFORMAÇÃO DOS ITENS SEPARADOS
plot(tri2, type = "IIC",items = 1,zrange = c(-10,10), main =
    "Curva de Informação para o item A combinado", xlab = "N
   ível de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 2,zrange = c(-3,3), main = "
   Curva de Informação para o item B combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 3,zrange = c(-6,6), main = "
   Curva de Informação para o item C combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 4,zrange = c(-6,6), main = "
   Curva de Informação para o item D combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 5,zrange = c(-6,6), main = "
   Curva de Informação para o item E combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 6,zrange = c(-6,6), main = "
   Curva de Informação para o item F combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 7,zrange = c(-6,6), main = "
   Curva de Informação para o item G combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 8,zrange = c(-6,6), main = "
   Curva de Informação para o item H combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 9,zrange = c(-6,6), main = "
   Curva de Informação para o item I combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 10,zrange = c(-10,10), main
   = "Curva de Informação para o item J combinado", xlab = "
   Nível de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 11,zrange = c(-6,6), main =
    "Curva de Informação para o item K combinado", xlab = "Ní
   vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 12,zrange = c(-6,6), main =
```

```
    "Curva de Informação para o item L combinado", xlab = "Ní
    vel de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 13,zrange = c(-10,10), main
    = "Curva de Informação para o item M combinado", xlab = "
    Nível de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 14,zrange = c(-10,10), main
    = "Curva de Informação para o item N combinado", xlab = "
    Nível de Comportamento Antissocial", ylab = "Informação")
plot(tri2, type = "IIC",items = 15,zrange = c(-10,10), main
    = "Curva de Informação para o item O combinado", xlab = "
    Nível de Comportamento Antissocial", ylab = "Informação")

#Cálculo da correlação entre os itens antes da TRI
correlacoes = cor(df)
write.csv(correlacoes, "correlacoes.csv" ,row.names = T)

#Calculando os estimadores de comparadores de modelos (AIC E
    BIC)
AIC(tri2)
BIC(tri2)
```