

# **A TRAJETÓRIA DE UMA MEDIDA PARA CLIMA ESCOLAR “SENSO DE SEGURANÇA” NA REDE ESTADUAL DE SÃO PAULO ENTRE 2008-2013 E 2018-2019: uma análise do clima escolar pela percepção dos estudantes dos anos finais do Ensino Fundamental e 3ª série do Ensino Médio**

Thiago Fernando Ferreira Costa\*

## **RESUMO**

Este artigo tem por objetivo elaborar uma medida para o clima escolar “senso de segurança” para a Rede de Ensino do Estado de São Paulo, a partir dos dados divulgados dos questionários entre 2008 a 2013 e 2018 a 2019. A elaboração desta medida exigiu que fossem feitas análises de dimensionalidade através da Análise Fatorial Exploratória (AFE) e uma Análise Fatorial de Informação Completa (AFIC), utilizando para esse fim, a Teoria da Resposta ao Item (TRI). Para a constituição da medida, o modelo de Samejima da TRI através do pacote mirt da linguagem R foi utilizado, além da função para análise de múltiplos grupos envolvendo todos os anos de aplicação. Com a definição da medida, as categorias foram ancoradas, possibilitando a interpretação pedagógica da escala. Estes procedimentos possibilitaram que a trajetória da medida do clima escolar “senso de segurança” fosse constituída e interpretada.

Palavras-chave: Clima escolar; Teoria da Resposta ao Item; Análise Fatorial; Análise Fatorial de Informação Completa; Escala Senso de Segurança.

## **1 INTRODUÇÃO**

A prática em torno das avaliações externas em larga escala no Brasil, tem demonstrado que a aplicação de questionários de contexto se faz necessária, pois a compreensão do contexto em que essas avaliações são aplicadas é de suma importância, principalmente, para que injustiças em relação à medida, possam ser dirimidas. No caso dos questionários de contexto do Sistema de Avaliação do Rendimento do Estado de São Paulo (Saresp), os instrumentos passaram a ser aplicados em 2008, com a participação dos estudantes da 2ª, 4ª, 6ª e 8ª séries do Ensino Fundamental e da 3ª série do Ensino Médio, os professores de língua portuguesa, matemática, ciências (física, química e biologia no Médio), diretores e supervisores de escola. Todavia, esses questionários foram aplicados até o ano de 2013, retornando em 2018 e se mantendo a aplicação em 2019.

Na última década houve um grande avanço nas técnicas estatísticas e de análise dos questionários, com a introdução dos Fatores Associados<sup>1</sup>, o contexto das avaliações externas passou a servir de base na compreensão dos principais fatores que interferem nas proficiências estimadas. Esse entendimento foi fundamental para que os questionários passassem a ocupar um papel de destaque nas avaliações.

Entretanto, os questionários aplicados pelo Saeb, por exemplo, não divulgam estudos que envolvam os Fatores Associados, logo, permanecendo na lógica do questionário descritivo, que pouco contribui com o entendimento dos fatores que afetam nas medidas de proficiência. Contudo, muitas avaliações estaduais e municipais, realizam esses estudos, mas, acabam divulgando somente em sua própria rede, o que inviabiliza o uso de forma mais geral por parte dos pesquisadores, jornalistas e sociedade em geral.

\* Matemático pela Universidade Bandeirante de São Paulo, Pedagogo pela Universidade de São Paulo, Mestrando em Educação pela Universidade de São Paulo. E-mail: thiago.costa@usp.br.

<sup>1</sup> Os Fatores Associados de maior destaque no Saresp, são: Nível Socioeconômico (NSE) das famílias dos estudantes, Nível de Qualidade da Escola na percepção do estudante, Nível de Bullying sofrido pelo estudante, Nível do Relacionamento Escolar na percepção do próprio e Nível de Participação dos Pais nos Estudos dos filhos.

No entanto, a maneira como os questionários passaram a ser constituídos, proporcionou a definição de construtos mais específicos ao contexto escolar, como por exemplo, o clima escolar. O Instituto Nacional de Estudos e Pesquisa Educacionais Anísio Teixeira (Inep) divulgou no documento LEIA-ME Saeb 2005, as informações sobre o Sistema de Avaliação da Educação Básica (Saeb) e os Questionários de Contexto, indicando, pela primeira vez, a utilização de tais questionários para o estudo sobre os Fatores Associados, dando um novo contorno aos questionários. Em relação ao construto “clima escolar”, por mais que conste registros na literatura, segundo Moro (2018, p.49), desde o século 19, tem feito com que pesquisadores e profissionais da Educação, se dediquem a entendê-lo, com a premissa de que um ambiente escolar deve ser positivo.

Outrossim, o Saresp apresenta desde 1996, uma estrutura de instrumentos baseada no Saeb, incluindo, a utilização da Teoria da Resposta ao Item (TRI). Já os Questionários de Contexto foram introduzidos somente na aplicação de 2008, com uma proposta que incluía o estudo dos Fatores Associados, como também, nos anos finais do Ensino Fundamental e na 3ª série do Ensino Médio, o construto sobre o clima escolar. Porém, os questionários aplicados no âmbito do Saresp eram extensos, principalmente, os dos professores e gestores, algo que destoava do convencionado no Saeb.

Nesse contexto, o intuito desse artigo é realizar um estudo sobre a medida de clima escolar da rede estadual de São Paulo, com base nos questionários do Saresp. Por conta de algumas descontinuidades de aplicação dos questionários, iremos realizar um recorte temporal, dividindo a análise em dois momentos, porém, compreendendo que esses recortes devem estar na mesma escala de medida para que haja uma comparação longitudinal.

Sendo assim, o objetivo geral desse artigo se define da seguinte forma: “constituir uma medida do clima escolar “senso de segurança” na rede estadual de São Paulo, nos períodos de 2008 – 2013 e 2018 – 2019, incluindo a criação de uma escala, sua interpretação e a análise dos resultados da rede e suas escolas no geral.

A prática profissional na Secretaria Municipal de Educação de São Paulo e a participação no Grupo de Estudos e Pesquisa em Avaliação Educacional (Gepave) da Faculdade de Educação da Universidade de São Paulo (FEUSP), fazem com que essa abordagem em constituir uma medida de maneira longitudinal seja possível. Principalmente, por ter se dedicado nos últimos quatro anos, aos estudos sobre a constituição de medidas com base na Teoria da Resposta ao Item (TRI).

Esse problema de pesquisa pode ser considerado de grande relevância, pois cada vez mais, o clima escolar, em suas diversas dimensões, têm sido tema de debates, reuniões e publicações, principalmente, por contribuir com aquilo que é central em qualquer sistema de ensino, a melhoria na aprendizagem dos estudantes. Na rede municipal de São Paulo, o clima escolar tem sido abordado como uma medida que pode ser acompanhada ao longo do tempo, com a utilização de uma escala interpretada e que, em muito, pode ajudar as equipes gestoras, professores e comunidade escolar na compreensão dessa medida.

Como já exposto, a pesquisa tem a rede estadual de São Paulo como foco para o estudo, considerando os 7º e 9º anos do EF<sup>2</sup> e a 3ª série do EM e os períodos entre 2008-2013 e 2018-2019, como recorte temporal<sup>3</sup>. A Secretaria Estadual de Educação disponibilizou os questionários e as respectivas bases de dados em uma plataforma<sup>4</sup>. Os itens que fazem parte da definição do construto de clima escolar “senso de segurança” estão em anos diferentes, mesmo quando esse construto não foi definido, compõem a constituição da medida. O maior desafio nesse ponto, são as quantidades de itens e o não ajuste à medida, pois a simples definição e

<sup>2</sup> Considerando que 7º ano equivale a 6ª série do EF e 9º ano equivale a 8ª série do EF.

<sup>3</sup> O 7º ano esteve presente nas aplicações somente até 2013. Em 2018 e 2019 não consta nas bases.

<sup>4</sup> Link: <https://dados.educacao.sp.gov.br/dataset/question%C3%A1rios-saresp> (acessado em 20/12/2020 às 17h09min.)

composição de itens em torno de um construto, não garante que eles tenham esse comportamento após a aplicação.

Para a definição de um único construto (unidimensionalidade) iremos realizar uma Análise Fatorial Exploratória e, como ampliação, uma Análise Fatorial de Informação Completa, que é baseada na TRI. Após definirmos os itens que compõem o construto de nosso interesse, seguiremos para uma análise da medida desse construto, com a utilização do modelo de Samejima da TRI, onde, segundo Andrade et.al (2000, p. 19) “assume que as categorias de resposta de um item podem ser ordenadas entre si”. Dessa maneira, as categorias vão obter os respectivos parâmetros de “dificuldade” e, posteriormente, serem ancoradas na escala do clima escolar “senso de segurança”.

O artigo está estruturado em cinco seções, incluindo a conclusão. A primeira seção apresenta a estrutura dos questionários, a partir dos itens que compõem o construto clima escolar “senso de segurança”. A segunda seção evidencia a definição de construto e a análise de dimensionalidade. Na terceira seção, o foco está na medida relacionada ao clima escolar “senso de segurança”, descrevendo toda a análise pela TRI e a construção da escala de medida e sua descrição pedagógica. Já na quarta seção traz a exploração dos resultados da medida do clima escolar ao longo dos anos em que os questionários foram aplicados. Por fim, a quinta sessão, apresenta as considerações finais e indicações de possíveis estudos posteriores.

## 2 ORGANIZAÇÃO E INFORMAÇÕES DESCRITIVAS SOBRE OS QUESTIONÁRIOS

Os questionários de contexto do Saresp tiveram sua primeira aplicação em 2008, com instrumentos voltados para o levantamento de dados a respeito de: estudantes (2<sup>a</sup>, 4<sup>a</sup>, 6<sup>a</sup>, 8<sup>a</sup> séries do ensino fundamental e 3<sup>a</sup> série do ensino médio), professores (Língua Portuguesa, Matemática e Ciências), diretores e supervisores, que faziam parte da Rede Estadual. Além de interesse em mapear o Nível Socioeconômico dos estudantes, a Secretaria Estadual de Educação (SEE), através de seu site oficial<sup>5</sup>, divulgou a informação, que por meio dos questionários de contexto, buscou realizar uma escuta da Rede sobre aprendizagem, ensino e os fatores associados à aprendizagem, com intuito de obter uma melhora da Educação Paulista.

A partir dos questionários de contexto de 2018 e 2019, onde pela primeira vez, conforme o que pôde ser constatado nos questionários divulgados pela SEE, há uma definição do que seriam os itens que compõem o construto clima escolar, porém, sem definir, de forma explícita, em quais dimensões os itens se enquadram. Com base nessa definição, fizemos um cotejamento entre os oito questionários, com o intuito de criarmos uma escala única. Dessa maneira, identificamos os itens que mais se assemelhassem com as aplicações de 2018 e 2019. Essa mesma estrutura de itens encontrada em 2008 se manteve até 2013. Sendo assim, neste artigo iremos nos referir como **Momento 1**, o período entre os anos de 2008 a 2013. É importante destacar que na plataforma em que os dados estão disponibilizados, não encontramos bases de dados para os anos entre 2014 e 2017. Consequentemente, não iremos considerá-los neste artigo. Já os dados de 2018 e 2019 estão disponíveis e iremos definir como **Momento 2**,

Os questionários de contexto do **Momento 2** recebem um acréscimo de 9 itens à estrutura anterior, como também, já anunciado acima, uma delimitação clara do construto clima escolar no próprio questionário. Essas atitudes facilitam para os pesquisadores e sociedade civil identificarem qual é o norte que a SEE está dando para a definição dos construtos no questionário. Contudo, também é possível perceber que o clima escolar destacado nos questionários de contexto do **Momento 2** restringem os itens à uma dimensão que podemos

<sup>5</sup> Link: <https://www.educacao.sp.gov.br/saresp-2008-tera-mega-questionario-sobre-aprendizagem-para-professores/> (acessado em 21/12/2020 às 9h17min.)

classificar como sendo de “relações humanas”. Essa classificação se baseia, em resumo que, na afirmação de MORO (2018, p.48)

[...] podemos afirmar que o clima de uma escola emerge de uma avaliação subjetiva e compartilhada pelos atores escolares, em razão das vivências cotidianas no contexto educacional, em relação às diferentes dimensões da instituição (normas, objetivos, relações humanas, organização e estruturas física, pedagógica e administrativa).

Nesse artigo, em ambos os momentos, iremos nos ater aos questionários aplicados aos estudantes da 6ª série (7º ano), 8ª série (9º ano) e 3ª série do ensino médio. Esse recorte se impõe pelo fato de não termos itens que abordem, mesmo que tangencialmente, o construto ora analisado, na 2ª série (3º ano) e 4ª série (5º ano). Todavia, merece destaque o fato de essa falta de itens nos anos iniciais ocorrer no **Momento 1**, pois no **Momento 2** ela não ocorre, pois o questionário de contexto é o mesmo para todos os anos escolares que tiveram aplicação do Saresp<sup>6</sup>.

Em relação ao quantitativo de estudantes considerados na análise<sup>7</sup> por ano de aplicação, a Figura 1, apresenta graficamente esses quantitativos. É possível observar que os anos de 2008 e 2018 apresentam um quantitativo expressivamente menor em comparação aos demais anos de aplicação. Não conseguimos compreender ou encontrar explicações para esse quantitativo menor, contudo, em 2018, a base de dados trazia estudantes de praticamente todos os anos escolares. Esse fato nos chamou atenção, mas para mantermos a análise com os mesmos anos escolares, realizamos um filtro na base. Não é de nosso conhecimento que existam problemas de inconsistências nas bases de 2008 e 2018.

Figura 1 – Gráfico com o quantitativo dos estudantes considerados nas análises



Fonte: Elaborado pelo autor (2021)

### 3 CLIMA ESCOLAR

A concepção de clima escolar por si só, já nos remete a algo que se apresenta como complexo e que muito provavelmente não esteja restrita a somente uma dimensão, logo, com um caráter polissêmico. Nessa perspectiva, seria algo complexo tentarmos definir o que seria o

<sup>6</sup> Para mais detalhes sobre esses itens que vão compor a medida, consultar a **Tabela 4**, como também, o site dos “Dados Abertos da Educação” mencionados anteriormente neste artigo.

<sup>7</sup> Consideramos somente os estudantes que responderam todos os itens que compõem o construto clima escolar. Essa decisão está ancorada no quantitativo de itens que fazem parte do construto, como são somente dez itens, sendo 7 no **Momento 1**, 8 no **Momento 2** e 4 que interligam os dois momentos. Optamos por não perdemos informação com dados faltantes.

clima escolar a partir de um instrumento já construído e aplicado. Sendo assim, partimos do que nos foi dado, a definição de clima escolar por parte da própria SEE.

### 3.1 DEFINIÇÃO DO CONSTRUTO DO CLIMA ESCOLAR

Observando o que foi apresentado como clima escolar, verificamos que os itens estão completamente relacionados às relações daqueles que compõem o dia a dia da escola. Contudo, as relações entre os estudantes, ao nosso ver, ganham maior destaque, pois são o foco de um bloco inteiro e mais alguns itens de um segundo bloco do questionário. Para além de um item que busca trazer uma reflexão por parte do estudante em relação ao seu futuro estudantil. Contudo, as relações entre os estudantes identificadas nestes itens, estão direcionadas para uma espécie de violência, que pode ser física ou psicológica.

Essa violência exposta acima, é classificada como “Segurança social e emocional” por Cohen et. al. (2008), onde a descrição apresentada pelos pesquisadores é “Estudantes se sentem seguros em relação a abuso verbal, intimidação, provocação e exclusão”. Podemos destacar, na mesma perspectiva, outras classificações em relação a segurança que fazem parte dos itens que compõem o construto que estamos analisando, como por exemplo, “Segurança física”. Segundo Cohen et.al. (2008), essa classificação é descrita como: “Educadores e estudantes sentem-se seguros, quanto a danos físicos na escola”.

Sendo assim, nos parece mais significativo descrevermos o construto analisado por nós como “Senso de segurança”, conforme Cohen et. al (2008) estabeleceram em suas pesquisas. Na parte de ancoragem e descrição da escala, iremos compreender como que essa medida se estabeleceu e o significado descritivo de cada nível.

### 3.2 DIMENSIONALIDADE

A dimensionalidade em relação aos itens, pressupõem estarmos medindo a mesma coisa, logo, como sinaliza o professor Pasquali (2013, p.114) sobre a unidimensionalidade, “[...] ela faz a suposição que eles são somáveis e isto faz sentido somente se eles se referem à mesma coisa, pois não dá para somar alhos e bugalhos”. Sendo assim, pouco importa se iremos realizar a análise estatística da medida com base na Teoria Clássica dos Teste ou na Teoria da Resposta ao Item, ambas se baseiam em dimensionalidade única, portanto, unidimensionalidade.

Existem diversas técnicas para analisar essa dimensionalidade, porém, neste artigo, iremos utilizar a Análise Fatorial Exploratória (AFE) e a Análise Fatorial de Informação Completa (AFIC ou na sigla em inglês FIFA – Full Information Factorial Analysis). Ambas as técnicas consistem em reduzir o conjunto de itens à uma única dimensão ou fator. A AFE está no bojo das técnicas de Análises Fatoriais (AF), que por sua vez são definidas da seguinte forma por Pasquali (2013, p.116):

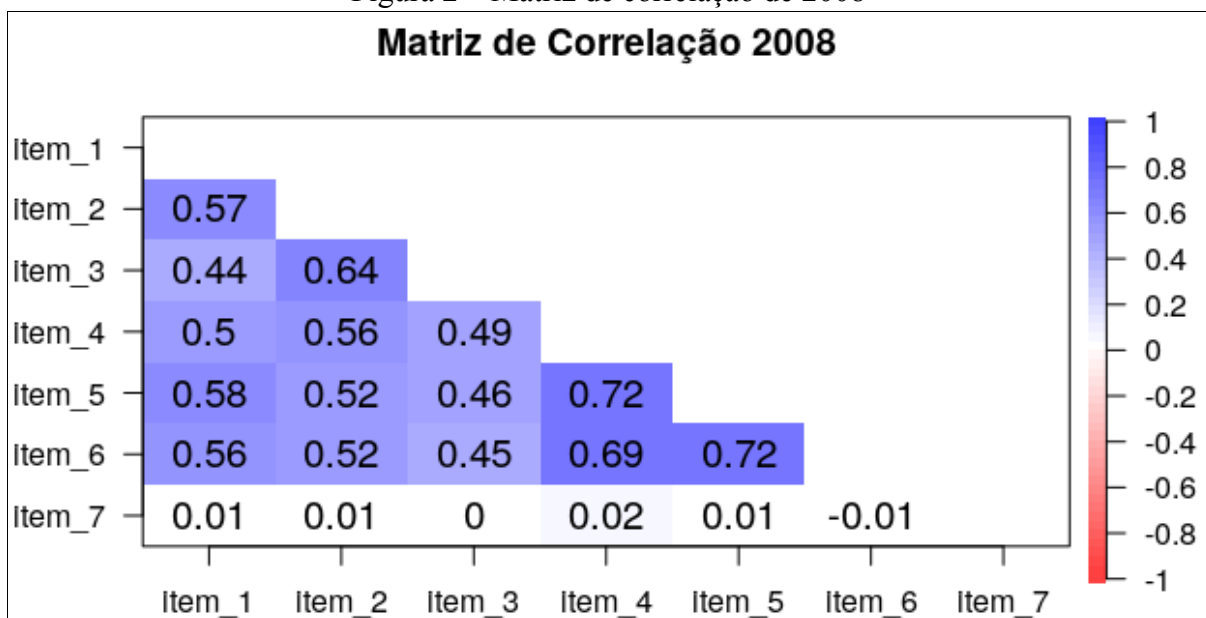
A análise fatorial tradicional, abreviada como AF, consiste numa série de técnicas estatísticas que trabalha com análises multivariadas e matrizes. A matriz que a AF trabalha é tipicamente uma matriz de intercorrelações entre uma série de variáveis ou itens. A análise que ela faz consiste em verificar se uma série de itens pode ser reduzida idealmente a uma única dimensão ou variável, que ela chama de fator, com o qual todas as variáveis da série estão relacionadas.

A Análise Fatorial de Informação Completa está baseada na TRI e tem um potencial maior para com a análise de traços latentes. Segundo Pasquali (2013, p.118): “Embora laborioso de um ponto de vista computacional [...] no presente, parece ser o melhor método para decidir

a unidimensionalidade [...]”. Contudo, preferimos realizar ambas as análises, muito mais para termos uma redundância, do que propriamente uma comparação.

Antes de iniciarmos as Análises Fatoriais Exploratórias e de Informação Completa, realizamos as análises de correlação entre as variáveis dos questionários por ano de aplicação. Na Figura 2, é apresentada a Matriz de Correlação de 2008. Nela é possível constatar que o item 7, na estrutura do questionário do **Momento 1**, apresenta baixos valores de correlação com os demais itens. Isso já pode indicar que esse item poderá ser retirado da análise, pois, segundo Matos (2019, p.42) que realiza um cotejamento entre diversos autores sobre os valores das correlações, indica que estas por sua vez, devem estar acima do valor de 0,3.

Figura 2 – Matriz de correlação de 2008



Fonte: Elaborado pelo autor (2021).

Na sequência, dois testes estatísticos são importantes para o processo da AFE, o “Teste de Bartlett” e o “Teste de Kaiser-Meyer-Olkin (KMO)”. O primeiro, tem o intuito de verificar se a matriz de correlação se reduz à uma matriz identidade. Já o segundo, representa uma proporção que varia entre 0 e 1, e indica o quanto que a variância das variáveis pode ser explicada pelo traço latente. O Teste de Bartlett realizado com os dados de 2008 apresenta um valor  $p$  abaixo de 0,05, logo, segundo quadro apresentado por Matos (2019, p.44) devemos rejeitar a hipótese nula, que neste caso, significa que a matriz é diferente de uma matriz identidade. Já o valor para o teste KMO foi igual a 0,86 que podemos considerar como “ótimo”. Ainda sobre o KMO, temos valores para cada um dos itens. O único item com valor abaixo de 0,8 e que está fora do que, segundo a tabela citada acima, seria o mínimo aceitável para esse teste, é o item 7, com um valor igual a 0,26.

Os resultados apresentados na Matriz de Correlação e dos testes acima, nos dão o aval para prosseguirmos na AFE, com a extração dos fatores. Para tanto, iremos utilizar o método de extração pela máxima verossimilhança através da função “fa” do pacote psych do R. Na coluna “fa\_psych” da Tabela 1, podemos observar que as cargas fatoriais para os itens de 1 a 6 são acima de 0,6. Já o item 7 apresenta carga fatorial praticamente igual a 0. A proporção de variância acumulada, com o item 7 incluso, tem 0,51 porcentos considerando um único fator. A análise pela Informação Total apresentou valores próximos do que os encontrados com a função “fa”. Na coluna “info\_total” da Tabela 1, podemos observar que o item 7 também apresenta valores mais baixos que os demais itens, neste caso, sendo negativo.

Tabela 1 – Extração dos fatores de 2008

Item	Extração do fatores 2008 – Rodada 1			Extração do fatores 2008 – Rodada 2	
	fa psych	info total	comunalidades	fa psych	info total
1	0,68	0,72	0,47	0,68	0,72
2	0,69	0,77	0,48	0,69	0,77
3	0,61	0,67	0,37	0,61	0,67
4	0,82	0,84	0,68	0,82	0,84
5	0,84	0,83	0,71	0,84	0,83
6	0,82	0,81	0,67	0,81	0,81
7	0,01	-0,001	0	-	-

Fonte: Elaborado pelo autor (2021).

Antes de realizarmos a exclusão do item 7 do modelo, verificamos os valores de comunalidades. Os itens apresentaram os seguintes valores, respectivamente, 0,466; 0,479; 0,372; 0,679; 0,712; 0,674; 0,000. Novamente, podemos observar que o item 7 realmente destoa dos demais e deve, com base nos valores baixos da carga fatorial e comunalidade, ser retirado da análise.

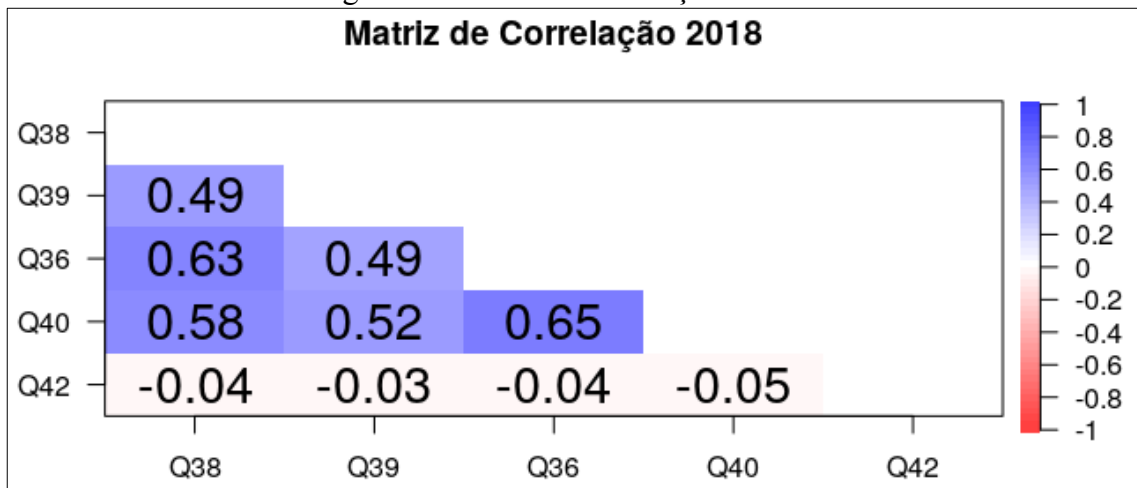
Com a retirada do item 7 do modelo, os valores apresentados para “fa\_psych” e “info\_total” na rodada 2 da tabela 1, não apresentam nenhuma diferença em relação aos apresentados na rodada 1. Contudo, a variância explicada tem um aumento, no modelo ajustado pela função “fa”, o valor saiu de 0,48 para 0,56. Já a análise pela Informação Total, o valor era de 0,51 e passa a ser 0,60, um crescimento considerável. É possível perceber que realmente podemos considerar um modelo com um único fator, algo que atende um dos pressupostos mais importantes da TRI, a unidimensionalidade.

O instrumento de 2018 apresenta uma diferença considerável em relação aos instrumentos do **Momento 1**, porém, quatro itens são praticamente iguais<sup>8</sup>, o que permite realizarmos a comparação ao longo dos anos. Nesse sentido, realizamos uma AFE com o intuito de observar se os itens se relacionam da mesma maneira que em 2008 com o construto Clima Escolar.

Como à análise de 2008, começamos 2018 analisando as correlações das variáveis, Figura 3, que nos apresenta a “Matriz de Correlações”. Nela é possível observar que o item “Q42”, equivalente ao item 7 de 2008, tem valores negativos para suas correlações com os demais itens. Algo que sugere uma possível retirada dele do modelo. Os testes estatísticos de Bartlett e KMO apresentam resultados dentro do esperado para que a AFE possa prosseguir. No teste KMO algo inusitado ocorreu, pelo menos se comparado ao que ocorreu no teste com os dados de 2008, o item equivalente ao 7, não obteve um valor inapropriado, pelo contrário, neste caso ele obteve 0,83, que segundo a tabela citada anteriormente, considera como “ótimo”. Contudo, essa classificação não significa que este item pode explicar a variância do mesmo fator que os demais, mas, que pode explicar a variância de um fator.

<sup>8</sup> Existe uma pequena modificação na escrita, porém, o sentido do item permanece o mesmo.

Figura 3 – Matriz de correlação de 2018



Fonte: Elaborado pelo autor (2021).

Com bases nesses resultados prosseguimos para a extração dos fatores. Como em 2008, iremos utilizar o método de extração pela máxima verossimilhança, a partir da análise com a função “fa” do pacote “psych” e a de “Informação Total” com o pacote “mirt”. Entretanto, antes de partirmos de fato às análises, é importante destacar que a proporção acumulada de explicação nos mostra que somente com dois fatores teremos 0,76 por cento de explicação, algo que já nos indica a real possibilidade de termos dois fatores.

A Tabela 2 traz os dados das rodadas 1 e 2. Na rodada 1, com todos os itens, podemos observar que de fato o item “Q42” não apresenta uma relação com o mesmo fator que os demais itens. Tanto a análise pela função “fa” como a de “Informação Total” apresentam valores praticamente nulos para essa relação do item com o fator comum aos demais. Isso também pode ser observado com os valores da variável “comunalidades”, o item “Q42” apresenta 0,00, quando os demais variam de 0,45 a 0,69. Logo, com base nesses dois critérios, carga fatorial e comunalidade, o item “Q42” foi retirado do modelo e ele foi ajustado novamente.

Tabela 2 - Extração dos fatores de 2018

Extração do fatores 2018 Geral – Rodada 1				Extração do fatores 2018 Geral – Rodada 2	
Item	fa psych	info total	comunalidades	fa psych	info total
Q34	0,55	0,71	0,3	0,55	0,71
Q35	0,73	0,84	0,54	0,75	0,88
Q36	0,64	0,82	0,4	0,64	0,82
Q37	0,62	0,73	0,38	0,62	0,73
Q38	0,77	0,87	0,59	0,79	0,89
Q39	0,54	0,65	0,3	0,54	0,65
Q40	0,57	0,77	0,33	0,58	0,77
Q41	0,64	0,76	0,41	0,64	0,75
Q42	-0,03	-0,03	0	-	-
Q43	0,23	0,34	0,05	-	-
Q44	0,26	0,36	0,07	-	-
Q45	0,19	0,27	0,04	-	-
Q46	0,2	0,3	0,04	-	-
Q47	0,23	0,33	0,05	-	-
Q48	0,2	0,31	0,04	-	-
Q49	0,18	0,26	0,03	-	-

Fonte: Elaborado pelo autor (2021).



Os valores das cargas fatoriais permanecem os mesmos, a exemplo do que ocorreu em 2008, como também, a proporção explicada salta de 0,56 para 0,70, algo que realmente demonstra que os quatro itens restantes se relacionam com um fator único. A comunalidade também não sofre nenhuma mudança com a retirada, porém, a variância explicada pelo modelo apresenta importantes ganhos. A variância do modelo ajustado pela função “fa” apresenta um salto de 0,12, saindo de 0,48 para 0,6. Já a de “Informação Total” sai de 0,49 para 0,62, algo muito próximo da função anterior. Todos esses resultados nos demonstram que é possível considerarmos um único fator quando considerarmos os itens em comum, algo que nos credencia a utilizarmos a TRI nas análises.

Entretanto, antes de partirmos para o próximo capítulo deste artigo, onde a análise da TRI será apresentada. Precisamos realizar uma AFE com todos os itens de 2018 que estão classificados no questionário como sendo do construto Clima Escolar. Essa análise será útil para sabermos se no **Momento 2**, poderemos ampliar os itens que compõem o instrumento de medida. De modo geral, ambas as análises indicam existirem dois construtos. A Tabela 3 nos mostra que realmente as cargas fatoriais para os itens de “Q34” a “Q41” são maiores, pois estamos fixando somente um fator. Os valores para as comunalidades também demonstram que os itens de “Q42” até “Q49” são praticamente iguais a 0. É importante destacar que o item “Q42” equivalente ao item 7 em 2008, é o único item com carga negativa.

Tabela 3 – Extração dos fatores dos anos demais anos de aplicação

Extração do fator nos demais anos				
Ano	Item	fa psych	info total	comunalidades
2009	1	0,57	0,7	0,33
2009	2	0,61	0,76	0,37
2009	3	0,54	0,67	0,29
2009	4	0,57	0,84	0,33
2009	5	0,7	0,83	0,49
2009	6	0,66	0,81	0,44
2010	1	0,58	0,72	0,34
2010	2	0,62	0,77	0,38
2010	3	0,54	0,67	0,29
2010	4	0,51	0,83	0,26
2010	5	0,7	0,83	0,49
2010	6	0,66	0,83	0,44
2011	1	0,56	0,7	0,31
2011	2	0,58	0,74	0,34
2011	3	0,5	0,64	0,25
2011	4	0,53	0,81	0,29
2011	5	0,71	0,84	0,51
2011	6	0,68	0,83	0,46
2012	1	0,57	0,2	0,33
2012	2	0,6	0,76	0,35
2012	3	0,52	0,69	0,27
2012	4	0,53	0,84	0,28
2012	5	0,71	0,84	0,5
2012	6	0,68	0,84	0,47
2013	1	0,56	0,71	0,56
2013	2	0,59	0,77	0,59
2013	3	0,51	0,67	0,51
2013	4	0,58	0,86	0,58
2013	5	0,72	0,85	0,72
2013	6	0,7	0,84	0,7
2019	1	0,61	0,75	0,37
2019	2,3,4	0,53	0,67	0,28
2019	5	0,69	0,83	0,47
2019	6	0,65	0,83	0,42

Fonte: Elaborado pelo autor (2021).

A análise pela “Informação Total” revela cargas um pouco mais elevadas, contudo, confirma que estamos nos referindo a mais de um fator, pois os valores para os itens do “Q42” ao “Q49” são baixos ou negativo. A variância explicada com o modelo em que todos os itens estão presentes explica somente 0,34 por cento. Já com a retirada dos itens que apresentaram cargas fatoriais baixas, a variância explicada passa para 0,60 por cento, algo que pode ser considerado como interessante para o ajuste do modelo.

Sendo assim, o instrumento analisado pela TRI contém os quatro itens que apresentaram resultados satisfatórios nas análises AFE e de Informação Total, mais, os quatro itens que compõem os questionários de 2018 e 2019, que também apresentaram resultados significativos. Para que não houvesse nenhum tipo de confusão em relação a estrutura dos itens que vamos utilizar no processamento, reorganizamos os nomes dos itens que estavam presentes nos questionários. A partir deste ponto, os itens serão denominados de item\_1 até item\_10, conforme o exposto na Tabela 4.

Tabela 4 – Estrutura dos questionários contextuais

Estrutura dos questionários contextuais								
Ano aplicação	Item	Código dos questionários	Desde o começo deste ano, quantas vezes os seguintes problemas aconteceram com você na escola?	Alternativa A	Alternativa B	Alternativa C	Alternativa D	Alternativa E
2008 a 2019	1	44.1*/49.1**/ Q38***	Alguém zombou ou insultou você.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2008 a 2019	2	44.2*/49.2**/ Q39***	Alguém estragou alguma coisa sua.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2008 a 2013	3	44.3*/49.3**	Alguém roubou alguma coisa sua.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2008 a 2013	4	44.4*/49.4**	Alguém tomou dinheiro ou alguma coisa sua usando a força ou ameaçando você.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2008 a 2019	5	44.5*/49.5**/ Q36***	Alguém ameaçou ou intimidou você.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2008 a 2019	6	44.6*/49.6**/ Q40***	Alguém já atacou ou agrediu fisicamente você.	Muitas vezes, quase sempre	Algumas vezes	Nunca	Não se aplica	Não se aplica
2018 e 2019	7	Q34***	Você já sofreu algum tipo de discriminação na escola?	Sim.	Não			
2018 e 2019	8	Q35***	Fui chamado de nomes feios por outros alunos	Sempre ou quase sempre	De vez em quando	Nunca		
2018 e 2019	9	Q37***	Outros alunos me deixaram de fora das coisas de propósito.	Sempre ou quase sempre	De vez em quando	Nunca		
2018 e 2019	10	Q41***	Outros alunos espalharam boatos desagradáveis ao meu respeito	Sempre ou quase sempre	De vez em quando	Nunca		
*Questionário do 9º ano (referência 2008) **Questionário da 3ª série (referência 2008) ***Questionário do EF e EM (referência 2018)								

Fonte: Elaborado pelo autor (2021).

## 4 MEDIDA DO CLIMA ESCOLAR

Os itens que compuseram a parte do instrumento que se relacionam com uma dimensão do clima escolar “senso escolar”, apresentam duas estruturas: uma com três opções de escolha, onde é possível ordená-las em: “Muitas vezes, quase sempre”; “Algumas vezes”; “Nunca” e outra com apenas duas opções, “sim” ou “não”. Essa ordenação nos possibilita escolhermos o modelo de Samejima da Teoria da Resposta ao Item para a análise da medida do clima escolar “senso escolar”.

### 4.1 ASPECTOS GERAIS DA TEORIA DA RESPOSTA AO ITEM (TRI)

A Teoria da Resposta ao Item (TRI) é um conjunto de modelos estatísticos que nos permitem realizarmos uma análise mais qualificada dos itens e medidas do traço latente. Segundo Muñiz (1997), a grande contribuição da TRI, “se centra na possibilidade de obter medições invariantes a respeito dos instrumentos utilizados e dos sujeitos implicados”. Sendo assim, o fato de a TRI possibilitar medições invariantes, faz com que o seu uso seja imprescindível a partir do momento que há o desejo de comparar medidas de traço latentes ao longo do tempo.

Em linhas gerais, a TRI pressupõe que existe uma relação entre a função das variáveis definidas para medir os itens e a probabilidade de os respondentes acertarem os mesmos ou marcarem uma determinada opção de escolha. Dessa maneira, quanto menor for a medida do traço latente estimada para o respondente, menor será sua probabilidade de acertar ou marcar determinados itens ou opções de escolha. Para maiores informações sobre a TRI, consultar Muñiz (1997), Andrade et.al (2000), Pasquali (1996) e Pasquali (2013), Birnbaum (1968) e Lord (1980).

### 4.2 ANÁLISE DE TRI PELO MODELO DE SAMEJIMA – ORGANIZAÇÃO DOS DADOS E CALIBRAÇÃO DOS ITENS

O modelo de Samejima (1969), ou melhor, Modelo de Resposta Gradual (MRG) amplia a capacidade de obter mais informação do que simplesmente os modelos que consideram as respostas como certas ou erradas. O modelo está baseado no modelo de 2 parâmetros, acrescentando uma variação no parâmetro “b” (dificuldade do item) para cada uma das categorias. Sendo assim, o modelo de Samejima é definido da seguinte forma:

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

Com  $i = 1, 0, \dots, I$ ,  $j = 1, 2, \dots, n$ , e  $k = 0, 1, \dots, s_i$ , onde:

$P_{i,k}^+(\theta_j)$  = probabilidade de o respondente com  $\theta_j$  marcar a categoria  $k$  ou superior;

$D$  = constante igual a 1 para o modelo logístico;

$a_i$  = parâmetro relacionado a “discriminação” do item;

$b_{i,k}$  = parâmetro onde a probabilidade de cada par de categoria é igual;

$\theta_j$  = medida do clima escolar “senso de segurança” para cada respondente.

É importante destacar que no caso dos itens dicotômicos o parâmetro “a” (discriminação do item) está relacionado a alternativa considerada o “gabarito”, já no MGR a discriminação de cada categoria depende tanto do parâmetro “a”, igual para todas as categorias do item, como das distâncias das categorias em relação as “dificuldades”. Para maiores informações sobre a discriminação, como também, sobre o próprio modelo, consultar Andrade et.al (2000).

Em nossa análise, os itens 1, 2, 3, 4, 5, 6, 8, 9 e 10 apresentam três categorias, definidas como: 0 - (Muitas vezes, quase sempre), 1 - (Algumas vezes) e 2 - (Nunca). Os dados do item

7 estavam definidos como 0 – (Não) e 1 – (Sim). Nos casos em que as bases de dados não estavam transformadas na definição acima, foram feitos ajustes para adequar. É importante ressaltar que os itens se referem a situações “negativas”, logo, quanto menos tais situações ocorrem, melhores são as relações internamente a escola.

Por 2008 ser o primeiro ano de aplicação, utilizamos os seus dados como referência para a escala. Entretanto, para diminuir a propagação de erro de medida, ao invés de realizarmos uma equalização a partir dos dados da referência com os demais, decidimos realizar um único processamento, onde todos os anos onde houve aplicação do questionário foi processado conjuntamente com os demais. Dessa maneira, conseguimos garantir uma escala única com uma menor probabilidade de ampliarmos as inconsistências causadas pela equalização em vários estágios.

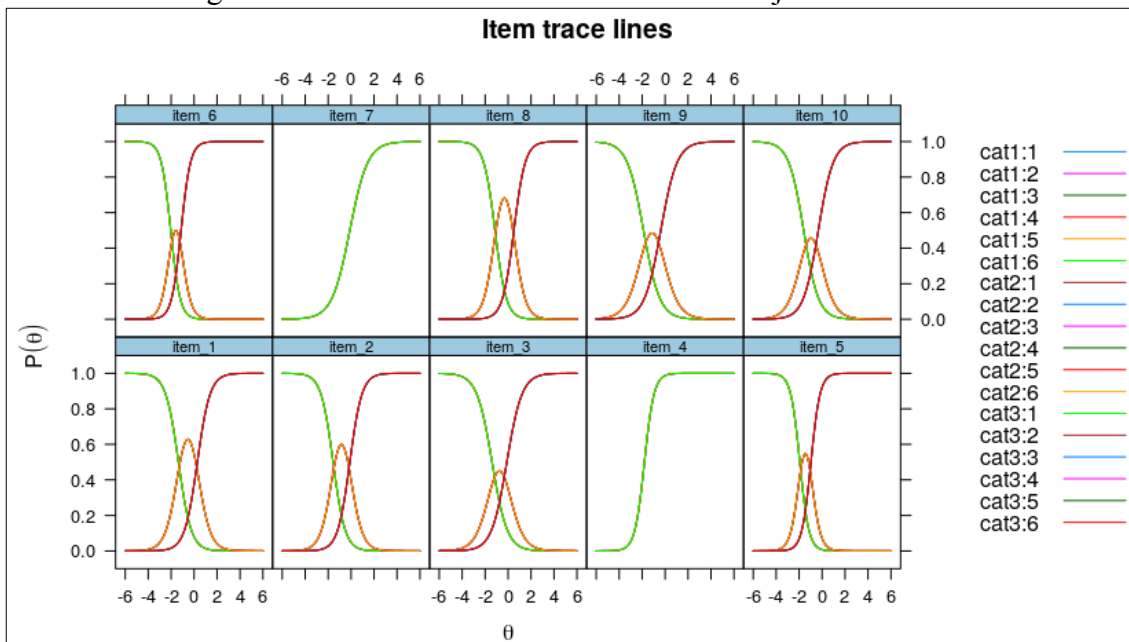
As bases de dados foram empilhadas, adicionando “NA” para os casos em que os itens não foram submetidos aos estudantes. Isso ocorreu nos itens 3 e 4 para os estudantes que responderam os questionários em 2018 e 2019, já para os estudantes que responderam de 2008 a 2013, os itens 7, 8, 9 e 10 receberam tal ação. Foi introduzida uma variável de “grupo”, para qualificar cada ano de aplicação como um grupo nas análises, o que viabiliza a utilização da metodologia de múltiplos grupos<sup>9</sup>.

O pacote mirt, já especificado acima quando fizemos a análise de Informação Total, nos permite realizarmos a análise de múltiplos grupos, a partir da função “multiplegroup”. Realizamos inicialmente a análise com os 8 grupos, porém, os primeiros resultados demonstraram haver alguma interferência inicial dos grupos de 2009 e 2019, pois os valores dos traços latentes estimados estavam muito superiores aos demais, o que nos levou a retirarmos inicialmente esses dois grupos. Após essa retirada os dados apresentavam uma distribuição aparentemente mais ajustadas. Esse procedimento nos garantiu colocarmos os grupos retirados na mesma métrica de os demais, conforme o descrito em Andrade et al.(2000, p. 85 e 86).

A figura 4, nos mostra as Curvas Características dos Itens (CCI), é possível observar que os itens 3 e 6 apresentam uma certa sobreposição das demais categorias, porém, como os valores são maiores que 0,4 entendemos que seria melhor deixarmos, porém, nos parece pertinente essa categoria “Algumas vezes” ser revista em aplicações futuras. Contudo, entendemos que neste caso a permanência contribuí mais do que atrapalha, pois adiciona mais informação, logo, os mantemos. Somente no item 4, após a primeira rodada, essa categoria não pode ser mantida, transformando o item em dicotômico. Assim, ficamos com os itens 4 e 7 sendo dicotômicos e os demais politômicos.

<sup>9</sup> Para maiores detalhes consultar Andrade et.al.(2000)

Figura 4 – Curva Característica do Item – Conjunto dos itens



Fonte: Elaborado pelo autor (2021).

A tabela 5 apresenta os parâmetros dos itens estimados na rodada 2 com os anos de 2008, 2010, 2011, 2012, 2013 e 2018. É possível observar que os valores para “b2” não ficam muito distantes de 0 positivamente, sendo o item 8 com o maior valor. Já para “b1” o item 6 apresenta o menor valor, em conjunto com o item 5, são os itens com menores valores para esse parâmetro. Foi verificado se houve DIF entre os grupos, utilizando a função DIF do próprio pacote “mirt”, porém, não houve a necessidade de realizar ajustes.

Tabela 5 – Parâmetros dos itens

Parâmetros dos itens			
itens	a1	b1	b2
item 1	1,7983	-1,3604	0,2817
item 2	1,8958	-1,5681	-0,1097
item 3	1,5098	-1,3962	-0,1116
item 4	2,3935	-1,7998	-
item 5	2,5261	-1,9410	-0,9696
item 6	2,4311	-2,0359	-1,1358
item 7	1,2852	-0,1108	-
item 8	2,0227	-1,1386	0,5052
item 9	1,4099	-1,8786	-0,3743
item 10	1,5402	-1,6061	-0,3280

Fonte: Elaborado pelo autor (2021).

#### 4.3 DEFINIÇÃO DA ESCALA

A definição da escala de “habilidade” possibilita irmos além da simples medida, tendo em vista que até o momento somente realizamos o movimento de colocarmos itens e respondentes na mesma métrica, porém, sem darmos significado para essa medida. Sendo assim, utilizamos os critérios definidos em Andrade et.al. (2000, p. 110) para estabelecermos as categorias âncoras. A tabela 6 apresenta as ancoragens das categorias através de três cores, sendo laranja para a menor categoria (“Muitas vezes, quase sempre” ou “Não” - nos itens dicotômicos), para a segunda categoria utilizamos o amarelo para representar (“Algumas vezes”

ou “Sim” - nos itens dicotômicos) e por fim, a cor verde para a maior categoria (“Nunca”)<sup>10</sup>. É importante destacar que não foi possível seguir os três critérios na maioria dos itens, sendo assim, quando isso não era possível, ancorávamos a partir do critério “probabilidade maior ou igual a 0,65”.

Para facilitar a compreensão e evitar estranhamentos, realizamos uma transformação linear na escala (0,1) para uma escala (5,1.5). Dessa maneira, evitamos valores negativos, algo que seria difícil de explicar e, sem sombras de dúvida, causaria dificuldades semelhantes em sua compreensão. É possível observar na tabela 6 que a menor categoria ficou limitada aos valores entre 2,75 e 3,125, indicando que os estudantes que apresentam estimativas para o traço latentes entre esses valores ou até menores, consideram que o “senso de segurança” apresenta problemas, pois, “muitas vezes, quase sempre”, eles são agredidos, xingados ou têm seus pertences ou roubados ou furtados.

Tabela 6 – ancoragem das categorias

Escala de Clima Escolar (Senso de segurança)													
Itens	2	2,38	2,75	3,13	3,5	3,88	4,25	4,63	5	5,38	5,75	6,13	6,5
Item 1													
Item 2													
Item 3													
Item 4													
Item 5													
Item 6													
Item 7													
Item 8													
Item 9													
Item 10													

Fonte: Elaborado pelo autor (2021).

Nesta mesma perspectiva, a categoria de cor amarela nos mostra que os estudantes que estão localizados entre 3,5 a 5,375, consideram que para a maior parte dos itens apresentados os problemas ocorrem “algumas vezes”. Já para a categoria verde, são mais marcadas pelos estudantes com valores maiores que 5,375. Para estes, as situações ocorrem com menor frequência, considerando o ambiente escolar com maior “senso de segurança”.

Tabela 7 – Interpretação dos níveis (descrição dos níveis)

Interpretação Pedagógica (Descrição da escala em alguns níveis)		
Níveis	Pontos na escala	Descrição - No ambiente escolar, os estudantes neste nível têm a percepção que:
1	Menores do que 2	as pessoas zombam delas, estragam e roubam suas coisas muitas vezes ou quase sempre e já tomaram dinheiro ou alguma coisa. Elas as ameaçam ou intimidam, atacam ou agridem fisicamente muitas vezes ou quase sempre e fazem discriminação. Como também, fazem xingamentos, excluem deliberadamente e espalham boatos desagradáveis ao seu respeito muitas vezes ou quase sempre.
2	2 até 2,375	as pessoas realizam as mesmas coisas que foram descritas no nível anterior com um abrandamento em relação as ameaças ou intimidações e ataques ou agressões físicas, que acabam ocorrendo algumas vezes.
3	2,375 até 2,75	ocorrem as mesmas coisas que os níveis anteriores com dois abrandamentos, um em relação ao fato de tomarem dinheiro não ter ocorrido e o outro, em relação ao fato de as pessoas lhe deixarem de fora das coisas deliberadamente.
4	2,75 até 3,125	ocorrem as mesmas coisas que os nos níveis anteriores com dois abrandamentos, um em relação as pessoas estragarem as coisas algumas vezes e outro, em relação as pessoas espalharem boatos ocorrem algumas vezes.
5	3,125 até 3,875	todas as coisas anteriores ocorrem de maneira mais branda, passando para algumas vezes ou não ocorrendo, porém, a discriminação continua ocorrendo.
6	3,875 a 4,625	ocorrem as mesmas coisas que o nível anterior, porém, como três abrandamentos. As ameaças, os ataques ou agressões e os xingamentos nunca terem ocorrido.
7	4,625 até 5	ocorrem as mesmas coisas que o nível anterior, ampliando os abrandamentos para nunca terem deixado fora das coisas deliberadamente ou espalharem boatos.
8	5 até 5,375	houve abrandamento em praticamente tudo o que foi perguntado, exceto, em relação ao fato de continuarem zombado algumas vezes.
9	Acima de 5,375	nunca ocorreram nenhuma das coisas que foram perguntadas.

Fonte: Elaborado pelo autor (2021).

<sup>10</sup> Nos itens de 2018 e 2019 existe uma mudança na escrita das categorias, porém o sentido permanece o mesmo. Consultar a Tabela 4 para verificar as diferenças e igualdades.

A tabela 7 nos mostra a “interpretação pedagógica” ou melhor, a descrição de alguns níveis que estabelecemos a partir da quantidade de quadraturas<sup>11</sup> e os limites impostos para as medidas. Sendo assim, após a ancoragem, retiramos da análise os pontos na escala que não tinham nenhuma categoria ancorada e nos dedicamos a descrever aqueles que as tinham. Dessa maneira, a descrição foi realizada para 9 níveis, com a medida para o clima escolar “senso de segurança” variando de 2 a mais de 5,75 pontos na escala.

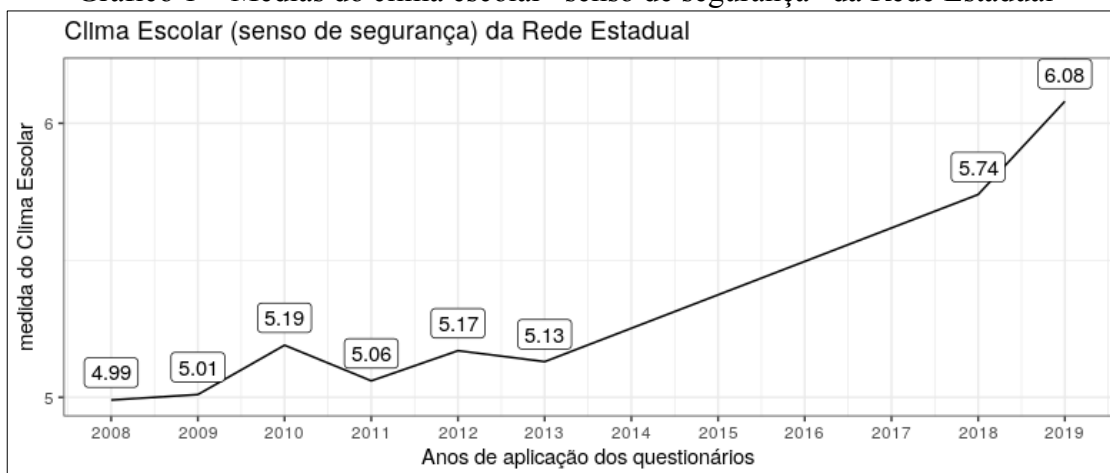
## 5 EXPLORANDO OS RESULTADOS DO CLIMA ESCOLAR “SENDO DE SEGURANÇA”

As análises realizadas até o momento deste artigo, possibilitam termos resultados de toda a série histórica do clima escolar “senso de segurança” da Rede Estadual de São Paulo. Esses dados podem revelar percepções importantes dos estudantes, ao mesmo tempo que colocam luz em situações de violência física e psicológica vividas por uma parcela desses mesmos estudantes. Contudo, não se esgotam as possibilidades de análises e avanço nas formas de definirmos as dimensões do clima escolar, principalmente, para termos uma “visão” mais qualificadas dessas percepções, inclusive, com a inclusão de outros participantes da convivência escolar. Nas próximas páginas iremos apresentar os resultados gerais do clima escolar “senso de segurança”.

### 5.1 RESULTADOS GERAIS DO CLIMA ESCOLAR “SENDO DE SEGURANÇA”

Para apresentar os resultados obtidos, criamos um gráfico com as médias do clima escolar “senso de segurança” por ano de aplicação dos questionários para a Rede Estadual. Esses resultados, apresentados no gráfico 1, nos mostram que entre os anos de 2008 e 2013 as médias para o clima escolar “senso de segurança” oscilam sem grandes magnitudes, contudo, é possível perceber um aumento de 0,14 pontos no período. Já ao avançarmos para os anos de 2018 e 2019<sup>12</sup>, podemos observar um aumento na média de cada ano, com destaque para a diferença entre os dois anos, que é igual a 0,34. Entre 2013 e 2018, houve um crescimento de 0,61 pontos, que representa 0,12 pontos se pensarmos um crescimento por ano.

Gráfico 1 – Médias do clima escolar “senso de segurança” da Rede Estadual



Fonte: Elaborado pelo autor (2021).

<sup>11</sup> Definimos como 20 a quantidade de quadraturas, porém, ao utilizarmos esse critério na ancoragem, percebemos que muitas ficavam sem nenhuma categoria fixada, logo, não colocamos na tabela 7.

<sup>12</sup> Aceitamos o erro conceitual no gráfico 1 ao mantermos a continuidade da curva, porém, entendemos ser mais explicativo desta maneira, pois indica o quanto houve de crescimento da medida. Logo, os pontos que compõem a curva que vai de 2013 até 2018 forma uma linha de tendência dado as médias dos dois anos citados.

É importante destacar que essas médias estão relacionadas com a percepção dos estudantes que responderam o questionário em cada ano de aplicação, como também, nas diferentes definições impostas em cada ano de aplicação dos questionários. No caso dos questionários do **Momento 1**, temos uma certa perda de informações, pois temos uma quantidade menor de itens comparada com a do **Momento 2**. Obviamente, o crescimento na média do clima escolar “senso de segurança” deve estar ligado aos esforços de aumentar a segurança nas escolas.

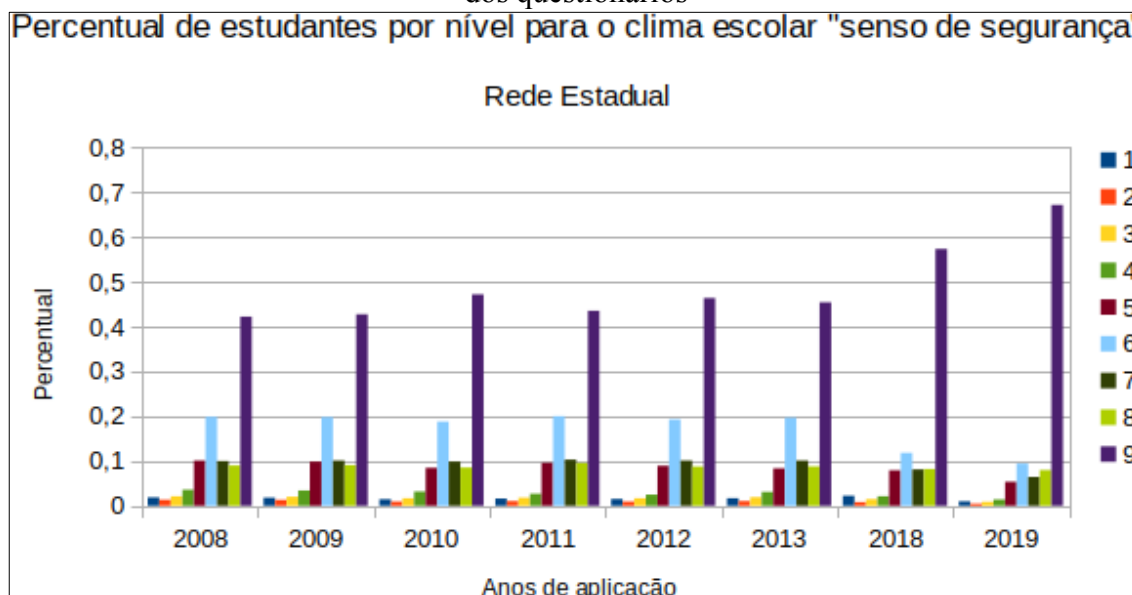
Para além dos resultados compostos pelas médias, fizemos a distribuição dos estudantes pelos 9 níveis definidos acima apresentados na tabela 8 e no gráfico 2 a seguir.

Tabela 8 – Distribuição dos estudantes nos níveis

Distribuição dos percentuais de estudantes em cada nível ao longo dos anos de aplicação – Rede Estadual								
níveis	2008	2009	2010	2011	2012	2013	2018	2019
1	0,02	0,02	0,01	0,02	0,02	0,02	0,02	0,01
2	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00
3	0,02	0,02	0,02	0,02	0,02	0,02	0,01	0,01
4	0,04	0,03	0,03	0,03	0,02	0,03	0,02	0,01
5	0,10	0,10	0,08	0,10	0,09	0,08	0,08	0,05
6	0,20	0,20	0,19	0,20	0,19	0,20	0,12	0,09
7	0,10	0,10	0,10	0,10	0,10	0,10	0,08	0,06
8	0,09	0,09	0,09	0,10	0,09	0,09	0,08	0,08
9	0,42	0,43	0,47	0,44	0,46	0,45	0,57	0,67

Fonte: Elaborado pelo autor (2021).

Gráfico 2: Distribuição dos percentuais de estudantes em casa nível nos anos de aplicação dos questionários



Fonte: Elaborado pelo autor (2021).

Podemos dizer que três pontos na tabela nos chamam a nossa atenção, a saber: Primeiro, por mais que as médias tenham aumentado ao longo dos anos, ainda temos uma variação ao longo dos anos de percentuais de estudantes de 0,01 a 0,02 no nível 1. Isto indica que temos, por exemplo, 6757 estudantes têm percepções de que “as pessoas zombam deles, estragam e roubam suas coisas muitas vezes ou quase sempre e já tomaram dinheiro ou alguma coisa. Elas as ameaçam ou intimidam, atacam ou agredem fisicamente muitas vezes ou quase sempre e fazem discriminação. Como também, fazem xingamentos, excluem deliberadamente e espalham boatos desagradáveis ao seu respeito muitas vezes ou quase sempre”. Algo que nos

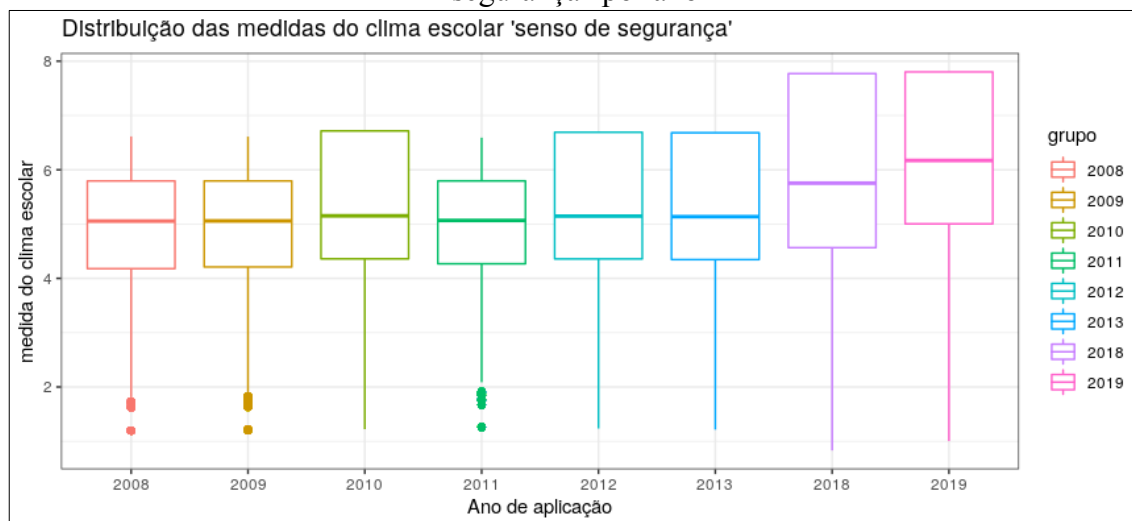


apresenta uma realidade “dura”, mesmo em um cenário onde de 2008 a 2019 houve um crescimento de 2,09 pontos na escala.

O segundo ponto, se refere ao aumento de percentuais nos níveis 5 e 6, diferente dos demais níveis. Já o terceiro ponto é o fato de termos os maiores valores percentuais no nível 9, isso indica que para a grande maioria dos estudantes, a percepção sobre o “senso de segurança” nas escolas é adequada, onde nunca situações que possam ferir física ou emocionalmente ocorram. Contudo, a quantidade de itens pode interferir negativamente na percepção de algumas situações, principalmente, de situações mais sutis ou menos agressivas, mas que podem diferenciar escolas que realizam intervenções pedagógicas para melhorar o clima escolar da escola em seus diversos aspectos. Isso seria resolvido com a introdução de uma quantidade maior de itens, principalmente, de itens que façam a escala ser ampliada de tamanho.

O gráfico 3, apresenta um diagrama de caixa, conforme Barbetta (2017, p.106 e 107) define a importância desse diagrama para “apresentar aspectos relevantes de uma distribuição de frequência”. É possível observar que as medianas para o **Momento 1** são praticamente iguais, porém, nos anos de 2010, 2012 e 2013. Contudo, nestes mesmos anos, as concentrações entre as medianas e o quartis superiores são maiores que nos demais anos do **Momento 1**. Já no **Momento 2** temos uma mediana maior em 2019, como também, uma maior concentração dos estudantes entre os quartis inferior e superior.

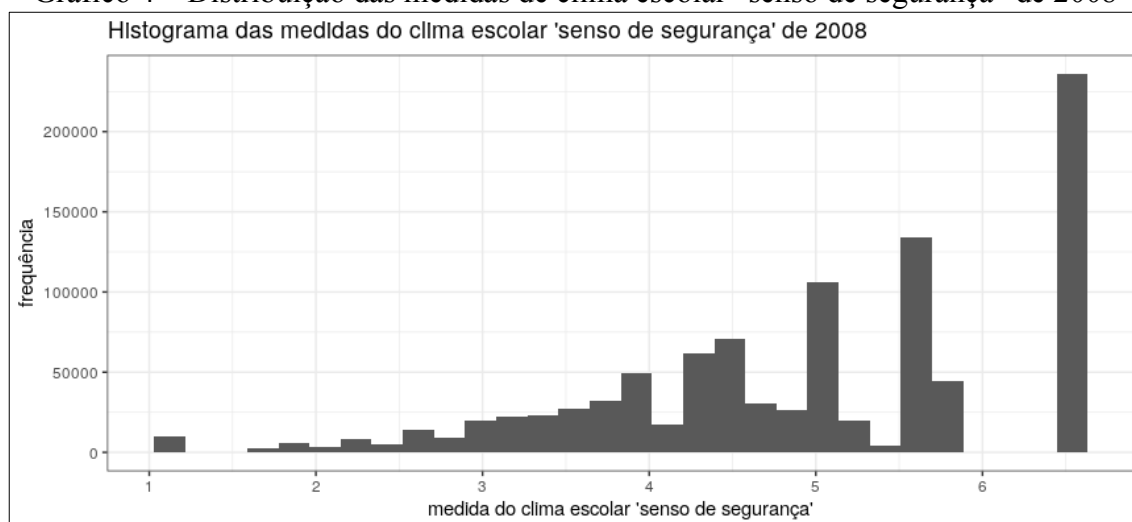
Gráfico 3 – Diagrama de caixa (boxplot) da distribuição da medida do clima escolar “senso de segurança” por ano



Fonte: Elaborado pelo autor (2021).

Um outro ponto importante a ser destacado, diz respeito ao fato de cinco dos oito anos analisados terem uma concentração dentro dos limites do quartil superior, ao mesmo tempo em que apresentam grande dispersão das medidas de clima escolar “senso de segurança” no extremo inferior. Entretanto, somente nos anos de 2008, 2009 e 2011, temos a presença de valores que ficam abaixo de  $Q_1 - 1,5 \cdot d_q$  e são conhecidos como outliers. Na perspectiva de ampliarmos estas análises, elaboramos dois gráficos de histograma, um para o ano de 2008 e o outro para o ano de 2019, com o intuito, a partir das distribuições observadas no gráfico 3, de aprofundarmos as análises sobre a distribuição das medidas de clima escolar “senso de segurança”.

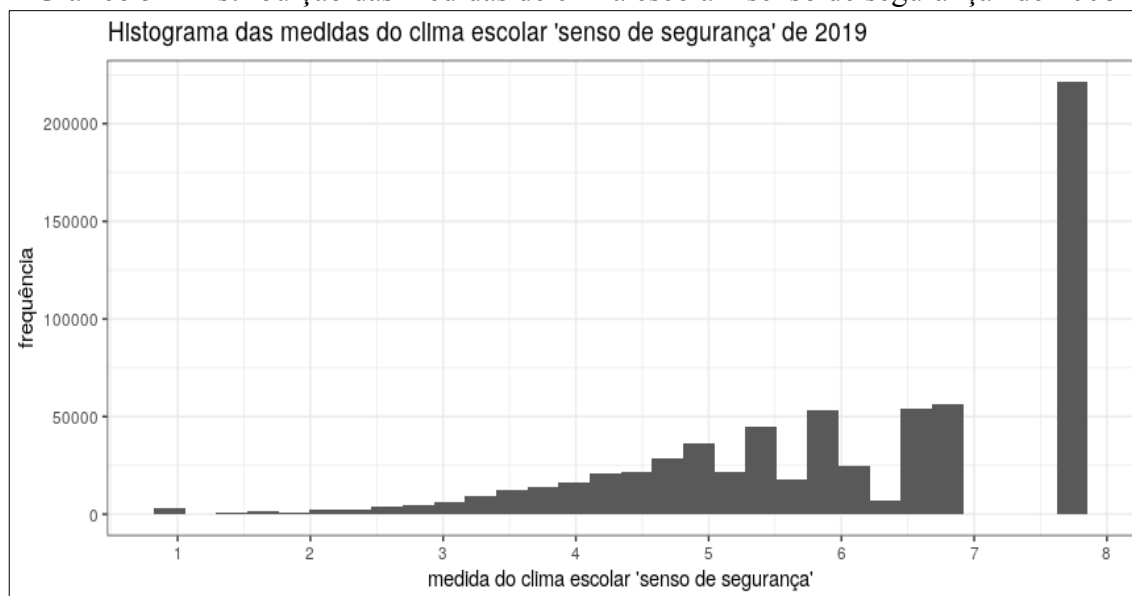
Gráfico 4 – Distribuição das medidas de clima escolar “senso de segurança” de 2008



Fonte: Elaborado pelo autor (2021).

Em 2008 as medidas do clima escolar “senso de segurança” apresentam uma concentração, se comparada com os pontos anteriores, maior de estudantes com medida acima de 6 pontos da escala, algo que indica uma percepção, por parte dos estudantes, de haver um bom “senso de segurança” nas escolas naquele ano. Contudo, não podemos deixar de notar que existe um pequeno grupo de estudantes com a medida completamente oposta às medidas deste grupo “superior”.

Gráfico 5 - Distribuição das medidas de clima escolar “senso de segurança” de 2008



Fonte: Elaborado pelo autor (2021).

Já no gráfico 5, temos a representação dos dados de 2019 onde é possível observar uma concentração maior de estudantes com medidas do clima escolar “senso de segurança” com valores maiores que 7,5 pontos na escala. Em contra partida, existe uma certa uniformidade entre os resultados dos estudantes que apresentam valores entre 5,5 e 7. Obviamente, como é possível constatar no gráfico 5, existem pontos entre o intervalo citado acima em que a quantidade de estudantes apresenta uma queda acentuada, contudo, se compararmos com os dados de 2008, podemos perceber uma tendência a uma quantidade constante por volta de 50.000 estudantes, diferente do que ocorre em 2008, em que as quantidades vão aumentando conforme os valores da medida aumenta.

## 6 CONCLUSÃO

A elaboração de uma medida sobre um instrumento que já foi constituído e aplicado se apresenta como um grande desafio, pois pouco ou quase nada pode ser feito para modificar discrepâncias encontradas nas análises estatísticas, causadas por problemas oriundos da formulação do instrumento. Contudo, percebemos que a medida de clima escolar ‘senso de segurança’ nos pareceu relevante e com grandes potencialidades, principalmente, se houver a possibilidade de se interferir no instrumento, com a ampliação da quantidade de itens e na diversificação deles, como também, na elaboração de uma Matriz de Referência para a elaboração dos itens.

Em relação ao aumento da quantidade de itens, esta, por sua vez, está intrinsecamente ligada a necessidade de se medir adequadamente, principalmente, ao longo de um grande período. Dessa maneira, ter poucos itens contribuí para criar incertezas que acabam interferindo na estimativa da medida, principalmente, se precisarmos ampliar ou modificar itens ao longo desse tempo. Já a diversificação surge da necessidade de termos uma medida melhor qualificada, com um rol maior de itens e com diversificação na maneira de se elaborar o item sobre determinados assuntos, incluindo itens que façam isso de maneira indireta.

Outro ponto importante é acrescentar outras dimensões sobre clima escolar, de uma maneira que se tenha um entendimento mais amplo do que venha ser o “clima escolar” e suas potencialidades e deficiências. Acreditamos que as relações dos estudantes com os próprios estudantes, com os professores, funcionários e direção, como também, as informações sobre infraestrutura e, até mesmo, a maneira como os professores realizam suas avaliações, podem ser úteis para o aprimoramento dessa polissemia que é o clima escolar.

Já em relação aos tramites de aplicação, julgamos ser importante estabelecer critérios claros para esse processo, como também, se necessário, retirar o procedimento de aplicação do mesmo dia da aplicação do Saresp, pois isso pode influenciar a dedicação dos estudantes em responder aos itens. Dessa maneira, teremos um instrumento com uma quantidade menor de variáveis externas que afetam a medida do clima escolar.

Por fim, existem outras possibilidades de análises a partir da estimação das medidas do clima escolar “senso de segurança”, como por exemplo, realizarmos uma análise de Fatores Associados, buscando compreender os fatores que podem interferir positivamente ou negativamente na percepção dos estudantes sobre essa dimensão do clima escolar. Contudo, essa possibilidade ficará como proposta para novos estudos.

## REFERÊNCIAS

- ANDRADE, Dalton Francisco de; TAVARES, Héilton Ribeiro; VALLE, Raquel da Cunha. **Teoria da Resposta ao Item: conceitos e aplicações**. São Paulo: Associação Brasileira de Estatística, 2000.
- BARBETTA, Pedro Alberto. **Estatísticas aplicadas às Ciências Sociais**. 9ª edição. Florianópolis: Editora da UFSC, 2017.
- BIRNBAUM, Allan. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In LORD, Frederic M.; Lord & NOVICK, Melvin R. **Statistical Theories of Mental Test Scores**. Reading, MA: Addison-Wesley. 1968. páginas 395-464.
- COHEN, Jonathan; PICKERAL, Terry; McCLOSKEY, Molly. The challenge of assessing school climate. **Educational Leadership**. [s.l.: s.n.] v. 66, n. 4, dec.2008/jan.2009.
- LORD, Frederic. M. **Applications of Item Response Theory to Practical Testing Problems**. [s.n.]. New York: Routledge Taylor & Francis Group, 1980.
- MATOS, Daniel Abud Seabra; RODRIGUES, Erica Castilho. **Análise fatorial**. Brasília: Enap, 2019.
- MORO, Adriano. **A construção e as evidências de validade de instrumentos de medida para avaliar o clima escolar**. 2018. Tese (doutorado). Universidade Estadual de Campinas, Faculdade de Educação. Campinas, SP: [s.n], 2018.
- MUÑIZ, José. **Introducción a la Teoría de Respuesta a los Ítems**. Madrid: Ediciones Pirámide, 1997.
- PASQUALI, Luiz (Org.). **Teoria e Métodos de medida em ciências do comportamento**. Brasília: Laboratório de Pesquisa em Avaliação e Medida / Instituto de Psicologia / UNB: Inep, 1996.
- PASQUALI, Luiz. **Psicometria: teorias dos testes na psicologia e na educação**. 5ª Edição. Petrópolis: Vozes, 2013.
- SAMEJIMA, Fumi. **Estimation of Latent Ability Using a Response Pattern of Graded Scores**. Princeton: Educational Testing Service Princeton, 1968.