

**Autoavaliação Institucional da Universidade Federal de Santa Catarina:
uma análise por meio da Teoria de Resposta Ao Item**

Maria Luciana Biondo Silva

Universidade Federal de Santa Catarina (UFSC)

e-mail: marialucianabiondosilva@gmail.com

Rafael da Silva Fernandes

Universidade Federal Rural da Amazônia (UFRA)

e-mail: rafasfer2@ufra.edu.br

Antônio Cezar Bornia

Universidade Federal de Santa Catarina (UFSC)

e-mail: cezar.bornia@ufsc.br

Dalton Francisco de Andrade

Universidade Federal de Santa Catarina (UFSC)

e-mail: dalton.andrade@ufsc.br

Resumo

A busca pela qualidade é presente na realidade das instituições de ensino superior. O sistema nacional de avaliação da educação superior tem o propósito de assegurar um padrão de qualidade para a educação superior elencando diretrizes para que comissões próprias de avaliação realizem a autoavaliação institucional na comunidade acadêmica com o propósito de identificar o grau de concordância desta para com as dimensões do sistema nacional de avaliação da educação superior. Desta forma, o presente estudo tem por objetivo a validação e interpretação de uma escala de medida para mensurar a perspectiva de concordância ou discordância de pós-graduandos em relação à autoavaliação institucional da Universidade Federal de Santa Catarina. O instrumento composto por 16 itens foi aplicado em 2019 e 2.709 respostas foram analisadas. Para tanto realizará um comparativo com as informações obtidas pela análise da teoria da resposta ao item com as informações divulgadas pela comissão própria de avaliação no relatório de autoavaliação institucional de 2019. O resultado foi a validação e interpretação da escala de medida proporcionando mais informações para tomada de decisão e melhor entendimento acerca do grau de concordância dos discentes.

Palavras-Chave: Teoria da Resposta ao Item; Autoavaliação Institucional; Sistema Nacional de Avaliação da Educação Superior.

Abstract

The quest for quality is present in the reality of higher education institutions. The objective of national system of evaluation of higher education is to ensure a quality standard for higher education through the elaboration of guidelines so that its own evaluation committees carry out institutional self-evaluation in the academic community with the purpose of identifying the degree of agreement with the dimensions of national system of evaluation of higher education. Thus, the present study has the objective of validating and interpreting a measurement scale to measure the perspective of success or failure of postgraduate students in relation to the institutional self-evaluation of the Federal University of Santa Catarina. In 2019, the instrument composed of 16 items was applied and 2,709 responses were analyzed. For this, a comparison will be made between the information obtained when analyzing the answer theory on the item with the information disclosed by the evaluation committee itself in the AAI 2019 institutional self-assessment report. The result was the validation and interpretation of the measurement scale, providing more information for the make decisions. and a better understanding of the student's degree of agreement.

Keywords: Item Response Theory; Institutional Self-Assessment; National System of Evaluation of Higher Education.

1. Introdução.

A busca pela qualidade da educação superior no Brasil data de 1980, quando foram idealizados instrumentos para regulamentá-la e para avaliá-la. Um dos mecanismos mais recentes, e em vigor, é o Sistema Nacional de Educação Superior (SINAES), que consiste em um conjunto de critérios por meio dos quais as Instituições de Ensino Superior (IES) são avaliadas. Dentre este conjunto de avaliações está a Autoavaliação Institucional (AAI) (ARAUJO; OLIVEIRA DIAS, 2021; BRASIL, 2004a).

A AAI é um instrumento que pode ser utilizado pelas IES para que obter informações para a tomada de decisão que fomenta o seu planejamento (NUNES *et al.*, 2017). A construção do instrumento está relacionada ao conceito de qualidade orientada por padrões, tendo em vista que os processos auto avaliativos representam uma forma de repensar de maneira continuada todos os demais processos desde macro ao nível de tarefa da IES. No âmbito institucional, os produtos/serviços devem estar em conformidade com uma missão/visão declarada ou um conjunto de especificações, requisitos ou padrões, incluindo aqueles definidos por órgãos acreditadores e/ou reguladores (SCHINDLER *et al.*, 2015).

Desta forma, dado que a AAI deve ser conduzida pelas diretrizes do SINAES, instituído pela Lei nº 10.861/2004, compete a Comissão Própria de Avaliação (CPA) elaborar o instrumento de forma a obter a percepção da comunidade acadêmica acerca de elementos que sejam considerados críticos para a qualidade institucional (BERNARDES; ROTHEN, 2016).

No contexto da instituição analisada, os processos de avaliação interna, conduzidos pela CPA, constituem-se em partes essenciais do ciclo da melhoria contínua, sendo que o processo de autoavaliação institucional considera as dimensões do SINAES organizadas nos cinco eixos estabelecidos pela Nota Técnica INEP/DAES/CONAES n.º 065. Nesta perspectiva, a compreensão das autoavaliações como processos dinâmicos e contínuos tem como metas: 1) identificar os pontos fortes e fracos referentes às práticas institucionais; e 2) adequar-se cada vez mais aos cenários externo e interno em que se encontra a instituição (BRASIL, 2004b).

Como a comunidade acadêmica possui muito indivíduos, é imprescindível uma escolha adequada de um método estatístico específico para realizar as análises dos dados obtidos (ARRUDA *et al.*, 2019). Desta forma, dado que o propósito do instrumento de medida deve ser conduzido por um único objetivo de medição, ou seja, única e exclusivamente para medir o traço latente, o conteúdo e forma do teste devem ser especificadas para fornecer evidências e apoiar as interpretações para os usos pretendidos. Em outras palavras, quando as pontuações de um único teste interpretadas de forma diferente, para tomar diferentes tipos de decisão, é aceito que cada interpretação pretendida deve ser validada de forma independente (AERA *et al.*, 2014).

Dentre os métodos estatísticos para mensurar traços latentes a psicometria utiliza a Teoria Clássica das Medidas (TCM), também conhecida por Teoria Clássica dos Testes (TCT), assim como a Teoria de Resposta ao Item (TRI) (PASQUALI; PRIMI, 2003). Para os propósitos desta pesquisa, a TRI permite analisar o desempenho dos itens, assim como o grau de dificuldade de cada um, não sendo, portanto, considerados igualmente difíceis, pontos estes em que se diferencia da TCM. Além disso, na TRI a probabilidade de que um determinado indivíduo com um certo nível de uma característica específica tenha um nível desse traço é determinado pela verdadeira medida do traço da pessoa e dois parâmetros de item de mensuração, isto é, discriminação e dificuldade (ANDRADE *et al.*, 2000).

Neste âmbito, o escopo deste trabalho foi delimitado aos pós-graduandos da Universidade Federal de Santa Catarina (UFSC) no ano de 2019 e tem por objetivo a validação e interpretação de uma escala de medida para mensurar a perspectiva de concordância ou discordância de pós-graduandos em relação à autoavaliação institucional da UFSC.

2. Revisão da literatura.

Para uma melhor compreensão do leitor acerca da análise apresentada pelos autores, na revisão de literatura são apresentados os conhecimentos sobre a autoavaliação institucional, que é um pilar da gestão da qualidade do ensino superior, assim como da teoria de resposta ao item, amplamente aplicada na inteligência em dados em diversas áreas.

2.1. Instrumento de Autoavaliação Institucional (AAI).

A avaliação interna da IES é nominada Autoavaliação Institucional (AAI). Em 2004 o INEP elaborou e disponibilizou um roteiro para orientar a (AAI) das IES. Conforme consta neste roteiro, são etapas da AAI: [1] Preparação que compreende a constituição da CPA e o planejamento; [2] Desenvolvimento que compreende a execução; e [3] Consolidação que processa os dados obtidos na AAI para a elaboração do relatório que deverá ser divulgado (BRASIL, 2004b).

Desta forma, o instrumento de medida será representado por uma estruturação hierárquica composta de fatores primários sendo, cada um deles, desagregado em vários atributos de interesse ou itens de mensuração. No contexto deste trabalho, os fatores primários são os cinco eixos do SINAES, a saber: [I] Planejamento e Avaliação Institucional; [II] Desenvolvimento Institucional; [III] Políticas Acadêmicas; [IV] Políticas de Gestão; e [V] Infraestrutura Física.

O instrumento é composto por 16 itens com seis alternativas de respostas conforme a escala de concordância do tipo Likert, ordenadas da seguinte forma: [1] Discordo totalmente; [2] Discordo parcialmente; [3] Não concordo, nem discordo; [4] Concordo parcialmente; [5] Concordo totalmente; e [NA] “Desconheço/Não se aplica” para ausência de resposta. A Tabela 1 apresenta descrição de cada item e sua respectiva associação ao Eixo SINAES.

Tabela 1 – Itens elaborados pela CPA para a AAI aplicada aos discentes de pós-graduação.

Eixo	id	Descrição
I	1	Os resultados das avaliações internas (CPA) são discutidos em seu curso
II	2	O seu curso contribui para o desenvolvimento da Missão da UFSC.
II	3	A UFSC tem contribuído para a inclusão social, preservação do meio ambiente, desenvolvimento local e regional, e preservação da memória e patrimônio cultural.
III	4	As políticas de ensino e as ações acadêmico-administrativas para o seu curso são suficientes.
III	5	A política de pesquisa em relação ao seu curso é suficiente.
III	6	As informações sobre a instituição, tais como: divulgação dos cursos, da pesquisa e da extensão; existência de mecanismos de transparência institucional; ouvidoria; resultados das avaliações recentes; entre outros são acessíveis à comunidade EXTERNA.
IV	7	A atuação do Conselho Universitário (Cun) é eficiente.
IV	8	A atuação dos órgãos colegiados relacionados ao seu Programa é eficiente.
IV	9	As ações da coordenação do seu Programa são eficientes.
IV	10	Você conhece a previsão orçamentária e o cronograma de execução previstos no Plano de Desenvolvimento Institucional (PDI) 2015-2019 da UFSC.
IV	11	As fundações de apoio da UFSC contribuem adequadamente para o desenvolvimento das atividades de ensino, pesquisa e extensão.
V	12	A infraestrutura das salas de aula no que se refere à suficiência do espaço, disponibilidade de equipamentos audiovisuais, limpeza, iluminação, acústica, ventilação, conservação e comodidade é suficiente.
V	13	A infraestrutura dos laboratórios/áreas experimentais no que se refere a equipamentos e espaços físicos, acesso à internet, atualização de equipamentos e softwares, disponibilidade de insumos, apoio técnico e manutenção é suficiente.
V	14	O acervo disponível para a realização das suas atividades acadêmicas no tocante à qualidade e a quantidade do acervo disponível nas bibliotecas (central/setoriais/polo EaD), no acervo virtual e nas bases de dados é suficiente.

V	15	A qualidade dos serviços prestados nas bibliotecas no que se refere ao acesso via internet (consulta e reserva), informatização do acervo, bancos de dados, empréstimo e horário de funcionamento é suficiente.
V	16	A qualidade dos espaços de convivência e de alimentação no que se refere à dimensão, limpeza, iluminação, ventilação e conservação é suficiente.

Fonte: (BRASIL, 2019).

2.2. Teoria de Resposta ao Item (TRI).

Os modelos derivados da teoria de resposta ao item têm sido amplamente usados na literatura, nas mais diferentes áreas. Especificamente, com aplicações na área educacional, avaliações psicológicas e de saúde. Dada sua vantagem, em relação à teoria clássica de medida, suposições sólidas precisam ser realizadas quanto à sua aplicabilidade. Tipicamente, há duas premissas que precisam ser fortemente avaliadas: a unidimensionalidade e independência local (EDWARDS *et al.*, 2018).

A teoria de resposta ao item (TRI) é composta de um conjunto de modelos probabilísticos para mensuração de um traço latente representado por (θ) , que utiliza de um a três parâmetros para indicar a probabilidade de um respondente escolher uma determinada resposta em um determinado item para mensurar esse traço que não pode ser medido diretamente (LORD, 2012). Os itens são geralmente classificados como dicotômicos o que significa que existem apenas duas categorias de resposta (normalmente denominadas "correto" e "incorreto"), ou politômicos, o que significa que as respostas podem ser classificadas em várias categorias (BROSSMAN; LEE, 2010).

A TRI foi originalmente desenvolvida em psicometria para avaliação de habilidades, principalmente para fins educacionais, ou seja, analisar itens com respostas corretas e incorretas (DE AYALA, 2013). No entanto, o termo 'item' é genérico e, como tal, se aplica a declarações de atitude para as quais os respondentes são solicitados a indicar seu nível de concordância. A TRI modela a relação entre a probabilidade de endossar uma determinada resposta a um item e a pontuação de um indivíduo em um traço latente (subjacente não medido) (RECKASE, 2009).

Ao contrário do foco no nível de teste da teoria clássica da medida, a TRI concentra-se no desempenho dos itens. Assim, ao passo que a Teoria Clássica da Medida (TCM) assume que os itens do teste são réplicas uns dos outros (isto é, medidas paralelas), a TRI não assume que os itens do teste sejam igualmente difíceis. Além disso, na TCM é assumido que a pontuação obtida por uma pessoa em um teste é a soma de uma pontuação verdadeira e uma pontuação de erro, enquanto na TRI a probabilidade de que um determinado indivíduo com um certo nível de uma característica específica tenha um nível desse traço é determinado pelo verdadeiro nível do traço da pessoa e dois parâmetros de item de medição, isto é, discriminação e dificuldade (ANDRADE; TAVARES *et al.*, 2000).

O modelo de resposta gradual (MGR) de (SAMEJIMA, 1969) se utiliza das alternativas em escala *Likert* para ordenar as categorias de tal forma que seja possível estimar a categoria mais provável de ser escolhida pelo respondente em determinado intervalo de escore. Este modelo tenta obter mais informação das respostas dos indivíduos do que simplesmente se eles deram respostas corretas ou incorretas, ou ainda sim, ou não.

Neste modelo os escores das categorias dum item (i) são organizados em posição crescente de tal forma que é possível representa-los por $k = 0, 1, \dots, m_i$ onde $(m_i + 1)$ é o número de categorias do i -ésimo item (ANDRADE; TAVARES *et al.*, 2000). A probabilidade de um respondente j responder um valor ou outra mais alta do item i pode ser dada pela extensão do modelo logístico de 2 parâmetros da equação:

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

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 0, 1, \dots, m_i$, onde $b_{i,k}$ é o parâmetro de dificuldade da k -ésima categoria do item i .

Note que um item com $(m_i + 1)$ categorias, m_i valores de dificuldade necessitam ser estimados, além do parâmetro de discriminação do item. Assim, para cada item, o número de parâmetros

a ser estimado será dado pelo seu número de categorias de resposta. Se, por exemplo, tivermos um teste com I itens, cada um com $(m_i + 1)$ categorias de resposta, teremos então $[\sum_{i=1}^I m_i + I]$ parâmetros a serem estimados.

3. Metodologia.

Nesta seção será apresentada a identificação da oportunidade de pesquisa, a caracterização do estudo, procedimentos realizados pela IES e obtenção do banco de dados pelos autores, procedimentos metodológicos adotados pelos autores, procedimentos para interpretação pedagógica da escala.

3.1. Identificação da oportunidade de pesquisa.

A ideia do estudo surgiu quando a autora principal, que é mestranda na IES, recebeu uma mensagem para participar da AAI. Ao responder o instrumento de pesquisa, verificou que os itens elaborados permitiriam o processamento das informações por meio da resposta gradual de Samejima, que é um modelo da TRI.

Em consulta ao relatório com os resultados do ano de 2019, foi verificado que a CPA da IES em que a AAI foi aplicada executou os procedimentos teóricos e os procedimentos experimentais conforme preconizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) para atender ao determinado pelo SINAES. Ainda no referido relatório foi possível compreender que a metodologia utilizada para o processamento das respostas foi a prevista na TCM. A informação referente ao processamento das respostas foi confirmada por e-mail com CPA, quando se constatou que nem todos os procedimentos da TCM foram utilizados, mas que se tratava de um estudo de teor psicométrico, pois se buscava mensurar um traço latente, e que o modelo utilizado era o da TCM.

3.2. Caracterização do estudo.

O estudo é caracterizado como um estudo de caso, pois é aprofundado no contexto de uma IES pública federal situada na região sul do Brasil, quantitativa, pois os dados a serem analisados por meio de métodos estatístico, conforme a metodologia preconizada pela psicometria para mensuração de traço latente, que prevê procedimentos teóricos, experimentais e analíticos. Esses procedimentos são necessários para, respectivamente, elaborar os itens e a escala de resposta do instrumento de pesquisa, aplicar o referido instrumento na população-alvo, e processar e analisar as referidas informações utilizando procedimentos estatísticos (PASQUALI, 2010).

3.3. Procedimentos realizados pela IES e obtenção do banco de dados pelos autores.

A CPA da IES realizou os procedimentos teóricos e os procedimentos experimentais, o que significa, respectivamente, a elaboração dos itens no instrumento de pesquisa, e a coleta dos dados.

Os procedimentos teóricos são os necessários para a construção do instrumento de pesquisa e que consistem na definição do construto; na revisão de literatura acerca do construto, que auxiliará na definição da população-alvo; na definição constitutiva, que define o que é o construto teoricamente; na definição operacional, que define como o construto se manifesta na prática em termos de comportamento momento em que as dimensões do construto são também identificadas; na análise dos itens elaborada por juízes, que são especialistas no construto para o qual os itens foram elaborados; na análise semântica realizada pelo público-alvo para verificar se os itens são adequados para o entendimento da população-alvo (PASQUALI, 2010).

Os procedimentos experimentais dizem respeito à aplicação do instrumento de pesquisa na população-alvo. O planejamento da aplicação do instrumento, que define onde, quanto, como e quando o instrumento será aplicado, tanto o pré-teste, quando o instrumento definitivo é um procedimento experimental. Onde significa a instituição e o meio utilizados para aplicar o instrumento. Quanto diz respeito ao quantitativa de respondentes que receberão o instrumento para responder, ou seja, trata-se da quantidade amostral. Como consiste na forma de abordagem dos respondentes. Quando é a definição

do início e fim do período de coleta das respostas. Outro procedimento experimental é a execução da aplicação do instrumento (PASQUALI, 2010).

A CPA da IES, a pedido dos autores, forneceu o banco de dados anonimizado contendo 2.730 respostas de discentes da pós-graduação do ano de 2019. A amostra utilizada neste estudo compreende, portanto, um total de 2.730 discentes da pós-graduação da IES no ano de 2019, que participaram fornecendo os dados coletados de forma voluntária.

3.4. Procedimentos metodológicos realizados pelos autores.

Os pesquisadores ajustaram a numeração da escala de medida para leitura no software R, de forma a viabilizar o processamento da informação pelo programa, sem comprometer a fidedignidade do estudo. O ajuste consistiu em substituição da numeração das opções de repostas da escala *Likert* utilizada pela IES, sendo a opção menos favorável renomeada com zero, e as opções mais favoráveis, renomeadas por um, por dois, e assim sucessivamente. Além disto um processo prévio nos dados foi realizado para identificar respostas inconsistentes, por exemplo, quando um único respondente seleciona nos 16 itens a opção “NA”. Após o processamento, foram removidas 21 respostas e 2.709 respostas foram analisadas neste estudo. Caso outros pesquisadores tenham interesse em analisar os dados utilizados neste estudo, poderão contatar a CPA da IES para solicitar acesso ao referido banco de dados.

Os procedimentos analíticos, foram submetidos a um novo processamento pelos autores, conforme previsto na psicomетria para análise pela TRI. A seguir serão apresentadas, em tópicos, as análises realizadas pelos autores, assim como a indicação dos métodos utilizados.

Para a análise descritiva dos itens e estimativas de confiabilidade foram aplicados os seguintes métodos estatísticos: correlação de cada item com a pontuação total, não corrigida para sobreposição dos itens, a mesma correlação, porém, considerando a possibilidade de todos os itens serem padronizados, correlação total do item corrigida para sobreposição de itens e confiabilidade de escala e correlação total do item para este item em relação à escala sem este item, coeficiente alfa baseado nas covariâncias, coeficiente de alfa padronizado com base nas correlações, confiabilidade Lambda 6 de Guttman, análise pelo valor de Kaiser – Meyer – Olkin, e teste de Bartlett;

Para a análise fatorial e dimensionalidade foram aplicados os seguintes métodos estatísticos: análise paralela de informação restrita baseada na correlação tetracórica, análises de cargas fatoriais e comunalidade, *Root-mean-square error of approximation* (RMSEA), *Standardized root mean square residuals* (SRMSR), *comparative fit index* (CFI) e *Tucker Lewis Index* (TLI), “*Jackknife Slope Index*” (JSI): para modelos unidimensionais.

Para a calibração dos parâmetros dos itens foi aplicado o modelo de resposta gradual de Samejima usando o programa R para processamento. Foram inseridos no programa R uma planilha contendo os dados e uma sintaxe que contém os comandos para processamento, ou seja, realizar os procedimentos estatísticos que permitam a análise por meio do referido modelo da TRI;

3.5. Procedimentos para interpretação pedagógica da escala.

Para a elaboração da interpretação pedagógica da escala de medida foi analisado o posicionamento dos itens âncora para determinação dos níveis âncora, de forma a obter informação acerca do posicionamento das respostas de acordo com a métrica do traço latente. O agrupamento dos itens em níveis será demonstrado por meio da distribuição de frequência, assim como do posicionamento os itens nos níveis. A descrição dos níveis é a interpretação pedagógica da escala em que informa, em cada nível, como o traço latente é operacionalizado.

Para que os itens pudessem ser identificados, tanto no agrupamento de itens por nível, como na descrição da interpretação pedagógica da escala, foi adotada a seguinte como estrutura de nomenclatura I (que representa a palavra item) + número do item (para que se identifique a qual item a resposta pertence com dois dígitos. Como são dezesseis itens, este número poderá ser 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 15, ou 16) + b (é o parâmetro de dificuldade que está na mesma medida do traço latente) + número do parâmetro de dificuldade de acordo com a escala *Likert* (Como a escala *Likert* do instrumento de coleta de dados possui 5 categorias, são produzidos 4 parâmetros de dificuldade. Na escala os respondentes que estiverem abaixo do parâmetro de dificuldade 1, optaram pela categoria 1;

os que estiverem no parâmetro de dificuldade 1 e abaixo do parâmetro de dificuldade 2, optaram pela categoria 2; os que estiverem no parâmetro de dificuldade 2 e abaixo do parâmetro de dificuldade 3, optaram pela categoria 3; os que estiverem no parâmetro de dificuldade 3 e abaixo do parâmetro de dificuldade 4, optaram pela categoria 4; os que estiverem no parâmetro de dificuldade 4 e acima dele optaram pela categoria 5). Outro recurso visual utilizado pelos autores foi destacar essa nomenclatura por uma cor que facilita identificar a qual parâmetro de dificuldade ele pertence. Vermelho destaca o parâmetro de dificuldade 1. Amarelo realça o parâmetro de dificuldade 2. Verde claro é atribuído ao parâmetro de dificuldade 3. Verde escuro é usado para o parâmetro de dificuldade 4

Para melhor compreensão, será apresentado um exemplo da aplicação do padrão de nomenclatura para cada caso.

- a) A resposta no item 2 com a categoria 1 selecionada, em que consta “discordo totalmente”, não terá um parâmetro de dificuldade que o represente, mas será identificado, pois estará abaixo do parâmetro de dificuldade 1 do item 2;
- b) A resposta no item 2 com a categoria 2 selecionada, em que consta “discordo parcialmente”, terá a seguinte nomenclatura: I (que representa a palavra item) + número do item (que é 02, considerando que deve ter dois dígitos) + b (é o parâmetro de dificuldade que está na mesma medida do traço latente) + número do parâmetro de dificuldade que será o 1 e destacado na cor vermelho. Desta forma a nomenclatura padronizada será **I02b1**. A partir de **I02b1** e antes de **I02b2**, portanto, está a resposta que selecionou no item 02 “discordo parcialmente”;
- c) A resposta no item 5 com a categoria 3, em que consta “não concordo, nem discordo”, terá a seguinte nomenclatura: I (que representa a palavra item) + número do item (que é 05, considerando que deve ter dois dígitos) + b (é o parâmetro de dificuldade que está na mesma medida do traço latente) + número do parâmetro de dificuldade que será o 2 e destacado na cor amarelo. Desta forma a nomenclatura padronizada será **I05b2**. A partir de **I05b2** e antes de **I05b3**, portanto, está a resposta que selecionou no item 05 “não concordo, nem discordo”;
- d) A resposta no item 12 com a categoria 4, em que consta “concordo parcialmente”, terá a seguinte nomenclatura: I (que representa a palavra item) + número do item (que é 12, considerando que deve ter dois dígitos) + b (é o parâmetro de dificuldade que está na mesma medida do traço latente) + número do parâmetro de dificuldade que será o 3 e destacado na cor verde claro. Desta forma a nomenclatura padronizada será **I12b3**. A partir de **I12b3** e antes de **I12b4**, portanto, está a resposta que selecionou no item 12 “concordo parcialmente”;
- e) A resposta no item 16 com a categoria 5, em que consta “concordo totalmente”, terá a seguinte nomenclatura: I (que representa a palavra item) + número do item (que é 16, considerando que deve ter dois dígitos) + b (é o parâmetro de dificuldade que está na mesma medida do traço latente) + número do parâmetro de dificuldade que será o 4 e destacado na cor verde escuro. Desta forma a nomenclatura padronizada será **I16b4**. A partir de **I16b4**, portanto, está a resposta que selecionou no item 16 “concordo totalmente”.

A seguir serão apresentados os resultados obtidos após o processamento estatístico dos dados, assim como as respectivas análises e as discussões serão apresentadas a seguir.

4. Resultados e discussões.

4.1. Análise descritiva dos itens e estimativas de confiabilidade.

As estatísticas descritivas dos itens são apresentadas na Tabela 2 e incluem a quantidade de respostas válidas (*n*), a correlação de cada item com a pontuação total, não corrigida para sobreposição dos itens (*raw.r*), a mesma correlação, porém, considerando a possibilidade de todos os itens serem padronizados (*std.r*), correlação total do item corrigida para sobreposição de itens e confiabilidade de

escala ($r.cor$) e, por fim, a correlação total do item para este item em relação à escala sem este item ($r.drop$). Neste caso, todos os itens variam pouco em relação à média indicando boa confiabilidade dos itens. Além disso, dado que a categoria intermediária é (3 – Nem concordo, nem discordo), apenas o item 10 apresenta média de repostas inferior à 3.

Tabela 2 – Análise descritiva dos itens.

itens	n ¹	raw.r ²	std.r ³	r.cor ⁴	r.drop ⁵	média	Sd ⁶
item ₀₁	1789	0,703	0,690	0,664	0,639	3,335	1,475
item ₀₂	2490	0,532	0,581	0,539	0,506	4,694	0,640
item ₀₃	2533	0,627	0,651	0,616	0,586	4,447	0,766
item ₀₄	2604	0,759	0,753	0,741	0,705	3,875	1,051
item ₀₅	2593	0,729	0,727	0,710	0,672	3,926	1,093
item ₀₆	2445	0,738	0,726	0,703	0,679	3,706	1,215
item ₀₇	1891	0,723	0,733	0,713	0,687	3,736	1,033
item ₀₈	2361	0,724	0,739	0,727	0,687	4,078	1,011
item ₀₉	2629	0,705	0,724	0,710	0,667	4,286	0,970
item ₁₀	1395	0,608	0,601	0,562	0,541	2,887	1,456
item ₁₁	2139	0,701	0,704	0,676	0,653	3,868	1,041
item ₁₂	2456	0,687	0,651	0,633	0,603	3,380	1,316
item ₁₃	2265	0,718	0,689	0,676	0,647	3,347	1,258
item ₁₄	2518	0,659	0,654	0,624	0,595	4,047	1,023
item ₁₅	2434	0,598	0,611	0,575	0,542	4,437	0,794
item ₁₆	2383	0,702	0,678	0,652	0,628	3,551	1,223
Média	2308	0,6825	0,6813	0,6869	0,6275	3,851	1,085

Os valores de confiabilidade gerais foram avaliados, como o coeficiente alfa baseado nas covariâncias ($std_{\alpha} = 0,9233206$), coeficiente de alfa padronizado com base nas correlações ($raw_{\alpha} = 0,9186758$), confiabilidade Lambda 6 de Guttman ($G6_{smc} = 0,9294651$).

As análises univariadas, as estimativas de confiabilidade indicaram uma correlação moderada/alta dos itens com o traço latente sendo, de outra forma, um indicativo de unidimensionalidade nos dados a ser verificado na análise de dimensionalidade. Além disso, como $G6_{smc} > std_{\alpha} > raw_{\alpha}$, este fato indica cargas fatoriais próximas. A matriz de correlação dos itens foi considerada fatorável pois o valor de Kaiser – Meyer – Olkin ($KMO = 0,94 \geq 0,5$) e teste de Bartlett foi significativo (Bartlett's $K - squared = 3.184,6$ e $p - value \approx 2.2 e^{-16}$).

4.2. Análise fatorial e dimensionalidade.

A primeira avaliação é o método de análise paralela de informação restrita, baseada na correlação tetracórica. Na Figura 1 é possível observar que há um indicativo de 5 fatores e 2 componentes. O gráfico indica um traço latente forte sugerindo que o instrumento pode ser considerado unidimensional. Foi calculado o primeiro autovalor cujo valor é 8,722 e representa uma proporção acumulada da variação explicada de 54,5%.

¹Quantidade de repostas válidas.

²Correlação de cada item com a pontuação total, não corrigida para sobreposição dos itens.

³Correlação de cada item com a pontuação total, porém, considerando a possibilidade de todos os itens serem padronizados.

⁴Correlação total do item corrigida para sobreposição de itens e confiabilidade de escala.

⁵A correlação total do item para este item em relação à escala sem este item.

⁶Desvio-padrão.

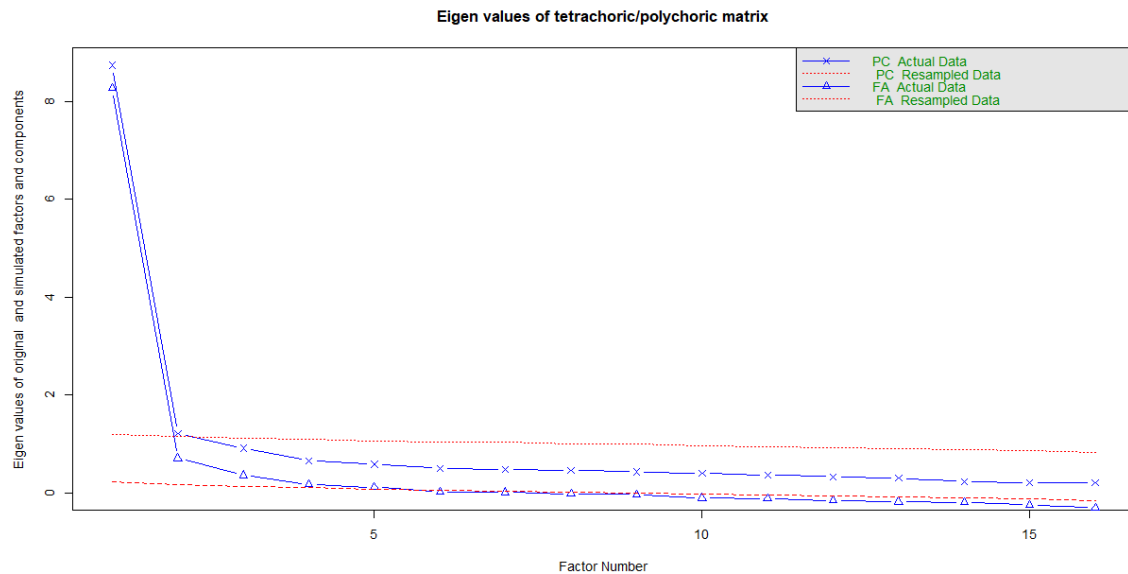


Figura 1 - Análise de componentes principais e análise fatorial.

Com o intuito de buscar sustentação para a hipótese de unidimensionalidade, é possível realizar análises de cargas fatoriais e comunalidade, tem seus resultados apresentados na Figura 2. Nota-se que todos os itens possuem uma carga fatorial (F1) maior que 0,30 para 1 dimensão, conforme sugerido por (PASQUALI, 2010), sendo a média das cargas fatoriais de 0,7362 e seus valores variando de 0,5928 (item_10) a 0,8159 (item_04). A comunalidade (h^2) também é aceitável neste conjunto de itens, variando entre 0,3514 (item_10) a 0,6657 (item_04) e média de 0,5451, indicando boa contribuição para a medida.

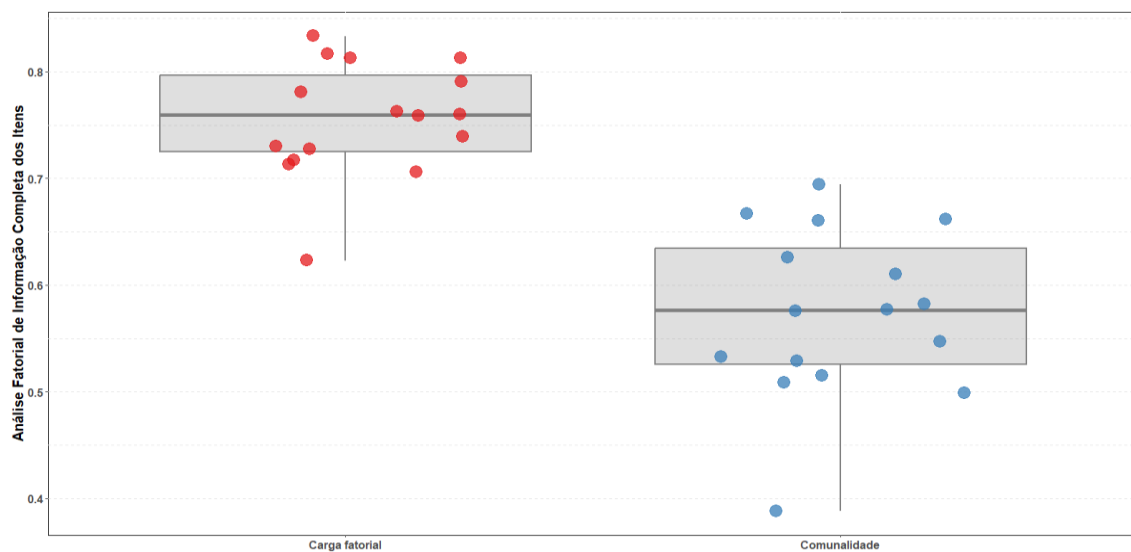


Figura 2 – Cargas fatoriais e comunalidades dos 16 itens do instrumento de medida para 1 e 2 dimensões.

Decidiu-se então avaliar a dimensionalidade pela aplicação de modelos multidimensionais, em particular, avaliando os dados para 1, 2 e 3 dimensões. A comparação dos modelos foi realizada pelas estatísticas RMSEA, SRMSR, CFI e TLI ao qual o modelo unidimensional obteve um RMSEA, TLI e CFI melhores (Tabela 3).

Tabela 3 – Análise de ajuste dos modelos.

Modelos	RMSEA ⁷	SRMSR ⁸	TLI ⁹	CFI ¹⁰
1 dimensão	0,122731	0,085406	0,903902	0,916715
2 dimensões	0,070286	0,066672	0,963046	0,972592
3 dimensões	0,070478	0,065532	0,968311	0,980194

Os modelos multidimensionais apresentam métricas que indicam uma qualidade de ajuste razoavelmente adequado, com ($RSMEA \leq 0.08$; $SRMSR \leq 0.08$; $CFI \geq 0.95$; $TLI \geq 0.95$). Já o modelo unidimensional possui o $RSMEA = 0,1227$ indicando um valor mais alto do erro quadrático médio de aproximação.

Além disso, a **suposição de independência local** foi avaliada pelo método “*Jackknife Slope Index*” (JSI): para modelos unidimensionais (EDWARDS; HOUTS *et al.*, 2018) e indicou que as respostas aos itens não estão relacionadas entre si, ou seja, são independentes.

Não obstante, é preciso considerar a presença de um fator dominante, explicando mais de 20% da variância dos dados e, segundo (RECKASE, 1979), permite a utilização de um modelo de TRI unidimensional, mesmo quando outros fatores estão presentes no construto. Além disso, em análises envolvendo a TRI, é comum manter itens que apresentam pouca representatividade e, de forma prática, não invalida a escolha modelo unidimensional.

4.3. Calibração dos parâmetros dos itens.

Os dados foram processados no programa R por meio de uma sintaxe estruturada para aplicar o MRG de Samejima para um processo chamado “calibração dos parâmetros dos itens” por meio do qual são definidos o parâmetro de discriminação, o parâmetro de dificuldade e o respectivo erro padrão associado a estimativa de cada parâmetro. No modelo de Samejima, utilizado para dados politômicos, cada item possui um parâmetro de discriminação (a) e parâmetro de dificuldade para estimar as probabilidades de cada categoria, ou categoria superior a ela. A primeira categoria comportará sempre a probabilidade de cem por cento de chance do respondente a escolher, de forma que os parâmetros de dificuldade serão estimados para as demais categorias, exceto a primeira. Como o instrumento de autoavaliação institucional em análise possui cinco categorias de resposta por item, foram estimados para cada item quatro parâmetros de dificuldade, respectivamente b_1 , para a segunda categoria, b_2 , para a terceira categoria, b_3 para a quarta categoria e b_4 para a quinta categoria.

A calibração dos parâmetros de discriminação e de dificuldade e seus respectivos erros padrão são apresentados na Figura 3. Para o parâmetro de discriminação todos os itens obtiveram valores $a > 1$ indicando que os respondentes variam adequadamente bem quanto ao seu traço latente. Além disso, os valores para os respectivos erros padrão variaram de 0,0619 (item₁₂) a 0,0903 (item₀₂) considerando que os valores são baixos e não apresentam alterações relevante, o que indica que as estimativas são consistentes.

⁷Root-mean-square error of approximation.

⁸Standardized root mean square residuals.

⁹Tucker Lewis Index.

¹⁰Comparative fit index.

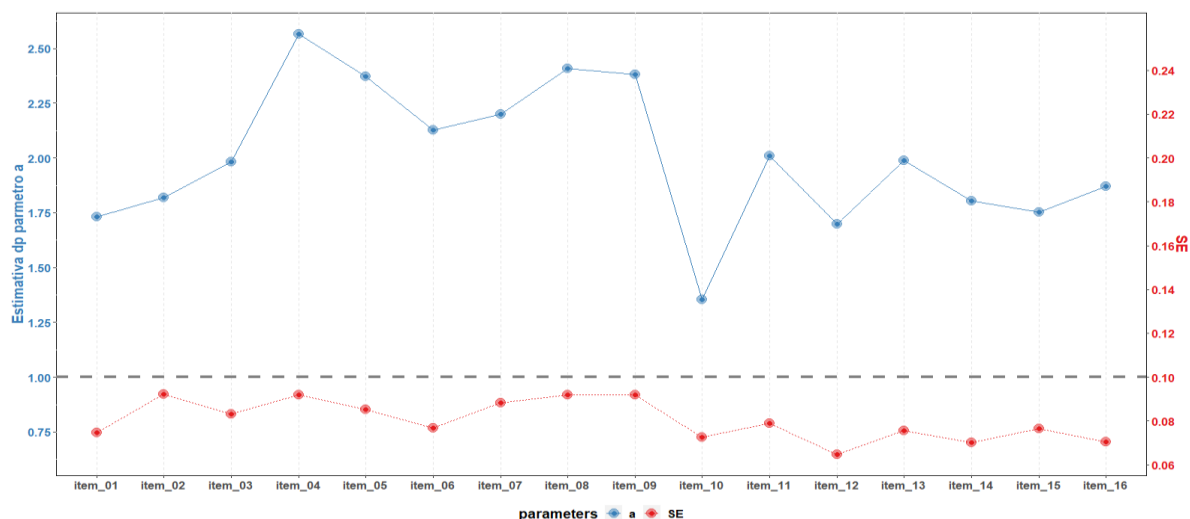


Figura 3 – Estimativas dos parâmetros de discriminação (a) de cada item e dos parâmetros de dificuldade (b) de cada categoria por item com o respectivo erro padrão (EP) escala (0,1).

Além da estimativa do parâmetro de discriminação, os parâmetros de dificuldade são apresentados na Figura 4. Nota-se que os itens (02, 03 e 15) são itens mais considerados mais discordantes e os itens (10, 12 e 13) mais concordantes na perspectiva dos discentes. Os erros padrões possuem valores aceitáveis que variam entre 0.05 e 0.15.

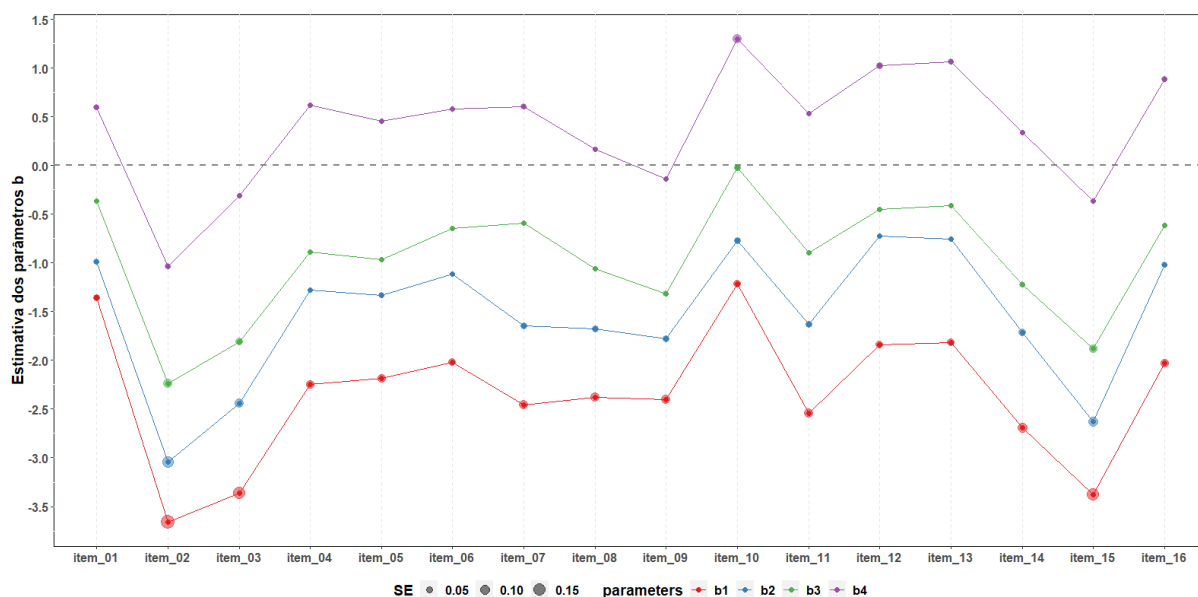


Figura 4 – Estimativa dos parâmetros de dificuldade.

Importante observar ainda o quão distantes os parâmetros de dificuldade de cada categoria estão distantes uns dos outros, pois se um parâmetro de uma categoria for muito próximo do de outra categoria, existe a possibilidade de que essa categoria nunca seja a mais provável de ser respondida. Por exemplo, o parâmetro b_1 e b_2 dos itens (01 e 10) são próximos, e suas diferenças são 0,3987 e 0,4825 respectivamente.

Nesse sentido, a análise das curvas características (CCI) são úteis ao fornecerem interpretações quanto a propensão dos respondentes a responderem uma determinada categoria em relação ao traço latente. Para um melhor entendimento P0 representa a categoria “Discordo totalmente”, P1 a categoria “Discordo parcialmente”, P2 a categoria “Não concordo, nem discordo”, P3 a categoria “Concordo parcialmente” e P4 a categoria “Concordo totalmente”.

Como exemplo, a Figura 5 mostra que o item 01 (“Os resultados das avaliações internas (CPA)

são discutidos em seu curso”), os respondentes com traço latente ($\theta < -0,705$) são mais propensos a discordar totalmente do item, no entanto, a curva referente a P1 é menor em todos os traços latentes, ou seja, essa categoria nunca é mais provável de ser respondida. Para P2 a propensão é maior num intervalo pequeno ($-0,705 \leq \theta < 0,665$) e, para P3 e P4 nota-se intervalos do traço latente onde as propensões são maiores para estas categorias. Como comparação, o item 10 (“*Você conhece a previsão orçamentária e o cronograma de execução previstos no Plano de Desenvolvimento Institucional (PDI) 2015-2019 da UFSC.*”), tem as curvas P1 e P2 as propensões são menores em toda a faixa do traço latente e, outra característica observada, é o deslocamento à direita indicando um item em que os respondentes tendem a concordar mais dificilmente, em detrimento ao item 01 que possui um deslocamento à esquerda.

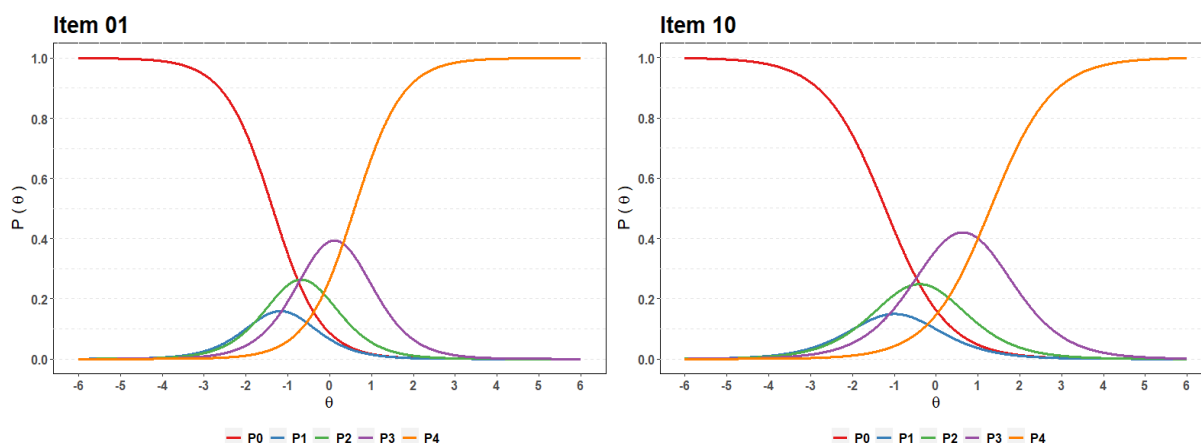


Figura 5 – Curva característica para os itens (01 e 10).

A curva de informação do teste permite identificar a região do traço latente em que as informações são mais consistentes, mais confiáveis, ou seja, em que existe um menor erro padrão associado. Na Figura 6 é possível verificar que o teste gera maior informação e menor erro no intervalo ($-4,7 < \theta \leq 2,6$), portanto é a área de maior precisão da escala que mensura o grau de concordância dos discentes da pós-graduação para com a IES em relação aos eixos do SINAES na AAI.

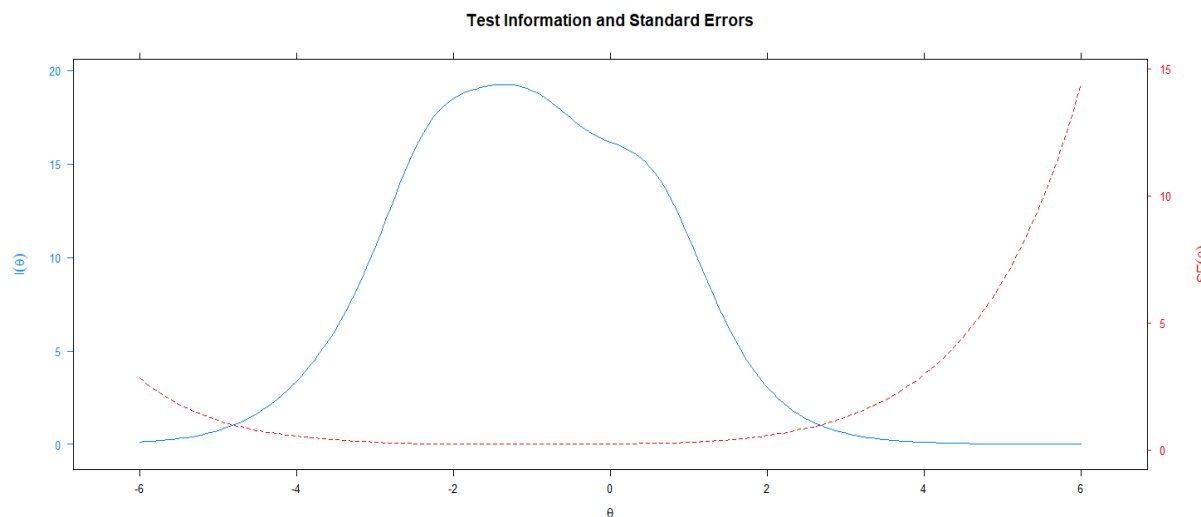


Figura 6 – Curva de Informação do Teste fornecida pela TRI.

4.4. Elaboração da interpretação pedagógica da escala de medida.

Todos os parâmetros de itens e traços latentes dos respondentes foram obtidos numa mesma métrica, representados na escala (0,1), contudo este estudo executou uma transformação linear para produzir a escala na métrica (300,30). Desta forma, a escala foi construída com média 300 e desvio

padrão 30, mas ainda respeitando as relações de ordem existentes entre seus pontos.

O primeiro procedimento para a definição da escala é a identificação dos níveis âncora, que são caracterizados por conjuntos de itens âncoras, que são os pontos de interpretação da escala. Dois critérios foram adotados para o posicionamento de cada item na escala: [1] itens que obtiveram estimativas do parâmetro de discriminação para $a \geq 1$ e; [2] uma categoria do item será âncora em um determinado nível de escala se, pelo menos, for respondido por mais de 60% dos discentes com o mesmo nível de traço latente.

A Figura 7 apresenta a distribuição de frequência dos traços latentes dos 2.709 discentes e o posicionamento de cada item por categoria na escala.

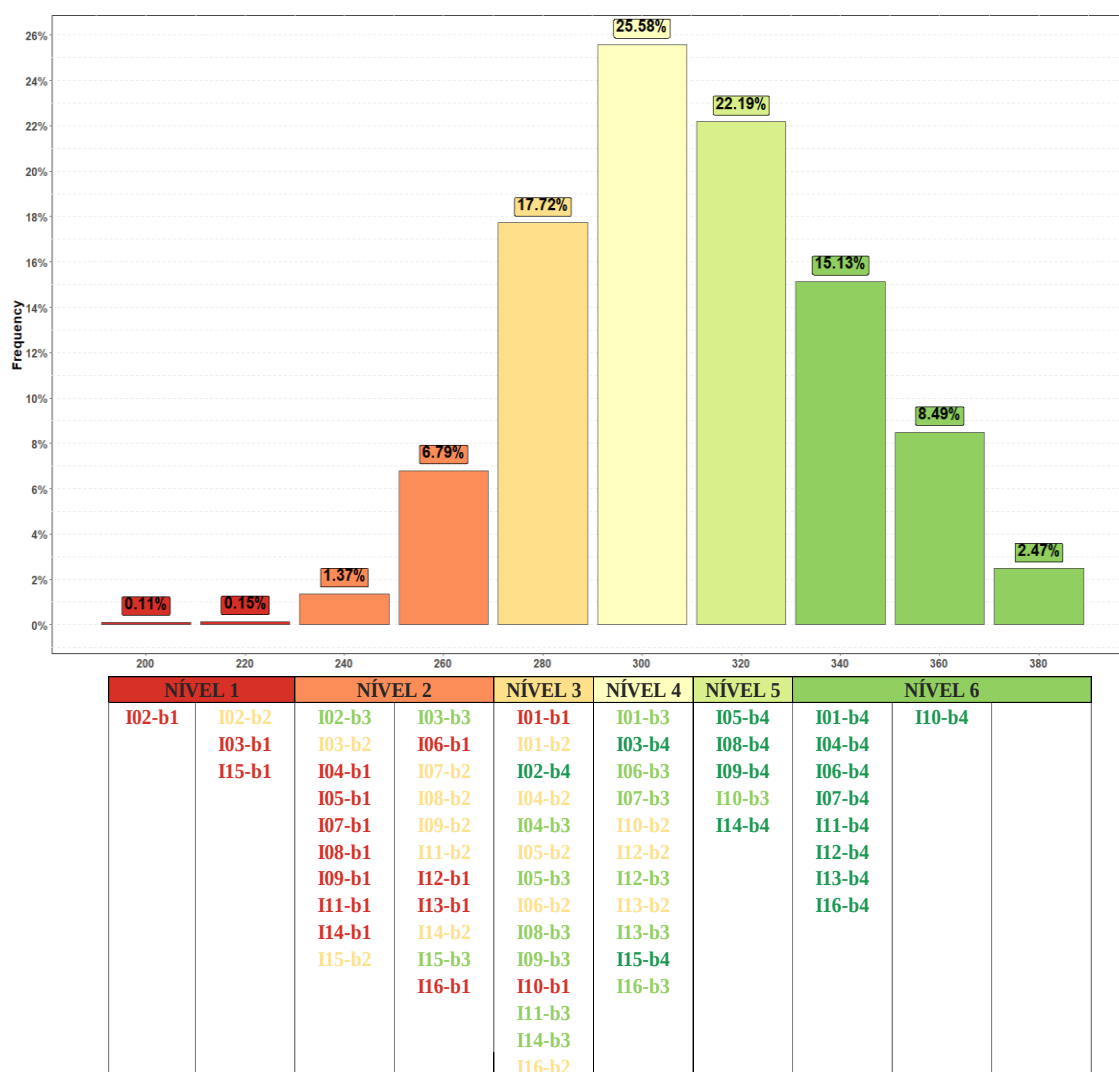


Figura 7 – Posicionamento das categorias dos itens e distribuição dos valores do traço latente dos respondentes na escala (300; 30).

Os valores abaixo de -230 foram atribuídos ao Nível 1 em que há predominância dos parâmetros b_1 e representam (0,26%) dos respondentes. Neste nível os respondentes tendem a discordar totalmente da maioria dos itens. Os valores entre $230 \leq \theta < 270$ foram agrupados no Nível 2 (8,16%) e estão associados a uma percepção de discordância parcial em grande parte dos itens, e começa a apresentar percepções neutras em alguns itens. O valor entre $270 \leq \theta < 290$ que compõe o Nível 3 (17,72%)

ocorre a presença da maioria dos itens com categoria (2 – discordo parcialmente) e (3 – não concordo, nem discordo). O Nível 4 (25,58%) trata-se de uma concordância intermediária e consolida uma maior parte dos itens na categoria (4 – concordo parcialmente). No Nível 6 (28,09%), o parâmetro b_4 aparece com maior frequência e indica que os respondentes concordam fortemente com a maioria dos itens.

A Tabela 4 apresenta a interpretação pedagógica da escala do grau de concordância da AAI.

Tabela 4 – Interpretação dos níveis de Escala.

Descrição dos Níveis	
NÍVEL 1	Neste nível os respondentes discordam fortemente da maioria dos itens avaliados e apresentam discordância parcial quanto a contribuição da UFSC para a inclusão social, preservação do meio ambiente, desenvolvimento local e regional e preservação da memória e patrimônio cultural (I03-b1¹¹), item do Eixo II - desenvolvimento institucional e quanto à qualidade dos serviços prestados nas bibliotecas no que se refere ao acesso via internet (consulta e reserva), informatização do acervo, bancos de dados, empréstimo e horário de funcionamento ser suficiente (I15-b1¹²), que pertence ao Eixo V - infraestrutura física. Ainda neste nível observa-se que os discentes discordam fortemente tendendo à indiferença no que tange à contribuição do seu curso para o desenvolvimento da missão da UFSC (I02-b1¹³ e I02-b2¹⁴), item que também está compreendido no Eixo II-desenvolvimento institucional.
NÍVEL 2	Neste segundo nível, os respondentes discordam parcialmente tendendo à indiferença (não concordando, nem discordando) com a maioria dos itens. Os respondentes relacionam a menor discordância, já que tendem à indiferença ao Eixo III – políticas acadêmicas, no que se refere às políticas de ensino e as ações acadêmico-administrativas serem suficientes (I04-b1¹⁵ e I04-b2¹⁶); e discordam parcialmente da política de pesquisa em relação ao seu curso ser suficiente (I05-b1¹⁷); assim como em relação às informações sobre a instituição, tais como: divulgação dos cursos, da pesquisa e da extensão; existência de mecanismos de transparência institucional; ouvidoria; resultados das avaliações recentes; entre outros serem acessíveis à comunidade externa (I06-b1¹⁸). Além disso, com relação ao Eixo IV – políticas de gestão, os respondentes transitam de uma discordância parcial para uma neutralidade (não discordam e nem concordam) quanto à eficiência quanto: a) à atuação do Conselho Universitário (Cun) (I07-b1¹⁹ e I07-b2²⁰); à atuação dos órgãos colegiados relacionados a seu programa (I08-b1²¹ e I08-b2²²); b) às ações da coordenação do seu programa (I09-b1²³ e I09-b2²⁴) e à contribuição das fundações de apoio da UFSC ao desenvolvimento das atividades de ensino, pesquisa e extensão (I11-b1²⁵ e I11-b2²⁶) os respondentes neste nível tendem perceber ineficiência.

¹¹ Item 03 com parâmetro de dificuldade b1.

¹² Item 15 com parâmetro de dificuldade b2.

¹³ Item 02 com parâmetro de dificuldade b1.

¹⁴ Item 02 com parâmetro de dificuldade b2.

¹⁵ Item 04 com parâmetro de dificuldade b1.

¹⁶ Item 04 com parâmetro de dificuldade b2.

¹⁷ Item 05 com parâmetro de dificuldade b1.

¹⁸ Item 06 com parâmetro de dificuldade b1.

¹⁹ Item 07 com parâmetro de dificuldade b1.

²⁰ Item 07 com parâmetro de dificuldade b2.

²¹ Item 08 com parâmetro de dificuldade b1.

²² Item 08 com parâmetro de dificuldade b2.

²³ Item 09 com parâmetro de dificuldade b1.

²⁴ Item 09 com parâmetro de dificuldade b2.

²⁵ Item 11 com parâmetro de dificuldade b1.

²⁶ Item 11 com parâmetro de dificuldade b2.

NÍVEL 2

Em relação à **Eixo V – Infraestrutura física**, os respondentes mantêm uma **discordância parcial** relacionado à infraestrutura das salas de aula (**I12-b1**²⁷), dos laboratórios/áreas experimentais (**I13-b1**²⁸) e à qualidade dos espaços de convivência e de alimentação (**I16-b1**²⁹), transição para uma neutralidade (**não discordam e nem concordam**) quanto ao acervo disponível para a realização das suas atividades acadêmicas (**I14-b1**³⁰ e **I14-b2**³¹) e uma **concordância parcial** caracterizada pela qualidade dos serviços prestados nas bibliotecas (**I15-b2**³² e **I15-b3**³³). Para o **Eixo II – desenvolvimento institucional**, os respondentes **concordam parcialmente** com a contribuição do curso para o desenvolvimento da Missão da UFSC (**I02-b3**³⁴) e com a contribuição da UFSC para a inclusão social, preservação do meio ambiente, desenvolvimento local e regional e preservação da memória e patrimônio cultural (**I03-b3**³⁵).

NÍVEL 3

Neste terceiro nível o respondente transita de uma **discordância parcial** para uma neutralidade (**não discordam e nem concordam**) no **Eixo I – planejamento e avaliação institucional** relacionado às discussões dos resultados das avaliações internas (CPA) em seu curso (**I01-b1**³⁶ e **I01-b2**³⁷). Além disso, **concordam fortemente** que o seu curso contribui para o desenvolvimento da Missão da UFSC (**I02-b4**³⁸). No **Eixo III – políticas acadêmicas** os respondentes **concordam parcialmente** que as políticas de ensino e as ações acadêmico-administrativas (**I04-b2**³⁹ e **I04-b3**⁴⁰) e a política de pesquisa (**I05-b2**⁴¹ e **I05-b3**⁴²) em relação ao seu curso são suficientes, ao passo que, em relação as informações sobre a instituição (**I05-b2**⁴³) os respondentes **não concordam e nem discordam (indiferentes)** sobre a acessibilidade à comunidade externa (**I06-b2**⁴⁴). Para o **Eixo IV - políticas de gestão**, os respondentes **concordam parcialmente** que a atuação dos órgãos colegiados (**I08-b3**⁴⁵) e as ações da coordenação (**I09-b3**⁴⁶) relacionados a seu Programa são suficientes, e as fundações de apoio da UFSC contribuem adequadamente (**I11-b3**⁴⁷). No entanto, os respondem **desconhecem parcialmente** a previsão orçamentária e o cronograma de execução previstos no Plano de Desenvolvimento Institucional (PDI) 2015-2019 da UFSC (**I10-b1**⁴⁸). Para o **Eixo V – Infraestrutura física**, os respondentes **concordam parcialmente** que o acervo disponível nas bibliotecas para a realização das suas atividades acadêmicas é suficiente (**I14-B3**⁴⁹) e são **neutros** quanto a qualidade dos espaços de convivência e de alimentação (**I16-b2**⁵⁰).

²⁷ Item 12 com parâmetro de dificuldade b1.

²⁸ Item 13 com parâmetro de dificuldade b1.

²⁹ Item 16 com parâmetro de dificuldade b1.

³⁰ Item 14 com parâmetro de dificuldade b1.

³¹ Item 14 com parâmetro de dificuldade b2.

³² Item 15 com parâmetro de dificuldade b2.

³³ Item 15 com parâmetro de dificuldade b3.

³⁴ Item 02 com parâmetro de dificuldade b3.

³⁵ Item 03 com parâmetro de dificuldade b3.

³⁶ Item 01 com parâmetro de dificuldade b1.

³⁷ Item 01 com parâmetro de dificuldade b2.

³⁸ Item 02 com parâmetro de dificuldade b4.

³⁹ Item 04 com parâmetro de dificuldade b2.

⁴⁰ Item 04 com parâmetro de dificuldade b3.

⁴¹ Item 05 com parâmetro de dificuldade b2.

⁴² Item 05 com parâmetro de dificuldade b3.

⁴³ Item 05 com parâmetro de dificuldade b2.

⁴⁴ Item 06 com parâmetro de dificuldade b2.

⁴⁵ Item 08 com parâmetro de dificuldade b3.

⁴⁶ Item 09 com parâmetro de dificuldade b3.

⁴⁷ Item 11 com parâmetro de dificuldade b3.

⁴⁸ Item 10 com parâmetro de dificuldade b1.

⁴⁹ Item 14 com parâmetro de dificuldade b3.

⁵⁰ Item 16 com parâmetro de dificuldade b2.

NÍVEL 4

Neste quarto nível os respondentes indicam **neutralidade** sobre conhecerem a previsão orçamentária e o cronograma de execução previstos no (PDI) 2015-2019 da UFSC (**I10-b2**⁵¹) que tange ao **Eixo IV - políticas de gestão**, sobre a infraestrutura das salas de aula no que se refere à suficiência do espaço, disponibilidade de equipamentos audiovisuais, limpeza, iluminação, acústica, ventilação, conservação e comodidade ser suficiente (**I12-b2**⁵²), e sobre a suficiência da infraestrutura dos laboratórios/áreas experimentais no que se refere a equipamentos e espaços físicos, acesso à internet, atualização de equipamentos e softwares, disponibilidade de insumos, apoio técnico e manutenção, que são itens por meio dos quais o **Eixo V – Infraestrutura física** é avaliado. **Concordância parcial** é manifestada pelos respondentes sobre os resultados das avaliações internas (CPA) serem discutidos em seu curso (**I01-b3**⁵³) que é item avaliativo do **Eixo I – planejamento e avaliação institucional**, assim como sobre as informações acerca da instituição serem acessíveis à comunidade EXTERNA (**I06-b3**⁵⁴) que tange ao **Eixo III – políticas acadêmicas**, sobre a eficiência da atuação do Conselho Universitário (Cun) (**I07-b3**⁵⁵) que diz respeito ao **Eixo IV - políticas de gestão**, e acerca de considerar a infraestrutura das salas de aula (**I12-b3**⁵⁶), dos laboratórios, áreas experimentais no que se refere a equipamentos e espaços físicos, acesso à internet, atualização de equipamentos e softwares, disponibilidade de insumos, apoio técnico e manutenção (**I13-b3**⁵⁷), a qualidade dos espaços de convivência e de alimentação no que se refere à dimensão, limpeza, iluminação, ventilação e conservação serem suficientes (**I16-b3**⁵⁸) que são itens que integram a avaliação do **Eixo V - Infraestrutura física**. Os respondentes **concordam** com a afirmação de que a UFSC tem contribuído para a inclusão social, preservação do meio ambiente, desenvolvimento local e regional e preservação da memória e patrimônio cultural (**I03-b4**⁵⁹) que aspecto que compõe o **Eixo II - desenvolvimento institucional** e que a qualidade dos serviços prestados nas bibliotecas no que se refere ao acesso via internet (consulta e reserva), informatização do acervo, bancos de dados, empréstimo e horário de funcionamento é suficiente (**I15-b4**⁶⁰) que concerne ao **Eixo V - infraestrutura física**.

NÍVEL 5

Neste quinto nível os respondentes **concordam parcialmente** que conhecem a previsão orçamentária e o cronograma de execução previstos no (PDI) 2015-2019 da UFSC (**I10-b3**⁶¹) que remete ao **Eixo IV - políticas de gestão**, e **concordam fortemente** que a política de pesquisa em relação ao seu curso é suficiente (**I05-b4**⁶²) item que avalia o **Eixo III - políticas acadêmicas**, assim como a atuação dos órgãos colegiados relacionados a seu programa (**I08-b4**⁶³), e as ações da coordenação do seu Programa (**I09-b4**⁶⁴) serem eficientes, itens estes abarcados no **Eixo IV - políticas de gestão**, como também que o acervo disponível para a realização das suas atividades acadêmicas no tocante à qualidade e a quantidade do acervo disponível nas bibliotecas (central/setoriais/polo EaD), no acervo virtual e nas bases de dados é suficiente; A qualidade dos espaços de convivência e de alimentação no que se refere à dimensão, limpeza, iluminação, ventilação e conservação é suficiente. (**I14-b4**⁶⁵) item avaliativo do **Eixo V - infraestrutura física**.

⁵¹ Item 10 com parâmetro de dificuldade b2.

⁵² Item 12 com parâmetro de dificuldade b2.

⁵³ Item 01 com parâmetro de dificuldade b3.

⁵⁴ Item 06 com parâmetro de dificuldade b3.

⁵⁵ Item 07 com parâmetro de dificuldade b3.

⁵⁶ Item 12 com parâmetro de dificuldade b3.

⁵⁷ Item 13 com parâmetro de dificuldade b3.

⁵⁸ Item 16 com parâmetro de dificuldade b3.

⁵⁹ Item 03 com parâmetro de dificuldade b4.

⁶⁰ Item 15 com parâmetro de dificuldade b4.

⁶¹ Item 10 com parâmetro de dificuldade b3.

⁶² Item 05 com parâmetro de dificuldade b4.

⁶³ Item 08 com parâmetro de dificuldade b4.

⁶⁴ Item 09 com parâmetro de dificuldade b4.

⁶⁵ Item 14 com parâmetro de dificuldade b4.

A partir do sexto nível os respondentes **concordam fortemente** que os resultados das avaliações internas (CPA) são discutidos em seu curso (**I01-b4⁶⁶**) que trata do **Eixo I – avaliação institucional**, assim como que as políticas de ensino e as ações acadêmico-administrativas para o seu curso são suficientes (**I04-b4⁶⁷**), e que as informações sobre a instituição são acessíveis à comunidade externa (**I06-b4⁶⁸**) questões estas que estão compreendidas no **Eixo III - políticas acadêmicas**, mantendo esse entendimento ainda acerca de que a atuação do Conselho Universitário (Cun) é eficiente (**I07-b4⁶⁹**), de que conhecem a previsão orçamentária e o cronograma de execução previstos no (PDI) 2015-2019 da UFSC (**I10-b4⁷⁰**), e de que as fundações de apoio da UFSC contribuem adequadamente para o desenvolvimento das atividades de ensino, pesquisa e extensão. (**I11-b4⁷¹**) questões estas compreendidas no **Eixo IV - políticas de gestão**, como também quanto à infraestrutura das salas de aula (**I12-b4⁷²**), a infraestrutura dos laboratórios/áreas experimentais no que se refere a equipamentos e espaços físicos, acesso à internet, atualização de equipamentos e softwares, disponibilidade de insumos, apoio técnico e manutenção (**I13-b4⁷³**), e a qualidade dos espaços de convivência e de alimentação no que se refere à dimensão, limpeza, iluminação, ventilação e conservação (**I16-b4⁷⁴**) serem considerados suficientes, itens comportados pelo **Eixo V - infraestrutura física**. Desta forma, a partir do nível 6, os respondentes passam a **concordar fortemente** com todos os eixos do SINAES.

5. Considerações finais.

O presente estudo se propôs a construir, por meio da TRI, uma escala de medida para mensurar a o instrumento de autoavaliação institucional desenvolvido pela IES sob a perspectiva de concordância ou discordância nos itens que compõe o instrumento. Além disso, validou-se as premissas da TRI de unidimensionalidade e a dependência local, sendo assim capaz de contribuir para a avaliação da qualidade de ensino usando as medidas psicométricas.

Adicionalmente, o estudo constatou que as informações divulgadas pela CPA no relatório de autoavaliação institucional acerca dos instrumentos aplicados na comunidade acadêmica se resumem ao score dos dados obtidos e dos achados principais. O processamento e análise dos dados proporcionado pela TRI permitiu identificar a possibilidade de melhorar os itens ao identificar que alguns deles possuem categorias de resposta que em nenhum momento possuem uma maior probabilidade de serem escolhidas. Também permitiu identificar que todos os itens possuem um bom parâmetro de discriminação, bons parâmetros de dificuldade, assim como baixos erros padrões associados.

A escala construída permite uma visualização mais apropriada acerca dos itens em que é mais difícil alcançar maior concordância, o que permite uma tomada de decisão mais precisa. Enquanto no relatório da CPA indica, por exemplo, que os discentes estão mais insatisfeitos acerca da infraestrutura, a análise da escala demonstra quais itens da infraestrutura causam maior discordância, no caso a infraestrutura das salas de aula e dos laboratórios, assim como a informação sobre a discordância acerca do conhecimento da informação referente ao orçamento, que tange ao teor do item 10 do instrumento da AAI.

Outro benefício está na identificação do grau de concordância dos discentes. Enquanto na metodologia atualmente adotada a CPA utiliza a média aritmética para o score, de forma que pode alocar a maioria dos respondentes em uma das categorias da escala *Likert* usada no instrumento de pesquisa, a TRI permite alocá-los em uma escala de medida de forma a auxiliar na tomada de decisão de quais eixos do SINAES demandam maior atenção. No caso a dimensão 10, sustentabilidade financeira, e a dimensão

⁶⁶ Item 01 com parâmetro de dificuldade b4.

⁶⁷ Item 04 com parâmetro de dificuldade b4.

⁶⁸ Item 06 com parâmetro de dificuldade b4.

⁶⁹ Item 07 com parâmetro de dificuldade b4.

⁷⁰ Item 10 com parâmetro de dificuldade b4.

⁷¹ Item 11 com parâmetro de dificuldade b4.

⁷² Item 12 com parâmetro de dificuldade b4.

⁷³ Item 13 com parâmetro de dificuldade b4.

⁷⁴ Item 16 com parâmetro de dificuldade b4.

7, infraestrutura física, foram as mais proeminentes. Mas é possível identificar nos níveis de maior concentração imediatamente inferiores outros pontos que podem promover a melhoria do grau de concordância dos discentes.

Foi, portanto, comprovado no estudo que a TRI tem um grande potencial para beneficiar o relatório produzido pela CPA permitindo fomentar o diagnóstico da atual situação da IES acerca das dimensões do SINAES por meio da aferição do grau de concordância dos discentes da pós-graduação. Uma limitação identificada foi o alto número de discentes que marcaram a opção desconheço/não se aplica, ou que deixaram de responder itens. Sugere-se que nas próximas aplicações da AAI seja estudada a possibilidade de não dar as opções de não responder, ou de responder desconheço e não se aplica, permitindo assim uma maior precisão dos dados. Para estudos futuros recomenda-se ampliar o escopo realizando a análise para os instrumentos aplicados em outros membros da comunidade acadêmica, realizando o comparativo com as informações constantes no relatório divulgado pela IES.

Referências.

AERA; APA; NCME. **Standards for Educational and Psychological Testing**. American Educational Research Association, 2014. 9780935302356.

ANDRADE, D. F. D.; TAVARES, H. R.; DA CUNHA VALLE, R. Teoria da Resposta ao Item: conceitos e aplicações. **ABE, Sao Paulo**, 2000.

ARAUJO, C. R. F. D.; OLIVEIRA DIAS, M. D. Políticas de avaliação e padrões de qualidade da educação superior no Brasil. **International Journal of Development Research**, 11, n. 11, p. 52284-52293, 2021.

ARRUDA, J. A. D.; PASCHOAL, T.; DEMO, G. Uso dos resultados da autoavaliação institucional pelos gestores da Universidade de Brasília. **Avaliação: Revista da Avaliação da Educação Superior (Campinas)**, 24, 2019.

BERNARDES, J. D. S.; ROTHEN, J. C. O campo da Avaliação da Educação Superior: foco na Comissão Própria de Avaliação. **2016**, 8, n. 23, p. 30, 2016-08-30 2016.

BRASIL. Lei nº 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior-SINAES e dá outras providências. Diário Oficial da União. Brasília 2004a.

BRASIL. Roteiro de Autoavaliação Institucional – Orientações Gerais. TEIXEIRA, I. N. D. E. E. P. E. A. Brasília: Inep/MEC: 41 p. 2004b.

BRASIL. Instrumento de Autoavaliação Institucional aplicado à Pós-Graduação. (CPA), U. F. D. S. C. C. P. D. A. Florianópolis 2019.

BROSSMAN, B. G.; LEE, W.-C. Book Review: Mark D. Reckase, Multidimensional Item Response Theory. New York, NY: Springer. 2009. 354 pp. **Applied Psychological Measurement**, 34, n. 8, p. 641-645, 2010.

DE AYALA, R. J. **The theory and practice of item response theory**. Guilford Publications, 2013. 1462514693.

EDWARDS, M. C.; HOUTS, C. R.; CAI, L. A diagnostic procedure to detect departures from local independence in item response theory models. **Psychological methods**, 23, n. 1, p. 138, 2018.

LORD, F. M. **Applications of item response theory to practical testing problems.** Routledge, 2012. 0203056612.

NUNES, E. B. L. D. L. P.; DUARTE, M. M. S. L. T.; PEREIRA, I. C. A. Planejamento e avaliação institucional: um indicador do instrumento de avaliação do SINAES. **Avaliação: Revista da Avaliação da Educação Superior (Campinas)**, 22, n. 2, p. 373-384, 2017.

PASQUALI, L. **Instrumentação psicológica-fundamentos e práticas.** Porto Alegre: ARTMED, 2010.

PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item: TRI. **Avaliação Psicológica: Interamerican Journal of Psychological Assessment**, 2, n. 2, p. 99-110, 2003.

RECKASE, M. D. Unifactor Latent Trait Models Applied to Multifactor Tests: Results and Implications. **Journal of Educational Statistics**, 4, n. 3, p. 207-230, 1979.

RECKASE, M. D. Multidimensional item response theory models. *In*: **Multidimensional item response theory**: Springer, 2009. p. 79-112.

SAMEJIMA, F. Estimation of latent ability using a response pattern of graded scores. **Psychometrika monograph supplement**, 1969.

SCHINDLER, L.; PULS-ELVIDGE, S.; WELZANT, H.; CRAWFORD, L. Definitions of Quality in Higher Education: A Synthesis of the Literature. **Higher Learning Research Communications**, 5, n. 3, p. 3-13, Sep 2015 2015.