

UNIVERSIDADE DA REGIÃO DE JOINVILLE – UNIVILLE
PROGRAMA DE PÓS-GRADUAÇÃO EM SAÚDE E MEIO AMBIENTE

**AVALIAÇÃO DAS CARACTERÍSTICAS DOS ITENS APLICADOS NO TESTE
DE PROGRESSO NOS ANOS DE 2016, 2017 E 2018 EM UM CONSÓRCIO
BRASILEIRO**

MSC. ADEMIR GARCIA REBERTI
Orientador Prof. Dr. Jean Carl Silva
Coorientador Prof. Dr. Olavo Franco Ferreira Filho

Joinville - SC
2020

ADEMIR GARCIA REBERTI

**AVALIAÇÃO DAS CARACTERÍSTICAS DOS ITENS APLICADOS NO TESTE
DE PROGRESSO NOS ANOS DE 2016, 2017 E 2018 EM UM CONSÓRCIO
BRASILEIRO**

Tese apresentada como requisito parcial para obtenção do título de Doutor em Saúde e Meio-Ambiente na Universidade da Região de Joinville. Orientador: Prof. Dr. Jean Carl Silva. Coorientador: Prof. Dr. Olavo Franco Ferreira Filho.

Joinville - SC

2020

Catálogo na publicação pela Biblioteca Universitária da Univille

Reberti, Ademir Garcia

R291a Avaliação das características dos itens aplicados no teste de progresso nos anos de 2016, 2017 e 2018 em um consórcio brasileiro / Ademir Garcia Reberti; orientador Dr. Jean Carl Silva; coorientador Olavo Franco Ferreira Filho. – Joinville: UNIVILLE, 2020.

64 p.: il.

Tese (Doutorado em Saúde e Meio Ambiente – Universidade da Região de Joinville)

1. Medicina – Testes de aptidão. 2. Escolas de medicina. 3. Medicina – Estudo e ensino. 4. Estudantes de medicina - Avaliação. I. Silva, Jean Carl (orient.). II. Ferreira Filho, Olavo Franco (coorient.). III. Título.

CDD 610.7

Termo de Aprovação

“Avaliação das Características dos Itens Aplicados no Teste de Progresso nos Anos de 2016, 2017 e 2018 em um Consórcio Brasileiro”

por

Ademir Garcia Reberti

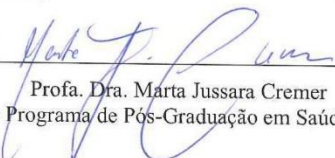
Tese julgada para a obtenção do título de Doutor em Saúde e Meio Ambiente, área de concentração Saúde e Meio Ambiente e aprovada em sua forma final pelo Programa de Pós-Graduação em Saúde e Meio Ambiente.



Prof. Dr. Jean Carl Silva
Orientador (UNIVILLE)

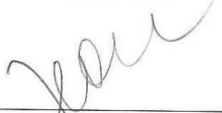


Prof. Dr. Olavo Franco Ferreira Filho
Coorientador (UEL)

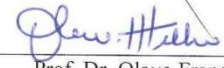


Profa. Dra. Marta Jussara Cremer
Coordenadora do Programa de Pós-Graduação em Saúde e Meio Ambiente

Banca Examinadora:



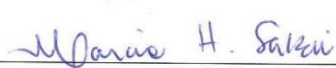
Prof. Dr. Jean Carl Silva
Orientador (UNIVILLE)



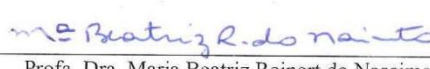
Prof. Dr. Olavo Franco Ferreira Filho
Coorientador (UEL)



Profa. Dra. Maria Inês da Rosa
(UNESC)



Profa. Dra. Marcia Hiromi Sakai
(UEL)



Profa. Dra. Maria Beatriz Reinert do Nascimento
(UNIVILLE)

Joinville, 26 de maio de 2020

AGRADECIMENTOS

Agradeço a Deus pela saúde e forças para chegar ao final deste longo processo, mantendo-me forte e exitoso neste projeto de pesquisa.

Sou grato à minha esposa e filho, pelo apoio que sempre me deram durante toda esta jornada, por terem compreendido as ausências, a falta de parceria durante este período e as noites insones.

Agradeço às instituições, aos seus alunos e aos seus professores, bem como aos amigos que compõem o NAPISUL II – um seminário de cooperação, ideias científicas e grandes amizades.

Sou grato pela confiança depositada na minha proposta de projeto pelo professor Dr. Jean Carl Silva, orientador do meu trabalho e por me manter motivado durante todo o processo.

Agradecimento especial ao professor Dr. Olavo Franco Ferreira Filho, coorientador, pela dedicação, competência, carinho e disponibilidade em analisar os dados deste trabalho e pelas valiosas contribuições dadas durante toda esta caminhada.

Gratulações afáveis aos grandes amigos entusiastas da educação médica: Carlos Eduardo Andrade Pinheiro, Lara Cristina Leite Guimarães Machado, Tatiane Muniz Barbosa.

Agradeço à aluna de graduação Nayme Hechem Monfredini que acreditou no projeto e se dedicou na busca pelos dados.

Por último, quero agradecer também à Universidade da Região de Joinville – UNIVILLE e aos professores do doutorado, pela elevada qualidade do ensino oferecido, aos colegas de turma pelo amável convívio e dias inesquecíveis.

“Imagine os receptores numa membrana celular, à espera de se ligarem a determinadas moléculas. Quando o contato acontece, uma sucessão de reações químicas desencadeia-se no interior da célula. Do mesmo modo, um médico precisa ter os seus receptores em alerta diante da pessoa que o procura. E o que liga o médico ao doente? São as Histórias.” (Rita Charon)

RESUMO

O Teste de Progresso (TP) consiste em uma avaliação objetiva, estruturada com questões de múltipla escolha, com objetivo de promover uma avaliação das competências cognitivas esperadas no final do curso de graduação. Este estudo teve como objetivo avaliar as características dos itens que podem influenciar na qualidade da elaboração dos testes e a determinação do seu impacto positivo e/ou negativo na construção do aprendizado. A metodologia foi dividida em dois momentos: no primeiro momento foram avaliados os dados psicométricos dos Testes de Progresso aplicados nos anos de 2016, 2017 e 2018 pelo consórcio NAPISUL II; no segundo momento as questões utilizadas nas mesmas provas foram avaliadas por três professores que realizaram a análise e classificação pedagógica dos itens de acordo com a Taxonomia de Bloom revisada. Os itens também foram distribuídos por áreas médicas, além da contagem do número de palavras dos enunciados e das alternativas. Concluiu-se que o teste foi mais consistente para os alunos dos dois últimos anos da graduação. Ou seja, alunos do internato apropriaram-se mais adequadamente ao item que lhes era cobrado. Percebeu-se que o tamanho do enunciado e das alternativas não influenciou no grau de dificuldade e na capacidade discriminativa. Não houve associação entre os níveis da Taxonomia de Bloom revisada e os índices psicométricos dos itens. Notou-se que os alunos dos dois últimos anos de graduação apreenderam melhor o conteúdo exigido na avaliação. Sugere-se que sejam formulados Testes de Progresso personalizados para cada ciclo da graduação (Ciclo Básico, Ciclo Clínico e Internato).

Palavras-chave: Teste de Progresso, Escola Médica, Características dos Itens, Taxonomia de Bloom.

ABSTRACT

The Progress Test (PT) consists of an objective assessment, structured with multiple choice questions, in order to promote an assessment of the cognitive skills expected at the end of the undergraduate course. This study aimed to evaluate the characteristics of the items that can influence the quality of the test design and the determination of its positive and/or negative impact on the construction of learning. The methodology was divided into two moments: in the first moment, the psychometric data of the Progress Tests applied in the years 2016, 2017 and 2018 were evaluated by the NAPISUL II consortium; in the second moment, the questions used in the same tests were evaluated by three teachers who performed the analysis and pedagogical classification of the items according to the revised Bloom Taxonomy. The items were also distributed by medical areas, in addition to counting the number of words in the statements and the alternatives. It was concluded the test was more consistent for students in the last two years of graduation. That is, boarding school students appropriated the item they were charged with more appropriately. It was noticed that the size of the statement and the alternatives did not influence the degree of difficulty and discriminative capacity. There was no association between the levels of revised Bloom Taxonomy and the psychometric indices of the items. It was noted that the students of the last two years of graduation learned better the content required in the assessment. It is suggested that personalized Progress Tests be formulated for each cycle of the undergraduate course (Basic Cycle, Clinical Cycle and Internship).

Keywords: Progress Test, Medical School, Item Characteristics, Bloom Taxonomy.

LISTA DE ABREVIATURAS

ABEM	Associação Brasileira de Educação Médica
CAAE	Certificado de Apresentação para Apreciação Ética
CNS	Conselho Nacional de Saúde
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
NAPISUL II	Núcleo de Apoio Pedagógico Interinstitucional Sul II
PR	Paraná
QME	Questões de múltiplas escolhas
SC	Santa Catarina
TP	Teste de Progresso
TCLE	Termo de Consentimento Livre e Esclarecido
TCT	Teoria Clássica dos Testes
TRI	Teoria de Resposta ao Item

LISTA DE QUADROS

Quadro 1: Taxonomia de Bloom Revisada.....	25
Quadro 2: Classificação das questões de acordo com índice de facilidade	33
Quadro 3: Classificação das questões de acordo com seu índice de discriminação (Ponto-bisserial)	34
Quadro 4: Número de participantes do TP subdivididos por ano de graduação e ano de realização do teste.	38
Quadro 5: Número de alunos respondedores do TP por ano.....	38

LISTA DE GRÁFICOS

Gráfico 1: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 1.....	42
Gráfico 2: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 2.....	42
Gráfico 3: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 3.....	43
Gráfico 4: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 1.....	44
Gráfico 5: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 2.....	44
Gráfico 6: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 3.....	45
Gráfico 7: Associação entre a capacidade discriminativa (correlação bisserial) do item e o número de palavras do enunciado de acordo com o Grupo 1.....	46
Gráfico 8: Associação entre a capacidade discriminativa (correlação bisserial) do item e o número de palavras do enunciado de acordo com o Grupo 2.....	46
Gráfico 9: Associação entre a capacidade discriminativa do item (correlação bisserial) e o número de palavras do enunciado de acordo com o Grupo 3	47

LISTA DE TABELAS

Tabela 1: Grau de confiabilidade encontrado nos itens aplicados nas provas realizadas nos anos de 2016, 2017 e 2018.....	39
Tabela 2: Grau de dificuldade encontrado nos itens aplicados nas provas realizadas nos anos de 2016, 2017 e 2018, com as porcentagens baseadas nos acertos dos itens.	39
Tabela 3: Capacidade discriminativa dos itens aplicados nas provas de 2016, 2017 e 2018 em relação aos grupos 1, 2 e 3; com porcentagens relativas.	40
Tabela 4: Tipos de erros pedagógicos encontrados nos itens das provas realizadas nos anos de 2016, 2017 e 2018.....	40
Tabela 5: Nível de Taxonomia de Bloom dos itens aplicados nas provas de 2016, 2017 e 2018.	41
Tabela 6: Taxonomia de Bloom versus a construção dos itens subdivididos em grandes áreas médicas nos Testes de Progresso aplicados em 2016, 2017 e 2018.	48

SUMÁRIO

1. INTRODUÇÃO	15
2. OBJETIVOS	17
2.1. OBJETIVO GERAL	17
2.2. OBJETIVOS ESPECÍFICOS	17
3. REVISÃO DA LITERATURA	18
3.1. TESTE DE PROGRESSO.....	18
3.2. NAPISUL II	20
3.3. TAXONOMIA DE BLOOM.....	22
3.4. MÉTODOS DE AVALIAÇÃO PARAMÉTRICA	26
3.4.1. Teoria Clássica dos Testes	26
3.4.2. Teoria De Resposta ao Item	26
3.4.3. Grau de confiabilidade do teste - Alfa de Cronbach	28
3.5. ELABORAÇÃO E REVISÃO DE ITENS.....	28
3.5.1. Matriz de referência	28
3.5.2. Estrutura do item de múltipla escolha.....	29
3.5.3. Elaboração do item.....	29
4. METODOLOGIA	31
4.1. DESENHO E CARACTERIZAÇÃO DO ESTUDO	31
4.2. LOCAL E PARTICIPANTES.....	32
4.2.1. Participantes do Teste de Progresso.....	32
4.2.2. Participantes da avaliação dos itens.....	33
4.3. ANÁLISE PSICOMÉTRICA DOS ITENS.....	33
4.4. ANÁLISE PEDAGÓGICA DOS ITENS.....	34
4.4.1. Erros pedagógicos.....	34
4.4.2. Contagem de palavras da questão	35

4.4.3. Complexidade das questões classificadas pela Taxonomia de Bloom revisada e as diferentes áreas de conhecimento do Teste de Progresso.....	36
4.4.4. Associação entre número de palavras do enunciado, alternativas, grau de dificuldade e capacidade discriminativa do item	37
4.5. CONSIDERAÇÕES ÉTICAS.....	37
5. RESULTADOS.....	38
5.1. PARTICIPANTES DO TESTE DE PROGRESSO	38
5.2. RESULTADOS DA ANÁLISE PSICOMÉTRICA DOS ITENS	39
5.2.1. Grau de dificuldade dos itens	39
5.2.2. Capacidade discriminativa.....	40
5.3. RESULTADOS DA ANÁLISE PEDAGÓGICA DOS ITENS.....	40
5.3.1. Erros pedagógicos.....	40
5.3.2. Aplicação da Taxonomia de Bloom Revisada	41
5.3.3. Associações.....	42
7. DISCUSSÃO	49
8. CONCLUSÃO	54
9. CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS.....	57
ANEXO A	61
ANEXO B	62
ANEXO C	63
ANEXO D	66

1. INTRODUÇÃO

A avaliação dos processos de gestão e formação educacional, bem como os mecanismos de ensino-aprendizagem são uma atividade fundamental para o planejamento e desenvolvimento da educação. Essa atividade permite acompanhar o processo de apropriação e complexificação do conhecimento, avaliando a evolução de aprendizagens relevantes e significativas para o desenvolvimento de habilidades e competências nos alunos (FERREIRA e WEVH, 2018).

O processo de avaliação busca compreender a adequação entre o desempenho do estudante e a fase em que se encontra matriculado, além de fomentar o aprendizado e embasar decisões que terão implicações em seu progresso, contribuindo para o controle da qualidade dos programas educacionais (BOLLELA, BORGES e TRONCON, 2018).

O método de avaliação cumulativa tem sido utilizado como ferramenta para orientar o comportamento de estudo dos alunos, uma vez que incentiva o tempo de autoestudo dos estudantes de maneira mais uniforme ao longo do curso (FERNANDES et al., 2018).

Os métodos de avaliação podem servir como ferramenta de melhoria do ensino. O processo de avaliar atua como regulador, mostrando falhas e possíveis soluções a obstáculos apresentados pelos estudantes, podendo gerar melhorias nas ferramentas didáticas (SCHUWIRTH e VAN DER VLEUTEN; 2012).

A validade de uma avaliação é garantida quando for possível demonstrar que o método escolhido avalia exatamente aquilo que se pretendia que o estudante aprendesse. Depende, então, da adequação do método à natureza do domínio ao qual se pretende avaliar (BOLLELA, BORGES e TRONCON; 2018).

As questões de múltipla escolha (QME) constituem um dos métodos mais utilizados em todo o mundo em exames destinados a avaliar habilidades cognitivas (BOLLELA, BORGES E TRONCON, 2018).

O Teste de Progresso (TP) consiste em uma avaliação objetiva, estruturada com 60 a 150 questões de múltipla escolha, com objetivo de promover uma avaliação das competências cognitivas esperadas no final do curso de graduação. Esse teste é aplicado a todos os discentes em uma mesma

data, de modo que se possa comparar os resultados entre as séries e analisar a performance evolutiva do conhecimento no decorrer do curso. A comparação dos resultados é possível com a equalização dos testes (MIRANDA JR. et al., 2018).

Com os resultados do Teste de Progresso constroem-se curvas de desempenho dos acadêmicos, que permitem identificar fragilidades e qualidades dos estudantes nas diversas áreas do conhecimento relacionadas ao curso (SAKAI et al., 2008). Esse *feedback* é fundamental para a construção de estratégias para melhorias da gestão do curso, bem como, da autoavaliação do aluno.

Ressalta-se a importância da qualidade e das boas práticas de elaboração dos itens, uma vez que esses fatores podem influenciar diretamente na habilidade de interpretar, avaliar, sintetizar e resolver questões complexas apresentadas aos estudantes de medicina. Para tal, o enunciado e as alternativas necessitam estar alinhados ao objetivo da questão, considerando as diretrizes curriculares nacionais e a matriz de referência de elaboração dos itens.

Visto que o Teste de Progresso, em medicina, surgiu como uma forma de medir o desenvolvimento dos acadêmicos em relação ao tempo decorrido de curso, questionam-se os parâmetros que podem influenciar na qualidade da elaboração dos testes e a determinação do seu impacto positivo e/ou negativo na construção do aprendizado.

2. OBJETIVOS

2.1. OBJETIVO GERAL

Avaliar as características dos itens dos Testes de Progresso aplicado nos anos de 2016, 2017 e 2018 em um consórcio de escolas médicas do Sul do Brasil.

2.2. OBJETIVOS ESPECÍFICOS

- Avaliar os dados psicométricos (grau de dificuldade e capacidade discriminativa) e o grau de confiabilidade da prova;
- Realizar a análise pedagógica dos itens aplicados de acordo com as regras estipuladas para o Teste de Progresso;
- Classificar os itens de acordo com a Taxonomia de Bloom revisada;
- Analisar a relação entre a Taxonomia de Bloom revisada e aplicada sobre os dados psicométricos dos itens;
- Quantificar o número de palavras do enunciado e alternativas de cada questão e sua influência sobre os dados psicométricos do teste;
- Verificar a distribuição da complexidade dos níveis da Taxonomia de Bloom revisada, aplicada nas questões das áreas médicas que integram a avaliação do Teste de Progresso.

3. REVISÃO DA LITERATURA

3.1. TESTE DE PROGRESSO

O desenvolvimento cognitivo é uma importante dimensão na formação do médico, sendo um processo contínuo de aquisição e consolidação de um conjunto de componentes necessários para domínio do conhecimento em uma ou mais áreas de desempenho (SAKAI et al., 2011).

A avaliação cumulativa pode ser utilizada como ferramenta para orientar o comportamento do ganho cognitivo dos alunos (FERNANDES et al., 2018). Independente da postura filosófica da avaliação, há um eixo comum entre as diversas concepções de avaliação educacional: associam-se a um processo interpretativo, supondo juízo de valor, qualidade e/ou mérito que tem por alvo diagnosticar e verificar a abrangência dos objetivos propostos no processo ensino-aprendizagem (RODRIGUES, 2006). Apesar das diferenças na formulação das questões, a essência das teorias é de melhorar o sistema de ensino e o aprendizado do aluno.

A técnica do Teste de Progresso (TP) foi inicialmente desenvolvida no início dos anos 1970, na *University Missouri-Kansas City School of Medicine*, nos Estados Unidos, e na Universidade de Limburg, na Holanda (SAKAI et al., 2008).

Consiste em uma avaliação com objetivo de promover uma avaliação das competências cognitivas esperadas no final do curso. O objetivo deste teste é comparar os resultados entre as séries e analisar a performance evolutiva do conhecimento no decorrer do curso (MIRANDA JR. et al., 2018). Não visa aprovar, selecionar ou classificar os estudantes, mas dimensionar o seu ganho de conhecimento cognitivo e constatar sua evolução individual no processo de construção de sua aprendizagem (SAKAI et al., 2008).

É um tipo de avaliação longitudinal do desenvolvimento cognitivo dos estudantes no decorrer do curso. Os testes aplicados pelas diversas escolas médicas apresentam diferenças, tais como: norma referenciada ou critério referenciado, as questões podem ser do tipo verdadeiro/falso ou múltipla escolha (VERHOEVEN et al., 2002), dentre outras.

O tempo máximo de prova é de 4 horas, sendo o tempo mínimo para a entrega da prova de 2 horas. Dependendo da instituição, a presença dos alunos pode ou não ser obrigatória para realização do teste (ABEM, 2017).

As questões do TP são elaboradas por uma equipe multidisciplinar, constituída por professores das instituições, baseado em uma matriz de referência que contempla as diretrizes nacionais para o sexto ano do curso de medicina. Cada questão possui de quatro a cinco alternativas possíveis nas respostas, com grau de dificuldade e capacidade discriminativa diferentes. Além disso, são divididas em seis áreas do conhecimento: Ciências Básicas, Cirurgia, Clínica Médica, Ginecologia e Obstetrícia, Pediatria e Saúde Coletiva. As questões podem ser contextualizadas em casos clínicos, de modo que os enunciados não exijam apenas a memorização do aluno, mas também seu raciocínio lógico (BALDIM, VICENTE e ARCURI, 2018).

Foi realizada uma revisão literária acerca do uso do TP por Schuwirth e Van der Vleuten (2012). Evidencia-se que a abordagem de avaliação longitudinal tem efeito positivo no comportamento de aprendizagem do aluno, desencorajando o aprendizado pré-provas, e consequentemente auxiliando no aprendizado acumulativo a longo prazo, o aprendizado duradouro. Supõe-se que os alunos experimentem menos estresse na realização do TP em comparação com os testes tradicionais, visto que um único resultado ruim não pode desfazer uma série de bons resultados. Os autores ressaltam também que o método longitudinal aumenta a confiabilidade do TP.

Segundo Wagener et al. (2015), o TP fornece aos alunos um *feedback* com relação ao nível de proficiência no decorrer de seus estudos. Foi aplicado o TP a acadêmicos do 1º ao 6º ano de medicina em sete escolas na Alemanha. O estudo observou que o número de respostas corretas aumentou linearmente com o avanço do nível do semestre letivo em que o aluno se encontrava matriculado.

Foi utilizada uma avaliação somativa de habilidades cognitivas para entender como o uso de boas práticas para elaboração de testes (questões) de múltipla escolha interferem na composição de exames, por Bollela et al. (2018) O autor cita que para a avaliação de habilidades referentes ao domínio cognitivo utilizam-se três níveis referenciais para elaboração de questões:

(a) básico, que exige conhecimento factual, envolvendo basicamente a memorização de fatos e sua recuperação; (b) intermediário, que envolve não só o conhecimento adquirido mas também as capacidades de compreensão, interpretação e aplicação daquele conhecimento para resolver problemas mais simples; (c) avançado, que envolve os níveis de análise, síntese e avaliação para propor soluções para problemas mais complexos como os que habitualmente se apresentam ao profissional da área da saúde. (BOLLELA et al., 2018).

O estudo concluiu que as Questões de Múltipla Escolha (QME) com somente uma alternativa correta têm inúmeras vantagens, mas que para aplicabilidade adequada é necessária atenção às boas práticas de confecção de questões e de construção dos exames, envolvendo recomendações relativas ao conteúdo e à forma dos testes, bem como à composição da matriz de avaliação ou tabela de especificação.

“No cenário de educação médica no Brasil, o Teste de Progresso vem sendo aplicado nas escolas médicas, de forma isolada ou em parceria, desde o final da década de 1990” (SAKAI et al., 2011).

Há uma escassez na literatura sobre o tema Teste de Progresso no que tange a relação entre o TP e a elaboração dos itens.

3.2. NAPISUL II

A Associação Brasileira de Educação Médica (ABEM) divide-se em oito regionais com o objetivo de promover a interação mais próxima e flexível com os associados, abarcando a diversidade de contextos e situações existentes em todo o Brasil, sendo que cada uma delas tem um coordenador e toda a equipe técnica necessária. As escolas médicas de Santa Catarina (SC) e Paraná (PR) pertencem à Regional ABEM SUL II.

Os membros da ABEM SUL II, preocupados com a excelência da educação médica, sentiram necessidade de implantar um processo avaliativo processual, a exemplo do que já estava ocorrendo em outro núcleo no Brasil, e, com o apoio da ABEM criaram o consórcio do Núcleo De Apoio Pedagógico Interinstitucional Sul II (NAPISUL II). Estabelecido em agosto de 2010, tem o NAPISUL II o propósito específico de formulação, implementação e análise do Teste de Progresso (ROSA et al.; 2016).

Após a criação do regimento do Núcleo, os seus componentes se preocuparam em elaborar, tendo como base as diretrizes nacionais para o curso de medicina, as matrizes de referência de conteúdos para cada uma das grandes áreas da prova: Ciências Básicas, Clínica Cirúrgica, Clínica Médica, Ginecologia e Obstetrícia, Pediatria e Saúde Coletiva. A partir destas matrizes, cada Instituição deve formular dez questões, com quatro alternativas, de cada área, com todo o rigor pedagógico para a formulação de itens.

No início do ano letivo, os representantes de cada grande área, de cada instituição, reúnem-se, geralmente em Curitiba ou Florianópolis, para escolherem os melhores itens que comporão as questões da prova final. Os itens escolhidos são enviados a um avaliador externo, que formata e escolhe as 120 questões da prova que será aplicada.

O teste final é aplicado a todos os estudantes do primeiro ao sexto ano, tendo no mínimo duas horas e no máximo quatro horas de duração. A avaliação é confidencial, sigilosa e todos os alunos do núcleo fazem a prova no mesmo dia e horário.

O gabarito comentado é disponibilizado para cada aluno e para a instituição em 24 a 48 horas, de forma online.

A análise dos resultados obtidos na realização da prova é realizada por um avaliador externo, utilizando métodos estatísticos e psicométricos: a Teoria Clássica dos Itens (TCT) e a Teoria de Resposta ao Item (TRI). Em 30 dias, os alunos recebem sua devolutiva individual de forma sigilosa, demonstrando em uma tabela o seu desempenho cognitivo em cada grande área e na prova como um todo. Para que o aluno possa analisar seu desempenho, é fornecida a sua média de acertos comparada à média dos demais participantes de seu mesmo ano de graduação.

A cada instituição é fornecida a média de seus alunos comparada à média geral dos alunos de todas as instituições participantes. A partir desta devolutiva, as escolas detêm a possibilidade de estabelecer estratégias de melhorias para sua instituição.

A partir da interação dos coordenadores/professores de cada grande área das diferentes instituições é possível gerar uma discussão a respeito dos temas mais relevantes, o que possibilita desenvolver habilidades de formulação de itens/questões pedagogicamente coerentes, melhorando a qualidade da

avaliação. Ademais, a reunião permite a interação de professores das diferentes instituições, possibilitando trocas de experiências pedagógicas. Esta interação resulta na adequação do conteúdo que é abordado em cada disciplina nas diferentes escolas.

Dentre as vantagens para a instituição, ressalta-se a análise do conteúdo da sua matriz curricular e a evolução do desempenho de seus acadêmicos. Além de permitir intervenções simultâneas, com ações para melhoria contínua do aluno e do curso. Sendo assim, é um bom indicador de autoavaliação para ambos.

A realização da prova gera um certificado de participação ao aluno, permitindo que este seja somado às suas horas complementares, além de servir como um simulado preparativo para a prova de residência médica e autoavaliação sobre seu desempenho acadêmico. Algumas instituições disponibilizam outros ganhos aos participantes, como por exemplo, pontuar com um adicional de nota em um módulo ou disciplina de escolha do estudante, quando o aluno fica acima da média do consórcio NAPISUL II.

Para maior adesão e realização adequada da prova, é fornecido um material ao coordenador de cada instituição para que seja apresentado aos alunos, explicando a importância e as vantagens da realização do TP, tanto para a instituição quanto para o acadêmico individualmente.

A realização do teste gera uma precificação financeira para as instituições, tendo em média atualmente o custo de R\$ 10,50 reais por acadêmico. Este conjunto de ações, alinhado às estratégias da ABEM, visa à melhoria e aprimoramento da educação médica por meio da união das escolas médicas e de políticas de saúde públicas, influenciando na formação e prestação de serviços de saúde no território brasileiro.

3.3. TAXONOMIA DE BLOOM

Segundo Conklin (2005), a Taxonomia de Bloom e sua classificação hierárquica dos objetivos de aprendizagem tem sido uma das maiores contribuições acadêmicas para educadores que, conscientemente, procuram meios de estimular, nos seus discentes, raciocínio e abstrações de alto nível

(*higher order thinking*) sem distanciar-se dos objetivos instrucionais previamente propostos.

Um dos motivos pelo qual a taxonomia proposta por Bloom tornou-se importante, e trouxe significativas contribuições à área acadêmica, foi o fato de que antes da década de 50 um dos grandes problemas na literatura educacional era a falta de consenso com relação a determinadas palavras usualmente relacionadas à definição dos objetivos instrucionais como, por exemplo, o verbo conhecer era utilizado com o sentido de ter consciência, saber da existência ou para expressar domínio de um determinado assunto (CONKLIN, 2005).

“De acordo com a taxonomia original, de 1956, conhecimento envolve: a) habilidade de lembrar especificidades e generalidades de métodos, procedimentos, padrões e instruções; e b) habilidade de achar, no problema proposto, sinais, dicas, pequenas informações que efetivamente tragam à consciência o aprendizado prévio adquirido, ou seja, conhecimento é o que é lembrado. Essas duas definições, por si só, já correspondem respectivamente à diferença significativa entre processo e produto” (ANDERSON, 1999 apud FERRAZ e BELHOT, 2010).

O que interessava para Bloom et al. (1956), era proporcionar uma ferramenta prática e útil que fosse coerente com as características dos processos mentais superiores (nível de conhecimento e abstração complexa) do modo como eram consideradas e conhecidas.

A taxonomia trouxe a possibilidade de padronização da linguagem no meio acadêmico e, com isso, também novas discussões ao redor dos assuntos relacionados à definição de objetivos instrucionais. Neste contexto, instrumentos de aprendizagem puderam ser trabalhados de forma mais integrada e estruturada, inclusive considerando os avanços tecnológicos que podiam prover novas e diferentes ferramentas para facilitar o processo de ensino e aprendizagem (FERRAZ e BELHOT, 2010).

A Taxonomia de Bloom avalia o domínio cognitivo e é estruturada em níveis de complexidade crescente – do mais simples ao mais complexo – e isso significa que, para adquirir uma nova habilidade pertencente ao próximo nível, o aluno deve ter dominado e adquirido a habilidade do nível anterior (FERRAZ e BELHOT, 2010).

Só após conhecer um determinado assunto alguém poderá compreendê-lo e aplicá-lo. Nesse sentido, a taxonomia proposta não é apenas um esquema para classificação, como também uma possibilidade de organização hierárquica dos processos cognitivos de

acordo com níveis de complexidade e objetivos do desenvolvimento cognitivo desejado e planejado (FERRAZ e BELHOT, 2010).

Propõem-se, por Bloom et al. (1956), níveis de domínio cognitivo, o que ficou conhecido como Taxonomia de Bloom. Os níveis são: 1) Conhecimento, 2) Compreensão, 3) Aplicação, 4) Análise, 5) Síntese, 6) Avaliação.

Quarenta anos depois de ter sido divulgada, Lori Anderson publicou, em 1999, um significativo trabalho de retrospectiva da utilização da taxonomia e, no mesmo ano, um grupo de especialistas se encontrou em Syracuse, Nova Iorque, para discutir a possibilidade de rever os pressupostos teóricos da Taxonomia de Bloom, uma vez que novos conceitos, recursos e teorias foram incorporados ao campo educacional, avanços psicopedagógicos e tecnológicos ocorreram, e diversas experiências de sucesso no uso efetivo da taxonomia foram publicadas (FERRAZ e BELHOT, 2010).

Esse grupo de especialistas (psicólogos, educadores, especialistas em currículos, testes, avaliação etc.) supervisionado por David Krathwohl, participante do desenvolvimento da Taxonomia original no ano de 1956, e, no ano de 2001, o relatório da revisão foi publicado num livro intitulado *A taxonomy for learning, teaching and assessing: a revision of Bloom's taxonomy for educational objectives* (ANDERSON et al., 2001).

A partir da discussão acerca da taxonomia e da observação da prática educacional de como educadores definem seus objetivos gerais e específicos de disciplinas/cursos, Krathwohl et al. (2002) começaram a perceber que mudanças na taxonomia original seriam necessárias e o primeiro ponto analisado estava relacionado à questão do verbo e sua associação direta com o objetivo cognitivo, avaliação do objetivo e desenvolvimento de competências.

Ao analisar a relação direta entre verbo e substantivo, conclui-se que verbos e substantivos deveriam pertencer a dimensões separadas na qual os substantivos formariam a base para a dimensão conhecimento (o que) e os verbos para a dimensão relacionada aos aspectos cognitivos (como) (KRATHWOHL et al., 2002).

Na atual Taxonomia de Bloom a base das categorias foi mantida, continuam existindo seis categorias, o nome da taxonomia continua sendo o mesmo (eventualmente pode aparecer com a expressão “revisada” adicionada a ele), entretanto, ao separar, conceitualmente, o conhecimento do processo cognitivo, ocorreram as seguintes mudanças nos níveis da Taxonomia: 1)

Lembrar, 2) Entender, 3) Aplicar, 4) Analisar, 5) Sintetizar, 6) Criar (KRATHWOHL et al., 2002).

Segundo Bloom et al. (1956) vários pesquisadores utilizam a terminologia Taxonomia para realizar classificações estruturadas e orientadas para definir algumas teorias instrucionais. Segundo Conklin (2005), a Taxonomia de Bloom e sua classificação hierárquica dos objetivos de aprendizagem tem sido uma das maiores contribuições acadêmicas para educadores que, conscientemente, procuram meios de estimular, nos seus discentes, raciocínio e abstrações de alto nível sem distanciar-se dos objetivos instrucionais previamente propostos.

A Taxonomia de Bloom revisada é apresentada no quadro a seguir.

Quadro 1: Taxonomia de Bloom Revisada

1. Lembrar	Relacionado a reconhecer e reproduzir ideias e conteúdos. Reconhecer requer distinguir e selecionar uma determinada informação e reproduzir ou recordar está mais relacionado à busca por uma informação relevante memorizada. Representado pelos seguintes verbos no gerúndio: Reconhecendo e Reproduzindo.
2. Entender	Relacionado a estabelecer uma conexão entre o novo e o conhecimento previamente adquirido. A informação é entendida quando o aprendiz consegue reproduzi-la com suas “próprias palavras”. Representado pelos seguintes verbos no gerúndio: Interpretando, Exemplificando, Classificando, Resumindo, Inferindo, Comparando e Explicando.
3. Aplicar	Relacionado a executar ou usar um procedimento numa situação específica e pode também abordar a aplicação de um conhecimento numa situação nova. Representado pelos seguintes verbos no gerúndio: Executando e Implementando.
4. Analisar	Relacionado a dividir a informação em partes relevantes e irrelevantes, importantes e menos importantes e entender a inter-relação existente entre as partes. Representado pelos seguintes verbos no gerúndio: Diferenciando, Organizando, Atribuindo e Concluindo.
5. Avaliar	Relacionado a realizar julgamentos baseados em critérios e padrões qualitativos e quantitativos ou de eficiência e eficácia. Representado pelos seguintes verbos no gerúndio: Checando e Criticando.
6. Criar	Significa colocar elementos junto com o objetivo de criar uma nova visão, uma nova solução, estrutura ou modelo utilizando conhecimentos e habilidades previamente adquiridos. Envolve o desenvolvimento de ideias novas e originais, produtos e métodos por meio da percepção da interdisciplinaridade e da interdependência de conceitos. Representado pelos seguintes verbos no gerúndio: Generalizando, Planejando e Produzindo.

Fonte: Ferraz e Belhot (2010).

3.4. MÉTODOS DE AVALIAÇÃO PARAMÉTRICA

3.4.1. Teoria Clássica dos Testes

Visando a melhoria da qualidade das provas são estudados métodos psicométricos de avaliações, dentre eles: a Teoria Clássica dos Testes. As avaliações psicométricas dos testes passaram por diferentes fases, tendo início por volta da década de 1880 com Galton, atravessando as eras de Cattell, na década de 1890, de Binet, na década de 1900, a era dos testes de inteligência, entre 1910 e 1930, a da análise fatorial e a era da sistematização, entre 1940 e 1980.

A partir da era de Binet, com os trabalhos de Spearman relacionados à correlação, desenvolve-se a Teoria Clássica dos Testes (PASQUALI, 1997). Nela se basearam a maioria dos métodos operacionais utilizados para avaliar as duas principais propriedades psicométricas dos instrumentos: a validade e a confiabilidade. A validade de um instrumento pode ser definida como sua capacidade de realmente medir aquilo a que se propõe a medir (PORTNEY e WATKINS, 1993; PASQUALI, 1997).

A TCT busca explicar o resultado final total, considerando a soma das respostas dadas ao teste. Esse modelo visa analisar o grau de confiabilidade da prova e dois parâmetros do item: o grau de dificuldade e sua capacidade discriminativa. A análise se baseia no escore total de acertos e erros, não considerando as habilidades latentes dos estudantes e o acerto casual. Sendo assim, avalia os testes de qualidade, avaliando a correlação entre o acerto em determinado item e o número total de acertos na prova (SAKAI et al., 2011).

3.4.2. Teoria De Resposta ao Item

A Teoria de Resposta ao Item passa a ser conhecida, principalmente, a partir do ano de 1968 com o trabalho de Lord e Novick intitulado “*Statistical Theories of Mental Tests Scores*”. Desde então, inúmeras aplicações têm sido exploradas para a teoria: criação de bancos de itens, avaliação adaptativa

computadorizada, equalização de provas, avaliação de mudança cognitiva, dentre outras. A TRI tem como unidade de análise o item, formalizando a relação que existe entre a probabilidade de acertar o item e a capacidade latente requerida na resolução do mesmo. Quanto maior a capacidade de um sujeito (traço latente), maior será a probabilidade de que este sujeito acerte um determinado item que meça este construto (VLEUTEN et al., 2018).

Embretson (1996) e Andrade (2001) enfatizam que a Teoria de Resposta ao Item é o modelo mais recomendado para ser aplicado nas avaliações dos testes objetivos, em comparação com a Teoria Clássica dos Itens,

por considerar as habilidades latentes dos estudantes e por analisar os parâmetros por item e não a prova como um todo. A TRI é, também, a metodologia que melhor estima a capacidade de determinado item ter sido acertado ao acaso. No TP, este é um fator de grande importância, pois, por se tratar de um teste elaborado em nível de sexto ano, do curso de medicina, e que envolve diversas competências para a resolução dos itens, tal como o raciocínio clínico, é provável que os estudantes iniciantes venham a “chutar” grande parte dos itens (SAKAI et al., 2011).

A TRI tem sido considerada como um marco para a psicometria moderna, apresentando vantagens em relação a Teoria Clássica dos Testes, considerando a virtual invariância dos parâmetros dos itens em relação à amostra, estimação mais precisa e interpretável nível de habilidade dos indivíduos e procedimentos de equalização de testes mais eficientes (NUNES e PRIMI, 2005).

“A Teoria de Resposta ao Item não está interessada no escore total de um teste, mas sim no resultado obtido por cada um dos itens que compõem o teste” (SAKAI et al., 2011). Tem como objetivo determinar a probabilidade e quais fatores afetam a probabilidade do estudante acertar ou errar um item de determinado teste. São considerados três parâmetros do item: grau de dificuldade, capacidade discriminativa e probabilidade de acerto casual. Também considera as habilidades latentes do indivíduo, procurando analisar itens de qualidade (SAKAI et al., 2011).

Essa teoria é utilizada na análise psicométrica do TP, não tendo sido aplicada na presente tese, visto que não contempla os objetivos propostos.

3.4.3. Grau de confiabilidade do teste - Alfa de Cronbach

“Em 1951, L. J. Cronbach publica um artigo onde discute os problemas associados à estimação da consistência interna de uma escala ou teste e as propostas de outros autores para o seu cálculo” (MAROCO e GARCIA-MARQUES, 2006). O autor considerou as derivações anteriores de Kuder-Richarson e Guttman, e assumiu os mesmos pressupostos, mas sem limites no padrão de classificação dos itens, formalizando uma proposta de estimativa de consistência interna a partir das variâncias dos itens e dos totais do teste por sujeito, que ficou conhecida como o índice alfa de Cronbach (MAROCO e GARCIA-MARQUES, 2006).

O alfa de Cronbach é uma ferramenta estatística que quantifica, numa escala de 0 a 1, a confiabilidade de um questionário. O valor mínimo aceitável para se considerar um questionário confiável é 0,7. O coeficiente é uma propriedade inerente do padrão de resposta da população estudada, não é uma característica da escala por si só, ou seja, o valor de alfa sofre mudanças segundo a população na qual se aplica a escala (STREINER, 2003).

3.5. ELABORAÇÃO E REVISÃO DE ITENS

3.5.1. Matriz de referência

A matriz de referência é o instrumento norteador para a construção de itens. É importante destacar que a matriz de referência não se confunde com o currículo, que é muito mais amplo. Ela é, portanto, uma referência para aqueles que irão participar do teste, garantindo transparência ao processo, permitindo-lhes uma preparação adequada, bem como para a análise dos resultados do teste aplicado (BRASIL, INEP, 2010).

No NAPISUL II, a matriz de referência foi elaborada de maneira a refletir os conteúdos da diretriz curricular nacional para o ensino médico, de modo a integrar o modelo multidisciplinar dentro das cadeiras médicas e com a contribuição interinstitucional dos integrantes do consórcio.

3.5.2. Estrutura do item de múltipla escolha

O item consiste na unidade básica de um instrumento de coleta de dados, que pode ser uma prova, um questionário. Nos testes educacionais, item pode ser considerado sinônimo de questão, termo mais popular e utilizado com frequência nas escolas (BRASIL, INEP, 2010).

Especialistas reconhecem que os instrumentos de medida educacional possuem vantagens e limitações, sejam eles objetivos ou não. No entanto, ambos podem ser utilizados, indistintamente, para medir os mesmos aspectos do desempenho acadêmico. É assim que os itens objetivos permitem verificar tanto comportamentos simples, de memorização ou reconhecimento, como comportamentos mais complexos, envolvendo compreensão, aplicação, análise, síntese e avaliação (BRASIL, INEP, 2010).

“O item deve ser estruturado de modo homogêneo, de maneira que explicita a situação-problema. Para tanto, devem ser observadas a coerência e a coesão entre as suas partes (texto-base, enunciado e alternativas)” (BRASIL, INEP, 2010).

3.5.3. Elaboração do item

Primeiramente, é necessário selecionar a habilidade da Matriz de Referência a ser cobrada. Em seguida é construída uma situação-problema. Se utilizar um texto-base, dar preferência para temas e textos atuais e adequados ao público-alvo. O formulador do item deve evitar temas que suscitem polêmicas, além de evitar enunciados e alternativas que possam induzir ao erro (“pegadinhas”) (BRASIL, INEP, 2010).

Recomenda-se utilizar termos impessoais como: “considere-se”, “calcule-se”, “argumente-se”, dentre outros. Não é adequado o uso de termos como: “falso”, “exceto”, “incorreto”, “não”, “errado”, “sempre”, “nunca”, “todo”, “totalmente”, “absolutamente”, “completamente”, “somente”, dentre outros.

Não se recomenda o uso de sentenças como: “Pode-se afirmar que”, “É correto afirmar que”, dentre outros (BRASIL, INEP, 2010).

Para construção das alternativas, é importante manter a coerência com o enunciado, manter a independência das alternativas umas das outras, de maneira que não sejam excludentes, negando informações do texto, nem semanticamente muito próximas. É contraindicado o uso de: “todas as

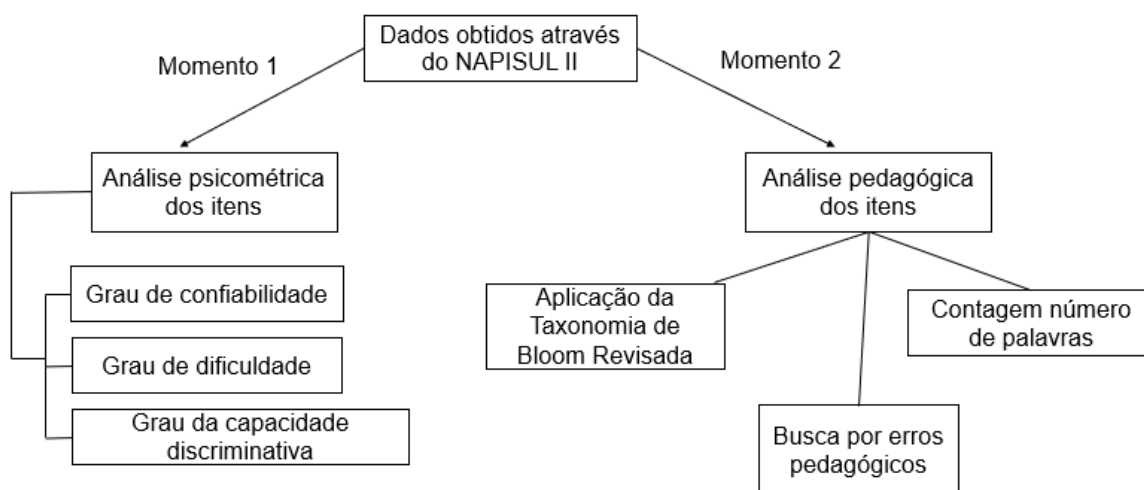
anteriores”, “nenhuma das anteriores”. O gabarito deve ser exposto de forma clara, com uma única alternativa correta e não deve ser mais atrativo que os distratores. “Os distratores não devem ser absurdos em relação à situação-problema apresentada” (BRASIL, INEP, 2010).

Para as alternativas é possível utilizar palavras ou frases-chave, evitando enfatizar detalhes irrelevantes. Aconselha-se evitar formulações que permitam o acerto por exclusão e afirmações demasiadamente longas (BRASIL, INEP, 2010).

4. METODOLOGIA

A metodologia foi dividida em dois momentos: no primeiro momento foram avaliados os dados psicométricos dos Testes de Progresso aplicados nos anos de 2016, 2017 e 2018 pelo consórcio NAPISUL II. No segundo momento, foi realizada a contagem do número de palavras dos enunciados e das alternativas – para isso foi utilizado o software Word 2016. Ainda neste momento, as questões utilizadas nestas provas, foram avaliadas por três professores que realizaram a análise pedagógica dos itens e os classificaram de acordo com a Taxonomia de Bloom revisada, conforme as grandes áreas médicas contempladas nas provas.

Figura 1: Fluxograma dos dois momentos metodológicos do estudo



Fonte: do autor.

4.1. DESENHO E CARACTERIZAÇÃO DO ESTUDO

O presente estudo possui o desenho de três recortes transversais de anos distintos. O estudo avaliou as características dos itens do Teste de Progresso aplicado nos anos de 2016, 2017 e 2018, pelo NAPISUL II. Optou-se pela análise dos três anos mais recentes da aplicação das provas e com maior número de alunos participantes, visto que com a avaliação de apenas um ano de realização da prova haveria possibilidade de acentuação de vieses.

Para isso utilizou-se a análise estatística psicométrica da Teoria Clássica dos Testes e grau de confiabilidade da prova, além da realização da análise pedagógica, o que classificou os itens e sua distribuição por áreas de acordo com a Taxonomia de Bloom revisada, além da realização de contagem de número de palavras de cada item (enunciado e alternativas).

4.2. LOCAL E PARTICIPANTES

4.2.1. Participantes do Teste de Progresso

As provas foram aplicadas nos anos de 2016, 2017 e 2018 pelas escolas Médicas vinculadas ao NAPISUL II. Os participantes foram os alunos do primeiro ao sexto ano de medicina, isto é, os acadêmicos matriculados no curso.

Participaram do Teste de Progresso em 2016, 2017 e 2018, respectivamente 3810, 4173 e 5093 alunos. Foram excluídos da análise os alunos que deixaram mais de 50% dos itens (mais de 60 questões) em branco ou assinalaram mais de 50% dos itens (mais de 60 questões) em uma única alternativa. Sendo assim, o número total de alunos incluídos foi de 12982, e 94 excluídos.

Os alunos analisados foram alocados em três grupos: Grupo 1 (1º ao 2º ano de graduação, denominado ciclo básico); Grupo 2 (3º e 4º ano de graduação, denominado ciclo clínico); e Grupo 3 (5º e 6º ano de graduação, denominado internato).

Esses grupos foram assim divididos de acordo com as diferentes etapas da graduação do curso médico, seguindo as recomendações da diretriz curricular nacional para o curso de medicina.

A divisão dos grupos permitiu análise estatística homogênea para os conteúdos e respondentes.

4.2.2. Participantes da avaliação dos itens

Para a análise pedagógica e classificação dos itens de acordo com a Taxonomia de Bloom, foram convidados três professores com ampla experiência em educação médica e elaboração de itens.

4.3. ANÁLISE PSICOMÉTRICA DOS ITENS

O grau de dificuldade (ou índice de facilidade) foi avaliado verificando-se o percentual de acerto de cada questão da prova e classificado de acordo com normas do MEC/INEP apresentadas no quadro a seguir, segundo a Teoria Clássica dos Testes. Para avaliar as diferenças de grau de dificuldade entre os três grupos foi utilizado o Teste de Qui-quadrado, a partir do software EpiInfo for Windows versão 3.5.4.

Quadro 2: Classificação das questões de acordo com índice de facilidade

Índice de facilidade	Classificação
Maior ou igual a 0,86	Muito fácil
0,61 a 0,85	Fácil
0,41 a 0,60	Médio
0,16 a 0,40	Difícil
Menor ou igual a 0,15	Muito difícil

Fonte: MEC/INEP/DAES – Enade 2016.

O grau de capacidade discriminativa dos itens de múltipla escolha aplicado no TP deve ter um mínimo de poder de discriminação. Para ser considerada apta a avaliar os alunos do curso de Medicina, uma questão deve ser mais acertada por alunos que tiveram bom desempenho do que pelos que tiveram desempenho ruim. Um índice que mede essa capacidade das questões, que é escolhido pelo MEC/INEP e aplicado no ENADE, é o denominado correlação ponto-bisserial. Para avaliar as diferenças de capacidade discriminativa entre os três grupos foi utilizado o Teste de Qui-quadrado, a partir do software EpiInfo for Windows versão 3.5.4..

Quadro 3: Classificação das questões de acordo com seu índice de discriminação (Ponto-bisserial)

Índice de discriminação	Classificação
Maior ou igual a 0,40	Muito bom
0,30 a 0,39	Bom
0,20 a 0,29	Médio
Menor ou igual a 0,19	Fraco

Fonte: MEC/INEP/DAES – Enade 2016.

O grau de confiabilidade de uma prova mede a capacidade daquele instrumento em inferir ou medir aquilo a que realmente se propõe, conferindo relevância à prova. O Coeficiente Alfa de Cronbach foi utilizado como parâmetro de medida de confiabilidade do Teste de Progresso.

Para o cálculo do grau de dificuldade e capacidade discriminativa do item, bem como para o cálculo do Coeficiente Alfa de Cronbach foi utilizado o programa Iteman for Windows – versão 3.5.

4.4. ANÁLISE PEDAGÓGICA DOS ITENS

Para a análise pedagógica dos itens foram convidados três professores com ampla experiência em educação médica e elaboração de itens dentre os docentes das instituições participantes do consórcio NAPISUL II. A análise pedagógica dos itens foi realizada, num primeiro momento, de maneira independente. Em seguida, os professores realizaram uma reunião presencial para chegar a um consenso sobre os itens discordantes.

4.4.1. Erros pedagógicos

Para classificar os erros pedagógicos encontrados nos itens dos três anos de TP, os professores seguiram a seguinte legenda:

- 0 - Sem erro;
- 1 - Item sem enunciado ou texto base;
- 2 - Tamanho heterogêneo das alternativas;

- 3 - Conteúdo heterogêneo das alternativas;
- 4 - Enunciados extensos com informações desnecessárias;
- 5 - Presença de palavras excludentes da veracidade (nunca, sempre, só, somente, exclusivamente, dentre outros) nas alternativas;
- 6 - Presença de erros gramaticais;
- 7 - Presença de convergência (solicitar duas informações na mesma questão e de maneira desnecessária);
- 8 - Outros (o professor descreveu qual erro encontrou).

O item pôde ter mais de um erro pedagógico.

A classificação dos itens de acordo com a Taxonomia de Bloom revisada (FERRAZ e BELHOT, 2010), foi realizada pelos mesmos professores qualificados que fizeram a análise pedagógica, que trabalharam, num primeiro momento, de maneira independente. Em seguida, os professores realizaram uma reunião presencial para chegar a um consenso sobre os itens discordantes.

Os docentes seguiram uma legenda numérica conforme os seus níveis:

- 1 - Lembrar;
- 2 - Entender;
- 3 - Aplicar;
- 4 - Analisar;
- 5 - Avaliar;
- 6 - Criar.

4.4.2. Contagem de palavras da questão

Foi contado o número de palavras de cada questão, utilizando o software Microsoft Word 2016. Em seguida, foi realizada uma análise para verificar a associação entre o número de palavras do enunciado e alternativas, e o grau de dificuldade e capacidade discriminativa do item. A hipótese analisada é de que o maior número de palavras leva a uma menor porcentagem de acertos e a uma

menor capacidade discriminativa. Foi aplicado o teste de Qui-Quadrado utilizando o programa Epilnfo for Windows versão 3.5.4.

Para a análise do número de palavras do enunciado, foram divididos três grupos:

- n1: até 80 palavras no enunciado
- n2: de 81 a 120 palavras no enunciado
- n3: mais de 120 palavras no enunciado

Estes grupos foram escolhidos de acordo com a mediana do número de palavras encontradas no enunciado dos três testes e seus quartis.

Para a análise do número de palavras das alternativas, foram divididos três grupos:

- na1: até 30 palavras nas alternativas
- na2: de 31 a 60 palavras nas alternativas
- na3: mais de 60 palavras nas alternativas

Estes grupos foram escolhidos de acordo com a mediana do número de palavras encontradas nas alternativas dos três testes e seus quartis.

4.4.3. Complexidade das questões classificadas pela Taxonomia de Bloom revisada e as diferentes áreas de conhecimento do Teste de Progresso

Para a análise da distribuição da complexidade dos níveis da Taxonomia de Bloom revisada e aplicada nas questões das áreas médicas, que integram o TP, foi utilizado o teste de Qui-quadrado, utilizando o programa Epilnfo for Windows versão 3.5.4.

Dos 360 itens analisados, em 2016 foi anulada uma questão de Ciências Básicas, em 2017 foi anulada uma questão de Clínica Cirúrgica e uma de Clínica Médica e em 2018 foi anulada uma questão de Ciências Básicas, uma de Pediatria e duas de Saúde Coletiva. As sete questões anuladas apresentaram duas alternativas corretas ou duas alternativas idênticas.

4.4.4. Associação entre número de palavras do enunciado, alternativas, grau de dificuldade e capacidade discriminativa do item

Para testar a hipótese de que quanto maior o número de palavras do enunciado e das alternativas, maior a dificuldade e melhor a discriminação do item, foi realizada regressão linear simples, utilizando como variável dependente o grau de dificuldade e a capacidade discriminativa do item, e como variável independente o número de palavras do enunciado e das alternativas. Utilizou-se para isso a ferramenta de análise de dados do Excel 2010.

4.5. CONSIDERAÇÕES ÉTICAS

Foi obtida permissão das instituições que compõem o núcleo para que os dados fossem fornecidos ao pesquisador, garantindo-se o anonimato das escolas e dos estudantes (ANEXO A). As provas foram aplicadas de forma independente deste estudo, justificando a Dispensa do Termo de Consentimento Livre e Esclarecido (TCLE) das escolas médicas e dos alunos envolvidos (ANEXO B).

Os dados obtidos estão sob guarda e posse do autor do estudo, e serão armazenados por 5 anos, conforme a Resolução CNS 466/12, item XI.2. Após o período, todo material será triturado e descartado.

O estudo passou pelo Comitê de Ética em Pesquisa da Universidade da Região de Joinville, sendo aprovado sob o parecer consubstanciado de número: 3.832.259, CAAE 24357419.1.0000.5366 (ANEXO C).

5. RESULTADOS

5.1. PARTICIPANTES DO TESTE DE PROGRESSO

Quadro 4: Número de participantes do TP subdivididos por ano de graduação e ano de realização do teste.

	GRUPO 1		GRUPO 2		GRUPO 3		Total
	1º ano	2º ano	3º ano	4º ano	5º ano	6º ano	
2016	799	817	687	536	544	410	3793
2017	828	826	790	638	562	502	4146
2018	1148	914	827	807	741	606	5043

Fonte: do autor.

Foram excluídos da análise 17 (0,45%) alunos em 2016; 27 (0,65%) alunos em 2017 e 50 (0,98%) alunos em 2018. Salienta-se que a grande maioria dos alunos eliminados eram dos primeiros quatro anos do curso.

Quadro 5: Número de alunos respondedores do TP por ano

Nº de alunos	Total	Excluídos	Analizados
2016	3810	17	3793
2017	4173	27	4146
2018	5093	50	5043
Total	13076	94	12982

Fonte: do autor.

5.2. RESULTADOS DA ANÁLISE PSICOMÉTRICA DOS ITENS

Para análise dos dados foi estabelecida uma divisão dos alunos por grupos a partir do ciclo de graduação ao qual pertenciam.

Tabela 1: Grau de confiabilidade encontrado nos itens aplicados nas provas realizadas nos anos de 2016, 2017 e 2018.

Grau de confiabilidade	Grupo 1	Grupo 2	Grupo 3
2016	0,563	0,749	0,847
2017	0,633	0,765	0,874
2018	0,554	0,717	0,806

Fonte: do autor.

Salienta-se que o Grupo 1 apresentou coeficientes de confiabilidade não aceitáveis; o Grupo 2 apresentou coeficientes confiáveis, porém baixos; o Grupo 3 apresentou coeficientes satisfatórios.

5.2.1. Grau de dificuldade dos itens

Tabela 2: Grau de dificuldade encontrado nos itens aplicados nas provas realizadas nos anos de 2016, 2017 e 2018, com as porcentagens baseadas nos acertos dos itens.

Grau de dificuldade	Grupo 1 (%)	Grupo 2 (%)	Grupo 3 (%)	Total (%)	Valor de p
Muito fácil	3 (0,8)	17 (4,8)	32 (9,1)	52(4,9)	<0,001
Fácil	51 (14,4)	100 (28,3)	129 (36,5)	280(26,4)	<0,001
Média	96 (27,2)	98 (27,8)	117 (33,1)	311(29,4)	>0,05
Difícil	169 (47,9)	125 (35,4)	69 (19,5)	363(34,3)	<0,001
Muito difícil	34 (9,6)	13 (3,7)	6 (1,7)	53(5,0)	<0,001
Total	353 (100)	353 (100)	353 (100)	1059 (100)	

P estimado pelo Qui-quadrado.

Fonte: do autor.

Somente as questões de média dificuldade não apresentaram frequência diferente de acordo com o grupo.

Para o Grupo 3 (internato) o número de questões que passaram a ser consideradas como muito difíceis diminuiu significativamente em relação ao número de questões consideradas muito difíceis para o Grupo 1 e 2.

5.2.2. Capacidade discriminativa

Tabela 3: Capacidade discriminativa dos itens aplicados nas provas de 2016, 2017 e 2018 em relação aos grupos 1, 2 e 3; com porcentagens relativas.

Capacidade discriminativa	Grupo 1 (%)	Grupo 2 (%)	Grupo 3 (%)	Total (%)	Valor de p
Fraca	233 (66,0)	175 (49,6)	133 (37,7)	541 (51,1)	<0,001
Média	27 (7,7)	36 (10,2)	77 (21,8)	140 (13,2)	<0,001
Boa	90 (25,5)	137 (38,8)	102 (28,9)	329 (31,1)	<0,001
Muito boa	3 (0,8)	5 (1,4)	41 (11,6)	41 (4,6)	<0,001
Total	353 (100)	353 (100)	353 (100)	1059 (100)	

P estimado pelo Qui-quadrado.

Fonte: do autor.

Observou-se que nos Grupos 1 e 2 (etapas iniciais) há um número muito maior de questões com discriminação fraca ou muito fraca quando comparada ao Grupo 3. Ao mesmo tempo, existe um número muito maior de questões consideradas com capacidade discriminativa muito boa no grupo avançado quando comparado aos grupos das etapas iniciais.

5.3. RESULTADOS DA ANÁLISE PEDAGÓGICA DOS ITENS

5.3.1. Erros pedagógicos

Tabela 4: Tipos de erros pedagógicos encontrados nos itens das provas realizadas nos anos de 2016, 2017 e 2018.

Tipo de erros Pedagógicos	Ano			Total
	2016	2017	2018	
0	110	114	107	331
2	3	0	2	5
3	0	1	0	1
4	0	0	1	1
5	3	1	3	7
7	2	2	3	7
2 e 5	1	0	0	1
Total Geral	119	118	116	353
Questões anuladas	1	2	4	

Uma questão possuía dois erros pedagógicos simultaneamente.

Os erros pedagógicos 1 e 6 não foram encontrados

Fonte: do autor.

Das 353 questões aplicadas, somente 23 (6,5%) tinham erros pedagógicos, sendo os principais erros: presença de palavras excludentes, presença de convergência e conteúdo heterogêneo das alternativas.

5.3.2. Aplicação da Taxonomia de Bloom Revisada

Tabela 5: Nível de Taxonomia de Bloom dos itens aplicados nas provas de 2016, 2017 e 2018.

Nível da Taxonomia de Bloom	Ano			Total
	2016	2017	2018	
	n (%)	n (%)	n (%)	
Nível 1: Lembrar	64 (53,8)	50 (42,4)	48 (41,4)	162
Nível 2: Entender	21 (17,6)	20 (16,9)	24 (20,7)	65
Nível 3: Aplicar	18 (15,1)	28 (23,7)	21 (18,1)	67
Nível 4: Analisar	15 (12,6)	20 (16,9)	21 (18,1)	56
Nível 5: Avaliar	1 (0,8)	0 (0)	2 (1,7)	3
Total Geral	119	118	116	353
Questões anuladas	1	2	4	
	χ^2	8,62		
	p	0,374		

Não houve nenhuma questão classificada como Nível 6 da Taxonomia de Bloom

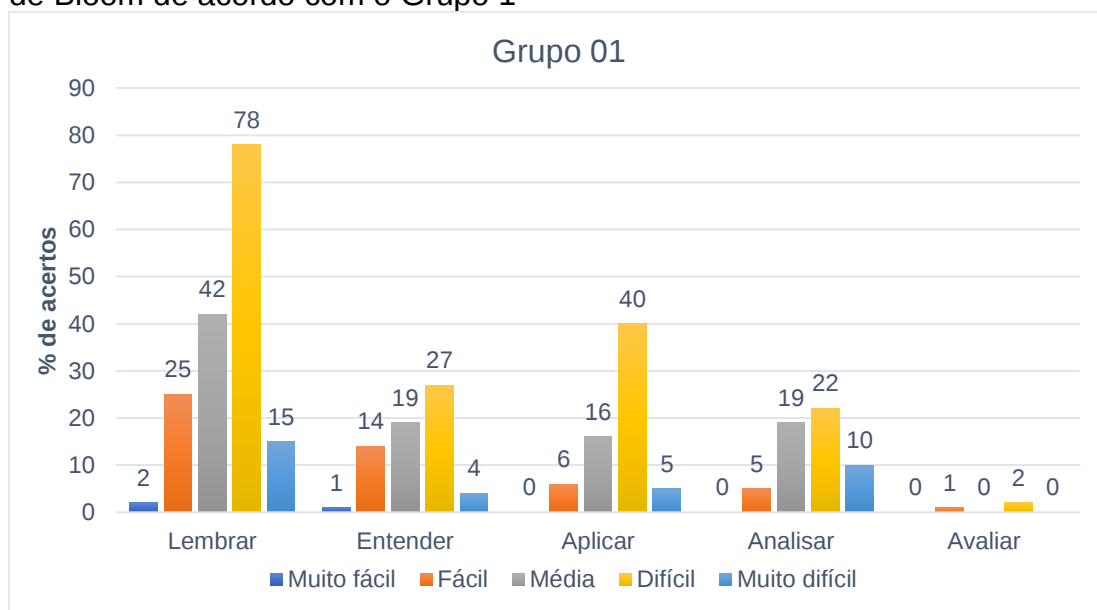
Fonte: do autor.

Apesar do qui-quadrado não apresentar diferença estatística significativa, podemos observar que a grande maioria dos itens (227 – 64,3%) exigiram baixos níveis cognitivos na Taxonomia de Bloom, ou seja, abordando os primeiros níveis da taxonomia (1 -Lembrar e 2 - Entender).

5.3.3. Associações

Foram realizadas associações entre os dados psicométricos e Taxonomia de Bloom revisada. Para isso foi utilizado o método qui-quadrado.

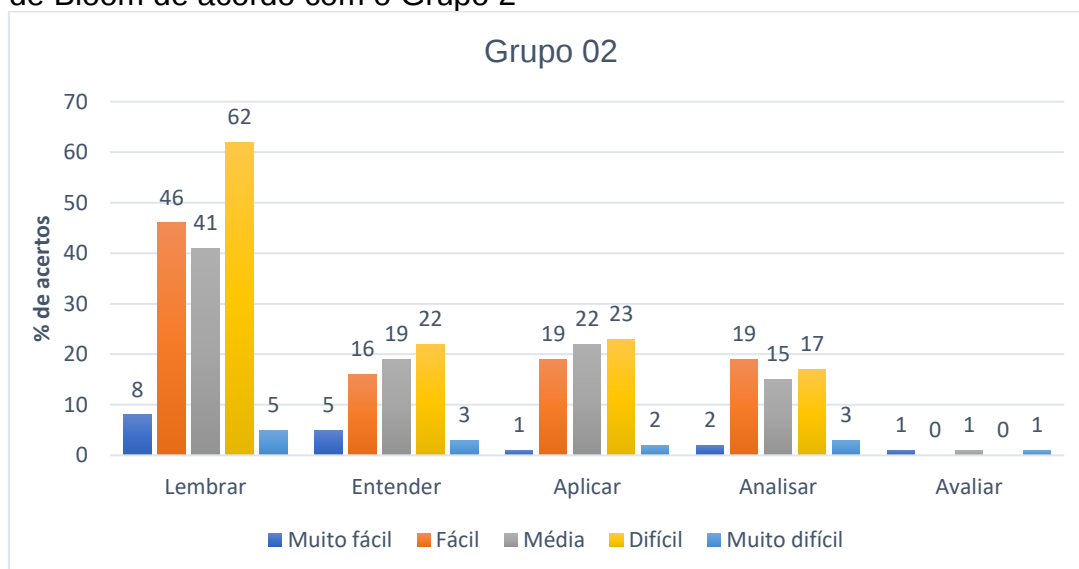
Gráfico 1: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 1



Qui-quadrado=18,6; p=0,280

Fonte: do autor.

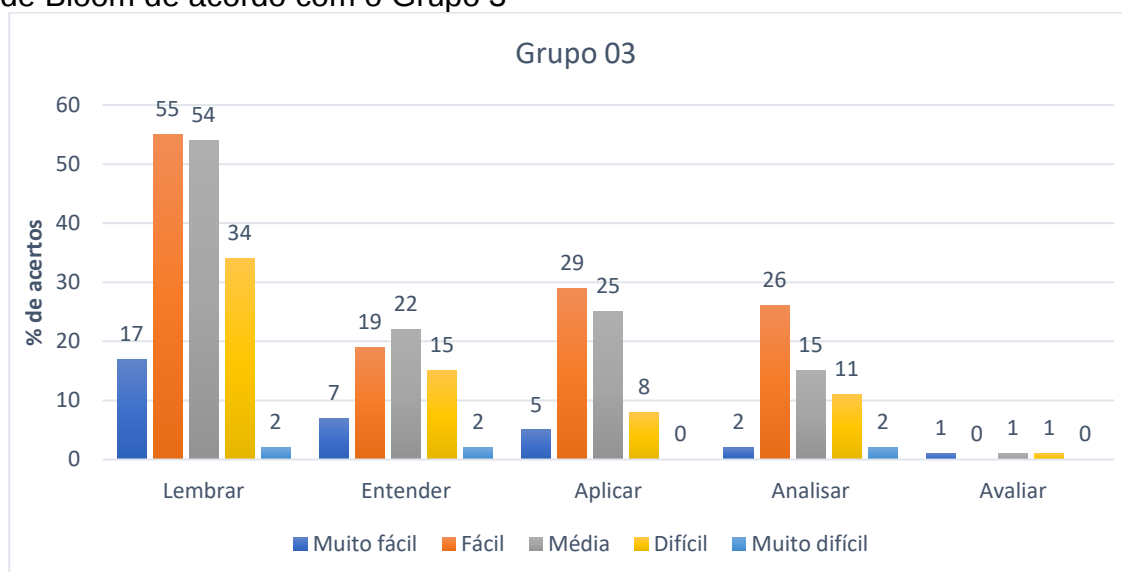
Gráfico 2: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 2



Qui-quadrado=12,6; p=0,705

Fonte: do autor.

Gráfico 3: Associação entre o grau de dificuldade (% de acertos) e a Taxonomia de Bloom de acordo com o Grupo 3



Qui-quadrado=16,4; $p=0,423$

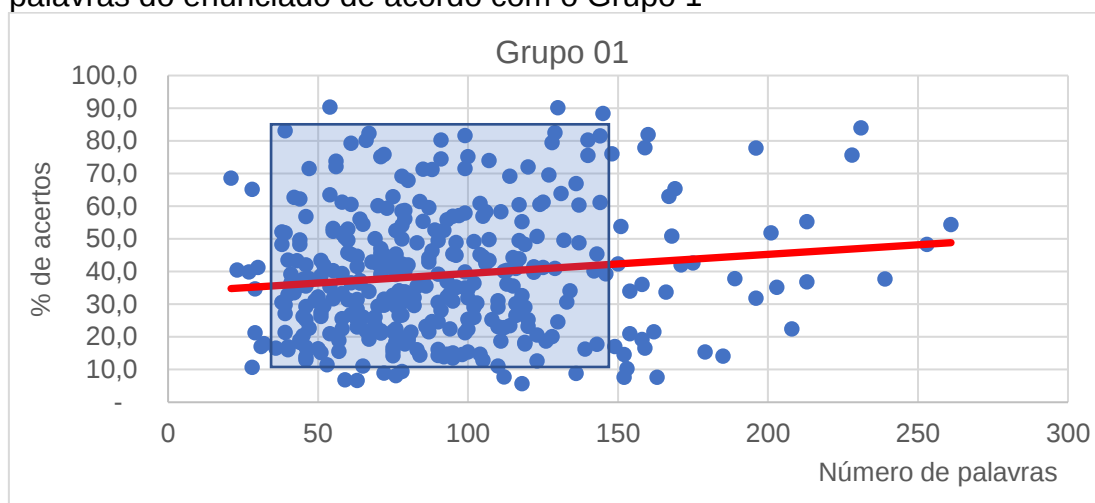
Fonte: do autor.

Quando comparado o grau de dificuldade dos itens com a Taxonomia de Bloom aplicada aos mesmos itens, não houve diferença significativa quando aplicado o teste do qui-quadrado. Visto que não houve associação entre o aumento da dificuldade e o grau de complexidade da questão para os três grupos.

Quando buscada a associação entre capacidade discriminativa do item e a Taxonomia de Bloom, também não foi encontrada diferença significativa. Isto é, não houve associação entre melhora da discriminação do item e grau de complexidade da questão.

Foram realizadas associações entre os dados psicométricos e o número de palavras do enunciado e das alternativas. Para isso foi utilizada a regressão linear simples, que engloba o coeficiente de determinação (R-quadrado). Foram plotados gráficos de dispersão, com destaque para a linha de tendência. Os dados devem convergir o mais próximo da linha de tendência para demonstrar que há correlação entre os fatores analisados.

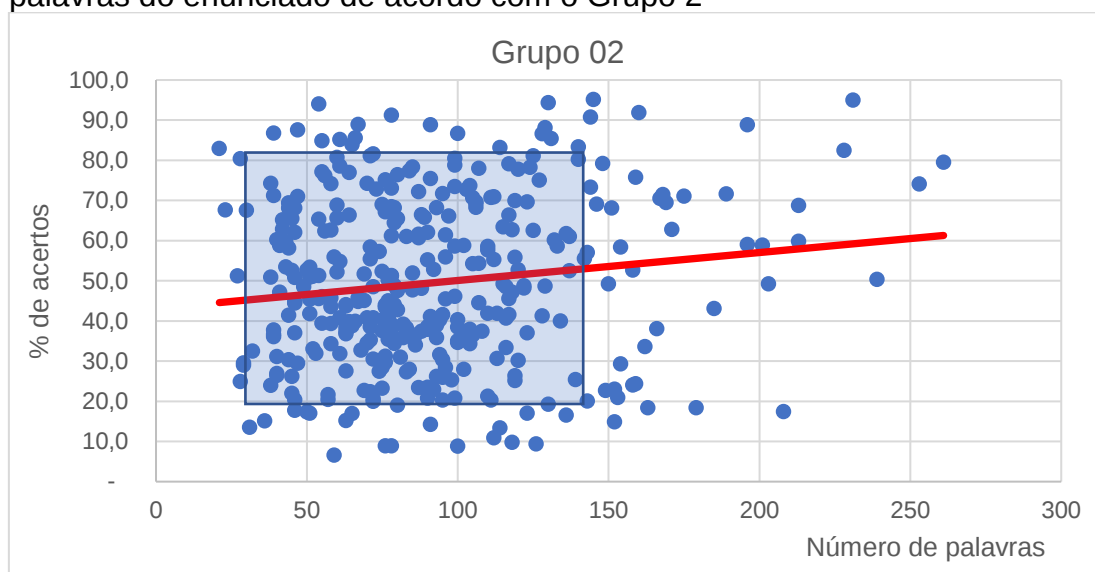
Gráfico 4: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 1



$F=0,015$; $p=0,106$; $R\text{-quadrado}=0,016$

Fonte: do autor.

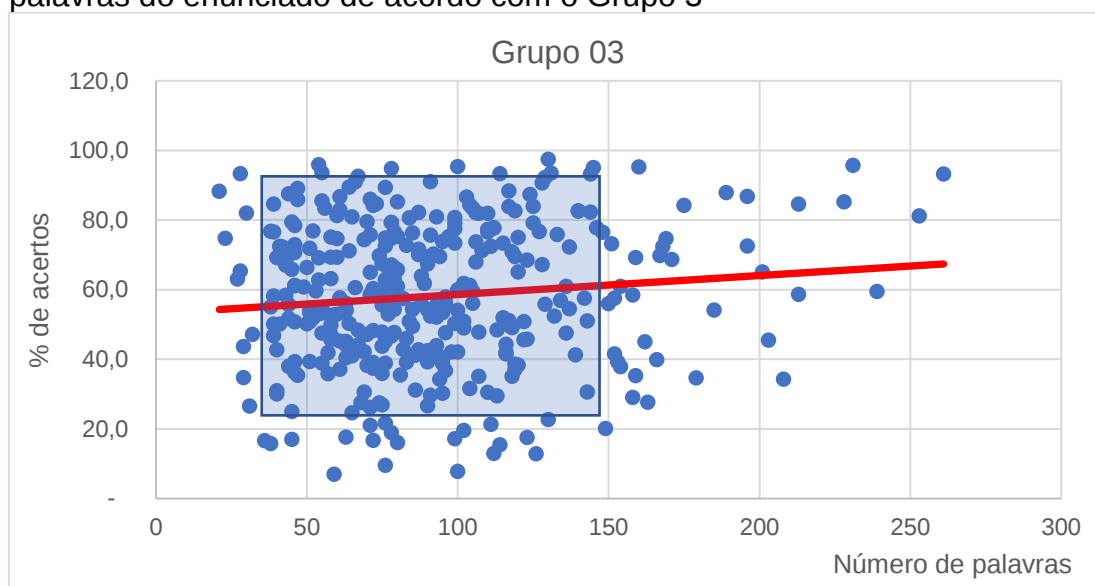
Gráfico 5: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 2



$F=0,008$; $p=0,121$; $R\text{-quadrado}=0,019$

Fonte: do autor.

Gráfico 6: Associação entre o grau de dificuldade (% de acertos) e o número de palavras do enunciado de acordo com o Grupo 3



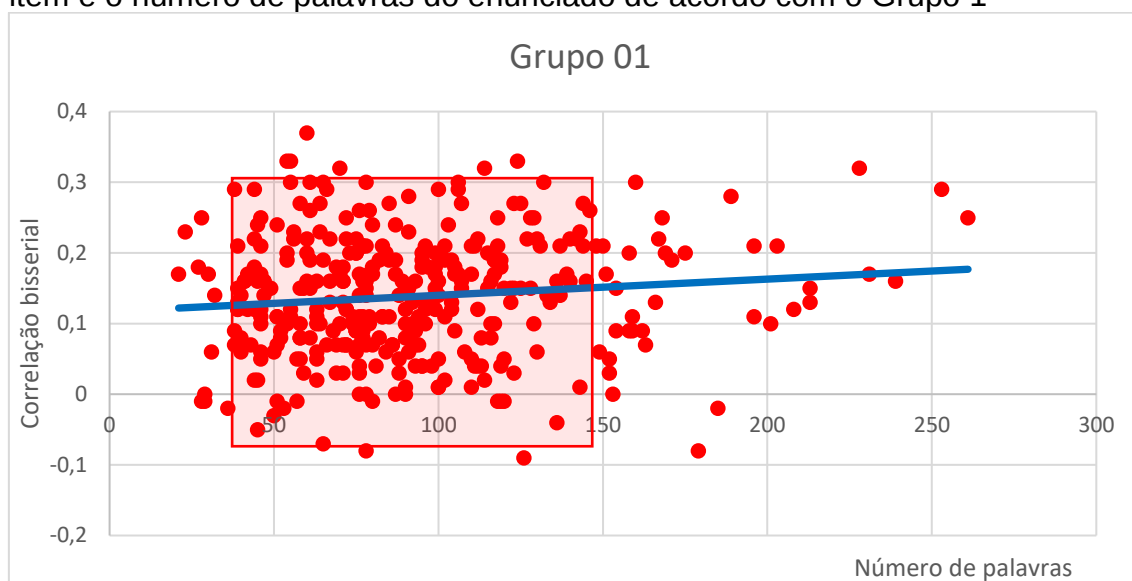
$F=0,035$; $p=0,105$; $R\text{-quadrado}=0,012$

Fonte: do autor.

Identificou-se que os dados não convergiram para a linha de tendência, o que demonstra resíduos estatísticos, logo, o eixo x não explica o eixo y.

Com relação à associação entre o grau de dificuldade do item e o número de palavras do enunciado da questão, o Gráfico 6 mostra que essa associação não existiu para o Grupo 3. A regressão linear simples usando como variável de desfecho o grau de dificuldade (% de acerto em cada item) e o número de palavras do enunciado como variável dependente mostrou que o modelo estava bem ajustado, porém a variável “número de palavras do enunciado” não explica o grau de dificuldade no item para qualquer um dos grupos.

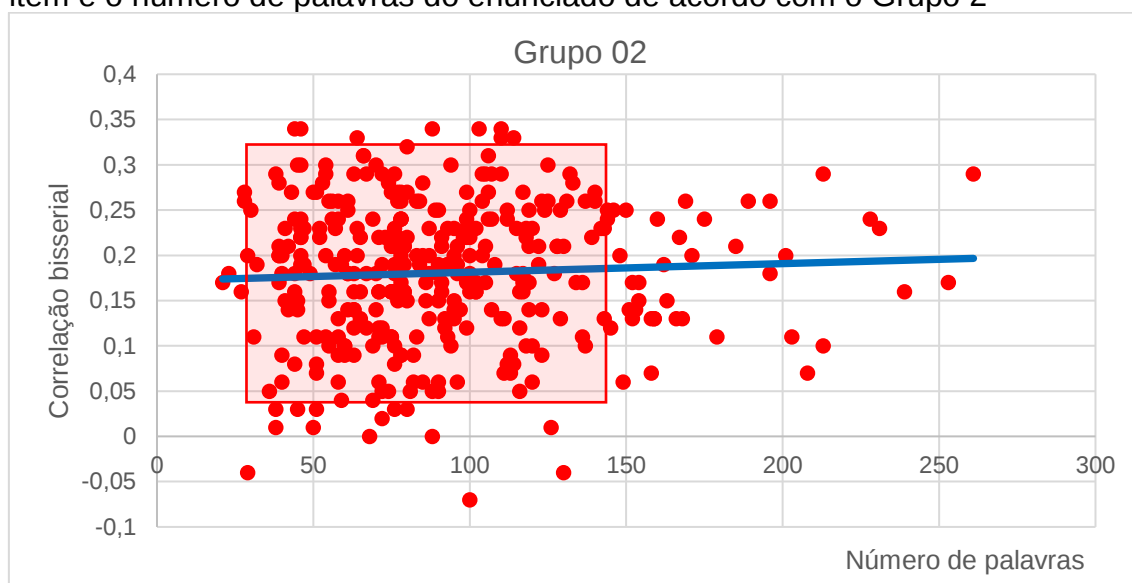
Gráfico 7: Associação entre a capacidade discriminativa (correlação bisserial) do item e o número de palavras do enunciado de acordo com o Grupo 1



$F=0,334$; $p=0,0000$; $R\text{-quadrado}=0,012$

Fonte: do autor.

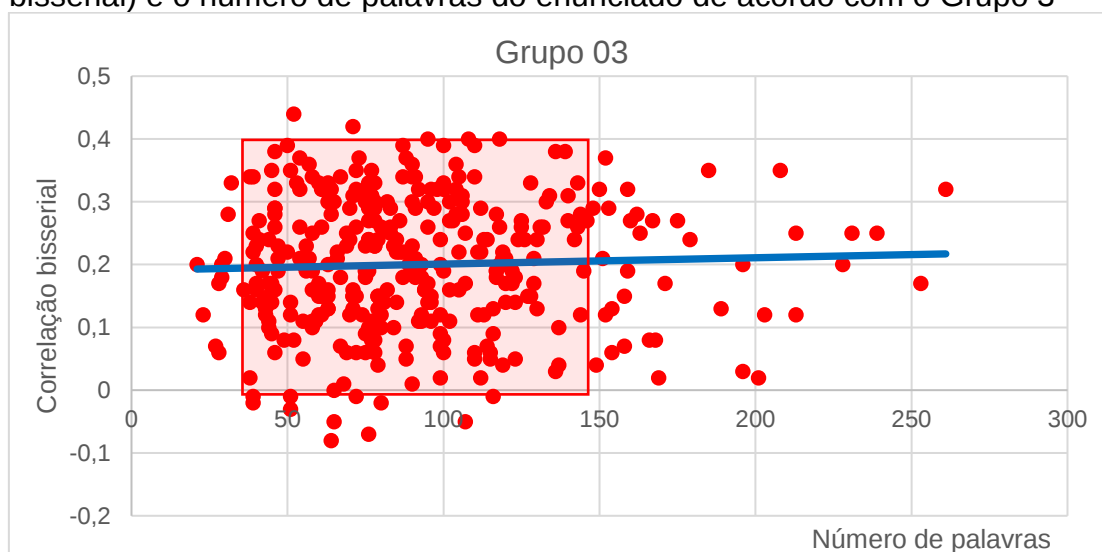
Gráfico 8: Associação entre a capacidade discriminativa (correlação bisserial) do item e o número de palavras do enunciado de acordo com o Grupo 2



$F=0,334$; $p=0,0000$; $R\text{-quadrado}=0,002$

Fonte: do autor.

Gráfico 9: Associação entre a capacidade discriminativa do item (correlação bisserial) e o número de palavras do enunciado de acordo com o Grupo 3



$F=0,443$; $p=0,000$; $R\text{-quadrado}=0,011$

Fonte: do autor.

Com relação à associação entre a capacidade discriminativa do item e o número de palavras do enunciado da questão, o Gráfico 9 mostra que essa associação não existiu para o Grupo 3. A regressão linear simples usando como variável de desfecho a capacidade discriminativa do item (Ponto-bisserial) e o número de palavras do enunciado como variável dependente mostrou que o modelo não estava bem ajustado e, portanto, uma variável não explica a outra.

Também não foi encontrada associação entre o número de palavras das alternativas e o grau de dificuldade e a capacidade discriminativa da questão.

Tabela 6: Taxonomia de Bloom versus a construção dos itens subdivididos em grandes áreas médicas nos Testes de Progresso aplicados em 2016, 2017 e 2018.

Área	Nível da Taxonomia de Bloom revisada					Total
	1	2	3	4	5	
Ciências Básicas	38 (64,4)	13 (22,0)	3 (5,1)	5 (8,5)	0	59
Clínica Cirúrgica	23 (39,0)	6 (10,2)	20 (33,9)	9 (15,3)	0	58
Clínica Médica	21 (35,6)	12 (20,3)	13 (22,0)	13 (22,0)	0	59
Ginecologia e Obstetrícia	31 (52,5)	7 (11,9)	10 (16,9)	12 (20,3)	0	60
Pediatria	26 (44,1)	8 (13,6)	14 (23,7)	10 (16,9)	1 (1,7)	59
Saúde Coletiva	23 (39,0)	19 (32,2)	7 (11,9)	7 (11,9)	2 (3,4)	58

Obs: Em 2016 foi anulada uma questão de Ciências Básicas; em 2017 foi anulada uma questão de Clínica Cirúrgica e uma de Clínica Médica; em 2018 foi anulada uma questão de Ciências Básicas, uma de pediatria e duas de Saúde Coletiva. As questões foram anuladas por apresentarem duas alternativas corretas ou duas alternativas idênticas.

Qui-quadrado=46,3 e $p=0,0007$.

Fonte: do autor.

Percebe-se relação das áreas do conhecimento com os diferentes graus de complexidade da Taxonomia de Bloom. As áreas de Saúde Coletiva e Ciências Básicas possuem maior acúmulo de questões nos níveis 1 (Lembrar) e 2 (Entender) da Taxonomia de Bloom. Além disso, percebeu-se uma distribuição mais homogênea dos itens nas outras áreas.

Foi observado que a maioria dos itens foi construída exigindo do aluno conteúdos de memorização.

7. DISCUSSÃO

O Teste de Progresso nos moldes atuais de sua aplicação parece não ser adequado para avaliar o desempenho cognitivo dos estudantes nas fases iniciais da graduação de medicina.

É importante acompanhar o desenvolvimento das habilidades básicas necessárias para o bom desempenho acadêmico dos estudantes, além de avaliar a qualidade do ensino, criar oportunidades de conscientização sobre as qualidades e fraquezas dos estudantes, visando ao planejamento de ações futuras que desenvolvam neles essas habilidades básicas (VENDRAMINI e DIAS, 2005).

A validade de uma avaliação estará garantida quando for possível demonstrar que o método escolhido avalia exatamente aquilo que se pretendia que o estudante aprendesse, o que indica que os resultados da sua aplicação podem ser generalizados. Melhor dizendo, a validade depende da adequação do método à natureza do domínio que se pretende avaliar, bem como se é coerente com o processo formativo (BOLLELA, BORGES e TRONCON, 2018).

Durante as revisões na elaboração dos itens é importante verificar a “pertinência do tema (validade), a redação e o formato de cada questão e se todas as recomendações de boas práticas foram atendidas” (BOLLELA, BORGES e TRONCON, 2018).

Neste estudo, alguns alunos precisaram ser excluídos, uma vez que foi detectado algum problema em seus gabaritos. A exclusão se deu por um aspecto visual do preenchimento da folha de respostas, ou seja, deixaram as respostas em branco ou assinalaram a mesma alternativa em mais de 50% dos itens. Visualmente, não foi possível identificar se houve marcação aleatória dos itens que não demonstravam estes padrões visuais. De acordo com a folha de respostas, este tipo de marcação poderia ter sido identificada se tivéssemos empregado a Teoria de Resposta ao Item, porém esta não era objeto deste estudo, podendo ser explorada em futuros estudos.

No que tange ao grau de confiabilidade, quanto mais o coeficiente se aproximar do valor 1, mais próximo do ideal, porém coeficientes acima de 0,85 são considerados bastante confiáveis. Observou-se então, nos três grupos e nos três anos de aplicação da prova, que o Grupo 3 (internato) apresentou grau de confiabilidade superior aos demais.

Segundo Bollela, Borges e Troncon (2018), “a confiabilidade constitui atributo que se relaciona à precisão, objetividade e consistência do método,

características que, quando presentes, indicam que os resultados da sua aplicação são reprodutíveis”. Muitos fatores influenciam a validade da avaliação: elaboração dos itens, aplicação e correção do exame. Os critérios de correção devem estar estabelecidos de forma clara e padronizada.

Visto que o conteúdo cobrado no Teste de Progresso é correspondente ao conhecimento necessário a um aluno do ciclo de internato da graduação médica, percebe-se que de fato, a análise da prova é mais confiável para os alunos dos dois últimos anos do curso.

O grau de confiabilidade foi encontrado pelo cálculo do Coeficiente de Alfa Cronbach, e este foi nitidamente baixo para as primeiras fases do curso de medicina (Ciclo Básico e Ciclo Clínico); já para a fase final (Internato), os itens apresentaram bom nível do grau de confiabilidade, o que é lógico pois a prova é elaborada para alunos da fase final do curso. Sendo assim, muitos assuntos cobrados ainda não foram apresentados para os alunos das fases iniciais.

O grau de dificuldade é um índice de proporção de respostas corretas do item encontrado por análise estatística, ou seja, ele traduz o grau de dificuldade encontrado pelos grupos. O grau de dificuldade é utilizado na tentativa de equilibrar a prova, com proporções de questões fáceis, moderadas e difíceis (BOLLELA, BORGES e TRONCON, 2018).

Semelhante ao grau de confiabilidade, o grau de dificuldade da prova mostrou-se bastante heterogêneo entre os três grupos, com itens de maior dificuldade para alunos das primeiras fases e itens de maior facilidade para alunos da última fase. Isso mostra que o grau de dificuldade é maior para os alunos das fases iniciais, pois estes não foram apresentados ainda a estes conteúdos.

A medida da capacidade discriminativa foi utilizada neste trabalho com objetivo de demonstrar os itens que mensuravam o ganho de cognição dos alunos de um mesmo grupo. Para a análise da capacidade discriminativa, foi utilizado o coeficiente bisserial, sendo este a medida da correlação entre o escore e a resposta dicotomizada para cada categoria do item analisado.

Para as fases iniciais (Ciclo Básico e Ciclo Clínico) os itens apresentavam baixa capacidade discriminativa, enquanto para alunos da fase final (Internato) havia melhora desta discriminação. Isso demonstra novamente, que os alunos

das primeiras fases não se apropriaram dos conteúdos cobrados no teste até o momento da avaliação.

A Taxonomia de Bloom preconiza que quanto maior o nível taxonômico do item, melhor o nível de avaliação da metacognição. Analisando os resultados do nível da Taxonomia de Bloom revisada, observa-se predomínio maior nos níveis cognitivos mais simples, os quais necessitam de memorização, como relembrar e entender o item. Com esse resultado, é possível perceber que no decorrer dos três anos houve melhora na formulação dos itens de prova, como pôde ser demonstrado por um discreto aumento do nível taxonômico das questões no passar dos anos. Porém, ainda com predomínio nas questões de memorização.

Apesar do grupo de professores elaboradores das questões serem docentes capacitados, ainda é muito alta a porcentagem de itens que cobram níveis baixos na Taxonomia de Bloom.

A Taxonomia de Bloom baseia-se no domínio cognitivo, ou seja, conhecimento e raciocínio. Esse domínio engloba seis níveis de complexidade crescente (relembrar, entender, aplicar, analisar, avaliar e criar). Para avaliação destas habilidades utiliza-se três níveis referenciais para elaboração das questões: nível básico (conhecimento factual, memorização e relembrar), nível intermediário (envolve a capacidade de compreensão, interpretação e aplicação daquele conhecimento para resolver problemas mais simples) e nível avançado (envolve níveis de análise, síntese e avaliação para propor soluções para problemas mais complexos, como os que habitualmente se apresentam ao profissional da área da saúde) (BOLLELA, BORGES e TRONCON, 2018).

Para atender ao propósito de fornecer subsídios ao andamento do curso para os órgãos gestores, a avaliação deve ser bem estruturada, preferencialmente organizada por uma equipe multidisciplinar e com visão integral do currículo do curso. Aliado a isto, é importante ressaltar a estruturação dos itens e das alternativas nos testes de múltipla escolha para garantir fidedignidade dos resultados. Neste sentido, os enunciados dos testes não devem exigir do estudante apenas a memorização, mas principalmente abordar categorias taxonômicas mais elevadas, como a síntese ou avaliação (KRATHWOHL, 2002).

O número de itens com erros pedagógicos foi muito baixo, demonstrando que os elaboradores estão empenhados em elaborar bons itens, entretanto, a respeito do nível dos itens relacionado à Taxonomia de Bloom há maior predominância nos níveis de memorização (Lembrar e Entender).

A elaboração de questões de múltipla escolha de boa qualidade necessita seguir boas práticas e estas podem ter impacto educacional positivo, no entanto, requer treino e bastante dedicação. Outro aspecto que pode melhorar a qualidade do item é a revisão feita pelos elaboradores onde definem-se distratores homogêneos, erros gramaticais e de concordância, ausência de pergunta clara do enunciado do teste, dentre outros erros (BOLLELA, BORGES e TRONCON, 2018).

O número de palavras de cada questão não demonstrou impacto significativo sobre o desempenho dos acadêmicos. No entanto, é necessário ressaltar que a extensão dos conteúdos cobrados, assim como o cálculo de tempo designado para realização de cada tarefa deve ser levado em consideração visto o tempo limite de prova. Bollela, Borges e Troncon (2018) ressaltam que as condições de aplicação do método escolhido na avaliação, em termos de tempo destinado para as tarefas, respostas às questões e ambiente físico, devem ser também analisados para melhor efetividade.

Os enunciados devem ser elaborados de maneira objetiva, definindo claramente o que está sendo solicitado. É fundamental incluir todas as informações necessárias para resolução. O excesso ou escassez de informações pode dificultar a compreensão do problema e ainda conduzir a uma dupla interpretação (SILVA, 2011).

Os testes de múltipla escolha possuem fatores limitantes que devem ser levados em consideração. É importante lembrar que é impossível utilizar este recurso de maneira isolada e exclusiva, visto que as profissões de saúde abrangem problemas e situações para os quais não existe uma única solução correta ou inexistente consenso bem firmado sobre o que é mais adequado (BOLLELA, BORGES e TRONCON, 2018).

Sobre as áreas médicas a serem cobradas no TP vale ressaltar que todos os itens necessitam ter “coerência com o currículo nuclear do curso e com as diretrizes curriculares nacionais para o curso de Medicina” (SAKAI, FERREIRA FILHO e MATSUO, 2011). Na presente tese, foi perceptível que as áreas

médicas de Ciências Básicas e Saúde Coletiva apresentaram níveis mais baixos na Taxonomia de Bloom (Lembrar e Entender), isso se dá pela dificuldade de elaborar casos clínicos nessas áreas, visto que estas cobram conhecimentos básicos.

Atualmente, os testes estão sofrendo significativas revisões, bem como novos instrumentos estão sendo criados na área educacional. Para uma adequada avaliação cognitiva é necessário verificar qual conhecimento está sendo transmitido durante o curso universitário e como este conhecimento está sendo assimilado pelos alunos (VENDRAMINI e DIAS, 2005).

Para avaliação do crescimento cognitivo utiliza-se um método estatístico de equivalência de escores, servindo “para comparar escores em formas diferentes de teste, ajustando níveis diferentes de dificuldade e de conteúdo dos testes” (SAKAI, FERREIRA FILHO e MATSUO, 2011). Essa avaliação poderia ser melhor empregada se o TP fosse customizado de forma que cada ciclo da graduação fosse avaliado conforme seu conhecimento cognitivo necessário para a fase em que se encontra no curso.

A construção do ensino médico pode ser apresentada em: ciclo básico, ciclo clínico e internato. Este modelo é baseado no currículo flexneriano que, apesar de ter sofrido algumas alterações ao longo do tempo, não mudou significativamente, preservando suas características e mantendo a falta de integração entre os ciclos e conteúdos (TENÓRIO et al., 2016). Percebe-se um conteúdo teórico mais concentrado nos quatro primeiros anos do curso de medicina.

O TP tem sido utilizado pelas escolas médicas que implantaram mudanças curriculares, com currículos baseados na comunidade, aprendizagem baseada em problemas e currículos orientados por competências (SAKAI et al., 2008). Apesar das escolas que compõem o NAPISUL II possuírem metodologia ativa de ensino, ainda assim, percebe-se a influência do currículo flexneriano. Constatam-se também, atualmente, novos esforços a respeito da metodologia de ensino, incentivando metodologias ativas e com currículo centrado no estudante.

8. CONCLUSÃO

Quanto ao grau de confiabilidade, grau de dificuldade e capacidade discriminativa encontrados nos itens aplicados no TP de 2016, 2017 e 2018, conclui-se que o teste foi mais consistente para os alunos dos dois últimos anos da graduação. Ou seja, os alunos do internato conseguiram compreender melhor o item e o que lhes era cobrado.

Acerca da classificação dos itens, notou-se que a maioria das questões analisadas estavam compreendidas nos níveis mais baixos da Taxonomia de Bloom revisada, sendo mais prevalentes os níveis referentes à lembrar, entender e aplicar.

Neste estudo, a análise dos dados psicométricos (grau de dificuldade e capacidade discriminativa) não apresentou relação com os níveis da Taxonomia de Bloom revisada, portanto, não houve associação entre estes dados com o nível de complexidade dos itens aplicados.

Percebeu-se que o tamanho do enunciado e das alternativas não influenciou no grau de dificuldade e na capacidade discriminativa.

As áreas médicas abordadas na avaliação que apresentaram níveis mais baixos na Taxonomia de Bloom revisada foram as áreas de: Ciências Básicas e Saúde Coletiva. Sendo assim, exigiram mais a memorização e o entendimento do item (Lembrar, Entender).

Foi possível constatar que os itens, em sua grande maioria, solicitaram conteúdos de níveis mais baixos na Taxonomia de Bloom (Lembrar e Entender).

Evidenciou-se com este estudo que os alunos do internato apreenderam melhor o conteúdo que estava sendo cobrado.

9. CONSIDERAÇÕES FINAIS

Esta tese teve como objetivo central avaliar as características das questões aplicadas no TP no curso de graduação de medicina em um consórcio de escolas médicas no Sul do Brasil. Com isso, considera-se que o objetivo foi alcançado, pois analisando o resultado destas características, o estudo apontou para uma inadequação do TP no modelo de sua atual aplicação. Este estudo surgiu com a intenção de nortear melhorias nesta ferramenta de avaliação.

Na atualidade, percebe-se uma preocupação constante com a melhoria e construção do conhecimento pedagógico. Em relação ao ensino-aprendizagem, notam-se constantemente modificações envolvendo novas metodologias de ensino, construção de novos conhecimentos científicos e tecnologias cada vez mais céleres e avançadas. Há, então, no contexto atual um movimento de mudança relacionado ao processo homem-máquina imerso em avanços acelerados do conhecimento técnico-científico.

Necessita-se, portanto, de uma avaliação cada vez mais precisa, fidedigna e ajustada para cada fase do aprendizado do curso médico, visando excelência. Para tanto, torna-se indispensável: avaliar, analisar e refletir sobre métodos empregados na mensuração do conhecimento de forma constante, além do *feedback* robusto às instituições e seus alunos. Neste cenário, o desenvolvimento do TP de forma personalizada para cada etapa da construção do aprendizado médico pode elevar esta ferramenta avaliativa a um patamar de excelência pedagógica.

Dentro do contexto do TP personalizado é fundamental ressaltar a importância prévia da calibração dos itens, montagem de um banco de itens, avaliação adaptativa computadorizada e equalização da prova, com o intuito de avaliar o ganho cognitivo no decorrer dos anos da graduação médica.

Além disso, a personalização do teste pode tornar a prova mais atrativa do ponto de vista dos estudantes, por maior identificação com o conteúdo que está sendo exigido, o que aumenta o grau de confiabilidade da avaliação. Isto permite inferir mudanças mais precisas e diagnósticas, tanto na autoavaliação do aluno, quanto nas estratégias a serem aplicadas pelo gestor no decorrer do curso.

Neste trabalho foram excluídos os gabaritos visualmente inadequados, como por exemplo, quando aluno marcou a mesma alternativa em mais de 50% da prova, criando um padrão visual. Sugere-se um novo estudo onde seja aplicada a Teoria de Resposta ao Item para que os critérios de exclusão sejam mais precisos e não apenas visuais como foi realizado neste trabalho.

Em relação às boas práticas de elaboração dos itens, propõem-se oficinas para o desenvolvimento docente. Estas oficinas devem abordar os temas de Taxonomia de Bloom revisada e *blueprinting*, para que sejam criados itens de avaliação com características de análise psicométrica e níveis taxonômicos mais elevados. Além disso, preconiza-se o fornecimento de um manual ou guia para elaboração adequada dos itens.

Com estes dados, propõe-se que o Teste de Progresso seja formulado de maneira personalizada para cada etapa da formação médica (Ciclo Básico, Ciclo Clínico e Internato). O objetivo do teste é avaliar o ganho de conhecimento no decorrer do tempo, e este também pode ser avaliado dentro da matriz curricular do período correspondente ao semestre/ano ao qual o aluno está matriculado.

REFERÊNCIAS

ABEM (Associação Brasileira de Educação Médica). **Avaliação do estudante: contribuição do Teste de Progresso**. ABEM, Brasília (DF), 2017. Disponível em: <<https://abem-educmed.org.br/acoes/avaliacao-do-estudante-contribuicao-do-teste-de-progresso/>>.

ANDERSON, L.W. **Rethinking Bloom's Taxonomy**: implication for testing and assessment. Columbia: University of South Carolina, 1999. (Report n. MF01/PC01).

ANDERSON, L.W. et. al. **A taxonomy for learning, teaching and assessing**: a revision of Bloom's Taxonomy of Educational Objectives. Nova York: Addison Wesley Longman, 2001. 336 p.

ANDRADE, D.F. Comparando desempenho de grupos de alunos por intermédio da teoria de resposta ao item. **Estudos em Avaliação Educacional**, Itatiba (SP), v.23, p.31-69, 2001.

BALDIM, T.L.; VICENTE, C.A.O.; ARCURI, M.B. O Teste de Progresso sob a visão do discente. **Rev da Faculdade de Medicina de Teresópolis**, Teresópolis (RJ), v.2, n.1, p.41-54, 2018.

BHAKTA, B.; TENNANT, A.; HORTON, M.; LAWTON, G.; ANDRICH, D. Using item response theory to explore the psychometric properties of extended matching questions examination in undergraduate medical education. **BMC Med Educ**, v.5, p.9, 2005.

BLOOM, B.S. et al. **Taxonomy of educational objectives**. New York: David McKay, 1956, v.1, p.262.

BOLLELA, V.R.; BORGES, M.C.; TRONCON, L.E.A. Avaliação Somativa de Habilidades Cognitivas: Experiência Envolvendo Boas Práticas para a Elaboração de Testes de Múltipla Escolha e a Composição de Exames. **Rev Bras de Educação Médica**, Ribeirão Preto (SP), v.42, n.4, p.74-85, 2018.

BRASIL. **Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Guia de elaboração e revisão de itens. Brasília (DF), 2010.

CONKLIN, J. A taxonomy for learning, teaching and assessing: a revision of Blooms's taxonomy of educational objectives. **Educational Horizons**, v.83, n.3, p.153-159, 2005.

CRONBACH, L.J.; GLESER, G.C.; NANDA, H.; et al. The dependability of behavioral measurements: Theory for generalizability of scores and profiles. **New York: John Wiley**, 1972.

EMBRETSON, S.E. The new rules of measurement. **Psychological Assessment**, v.8, n.4, p.341-9, 1996.

FERREIRA, A.R.L.; WEYH, C.B. Tecnologia da informação no processo da avaliação como estratégia para qualificar a prática docente. **Congresso Internacional de Educação e Tecnologias 2018**, CIET EnPED – UFSCar, p. 1-7.

FERRAZ, A.P.C.M.; BELHOT, R.V. Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. **Gest. Prod.**, São Carlos (SP), v. 17, n. 2, p.421-431, 2010.

FERNANDES, D.C.; NAGTEGAAL, M.; NOORDZIJ, G.; TIO, R.A. Cumulative assessment: does it improve students knowledge acquisition and retention? **Scientia Medica education in Health Sciences**, the Netherlands, v.28, n.4 (ID31880), 2018.

HAMAMOTO FILHO, P.T.; LOURENÇÃO, P.L.T.A.; VALLE, A.P. do; ABBADE, J.F.; BICUDO, A.M. A correlação entre a pontuação dos testes de progresso dos alunos e seu desempenho em um processo de seleção de residência. **Medical Science Educator**, v.29, n.1, p.1-5, 2019.

KRATHWOHL, D.R.A revision of Bloom's taxonomy: an overview. **Theory in Practice**, v. 41, n. 4, p. 212-218, 2002.

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

MIRANDA JUNIOR, U.J.P.; TATSCH, J.F.S.; BRAGA, C.S.; SPENGLER JUNIOR, S.L.; NOVAES, M.R.C.G. Avaliação Critério-Referenciada em Medicina e Enfermagem: Diferentes Concepções de Docentes e Estudantes de uma

Escola Pública de Saúde de Brasília, Brasil. **Rev Bras de Educação Médica**, Brasília (DF), v.42, n.3, p.67-77, 2018.

NUNES, C.H.S.S.; PRIMI, R. Impacto do tamanho da amostra na calibração de itens e estimativa de escores por teoria de resposta ao item. **Aval psicol**, Porto Alegre (RS), v.4, n.2, ISSN 2175-3431, 2005.

PASQUALI, L. **Psicometria: Teoria e aplicações**. Brasília, DF: Editora da Universidade de Brasília, 1997.

PERRENOUD, P. **As competências para ensinar no século XXI**: a formação dos professores e o desafio da educação. 1 ed, Porto Alegre (RS), ARTMED, 2002.

PORTNEY, L. G.; WATKINS, M. P. Power analysis and determination of sample size. *Foundations of Clinical Research: Applications to practice*. **Norwalk, CT: Applenton & Lange**, 1993.

PRIMI, R. Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes. **Avaliação psicológica**, Itatiba (SP), v.11, n.2, p.297-307, 2012.

RODRIGUES, M. Proposta de Análise de itens das Provas do Saeb sob a Perspectiva Pedagógica e a Psicométrica. **Estudos em Avaliação Educacional**, São Paulo (SP), v.34, p.43-78, 2006.

ROSA, M.I.; ISOPPO, C.C.; CATTANEO, H.D.; MADEIRA, K.; ADAMI, F.; FERREIRA FILHO, O.F. O Teste de Progresso como Indicador para Melhorias em Curso de Graduação em Medicina. **Rev Bras de Educação Médica**, Criciúma (SC) e Londrina (PR), v.41, n.1, p.58-68, 2017.

SAKAI, M.H.; FERREIRA FILHO, O.F.; ALMEIDA, M.J.; MASHIMA, D.A., MARCHESE, M.C. Teste de progresso e avaliação do curso: dez anos de experiência da medicina da Universidade Estadual de Londrina. **Rev Bras de Educação Médica**, Londrina (PR), v.32, n.2, p.254-263, 2008.

SAKAI, M.H.; FERREIRA FILHO, O.F.; ALMEIDA, M.J.; MATSUO, T. Avaliação do crescimento cognitivo do estudante de Medicina: aplicação do teste de equalização no Teste de Progresso. **Rev Bras de Educação Médica**, Londrina (PR), v.35, n.4, p.493-501, 2011.

SCHUWIRTH, L.W.T.; VAN DER VLEUTEN, C.P.M. The use of progress testing. **Perspect Med Educ**, Australia e The Netherlands, v.1, p.24-30, 2012.

SILVA, K.F. Inadequações em formulações de enunciados de atividades e avaliações escritas. **Facos educ.**, João Pessoa (PB), p. 58-68, 2011.

STREINER, D.L . Being inconsistent about consistency: when coefficient alpha does and doesn't matter. **Journal of Personality Assessment**, Toronto, Canadá, v.80, p.217-222, 2003.

TENÓRIO, L.P., ARGOLO, V. A., DE SÁ, H. P., DE MELO, E. V.; COSTA, E. F O. Saúde mental de estudantes de escolas médicas com diferentes modelos de Ensino. **Rev Bras de Educação Médica**, Rio de Janeiro (RJ), v.40, n.4, 2016.

VENDRAMINI, C.M.M., DIAS, A.S. Teoria de Resposta ao Item na análise de uma prova de estatística em universitários. **Psico-USF**, Itatiba (SP), v.10, n.2, p.201-210, 2005.

VERHOEVEN, B.H.; VERWIJNEM, G.M.; SCHERPBIER, A.J.; VAN DER VLEUTEN, C.P. Growth of medical knowledge. **Med Educ**, The Netherlands, v.36, n.8, p.711-7, 2002.

VLEUTEN, C.V.D.; FREEMAN, A.; COLLARES, C.F. Progress test utopia. **Perspect Med Educ**, The Netherlands e UK, v.7, p.136-138, 2018.

WAGENER, S.; MOLTNER, A.; TIMBIL, S.; GORNOSTAYEVA, M.; et al. Development of a competency-based formative progress test with student-generated MCQs: Results from a multi-centre pilot study. **GMS Z Med Ausbild**, Germany, v.32, n.4 (Doc46), 2015.

ANEXO A

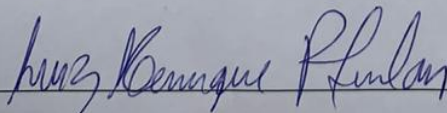
Carta de anuência do NAPISUL II

TERMO DE CONSENTIMENTO / CARTA DE ANUÊNCIA

Declaramos para fins de realização da pesquisa acerca do Teste de Progresso em Medicina, intitulada inicialmente como “Análise da qualidade dos itens aplicados no Teste do Progresso nos anos de 2016, 2017 e 2018 em um consórcio brasileiro” sob responsabilidade do pesquisador Ademir Garcia Reberti, orientado pelo Dr. Jean Carl Silva, que o Consórcio Napisul II (Núcleo de Apoio Pedagógico Interinstitucional Sul II), está ciente e de acordo com o fornecimento dos dados anonimizados, sem identificação das escolas médicas e dos alunos participantes para realização da referida pesquisa.

A equipe se compromete a submeter o projeto à Comissão de Ética e Pesquisa em Seres Humanos da Instituição proponente. Assim como o cumprimento do pesquisador aos requisitos das Resoluções do Conselho Nacional de Saúde e suas complementares, comprometendo-se utilizar os dados resultantes da pesquisa, exclusivamente para os fins científicos.

Florianópolis, em 02 / 08 / 2019.



Luiz Henrique Picolo Furlan – Coordenador Geral do Napisul II

ANEXO B

Carta de solicitação de dispensa de TCLE

**SOLICITAÇÃO DE DISPENSA DO TERMO DE CONSENTIMENTO LIVRE E
ESCLARECIDO - TCLE**

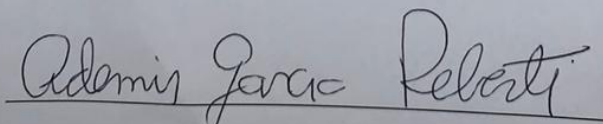
Prezado(a) Coordenador(a) CEP da Universidade da Região de Joinville

Eu, Ademir Garcia Reberti, portador do CPF nº 617.836.689-20, pesquisador responsável pelo projeto: **"ANÁLISE DA QUALIDADE DOS ITENS APLICADOS NO TESTE DE PROGRESSO NOS ANOS DE 2016, 2017 E 2018 EM UM CONSÓRCIO BRASILEIRO"** solicito perante este Comitê de Ética em Pesquisa a dispensa da utilização do Termo de Consentimento Livre e Esclarecido – TCLE para a coleta de dados, tendo em vista que o mesmo utilizará somente os dados obtidos a partir do estudo de material já coletado a partir dos Testes de Progresso aplicados nos anos de 2016, 2017 e 2018 pelo NAPISUL II.

Nestes termos, me comprometo a cumprir todas as diretrizes e normas reguladoras descritivas na Resolução CNS nº 466/12 e suas complementares.

Atenciosamente,

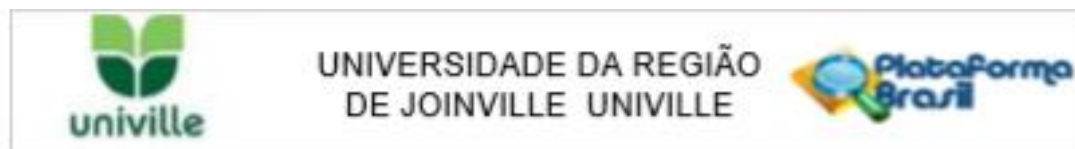
Joinville, 11 de outubro de 2019.



ADEMIR GARCIA REBERTI

ANEXO C

Parecer consubstanciado – Plataforma Brasil



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: ANÁLISE DA QUALIDADE DOS ITENS APLICADOS NO TESTE DE PROGRESSO NOS ANOS DE 2016, 2017 E 2018 EM UM CONSÓRCIO BRASILEIRO

Pesquisador: ADEMIR GARCIA REBERTI

Área Temática:

Versão: 2

CAAE: 24357419.1.0000.5366

Instituição Proponente: FUNDAÇÃO EDUCACIONAL DA REGIÃO DE JOINVILLE - UNIVILLE

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.832.259

Apresentação do Projeto:

Conforme exposto no parecer consubstanciado nº 3.755.251.

Objetivo da Pesquisa:

Conforme exposto no parecer consubstanciado nº 3.755.251.

 Avaliação dos Riscos e Benefícios:

Conforme exposto no parecer consubstanciado nº 3.755.251. O pesquisador responsável informa que os dados que serão utilizados nesta pesquisa já foram coletados em anos anteriores, estando sob posse e guarda do pesquisador responsável.

Comentários e Considerações sobre a Pesquisa:

Conforme exposto no parecer consubstanciado nº 3.755.251. O pesquisador responsável informa os participantes da pesquisa, a responsabilidade por posse e guarda dos dados e o envio de novo cronograma para o desenvolvimento desta pesquisa.

Considerações sobre os Termos de apresentação obrigatória:

Conforme exposto no parecer consubstanciado nº 3.755.251.

Recomendações:

Ao finalizar a pesquisa, o (a) pesquisador (a) responsável deve enviar ao Comitê de Ética, por meio do sistema Plataforma Brasil, o Relatório Final (modelo de documento na página do CEP no site da Univille Universidade).

Endereço: Rua Paulo Matschizi, nº 10, Bloco B, Sala 119, campus Bom Retiro

Bairro: Zona Industrial

CEP: 89.219-710

UF: SC

Município: JOINVILLE

Telefone: (47)3461-9235

E-mail: comitetic@univille.br

Segundo a Resolução 466/12, no item

XI- DO PESQUISADOR RESPONSÁVEL

XI.2 - Cabe ao pesquisador:

d) Elaborar e apresentar o relatório final;

Modelo de relatório para download na página do CEP no site da Univille Universidade.

Conclusões ou Pendências e Lista de Inadequações:

O projeto "ANÁLISE DA QUALIDADE DOS ITENS APLICADOS NO TESTE DE PROGRESSO NOS ANOS DE 2016, 2017 E 2018 EM UM CONSÓRCIO BRASILEIRO", de CAAE "24357419.1.0000.5366" teve sua(s) pendência(s) esclarecida(s) pelo(a) pesquisador(a) "ADEMIR GARCIA REBERTI", de acordo com a Resolução CNS 466/12 e complementares, portanto, encontra-se APROVADO.

Informamos que após leitura do parecer, é imprescindível a leitura do item "O Parecer do CEP" na página do Comitê no site da Univille, pois os procedimentos seguintes, no que se refere ao enquadramento do protocolo, estão disponíveis na página. Segue o link de acesso <http://www.univille.edu.br/pt-BR/a-univille/proreitorias/prppg/setores/area-pesquisa/comite-etica-pesquisa/status-parecer/645062>

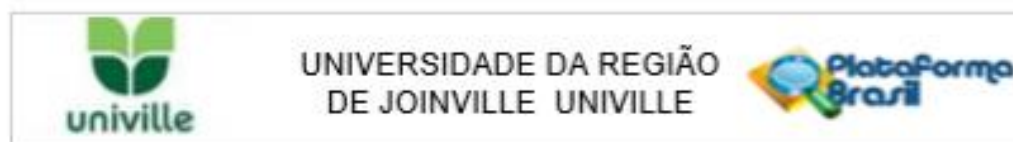
Considerações Finais a critério do CEP:

Diante do exposto, o Comitê de Ética em Pesquisa da Universidade da Região de Joinville - Univille, de acordo com as atribuições definidas na Res. CNS 466/12, manifesta-se pela aprovação do projeto de pesquisa proposto.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1431177.pdf	29/12/2019 16:45:46		Aceito
Projeto Detalhado / Brochura Investigador	ProjetoRebertICEP.docx	29/12/2019 16:45:34	NAYME HECHER MONFREDINI	Aceito
Outros	CARTA_RESPOSTA_AS_PENDENCIA	29/12/2019	NAYME HECHER	Aceito

Endereço: Rua Paulo Matschitzki, nº 10, Bloco B, Sala 119, campus Bom Retiro
Bairro: Zona Industrial **CEP:** 89.219-710
UF: SC **Município:** JOINVILLE
Telefone: (47)3451-9235 **E-mail:** comitetic@univille.br



Continuação do Parecer: 3.832.259

Outros	docx	16:44:32	MONFREDINI	Aceito
Outros	cartaanuenciapapisul.jpeg	24/10/2019 12:33:33	NAYME HECHEM MONFREDINI	Aceito
Outros	carta_de_esclarecimento_napisul.docx	24/10/2019 12:32:19	NAYME HECHEM MONFREDINI	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	dispensaTCLE.docx	24/10/2019 12:31:00	NAYME HECHEM MONFREDINI	Aceito
Folha de Rosto	folhaderostoassinada.pdf	24/10/2019 12:29:16	NAYME HECHEM MONFREDINI	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

JOINVILLE, 11 de Fevereiro de 2020

Assinado por:

Marcia Luciane Lange Silveira
(Coordenador(a))

Endereço: Rua Paulo Matschitzki, n° 10, Bloco B, Sala 119, campus Bom Retiro

Bairro: Zona Industrial

CEP: 89.219-710

UF: SC **Município:** JOINVILLE

Telefone: (47)3461-9235

E-mail: cometica@univille.br

ANEXO D

AUTORIZAÇÃO

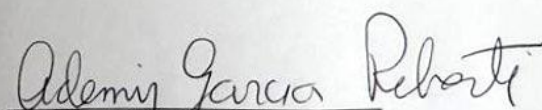
Nome do autor: Ademir Garcia Reberti

RG: 6.502.965-PR

Título da Tese: **Avaliação das características dos itens aplicados no Teste de Progresso nos anos de 2016, 2017 e 2018 em um consórcio brasileiro.**

Autorizo a Universidade da Região de Joinville – UNIVILLE, através da Biblioteca Universitária, disponibilizar cópias da tese de minha autoria.

Joinville, 09 de julho de 2020.


Assinatura do aluno