

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO EM CIÊNCIAS

Carlos Eduardo Andrade Pinheiro

**TESTE DE PROGRESSO, UMA AVALIAÇÃO SERIADA NOS CURSOS DE  
MEDICINA: ANÁLISE, MELHORIAS E SELEÇÃO DE BOAS QUESTÕES**

Porto Alegre

2023



Carlos Eduardo Andrade Pinheiro

**TESTE DE PROGRESSO, UMA AVALIAÇÃO SERIADA NOS CURSOS DE  
MEDICINA: ANÁLISE, MELHORIAS E SELEÇÃO DE BOAS QUESTÕES**

Tese apresentada no **Programa de Pós-Graduação em Educação em Ciências: Química da Vida e Saúde da Universidade Federal do Rio Grande do Sul** como requisito parcial para a obtenção do título de doutor em Educação em Ciências.

Orientador: Prof. Dr. Diogo Onofre Gomes de Souza

Porto Alegre

2023



## Dados Internacionais de Catalogação na Publicação

Pinheiro, Carlos Eduardo  
TESTE DE PROGRESSO, UMA AVALIAÇÃO SERIADA NOS  
CURSOS DE MEDICINA: ANÁLISE, MELHÓRIAS E SELEÇÃO DE  
BOAS QUESTÕES / Carlos Eduardo Pinheiro. -- 2023.  
79 f.  
Orientador: Diogo Onofre Gomes de Souza.

Tese (Doutorado) -- Universidade Federal do Rio  
Grande do Sul, Instituto de Ciências Básicas da Saúde,  
Programa de Pós-Graduação em Educação em Ciências:  
Química da Vida e Saúde, Porto Alegre, BR-RS, 2023.

1. Avaliação Educacional. 2. Educação Médica. 3.  
Teste de Progresso. 4. Teoria de Resposta ao Item. 5.  
Psicometria. I. Gomes de Souza, Diogo Onofre, orient.  
II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Carlos Eduardo Andrade Pinheiro

**TESTE DE PROGRESSO, UMA AVALIAÇÃO SERIADA NOS CURSOS DE  
MEDICINA: ANÁLISE, MELHORIAS E SELEÇÃO DE BOAS QUESTÕES**

Tese a ser apresentado no Programa de Pós- Graduação em Educação em Ciências do Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de doutor em Educação em Ciências.

Aprovado em: ..... de 2023

**BANCA EXAMINADORA**

---


Professor – Instituição

---

Professor – Instituição

---

Professor - Instituição



## RESUMO

O Teste de Progresso (TP), é um tipo de avaliação seriada dos alunos e dos cursos de medicina, largamente utilizado em vários países do mundo. No Brasil, ele é aplicado uma vez ao ano e a participação dos discentes não é obrigatória em todas as escolas. Este trabalho, através de três artigos científicos, sugere que o TP é um instrumento confiável para avaliar o domínio cognitivo dos alunos do último ano, porém, seus resultados são pouco confiáveis àqueles das fases iniciais dos cursos (1º artigo). Demonstra-se, pela primeira vez, que a falta de alunos ao Teste não se dá de forma aleatória, e sim, por uma característica que pode falsear os resultados para as escolas. Existe uma tendência de que o grupo de faltantes seja constituído por estudantes com pior desempenho acadêmico. Evidencia-se também que os homens faltam mais ao TP do que as mulheres (2º artigo). Para corrigir estas limitações, propõem-se que: a participação passe a ser obrigatória; e, que seja adotada uma nova forma de avaliação longitudinal, denominada de Teste de Progresso Customizado (TPC). Neste novo formato, a prova seria assim composta: metade das questões em comum a todas as etapas do Curso; e, as demais, específicas para as diferentes etapas. Usando a Teoria de Resposta ao Item, é apresentada uma nova técnica de seleção de questões de qualidade, para diferentes níveis de domínio cognitivo (3º artigo). Elas poderiam ser usadas no TPC, na constituição de bancos de dados institucionais, em Testes Adaptativos Computadorizados e mais variados usos. Espera-se com este trabalho contribuir com a qualidade do Teste de Progresso no Brasil, com a melhoria das escolas e da educação médica no Brasil.

**Palavras-chave:** Avaliação; Avaliação Educacional; Educação Médica; Avaliação Externa; Teste de Progresso; Imputação Múltipla de Dados Faltantes; Teoria de Resposta ao Item; Psicometria.

## ABSTRACT

The Progress Test (TP) is a type of serial assessment of students and medical courses, widely used in several countries around the world. In Brazil, it is applied once a year and student participation is not mandatory in all schools. This work, through three scientific articles, suggests that the TP is a reliable instrument to assess the cognitive mastery of final year students, however, its results are not very reliable for students in the initial stages of the courses (1st article). It demonstrates, for the first time, that the lack of students for the Test does not happen randomly, but due to a characteristic that can distort the results for schools. There is a tendency for the group of absent students to be made up of students with worse academic performance. It is also evident that men miss labor more than women (2nd article). To correct these limitations, it is proposed that participation becomes mandatory and that a new form of longitudinal assessment be adopted, called the Customized Progress Test (TPC). In this new format, the test would consist of half of the questions common to all stages of the Course and the rest specific to the different stages. Using Item Response Theory, a new technique for selecting quality questions for different levels of cognitive domain is presented (3rd article). They could be used in TPC, in the constitution of institutional databases, in Computerized Adaptive Tests and more varied uses. This work is expected to contribute to the quality of the Progress Test in Brazil, with the improvement of schools and medical education in Brazil.

**Keywords:** Assessment, Educational Assessment, Medical Education, External Assessment, Progress Test, Multiple Imputation of Missing Data, Item Response Theory, Psychometrics

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO COMPREENSIVA</b>	<b>1</b>
1.1	JUSTIFICATIVA	4
1.2	OBJETIVOS	5
1.2.1	Objetivo geral	5
1.2.2	Objetivos específicos	5
<b>2</b>	<b>ARTIGOS CIENTÍFICOS</b>	<b>7</b>
2.1	PRIMEIRO ARTIGO	8
2.1.1	introduction	9
2.1.2	Objectives	13
2.1.3	Methods	13
2.1.4	Results	15
2.1.5	Discussion	17
2.2	SEGUNDO ARTIGO	23
2.2.1	Introdução	25
2.2.2	Objetivos	27
2.2.3	Métodos	27
2.2.4	Resultados	30
2.2.5	Discussão	34
2.2.6	Conclusões e considerações finais	37
2.3	TERCEIRO ARTIGO	40
2.3.1	Introdução	42
2.3.2	Métodos	43
2.3.3	Resultados	47
2.3.4	Discussão	49
<b>3</b>	<b>CONCLUSÕES DA TESE E PERSPECTIVAS</b>	<b>52</b>

<b>REFERÊNCIAS .....</b>	<b>55</b>
<b>APÊNDICE A – PLANILHA COMPLETA DE POSICIONAMENTO DOS 66 ITENS, DE ÓTIMA E BOA QUALIDADE.....</b>	<b>63</b>
<b>ANEXO A – ARTIGO SUBMETIDO À REVISTA “ESTUDOS EM AVALIAÇÃO EDUCACIONAL”.....</b>	<b>65</b>
<b>ANEXO B – NORMAS DE PUBLICAÇÃO DA REVISTA “ESTUDOS EM AVALIAÇÃO EDUCACIONAL” .....</b>	<b>67</b>
<b>ANEXO C - SEGUNDO ARTIGO – PUBLICADO NA “Revista Brasileira de Educação Médica – RBEM” .....</b>	<b>71</b>





## 1 INTRODUÇÃO COMPREENSIVA

Tortuosos foram os caminhos que levaram ao desenvolvimento desta tese. Como professor universitário, com quarenta e um anos na docência, sempre chamou a atenção a variedade de domínio cognitivo (conhecimento teórico) dos formandos em medicina. Posteriormente, na Coordenação do Curso de Medicina (Dean of Medical School) da UFSC, pensou-se em como lidar com este fenômeno. Na busca por respostas pedagógicas, participou-se de congressos e outras iniciativas da Associação Brasileira de Educação Médica (ABEM). Com o passar do tempo, desenvolveu-se a paixão pela educação médica e avaliações educacionais de larga escala.

Neste contexto, em 2010 participou-se da criação de um Núcleo de Apoio Institucional Sul II (NAPISUL II), para formação de um consórcio de escolas e aplicação do Teste de Progresso (TP), na regional sul II da ABEM. O TP é um tipo de avaliação longitudinal seriado, que avalia tanto os alunos quanto as escolas de medicina. Ele é largamente utilizado em várias partes do mundo (VAN DER VLEUTEN, Cees e colab., 2018). O TP se propõe a avaliar o domínio cognitivo, conhecimento ou proficiência em medicina utilizando, principalmente, questões de múltipla escolha com resposta correta única.

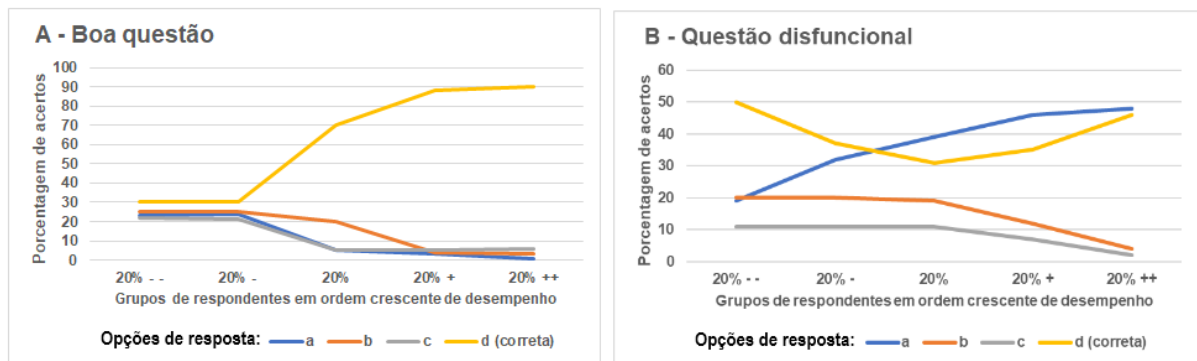
A construção de bons itens de múltipla escolha para avaliação do conhecimento em determinado assunto depende, inicialmente, de aspectos qualitativos. As questões devem estar alinhadas com os objetivos da aprendizagem, representar um domínio de interesse, estarem escritas de forma correta e clara mesmo para o segmento de menor domínio de conhecimento, sem erros de concordância ou construção que possam induzir ou confundir o raciocínio para resposta etc, etc. Simplificando, deve-se ter certeza que o item serve para avaliar o que se pretende, que tem uma construção semântica adequada e segue normas legitimadas para construção e aplicação (PASQUALI, 2017; ROBERTO BOLLELA e colab., 2018; TAVAKOL e DENNICK, 2017).

Sob o aspecto quantitativo, espera-se que os alunos com maior conhecimento acertem mais determinada questão de um teste, do que os alunos com menor domínio sobre o tema. Esta característica de um item, de diferenciar pessoas com distintos graus de conhecimento, é denominada de discriminação. Este atributo é indicativo da

qualidade de uma questão. A qualidade e o número de itens aumentam a confiabilidade dos resultados de uma avaliação cognitiva (KIBBLE, 2017; TAVAKOL e DENNICK, 2017; WRIGLEY e colab., 2012).

Ao se dividir os respondentes de uma prova em cinco grupos com graus crescente de conhecimento, numa boa questão de múltipla escolha, espera-se que as porcentagens de acertos, na única opção considerada correta, aumentem nos estratos com melhor desempenho, como se pode ver no exemplo A, da Figura 1.

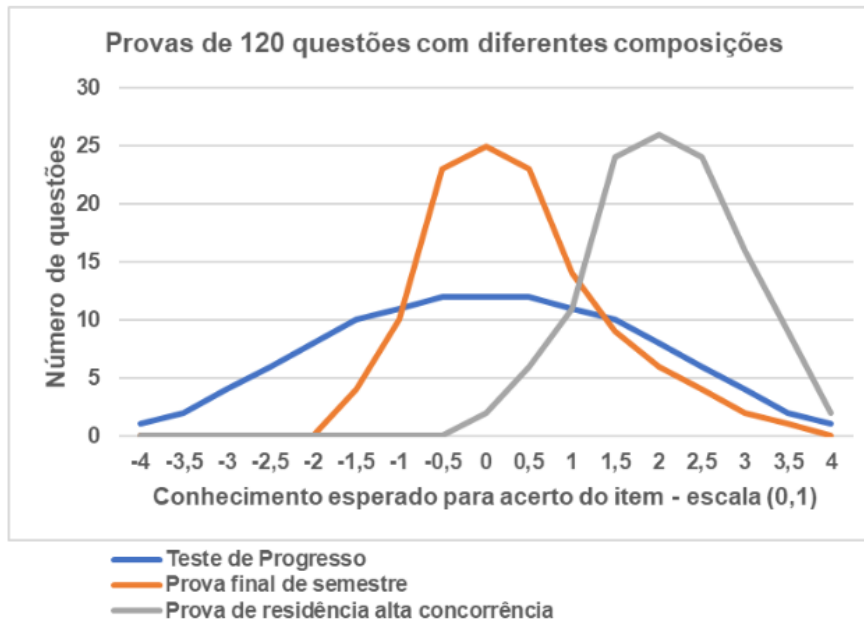
**Figura 1 – Representação gráfica de questões de múltipla escolha com uma opção correta de resposta.**



Mas, se o grupo de melhor desempenho apresentar uma proporção significativa de respostas em duas opções; se outro grupo de menor desempenho acertar mais a questão do que o grupo de melhor desempenho (exemplo B da Figura 1); ou, se não houver diferença na porcentagem de acertos, entre os grupos, diz-se que **a questão é disfuncional**. O item não está aferindo o conhecimento de forma adequada. As questões disfuncionais diminuem a confiabilidade dos resultados de uma prova.

Além do índice de discriminação, outra característica da questão que deve ser considerada na construção de uma avaliação, é seu grau de dificuldade. De acordo com os objetivos da prova, a porcentagem de itens fáceis, médios ou difíceis tem que variar. A distribuição das questões conforme seu grau de dificuldade deve espelhar a habilidade esperada dos alunos e o objetivo da prova (DE CHAMPLAIN, 2010; TAVAKOL e DENNICK, 2013), como exemplificado na FIGURA 2.

Figura 2 – Composição de provas com 120 questões de distintos graus de dificuldade apresentadas na escala (0,1), conforme os objetivos da avaliação.



Fonte: Pinheiro, Carlos Eduardo. Teste de Progresso... 2023. PPGECi - UFRGS

Numa prova de Teste de Progresso, em que se pretende aferir o conhecimento de todos os alunos, é importante que existam itens adequados para aferir toda a gama de níveis de conhecimento. Numa prova de final de semestre, o ideal é que a maioria das questões seja pertinente para medir o conhecimento esperado para aprovação, com maior proporção de itens perto da pontuação de corte. Entretanto, numa prova de residência médica de alta concorrência, o ideal é que as questões sejam majoritariamente adequadas para aferir conhecimento de alunos de alta *performance*.

Atualmente, as boas práticas da educação médica preconizam a elaboração de sistemas de avaliação, em que estes verifiquem tanto o domínio cognitivo, como as habilidades e atitudes dos alunos (NORCINI e colab., 2018). Uma das limitações do TP, é que ele se restringe a avaliar somente o domínio cognitivo. Entretanto, mesmo se tratando de somente uma das dimensões que deve ser avaliada, o conhecimento é básico, tanto para existência de habilidade como para o desenvolvimento de atitudes adequadas. A prática profissional, o exercício administrativo na educação, o pouco conhecimento sobre avaliação e a percepção do potencial pedagógico das avaliações seriadas, nos alunos e cursos, motivaram a elaboração de um estudo mais aprofundado do Teste de Progresso.

## 1.1 JUSTIFICATIVA

O Teste de Progresso, que se trata de uma avaliação com questões de múltipla escolha, é aplicado aos alunos das diferentes fases do curso, que respondem a uma mesma prova. As questões são casos ou situações clínicas, com enunciado não muito longo, dirigido ao conhecimento esperado do aluno ao finalizar o curso. Nota-se que os estudantes das fases iniciais se sentem pouco motivados em empregarem o esforço máximo para resolverem as questões, em matéria que nunca viram ou estudaram (ALBANESE e CASE, 2016). Isso provavelmente contribui para que os resultados não sejam muito confiáveis, para este segmento de alunos.

Escolas médicas de outros países usam esta forma de avaliação seriada (longitudinal), desde a década de 1970. Na Holanda, para aumentar a confiabilidade dos resultados, a prova se tornou obrigatória; passaram a pontuar os resultados no currículo escolar e a aplicar o TP, quatro vezes ao ano. No Brasil, em algumas escolas a participação no TP não é obrigatória. São poucas as escolas que pontuam seus resultados no histórico escolar e a prova, por motivos de custo, é aplicada somente uma vez ao ano. Atualmente, no país, existem 18 grupos ou consórcios de escolas que aplicam o Teste de Progresso (PALHARES NETO e colab., 2022).

O TP surgiu dentro de uma proposta de ensino construtivista. Ela estimula o uso de avaliações onde, a frequente estratégia de estudar muito, somente nas vésperas da avaliação, não seja a mais adequada. Inicialmente, o TP era somente uma avaliação formativa, em que cada aluno poderia avaliar as áreas de conhecimento que precisavam mais estudos. Com o passar dos anos, na Holanda, os resultados dos TPs passaram a pontuar no histórico escolar, tornando-se também uma forma de avaliação somativa.

Para os cursos, o TP sempre foi um bom instrumento de administração pedagógica. Ele possibilita comparação dos resultados de seus alunos, com os resultados das demais escolas participantes, podendo então corrigir áreas deficientes. Permite, ainda verificar quais estudantes precisam maior suporte pedagógico, nas diferentes etapas da graduação. Para avaliar os cursos no Brasil, além do TP, há outra forma de avaliação externa, o ENADE, que é aplicada pelo Ministério da Educação (MEC).

Desde 1996, no Brasil, uma autarquia do MEC aplica o “Exame Nacional de Desempenho dos Estudantes – ENADE”. Esta avaliação dos cursos superiores, não contribui para avaliação dos alunos durante a graduação (RISTOFF, 2022). Com este intuito, seguindo os preceitos do SUS de orientar a formação de mão de obra na saúde e, embutida no “Programa Mais Médicos para o Brasil”, em 2013, o MEC criou a Avaliação Nacional Seriada de Estudantes de Medicina – ANASEM (BRASIL / MINISTÉRIO DA EDUCAÇÃO, 2016).

Este programa propunha fazer avaliações seriadas, de todos os estudantes de medicina, que fariam provas ao final do 2º, 4º e 6º anos. Este autor teve participação na organização da primeira e única aplicação da ANASEM, em 2016, para os alunos do 2º ano. Infelizmente, por questões políticas, o programa foi descontinuado. Desta forma o Teste de Progresso, realizado por iniciativa de consórcios de escolas médicas e sob a tutela da ABEM, passou a ser a única forma de avaliação seriada dos alunos e cursos de medicina.

As questões da confiabilidade dos resultados para alunos de diferentes etapas dos cursos, do impacto da não participação de parte dos estudantes nos resultados do TP, da seleção de questões de qualidade, enfim, da busca por maneiras de melhorar a qualidade desta avaliação educacional seriada, justificam esta Tese.

## 1.2 OBJETIVOS

### 1.2.1 Objetivo geral

- Contribuir com o desenvolvimento de uma forma de avaliação longitudinal, dos cursos e dos estudantes de medicina, que sirva para melhorar a educação médica brasileira e possa ser usada como uma alternativa à realização de prova de licenciamento ao final dos cursos de graduação.

### 1.2.2 Objetivos específicos

Cada objetivo específico será desenvolvido por artigos:

- Analisar se o Teste de Progresso é adequado para avaliar os estudantes que estão cursando as diferentes etapas da graduação em medicina e propor mudanças nesta avaliação longitudinal de alunos e cursos – primeiro artigo, já submetido;

- Verificar se a ausência de estudantes ao Teste de Progresso ocorre de forma aleatória ou por alguma característica destes que poderia influenciar os resultados das provas – segundo artigo, já aceito;
- Demonstrar uma técnica, usando a Teoria de Resposta ao Item, de seleção de questões de qualidade, ordenadas segundo o grau de conhecimento exigido para uma resposta correta e escores dos alunos, para utilização em Testes de Progresso Customizados, na geração de banco de questões e em outras utilizações na educação médica – terceiro artigo, em fase final de elaboração.

# Primeiro artigo



**ACEITO PARA PUBLICAÇÃO...**

**O ARTIGO ESTÁ EM PROCESSO DE EDITORAÇÃO DO TEXTO PELA REVISTA.  
COMO EXIGE INEDITISMO NA PUBLICAÇÃO DAS SUBMISSÕES, ELE SERÁ  
AQUI DISPONIBILIZADO, APÓS PUBLICAÇÃO DO MESMO.**




## Segundo artigo



### *Teste de Progresso: ausência de alunos de Medicina é sugestiva de mau desempenho acadêmico*

Rev. bras. educ. med. 46 (4): e142, 2022 • <https://doi.org/10.1590/1981-5271v46.4-20220145>





## 2.2 SEGUNDO ARTIGO

### TÍTULO: “TESTE DE PROGRESSO: AUSÊNCIA DE ALUNOS DA MEDICINA É SUGESTIVA DE MAU DESEMPENHO ACADÊMICO”

#### RESUMO

**Introdução:** Não se sabe se a ausência de estudantes de medicina, ao Teste de Progresso (TP), se dá de forma aleatória ou por alguma característica sistemática dos alunos, o que poderia influenciar a representatividade dos resultados obtidos pelos participantes.

**Objetivos:** 1) comparar os índices de desempenho acadêmico, no Curso de Medicina da UFSC, dos alunos presentes e ausentes ao TP, em 2019; 2) propor uma maneira de estimar, a partir desses índices, quais seriam as notas dos faltantes se tivessem participado do TP; 3) identificar fatores associados à ausência ao TP.

**Método:** Foram comparadas as médias dos índices de desempenho acadêmico, globais e nas diferentes fases (semestres) dos grupos de alunos presentes e ausentes ao TP, utilizando teste *t de Student* para amostras independentes. Usando uma técnica de regressão linear foram imputadas as prováveis notas no TP ao grupo de alunos ausentes.

**Resultados:** As médias globais dos três indicadores acadêmicos foram significativamente menores nos alunos ausentes ao TP ( $p$  variando de  $< 0,03$  a  $< 0,0001$ ); em 10 das 11 fases (semestres) analisadas do curso, os indicadores acadêmicos dos faltosos foram piores do que dos presentes. A imputação de notas no TP aos ausentes permitiu verificar que existe correlação ( $R=0,62$ ) entre a porcentagem destes e a diferença de notas entre os grupos que realizaram e os que faltaram ao TP. Entre os alunos do gênero masculino, 25,8% não fizeram o TP, enquanto que no gênero feminino foram 16,6% (diferença com  $p < 0,01$ ).

**Conclusões:** A ausência de alunos ao TP não se dá de forma aleatória. Entre os faltosos há uma tendência sistemática de existirem alunos com piores índices de desempenho acadêmico. O uso de imputação múltipla de dados evidencia uma correlação entre a porcentagem de faltosos e a diferença na média da nota no TP,

deste grupo, comparada à média da nota dos participantes. A proporção de homens que faltaram ao TP foi significativamente maior do que a de mulheres.

**Palavras-chaves:** Educação Médica; Avaliação educacional; Teste de Progresso; Imputação Múltipla; Ausência de estudantes ao TP.

**“Progress Test: medical student absenteeism is suggestive of poor academic performance”**

## **SUMMARY**

**Introduction:** there is doubt whether the absence of medical students from the Progress Test (PT) is random or due to some systematic characteristic of the students, which could influence the representativeness of the results obtained by the participants.

**Objectives:** 1) to compare the academic performance indices, in the UFSC Medicine Course, of students present and absent from the TP in 2019; 2) propose a way of estimating, from these indices, what the absentee's grades would be if they had participated in the TP; 3) identify factors associated with absence from labor.

**Method:** using academic performance indices global and in the different phases (semesters) of the course, the means of the groups of students present and absent from the TP were compared, using Student's t test, for independent samples. Using a linear regression technique, the probable grades in the PT were attributed to the group of absent students.

**Results:** the global averages of the 3 academic indicators were significantly lower in students absent from the TP ( $p$  ranging from  $< 0.03$  to  $< 0.0001$ ); in 10 of the 11 phases (semesters) of the course analyzed, the academic indicators of absentees were worse than those present. The attribution of grades in the TP to absentees made it possible to verify that there is a correlation ( $R=0.62$ ) between the percentage of absentees and the difference in grades between those who perform and those who fail the test. Among male students, 25.8% missed the PT, while, among females, 16.6% were absent (difference with  $p < 0.01$ ).

**Conclusions:** the absence of students at TP does not happen randomly. Among absentees, there is a systematic tendency to have students with worse academic performance. The use of multiple imputation of data evidences a correlation between the percentage of absentees and the difference in the average of the grade in the PT, of this group, compared to the average of the participants' grade. The proportion of men who missed PT was significantly higher than that of women.

**Keywords:** Educational Assessment, Medical Education, Progress Test, Multiple Imputation, Absence of students at PT

### 2.2.1 INTRODUÇÃO

O Teste de Progresso (TP) é uma forma de avaliação seriada, longitudinal, do conhecimento de estudantes em que, periodicamente, os alunos de escolas médicas respondem, no mesmo dia, uma prova visando avaliar o conhecimento esperado ao final da graduação. Entre vários usos do TP no processo de ensino e aprendizagem, os resultados dessas provas servem para os alunos verificarem, durante a graduação, seu rendimento individual, e para cada escola avaliar os desempenhos dos seus discentes, comparando-os com as demais escolas participantes (HEENEMAN e colab., 2017; SCHÜTTPELZ-BRAUNS e colab., 2020; VAN DER VLEUTEN, Cees e colab., 2018). A proporção de estudantes que faltam ao TP varia muito entre as diferentes escolas. Não se sabe se as faltas ocorrem de forma aleatória ou por alguma característica sistemática dos faltantes, o que poderia afetar a representatividade global dos resultados do TP.

A omissão de informações sobre as características dos dados faltantes em uma avaliação, como o Teste de Progresso, pode gerar informações enviesadas (FERNÁNDEZ-ALONSO e colab., 2012; HAYATI REZVAN e colab., 2015; RAWLINGS e colab., 2017; VINHA, Luis Gustavo Do Amaral e LAROS, 2018). Se a ausência dos alunos se dá de forma aleatória, os resultados dessa avaliação podem ser representativos da escola; mas se a falta ocorrer, sistematicamente, de forma mais predominante, por exemplo, entre alunos de alto ou baixo desempenho no curso de medicina, então, é provável que o resultado do TP não seja verdadeiramente representativo da escola (FERNÁNDEZ-ALONSO e colab., 2012).

De forma independente, ou com apoio da Associação Brasileira de Educação Médica (ABEM), grupos ou consórcios regionais de escolas médicas elaboram e aplicam um Teste de Progresso ao ano (BICUDO e colab., 2019). No Brasil, a participação dos alunos no TP não é obrigatória em todas as escolas (SAKAI e colab., 2008; TOMIC e colab., 2005). Em outros países, essa avaliação é aplicada duas a quatro vezes ao ano. Na Holanda, por exemplo, a prova é aplicada quatro vezes ao ano e a participação é obrigatória; os resultados constam no currículo individual de cada estudante e os TPs servem como substituto ao exame de licenciamento profissional (KARAY e SCHAUBER, 2018; VAN DER VLEUTEN, Cees e colab., 2018).

Embora possa ter influência nos resultados das escolas, as causas do absenteísmo ao TP no Brasil, as características dos alunos faltantes e o impacto da proporção de ausentes no resultado do TP não são estudados. Existe pouca valorização sobre a existência, ou não, de algum viés nos resultados, em função de alguma(s) característica(s) sistemática(s) dos alunos faltantes. Existe, também, pouca preocupação de como lidar com os dados ausentes, mesmo existindo técnicas estatísticas, com robusta literatura, que podem ser usadas para imputar valores pressupostos ao grupo de estudantes faltantes ao TP (PEDERSEN e colab., 2017; RUBIN, 1987; SINHARAY, 2021).

Algumas técnicas de imputação de dados aos ausentes, pela média dos presentes, são pouco concisas (VINHA, Luis Gustavo Do Amaral e LAROS, 2018), mas outras, usando regressão múltipla, são mais precisas (FERRÃO e colab., 2020; PEDERSEN e colab., 2017; SHRIVE e colab., 2006; VINHA, Luis Gustavo Do Amaral e LAROS, 2018). Essas técnicas são largamente usadas em planejamento, pesquisa e avaliação de desempenho educacional (FERNÁNDEZ-ALONSO e colab., 2012; FERRÃO e colab., 2020; GERACI e MCLAIN, 2018; GRUND e colab., 2021; PEDERSEN e colab., 2017; VINHA, Luis Gustavo Do Amaral e LAROS, 2018). As notas e indicadores de desempenhos acadêmicos dos estudantes podem ser usados para verificar se existe alguma diferença sistemática entre presentes e faltosos ao TP. Tais indicadores também podem ser usados para estimar e imputar prováveis notas no TP aos alunos faltosos, de modo a tornar o resultado desse tipo de aferição mais representativo da escola.

Para que avaliações periódicas, tipo TP, possam, no futuro: i) contribuir para melhorar políticas públicas de controle e de aprimoramento da qualidade da educação médica e ii) influenciar ou determinar o exercício da medicina pelos médicos formados,

é importante que seus resultados no TP sejam, verdadeiramente, representativos da formação de todos os estudantes das escolas médicas. Então, a questão dos alunos faltosos deve ser muito bem valorizada para consolidar a confiança (representatividade) nos resultados dos TPs e consolidar a legitimidade do impacto desse tipo de avaliação, no ensino médico e futuro exercício da medicina.

### **2.2.2 OBJETIVOS**

Os objetivos desse trabalho são:

1. Verificar se existe diferença, aleatória ou sistemática, no desempenho acadêmico entre os alunos que participaram e os que não participaram do TP;
2. Propor uma metodologia que permita estimar os escores no TP dos alunos faltantes, baseados em seus índices de desempenho acadêmicos;
3. Identificar possíveis fatores associados à não participação no TP.

### **2.2.3 MÉTODOS**

#### **População**

É formada por estudantes do Curso de Medicina da Universidade Federal de Santa Catarina (UFSC), escola pública situada em Florianópolis, capital de Santa Catarina, estado da região Sul do Brasil. Nessa escola, no segundo semestre de 2019, havia 625 alunos aptos a participarem do TP anual.

#### **Critério de inclusão e exclusão**

Os dados de desempenho acadêmico e resultado no TP, dos alunos regularmente matriculados, seriam elegíveis para análise. Entretanto, foram excluídos os referentes aos estudantes da primeira fase do curso (60 alunos), por ainda não terem índice semestral de desempenho acadêmico. Restaram então, os dados de 565 estudantes, elegíveis para serem analisados no presente estudo.

### **Delineamento utilizado**

Estudo populacional, transversal e histórico baseado em dados do sistema de controle acadêmico da UFSC (UNIVERSIDADE FEDERAL DE SANTA CATARINA/UFSC, 2021) e dos resultados no TP. Os indicadores de desempenho acadêmicos, global e nas diferentes fases (semestres) dos alunos que participaram do TP foram comparados aos indicadores daqueles que não participaram da avaliação.

Baseando-se nos indicadores acadêmicos e utilizando a técnica de imputação múltipla de dados aos faltantes, foram atribuídas notas no TP dos alunos faltantes para aprimorar a comparação entre os grupos “presentes” e “ausentes” (GERACI e MCLAIN, 2018; NUNES e colab., 2009; PEDERSEN e colab., 2017; RUBIN, 1987; VINHA, Luís Gustavo Do Amaral, 2016). A seguir, foram aferidas: associação entre gênero e ausência ao TP; relação, nas fases do curso, entre a porcentagem de alunos ausentes ao TP; e, a diferença de notas no Teste de Progresso 2019, entre os presentes e as notas imputadas aos ausentes.

### **Instrumental utilizado**

O Teste de Progresso, que foi organizado por 12 escolas de medicina da Regional Sul II (ABEM / ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO MÉDICA, [S.d.]) da Associação Brasileira de Educação Médica, é uma prova com 120 questões de múltipla escolha simples, com 4 opções de respostas, sem penalização de respostas erradas. Abrange o conteúdo teórico dos cursos de medicina no Brasil, segundo as Diretrizes Curriculares Nacionais para cursos de Medicina (BRASIL / MINISTÉRIO DE EDUCAÇÃO / CONSELHO NACIONAL DE EDUCAÇÃO, 2014); sendo voltada ao conhecimento esperado para os alunos ao terminarem o Curso. No Brasil, os cursos de medicina são compostos de 12 fases ou semestres letivos. No presente estudo, a prova foi aplicada em 2019, e somente os resultados da UFSC foram analisados.

A participação na prova não era obrigatória e não havia penalização aos não participantes. Os alunos que participassem de TP teriam esta atividade registrada, em seus históricos escolares, como atividade complementar.

## Variáveis

As variáveis analisadas nesse trabalho foram: os três indicadores de desempenho acadêmicos utilizados na UFSC, Índice de Matrícula (IM), Índice de Aproveitamento Acumulado (IAA) e Índice de Aproveitamento Proporcional (IAP); participação ou falta ao TP; as notas dos participantes no TP; qual a fase (semestre letivo) que estava sendo cursado por ocasião do TP; e, o gênero dos alunos.

O **IAA** é calculado cumulativamente a cada semestre, representado pelo resultado da divisão entre o somatório de pontos até então obtidos e a carga horária acumulada até o semestre em que o aluno está matriculado ( $IAA = \text{pontos obtidos} / \text{carga horária matriculada}$ ). O **IAP** é o mesmo que o anterior, mas, usando somente as disciplinas em que o aluno foi aprovado ( $IAP = \text{pontos obtidos, excluídas reprovações} / \text{carga horária matriculada, excluídas cargas das matérias reprovadas}$ ). O **IM** é obtido pela multiplicação do IAA pelo resultado da divisão da carga horária já cursada (CHC) e carga horária total do curso (Carga Horária Total):  $IM = IAA \times CHC / CHT$ . Quanto maior os valores desses índices, melhor é o desempenho do aluno. A **nota do TP** é dada pela porcentagem de acertos na prova.

## Análise estatística

As variáveis: indicadores de desempenho acadêmicos (IAA, IAP e IM), escores no Teste de Progresso 2019, fases no curso e gênero foram transcritas em planilha Excel (da Microsoft) e importadas para o SPSS *Statistics*, versão 20, da IBM.

Inicialmente, calculou-se a porcentagem de ausentes e, pelo *teste Z para duas proporções*, foi verificado se existe associação entre gênero e ausência/presença ao TP. Foram, então, comparadas as médias globais dos indicadores IM, IAA e IAP dos grupos de alunos presentes e ausentes ao TP, utilizando teste *t de Student* para amostras independentes. A seguir, as médias dos indicadores IM, IAA e IAP, dos presentes e dos ausentes, foram confrontadas em todas as fases, usando o mesmo procedimento estatístico. Como os resultados dos três indicadores são numerosos e foram muito semelhantes entre si, para análise da diferença das notas nas fases, somente os resultados do Índice de Matrícula (IM) serão apresentados como representativo dos mesmos.

Em sequência, foi utilizada regressão linear para imputar prováveis notas dos alunos ausentes no TP, com base em 10 imputações pelo método monotônico, usando a nota no TP como variável dependente e fase no curso, IM, IAA e IAP como

variáveis independentes. Esse método, conhecido como imputação múltipla de dados faltantes (FERRÃO e colab., 2020; GERACI e MCLAIN, 2018; PEDERSEN e colab., 2017; RUBIN, 1987; WALANI e CLELAND, 2015), assume o modelo conhecido como “*missing at random (MAR)*” para o desfecho, ou seja, no contexto da presente pesquisa, assume-se que as notas no TP dos alunos faltantes foram associadas às mesmas variáveis **independentes** daqueles que participaram do TP. Com as médias no TP dos alunos presentes e as prováveis médias dos alunos ausentes obtidas por imputação, calcularam-se as diferenças nas 11 fases e compararam-se essas médias com *teste t*.

Por regressão linear foi testada a relação das variáveis geradas, diferenças das notas no TP (entre os presentes e imputadas aos ausentes) e as porcentagens de alunos que faltaram à prova.

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa em Seres Humanos da UFSC, sob número 5.261.272/2022. Houve dispensa do Termo de Consentimento Livre Esclarecido (TCLE), pois os dados foram recebidos pelo único pesquisador que fez a análise destes, de forma anonimizada, sem identificação dos respondentes.

#### 2.2.4 RESULTADOS

No ano de 2019, dos 565 alunos da medicina UFSC que poderiam participar do TP, 443 (78,4%) compareceram e 122 (21,6%) não compareceram. Com relação ao gênero, entre os 306 alunos, 79 (25,8%) não estiveram presentes e entre as 259 alunas, 43 (16,6%) se fizeram ausentes. Constata-se que a proporção de homens que não compareceram ao TP 2019 é estatisticamente maior do que a proporção de mulheres ( $p < 0,01$ ).

As médias dos índices de desempenho acadêmico (IM, IAA e IAP), no curso de Medicina, no semestre da aplicação do TP, de estudantes de ambos os sexos, presentes e ausentes na prova Teste de Progresso, são mostradas na Tabela 1.



**Tabela 1. Comparação dos índices de despenho acadêmico (IM, IAA e IAP) obtidos no curso de medicina da UFSC dos alunos presentes e ausentes na prova Teste de Progresso 2019**

Indicadores acadêmicos	Presentes (n = 443)		Ausentes (n = 122)		Teste-t
	média	DP	média	DP	sign*
<b>IM</b>	<b>4.063</b>	2.702	<b>3.608</b>	1.885	0,03
<b>IAA</b>	<b>8.517</b>	655	<b>7.818</b>	1.049	> 0,01
<b>IAP</b>	<b>8.567</b>	532	<b>8.077</b>	543	> 0,01

IM = Índice de Matrícula; IAA = Índice Acumulado.

IAP = Índice de Aproveitamento Proporcional.

\*Teste-t: sign  $\leq$  0,05.

Fonte: elaborada pelos autores.

Os três índices de desempenho acadêmicos utilizados na UFSC indicam que os alunos que compareceram à prova TP, em 2019, tinham melhores desempenhos no curso quando comparados aos que não compareceram. Tais diferenças foram estatisticamente significativas para todos os três indicadores.

A mesma comparação, utilizando somente o “IM”, foi feita nas diferentes fases do Curso – Tabela 2. As comparações nas fases, utilizando IAA e IAP, apresentaram resultados semelhantes e não serão aqui apresentadas.

**Tabela 2. Comparação do indicador “Índice de Matrícula” (IM) no Curso de Medicina da UFSC dos alunos presentes e ausentes ao TP2019**

Fases no curso	Alunos presentes			Alunos ausentes			Diferença	sign*
	N	média	DP	N	média	DP		
<b>2</b>	51	<b>575</b>	259	7	<b>402</b>	166	<b>173</b>	ns
<b>3</b>	43	<b>1.220</b>	66	9	<b>1.197</b>	161	<b>23</b>	ns
<b>4</b>	41	<b>1.836</b>	148	9	<b>1.712</b>	193	<b>124</b>	<b>0,05</b>
<b>5</b>	45	<b>2.520</b>	79	5	<b>2.563</b>	164	<b>-43</b>	ns
<b>6</b>	43	<b>3.080</b>	409	19	<b>2.698</b>	202	<b>382</b>	> <b>0,01</b>
<b>7</b>	38	<b>3.659</b>	552	16	<b>3.387</b>	347	<b>272</b>	ns
<b>8</b>	26	<b>4.489</b>	602	21	<b>3.897</b>	301	<b>592</b>	> <b>0,01</b>
<b>9</b>	25	<b>5.583</b>	298	23	<b>5.113</b>	275	<b>470</b>	> <b>0,01</b>
<b>10</b>	45	<b>6.459</b>	253	4	<b>6.283</b>	446	<b>176</b>	ns
<b>11</b>	33	<b>7.918</b>	861	5	<b>6.985</b>	426	<b>933</b>	ns
<b>12</b>	53	<b>8.494</b>	401	4	<b>8.377</b>	533	<b>117</b>	ns
<b>Total</b>	<b>443</b>	<b>4.063</b>	2.702	<b>122</b>	<b>3.608</b>	1.885	<b>455</b>	<b>0,03</b>

\* sign: diferença entre médias, teste-t de Student,  $\leq$  0,05.

Fonte: elaborada pelo autor.

Em dez das onze fases (ou semestres), as médias de desempenho acadêmico IM foram maiores entre os grupos de alunos que compareceram ao TP, quando comparados aos grupos de ausentes; as diferenças foram significativas em quatro fases.

Usando os três indicadores de desempenho acadêmicos dos alunos faltantes, na fase que estava sendo cursada, por regressão linear, foi feita a imputação das prováveis notas no TP aos ausentes. As médias de escores no TP dos alunos presentes, as porcentagens e as médias de escores imputados aos alunos faltantes e as diferenças entre as médias dos presentes e imputadas em cada fase do Curso aparecem na Tabela 3.

**Tabela 3. Comparação, em cada fase do curso, das médias das notas no TP2019 dos alunos presentes e das notas imputadas aos ausentes; porcentagens de ausentes e diferenças entre as médias.**

Fases no curso	Todos alunos N	Presentes média no TP	Ausentes ao TP			Diferença entre médias	sign*
			N	%	média imputada		
2	58	36,9	7	12,1	33,3	3,6	ns
3	52	35,6	9	17,3	33,9	1,7	ns
4	50	40,3	9	18,0	36,0	4,3	0,04
5	50	44,5	5	10,0	45,3	-0,8	ns
6	62	49,0	19	30,6	43,8	5,2	0,01
7	54	42,7	16	29,6	39,2	3,5	ns
8	47	46,7	21	44,7	39,5	7,2	0,01
9	48	54,1	23	47,9	46,3	7,8	> 0,01
10	49	57,9	4	8,2	54,1	3,8	0,05
11	38	56,6	5	13,2	47,9	8,7	0,04
12	57	59,9	4	7,0	60,0	-0,1	ns
<b>Total</b>	<b>565</b>	<b>47,4</b>	<b>122</b>	<b>21,6</b>	<b>42,1</b>	<b>5,3</b>	<b>&gt; 0,01</b>

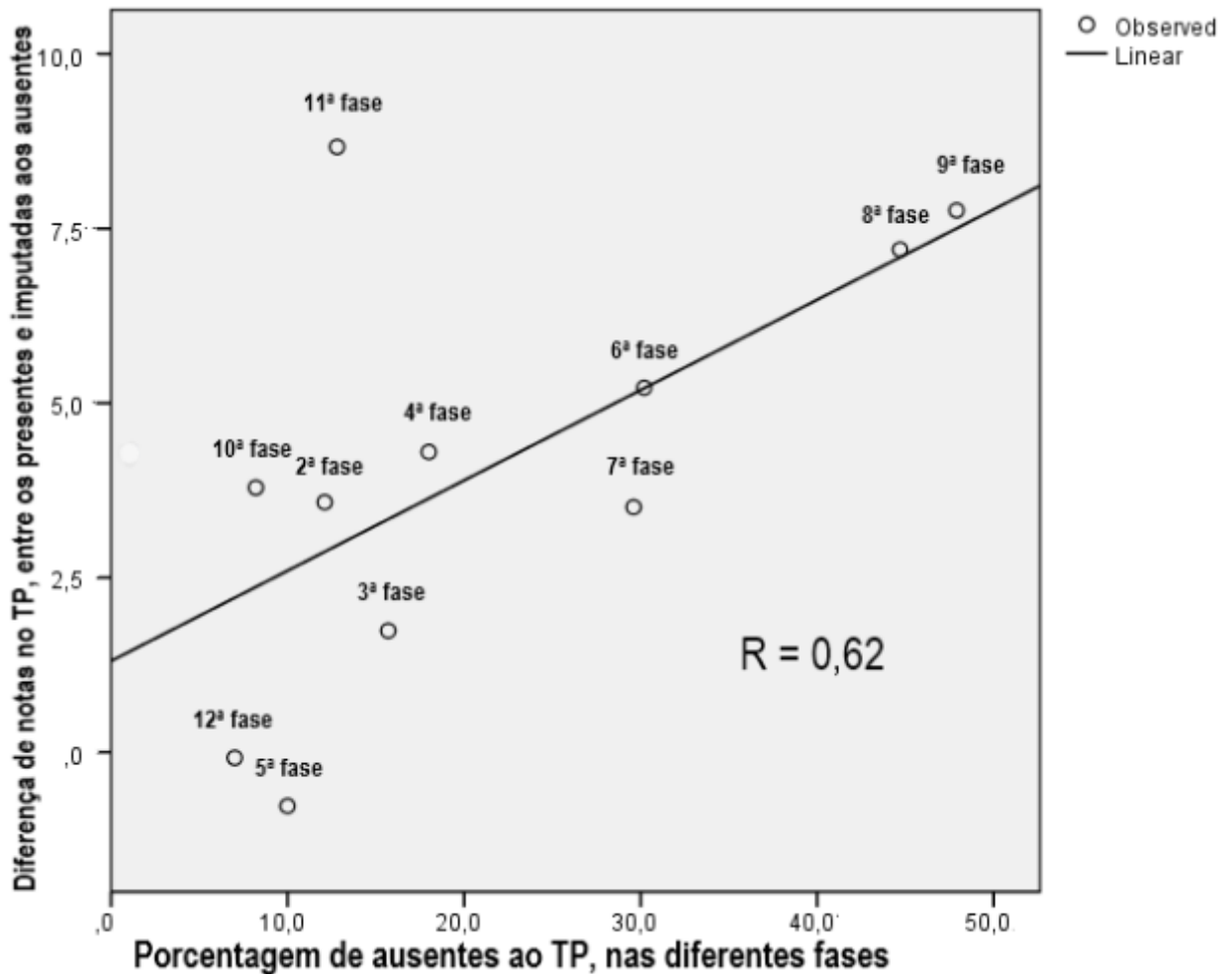
\* teste - t de Student: sign  $\leq$  0,05.

Fonte: elaborada pelos autores.

Em nove, das onze fases analisadas, as médias das notas no TP dos alunos presentes foram maiores do que as médias imputadas aos ausentes. Nessa análise, as diferenças significativas aumentaram para seis fases, com predominância nas fases finais do curso. Nas fases cinco e doze, com proporção de faltantes relativamente baixas, as notas imputadas aos ausentes ao TP foram maiores do que as notas dos participantes do TP, mas a diferença não tem significância estatística.

Com os dados, **porcentagem de faltantes e diferença das notas dos alunos no TP e as notas imputadas aos faltantes** descritos na Tabela 3, é possível verificar se há relação entre as duas variáveis (Figura 1).

Figura 1. Relação, nas fases do curso, entre a porcentagem de alunos ausentes ao TP e a diferença de notas, no Teste de Progresso 2019, entre os presentes e as notas imputadas aos ausentes



Fonte: elaborada pelos autores

A regressão mostrou que existe uma relação linear positiva entre as diferenças das notas dos alunos presentes e ausentes ao Teste de Progresso com a porcentagem de alunos ausentes ao TP. A correlação “ $R$ ” foi de 0,62 ( $p < 0,0001$ ) e coeficiente de determinação “ $R^2$ ” de 0,38 ( $p < 0,0001$ ). Ou seja, existe uma tendência de que, na medida em que aumenta a porcentagem de ausentes, maior será a diferença entre as notas dos presentes ao TP sobre as notas imputadas aos ausentes, como se vê na Figura 1.

### 2.2.5 DISCUSSÃO

No TP analisado, 21,6% de estudantes da UFSC faltaram à prova. Tal porcentagem foi maior do que as encontradas em outros estudos sobre TP no Brasil, que descrevem menos de 10% de faltantes (FERREIRA, 2019; ROSA e colab., 2017; SAKAI e colab., 2011). Na literatura internacional, um artigo holandês, que analisou durante 14 anos o uso do TP na pós-graduação, refere aumento, ao longo do período estudado, de 1% para 16% de residentes em radiologia que pedem dispensa da participação (RUTGERS e colab., 2018). Em relação à significativa maior proporção de alunos do gênero masculino faltantes ao TP 2019, não se encontram semelhantes dados na literatura nacional, nem na internacional, para comparar os resultados. Na literatura sobre educação médica, diferenças entre gêneros aparecem os homens tendo maior índice de absenteísmo nas aulas, o que está associado à maior reprovação (DUNN e colab., 2003; FREDJ e colab., 2016; HAKAMI, 2021) e ao menor desempenho educacional (FREDJ e colab., 2016; MACKENZIE e colab., 2016; MAHER e colab., 2013). As mulheres têm menor taxa de conflito dentro das escolas médicas (MAHER e colab., 2013).

As análises dos escores de três indicadores de desempenho acadêmico (IM, IAA e IAP), comparando os alunos que compareceram com os que faltaram ao TP, mostram que o grupo de faltosos tem pior desempenho acadêmico. Usando um desses indicadores, chamado Índice de Matrícula (IM), para analisar, separadamente, cada uma de 11 fases do curso, observa-se que em 10 fases o grupo que não compareceu teve pior desempenho acadêmico do que o do grupo que participou. Essa aferição sugere, fortemente, que a ausência ao TP não se dá de forma aleatória. O grupo de estudantes faltosos ao TP mostrou uma tendência sistemática de ser composto por alunos com menor desempenho acadêmico. Na literatura sobre o TP não se encontra outro trabalho que comparasse características de estudantes presentes e faltosos.

O achado de que as faltas ao TP não se dão de forma aleatória leva a duas reflexões. No plano individual, a literatura sobre educação sugere que características do aluno como absenteísmo nas aulas, isolamento social, depressão e dificuldade com o idioma são fatores de risco associados ao mau desempenho acadêmico (MAHER e colab., 2013). Como segunda reflexão, o atual trabalho sugere que

ausência ao TP também poderia ser sinalizador de dificuldade acadêmica, portanto, poderia, ou talvez, até deveria servir de alerta para intervenções precoces de apoio a esse grupo de estudantes, potencialmente com dificuldades, visando a inclusão pedagógica e social.

No plano institucional, se a falta ao TP se dá de forma sistemática, constituído majoritariamente por um grupo com menor desempenho acadêmico, então, com o aumento da proporção de faltosos, a média do desempenho no TP dos presentes tende a aumentar “aparentemente” o resultado da escola, tornando os resultados enviesados (FERNÁNDEZ-ALONSO e colab., 2012; FERRÃO e colab., 2020; SINHARAY, 2021; WANG e colab., 2021), como mostrado nesse estudo. Não há certeza, mas estudiosos que trabalham com estatísticas de dados faltantes sugerem em ausência sistemáticas que, quando a proporção de faltantes for maior do que 5 a 10% dos alunos de um curso, os resultados estarão enviesados (nesse estudo foi de 21,6%); quando, porém, as faltas se dão de forma aleatória, elas não afetam os resultados de maneira significativa (FERNÁNDEZ-ALONSO e colab., 2012; VINHA, Luís Gustavo Do Amaral, 2016).

Assim, é razoável sugerir que as instituições com 10% ou mais de alunos faltantes ao TP estabeleçam políticas para aumentar a participação de estudantes. Tornar a participação no TP obrigatória seria o primeiro passo. Pontuar a participação no TP, nos históricos escolares de cada aluno, seria outra ideia que poderia ser discutida. Para minimizar o viés causado pelos dados faltantes nos resultados das escolas, o uso de alguma estratégia estatística, como a imputação de dados, poderia aumentar a “representatividade” real dos resultados do TP.

A ocorrência de dados faltantes, mostrada em pesquisas clínicas ou educacionais, diminui o tamanho e a variabilidade da população, reduzindo o poder estatístico e a validade externa dos estudos (FERRÃO e colab., 2020; WALANI e CLELAND, 2015). No sentido inverso, a utilização de técnica de imputação múltipla de dados, para calcular qual seria a nota provável no TP dos alunos faltosos e a inclusão desses resultados, preserva o tamanho e a variabilidade das populações. Revisões sobre uso de imputação de dados faltantes em pesquisa educacional podem ser encontradas na literatura nacional (FERRÃO e colab., 2020; VINHA, Luis Gustavo Do Amaral e LAROS, 2018) e internacional (FERNÁNDEZ-ALONSO e colab., 2012; GERACI e MCLAIN, 2018; GRUND e colab., 2021).

De modo geral, os dados faltantes são classificados como: faltantes completamente ao acaso (Missing Completely at Random - MCAR), que não alteram resultados; faltantes ao acaso (Missing at Random – MAR), que podem, ou não, afetar os resultados (dependendo se as faltas ao acaso afetam de forma diferente diferentes extratos em estudo); faltantes não ao acaso (Missing Not at Random – MNAR), que afetam os resultados (RUBIN, 1987), como os encontrados neste estudo. Entretanto, após imputação de dados, observa-se que, especificamente na quinta e décima segunda fases, o desempenho dos faltosos foi melhor (mas não significativamente) do que o dos presentes ao TP. Então, não se pode afirmar que os ausentes são sempre os alunos de menor desempenho acadêmico.

A regressão linear entre as diferenças das notas no TP e a porcentagem de faltantes, vistos na Figura 1, mostram que existe uma relação entre porcentagem de faltantes e média no desempenho no TP. Quanto mais faltantes, maior a média da nota dos participantes no TP. Então, é razoável considerar que, quando se comparassem internamente os resultados de uma escola (SAKAI e colab., 2011), ou se comparassem os resultados entre diferentes escolas ou consórcios (BICUDO e colab., 2019), deveria se colocar a informação sobre a porcentagem de faltantes; também, que os dados faltantes fossem imputados, para que a expressão dos resultados esteja mais próxima da representatividade do desempenho global das escolas avaliadas.

Apesar de demonstrar que a falta ao TP não se dá de forma aleatória, e sim, sistemática, que existe relação entre a porcentagem de faltosos e a diferença entre notas dos presentes e notas imputadas aos ausentes e que isso afeta os resultados, o estudo não demonstra a partir de que porcentagem de faltantes os dados passam a não representar os valores reais da escola. Para verificar, com maior precisão, qual porcentagem afeta os dados, seriam necessárias populações muito maiores do que as usadas no presente estudo.

Outra limitação observada é que se trata de um trabalho unicêntrico, baseado somente em uma prova. No período de 2011 a 2018, os índices de absenteísmo ao TP, na UFSC, variaram de 8,0 a 37,7% (média = 18,2%; DP = 10,5). A porcentagem de ausentes de 21,6%, encontrada nesta investigação está próxima à média dos nove anos de TP na UFSC. Tradicionalmente, entre os alunos das sétimas e oitavas fases constavam as maiores proporções de ausentes. No ano de 2019, houve uma não usual ausência, de 47,9% dos alunos da nona fase. A tentativa de entender por que

isso aconteceu, e quem eram os faltosos na nona fase, originou o presente trabalho. Semelhante análise em relação aos faltosos deveria ser feita em outras escolas, para confirmar a validade externa dos achados deste estudo.

Atualmente, no Brasil, existem grupos dispostos a levantar a bandeira da implementação de exames de licenciamento para exercício da profissão (MOTA e colab., 2014; TRONCON, 2019). Essa discussão é calorosa e ocorre em muitos países do mundo, mesmo sem haver nenhuma pesquisa que mostre, de maneira inequívoca, a vantagem desse tipo de avaliação (BICA e KORNIS, 2020; VAN DER VLEUTEN, Cees, 2013). Como o exame de licenciamento ocorre ao final da graduação, tem-se receio de que a mesma recaia (tardamente) somente sobre o aluno e não tenha repercussão sobre as escolas e na qualidade do ensino. Com ou sem licenciamento, parece óbvio que avaliações seriadas, longitudinais, como o Teste de Progresso ou a Avaliação Nacional Seriada de Estudantes de Medicina – ANASEM (BRASIL / MINISTÉRIO DA EDUCAÇÃO, 2016), abandonada pelo governo federal, devam ser estimuladas e retomadas. Avaliações desde o início da graduação permitiriam ao aluno corrigir precocemente seus rumos e possibilitaria às escolas fazerem o mesmo. Poderiam contribuir sobremaneira com a qualidade do ensino médico.

## **2.2.6 CONCLUSÕES E CONSIDERAÇÕES FINAIS**

O atual estudo mostra que existe uma correlação positiva entre o gênero masculino e as faltas ao TP (homens faltaram mais); que as faltas não se deram de forma aleatória, e sim, sistematicamente, com o grupo de faltosos sendo composto, predominantemente, por alunos com menor desempenho acadêmico. Isso sugere que, a falta ao TP, possa ser usada como potencial marcador e alerta de discentes com mau desempenho para estabelecer políticas de estímulo e inclusão destes, desde as fases iniciais ou intermediárias dos cursos. Mostra também que existiu uma correlação entre porcentagem de ausentes e possível alteração do resultado “real” (representatividade global) do TP. A técnica de imputação múltipla de dados aos faltosos pode ser utilizada para estimar qual seriam as notas das escolas, caso todos os alunos participassem da prova.

Resta claro a importância de se incentivar estratégias que diminuam a porcentagem de faltosos ao TP no Brasil, de cuidar melhor desse grupo de estudantes

e de promover ações para aumentar a representatividade dos resultados dos TPs. Isso contribuiria para torná-los mais úteis e confiáveis, valorizando avaliações longitudinais dos alunos e dos cursos e, conseqüentemente, colaborando para políticas públicas que venham aprimorar o ensino médico em nosso país.





## Terceiro artigo



**ARTIGO PRONTO, SERÁ VERSADO PARA O INGLÊS PARA SUBMISSÃO EM REVISTA A1 INTERNACIONAL EM EDUCAÇÃO MÉDICA.**

**COMO AS REVISTAS EXIGEM INEDITISMO NAS SUBMISSÕES, ELE SERÁ INTEGRALMENTE DISPONIBILIZADO, APÓS SUA PUBLICAÇÃO**



## REFERÊNCIAS

ABEM / ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO MÉDICA. ABEM. **Teste de Progresso**. Disponível em: <<https://website.abem-educmed.org.br/projetos-e-acoas/teste-de-progresso/>>. Acesso em: 16 fev 2023.

ALBANESE, Mark e CASE, Susan M. Progress testing: Critical analysis and suggested practices. **Advances in Health Sciences Education**, v. 21, n. 1, p. 221–234, 2016.

ALBEKAIRY, Abdulkareem M e colab. Evaluation of the Potential of National Sharing of a Unified Progress Test Among Colleges of Pharmacy in the Kingdom of Saudi Arabia. **Advances in Medical Education and Practice**, v. 12, p. 1465–1475, 2021. Disponível em: <<https://doi.org/10.2147/AMEP.S337266>>.

ALI, Kamran e colab. Progress testing in undergraduate dental education: The Peninsula experience and future opportunities. **European Journal of Dental Education**, v. 20, n. 3, p. 129–134, 2016.

ALKHALAF, Zahraa S.A. e colab. Medical knowledge and clinical productivity: independently correlated metrics during radiology residency. **European Radiology**, v. 31, n. 7, p. 5344–5350, 2021. Disponível em: <<https://doi.org/10.1007/s00330-020-07646-3>>.

ANDRADE, D.F. e TAVARES, H.R. e VALLE, R. **Teoria de resposta ao item: conceitos e aplicações**. São Paulo: ABE – Associação Brasileira de Estatística, 2000.

ANDREW DANIELSON, Jared e colab. Status Quo of Progress Testing in Veterinary Medical Education and Lessons Learned. **Frontiers in Veterinary Science** | [www.frontiersin.org](http://www.frontiersin.org), v. 1, p. 559, 2020. Disponível em: <[www.frontiersin.org](http://www.frontiersin.org)>.

AUBIN, André e colab. Examinee Cohort Size and Item Analysis Guidelines for Health Professions Education Programs: A Monte Carlo Simulation Study. **Academic Medicine**, v. 95, n. 1, p. 151–156, 1 Jan 2020.

BICA, Rafael Barbosa e KORNIS, George Edward. Exames de licenciamento em Medicina – uma boa ideia para a formação médica no Brasil? **Interface - Comunicação, Saúde, Educação**, v. 24, p. 1–16, 2020.

BICUDO, Angélica Maria e colab. Teste de Progresso em Consórcios para todas as Escolas Médicas do Brasil. Consortia of Cross-Institutional Progress Testing for All Medical Schools in Brazil. **Revista Brasileira de Educação Médica**, v. 43, n. 4, p. 151–156, 2019.

BLAKE, Jennifer M e colab. Introducing progress testing in McMaster University's problem-based medical curriculum: Psychometric properties and effect on learning. **Academic Medicine**, v. 71, n. 9, p. 1002–1007, 1996.

BRASIL / MINISTÉRIO DA EDUCAÇÃO. ANASEM - **Avaliação Nacional Seriada dos Estudantes de Medicina** - Documento Básico. Brasília, Brasil: Inep/MEC. , 2016

BRASIL / MINISTÉRIO DE EDUCAÇÃO / CONSELHO NACIONAL DE EDUCAÇÃO. **Diretrizes Curriculares Nacionais do Curso de Graduação em Medicina**. 2014, p. 1–14.

CARVALHO, Annie Beatriz De e colab. Análise de conteúdo de duas avaliações externas brasileiras de cursos de medicina: Enade e Revalida. **Revista Brasileira de Educação Médica**, v. 45, n. 1, 2021.

CECILIO-FERNANDES, Dario. Implementando o teste adaptativo computadorizado. **Scientia Medica - Educação em Ciências da Saúde**, v. 29, n. 3, p. e34432, 2019.

CECILIO-FERNANDES, Dario e BICUDO, Angélica Maria e FILHO, Pedro Tadao Hamamoto. Progress testing as a pattern of excellence for the assessment of medical students' knowledge - Concepts, history, and perspective. **Medicina (Ribeirão Preto)**, v. 54, n. 1, p. e-173770, 2 Jun 2021.

COLLARES, Carlos Fernando. A short commentary about compliance to assumptions and choice of the model in item response theory. **Medical Teacher**. [S.l.]: Taylor & Francis. Disponível em: <<https://doi.org/10.1080/0142159X.2022.2101441>>. , 2022

COLLARES, Carlos Fernando e CECILIO-FERNANDES, Dario. When I say ... computerised adaptive testing. **Medical Education**, v. 53, n. 2, p. 115–116, 2019.

DAMAS, Bruna Rodrigues e MIRANDA, Gilberto j. Preparação da Instituição para o ENADE: importa? *In*: 3º Congresso de Contabilidade, 2019, Uberlândia, MG. **Anais [...]**. Uberlândia, UFU, 2019. p. 1-17. Disponível em: <[https://eventos.ufu.br/sites/eventos.ufu.br/files/documentos/030\\_artigo\\_completo.pdf](https://eventos.ufu.br/sites/eventos.ufu.br/files/documentos/030_artigo_completo.pdf)>.

DE CHAMPLAIN, André F Andre F. A primer on classical test theory and item response theory for assessments in medical education. **Medical Education**, v. 44, n. 1, p. 109–117, 2010.

DE SÁ, Marcos Felipe Silva e colab. The Individual Progress Test of Gynecology and Obstetrics Residents (TPI-GO): The Brazilian Experience by FEBRASGO. **Revista Brasileira de Ginecologia e Obstetricia**, v. 43, n. 6, p. 425–428, 2021. Disponível em: <<https://creativecommons.org/licenses/by/4.0/>>.

DOWNING, Steven M. Reliability : on the reproducibility of assessment data. **Medical Education**, v. 38, p. 1006–1012, 2004.

DUNN, Michelle C. e KADANE, Joseph B. e GARROW, John R. Comparing Harm Done by Mobility and Class Absence: Missing Students and Missing Data. **J. Educ.**

**Behav. Stat.**, v. 28, n. 3, p. 269–288, 2003.

EPSTEIN, Ronald M. Medical education - Assessment in medical education. **New England Journal of Medicine**, v. 356, n. 4, p. 387–396, 2007.

EVERETT, George Douglas e ISEA, Luis. Reporting of the Medical Licensing Examination. **JAMA**, v. 323, n. 2, p. 185–186, Jan 2020.

FERNÁNDEZ-ALONSO, Rubén e SUÁREZ-ÁLVAREZ, Javier e MUÑIZ, José. Imputación de datos perdidos en las evaluaciones diagnósticas educativas. **Psicothema**, v. 24, n. 1, p. 167–175, 2012.

FERRÃO, Maria Eugénia e PRATA, Paula e ALVES, Maria Teresa Gonzaga. Multiple imputation in big identifiable data for educational research: An example from the Brazilian Education assessment system. **Ensaio: aval. pol. públ. Educação**, v. 28, n. 108, p. 599–621, 2020.

FERREIRA, Ricardo Corrêa. **Relação entre o desempenho no Teste de Progresso e na seleção para Residência Médica**. 2019. Tese (Doutorado) - Faculdade de Ciências Médicas, Universidade Estadual de Campinas, Campinas, 2019.

FREDJ, Maha Ben e colab. Influencing factors of absenteeism among first cycle of medical students in the Faculty of Medicine of Sousse. **LA TUNISIE MEDICALE**, v. 97, n. 04, p. 519–524, 2016.

GERACI, Marco e MCLAIN, Alexander. Multiple imputation for bounded variables. **Psychometrika**, v. 83, n. 4, p. 919–940, 2018.

GRUND, Simon e LÜDTKE, Oliver e ROBITZSCH, Alexander. On the Treatment of Missing Data in Background Questionnaires in Educational Large-Scale Assessments: An Evaluation of Different Procedures. **Journal of Educational and Behavioral Statistics**, v. 46, n. 4, p. 430–465, 2021.

HAKAMI, Abdulrahim Refdan. Effect of absenteeism on the performance of medical sciences students: gender differences. **Medical Education (online)**, v. VOL. 26, n. 1, 2021. Disponível em: <<https://doi.org/10.1080/10872981.2021.1875531>>.

HAMAMOTO FILHO, Pedro Tadao e BICUDO, Angélica Maria. Implementation of the Brazilian National Network for Practices and Research with Progress Testing - BRAZ-NPT. **Revista Brasileira de Educação Médica**, v. 44, n. 3, p. e074, 2020a.

HAMAMOTO FILHO, Pedro Tadao e BICUDO, Angélica Maria. Improvement of Faculty's Skills on the Creation of Items for Progress Testing Through Feedback to Item Writers: a Successful Experience. **Revista Brasileira de Educação Médica**, v. 44, n. 1, p. e018, 2020b.

HAYATI REZVAN, Panteha e LEE, Katherine J. e SIMPSON, Julie A. The rise of multiple imputation: A review of the reporting and implementation of the method in medical research Data collection, quality, and reporting. **BMC Medical Research Methodology**, v. 15, n. 1, p. 1–14, 2015.

HEENEMAN, Sylvia e colab. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. **Medical Teacher**, v. 39, n. 1, p. 44–52, 2 Jan 2017.

HENNING, Marcus e PINNOCK, Ralph e WEBSTER, Craig. Does Progress Testing Violate the Principles of Constructive Alignment? **Medical Science Educator**, v. 27, n. 4, p. 825–829, 2017.

KARAY, Yassin e SCHAUBER, Stefan K. A validity argument for progress testing: Examining the relation between growth trajectories obtained by progress tests and national licensing examinations using a latent growth curve approach. **Medical Teacher**, v. 40, n. 11, p. 1123–1129, 2018.

KHEYAMI, Deena e colab. Item analysis of multiple choice questions at the department of paediatrics, Arabian gulf university, Manama1. **Sultan Qaboos University Medical Journal**, v. 18, n. 1, p. e68–e74, 2018.

KIBBLE, Jonathan D. Best practices in summative assessment. **Advances in Physiology Education**, v. 41, n. 1, p. 110–119, 2017.

LIMA, PRISCILA DA SILVA e colab. SysEnade - Análise das questões de provas do Enade organizadas pelos temas abordados. *In: VII Congresso Brasileiro de Informática na Educação (WCBIE) 2018, Fortaleza, CE. Anais [...].* Fortaleza, UFCE, 2018. p. 419.

MACKENZIE, R. K. e colab. Does the UKCAT predict performance on exit from medical school? A national cohort study. **BMJ open**, v. 6, n. 10, p. e011313, 2016.

MAHER, Bridget M. e colab. Medical school attrition-beyond the statistics a ten year retrospective study. **BMC Medical Education**, v. 13, n. 13, p. 1–16, 2013.

MOTA, Alice e colab. Exame do CREMESP como indicador da qualidade do ensino médico. **Revista Brasileira de Educação Médica**, v. 38, n. 1, p. 150–159, 2014.

NORCINI, John e colab. 2018 Consensus framework for good assessment. **Medical Teacher**, v. 40, n. 11, p. 1102–1109, 2 Nov 2018.

NUNES, Luciana Neves e KLÜCK, Mariza Machado e FACHEL, Jandyra Maria Guimarães. Uso da imputação múltipla de dados faltantes: uma simulação utilizando dados epidemiológicos. *Cadernos de Saúde Pública*, v. 25, n. 2, p. 268–278, 2009.

OLIVEIRA, Fabiana Aparecida Mayrink e colab. Percepção dos acadêmicos de Odontologia em relação ao teste de progresso. **Revista da ABENO**, v. 20, n. 2, p. 26–37, 28 Jul 2020.

PALHARES NETO, Aristides Augusto e colab. Teste de Progresso: avanços e perspectivas. **Revista Brasileira de Educação Médica**, v. 46, n. supl.1, p. e160, 2022.  
Disponível em:  
<<https://www.scielo.br/j/rbem/a/rhBgjSzGbNYRjk4BcGdmLKb/?lang=pt>>.

PASQUALI, Luiz. **Psicometria dos testes na psicologia e na educação**. Edição dig ed. Petrópolis, RJ: Vozes, 2017.

PEDERSEN, Alma B. e colab. Missing data and multiple imputation in clinical epidemiological research. **Clinical Epidemiology**, v. 9, p. 157–166, 2017.

PRIMI, Ricardo. Psicometria: fundamentos matemáticos da teoria clássica dos testes. **Avaliação Psicológica**, v. 11, n. 2, p. 297–307, 2012.

PUGH, Debra e REGEHR, Glenn. Taking the sting out of assessment: is there a role for progress testing? **Medical Education**, v. 50, n. 7, p. 721–729, 2016.

RAWLINGS, Andreea Monica e colab. Multiple imputation of cognitive performance as a repeatedly measured outcome. **European Journal of Epidemiology**, v. 32, n. 1, p. 55–66, 2017.

REBERTI, Ademir Garcia e colab. Teste de Progresso na Escola Médica: uma Revisão Sistemática acerca da Literatura Progress Test in Medical School: a Systematic Review of the Literature. **Revista Brasileira de Educação Médica**, v. 44, n. 1, p. e015, 2020.

RISTOFF, Dilvo. **Mitos e meias verdades: a educação brasileira sob ataque**. Florianópolis: Insular, 2022.

ROBERTO BOLLELA, Valdes e colab. Avaliação Somativa de Habilidades Cognitivas: Experiência Envolvendo Boas Práticas para a Elaboração de Testes de Múltipla Escolha e a Composição de Exames. **Revista Brasileira de Educação Médica**, v. 42, n. 4, p. 74–85, 2018.

ROSA, Maria Inês Da e colab. O Teste de Progresso como Indicador para Melhorias em Curso de Graduação em Medicina. **Revista Brasileira de Educação Médica**, v. 41, n. 1, p. 58–68, 2017.

RUBIN, D. **Multiple imputation for nonreponse in surveys**. New York: Wiley, 1987.

RUTGERS, D. e colab. Fourteen years of progress testing in radiology residency training: experiences from The Netherlands. **European Radiology**, v. 28, n. 5, p. 2208–2215, 2018.

SAHOO, Durgesh e SINGH, Rakesh. Item and distracter analysis of multiple choice questions (MCQs) from a preliminary examination of undergraduate medical students. **International Journal of Research in Medical Sciences**, v. 5, n. 12, p. 5351–5355, 2017.

SAKAI, Marcia Hiromi e colab. Teste de progresso e avaliação do curso: dez anos de experiência da medicina da Universidade Estadual de Londrina. **Revista Brasileira de Educação Médica**, v. 32, n. 2, p. 254–263, 2008.

SAKAI, Marcia Hiromi e FERREIRA FILHO, Olavo Franco e MATSUO, Tiemi.

Avaliação do crescimento cognitivo do estudante de medicina: aplicação do teste de equalização no teste de progresso. **Revista Brasileira de Educação Médica**, v. 35, n. 4, p. 493–501, 2011.

SARTOR, Laís Büttner e colab. Undergraduate Medical Student ' s Perception about the Progress Testing Percepção dos Acadêmicos de Medicina sobre o Teste de Progresso. **Revista Brasileira de Educação Medica**, v. 44, n. 2, p. e062, 2020.

SCHEFFER, Mario e colab. **DEMOGRAFIA NO BRASIL MÉDICA 2020**. São Paulo: SP:FMUSP/CFM, 2020. Disponível em: <[www.portalmedico.org.br](http://www.portalmedico.org.br)>.

SCHEFFER, Mário e DAL POZ, Mario. The privatization of medical education in Brazil: Trends and challenges. **Human Resources for Health**, v. 13, n. 96, p. 1–10, 2015.

SCHÜTTPELZ-BRAUNS, Katrin e colab. Comparison of the evaluation of formative assessment at two medical faculties with different conditions of undergraduate training, assessment and feedback. **GMS Journal for Medical Education**, v. 37, n. 4, p. 1–23, 2020.

SHRIVE, Fiona M e colab. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. **BMC Medical Research Methodology**, v. 6, n. 57, p. 1–10, 2006. Disponível em: <<http://www.biomedcentral.com/1471-2288/6/57>>.

SINHARAY, Sandip. Score Reporting for Examinees with Incomplete Data on Large-Scale Educational Assessments. **Educational Measurement: Issues and Practice**, v. 40, n. 1, p. 79–91, 2021.

SOUSA, Leandro Araujo De e BRAGA, Adriana Eufrásio. Teoria clássica dos testes e teoria de resposta ao item em avaliação educacional. **Revista de Instrumentos, Modelos e Políticas em Avaliação Educacional**, v. 1, n. 1, p. e020002, 2 Jan 2020. Disponível em: <<https://revistas.uece.br/index.php/impa/article/view/3523>>.

TAVAKOL, Mohsen e BRENNAN, Robert L. Medical education assessment: a brief overview of concepts in generalizability theory. **International Journal of Medical Education**, v. 4, p. 221–222, 9 Nov 2013.

TAVAKOL, Mohsen e DENNICK, Reg. Making sense of Cronbach's alpha. **International Journal of Medical Education**, v. 2, p. 53–55, 2011.

TAVAKOL, Mohsen e DENNICK, Reg. Psychometric evaluation of a knowledge based examination using Rasch analysis: An illustrative guide: AMEE Guide No. 72. **Medical Teacher**, v. 35, n. 1, p. e838–e848, 2013.

TAVAKOL, Mohsen e DENNICK, Reg. The foundations of measurement and assessment in medical education. **Medical Teacher**, v. 39, n. 10, p. 1010–1015, 3 Out 2017. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/0142159X.2017.1359521>>. Acesso em: 9 abr 2019.

TOMIC, Eliane R e colab. Progress Testing: evaluation of four years of application in the school of Medicine, University of São Paulo. **CLINICS**, v. 6060, n. 55, p. 389–396, 2005.

TREVISAN, Ligia Maria Vettorato e colab. Dimensionalidade e escala de proficiência em uma prova interdisciplinar. **Estudos em Avaliação Educacional**, v. 30, n. 74, p. 392, 30 Set 2019.

TRONCON, Luiz Ernesto. Licensing examinations - a necessary component to external evaluation of students and graduates of medical courses. **Interface - Comunicação, Saúde, Educação**, v. 24, p. e190576, 2019.

UNIVERSIDADE FEDERAL DE SANTA CATARINA / UFSC. **Sistema de Controle Acadêmico da Graduação - CAGR**. Disponível em: <<https://cagr.sistemas.ufsc.br/>>. Acesso em: 11 nov 2021.

VAN DER VLEUTEN, C P M e colab. Cross institutional collaboration in assessment: a case on progress testing. **Medical Teacher**, v. 26, n. 8, p. 719–725, 2004. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/15763876>>.

VAN DER VLEUTEN, Cees. Editorial - National licensing examinations and their challenges. **Journal of Health Specialties**, v. 1, n. 1, p. 1–2, 2013.

VAN DER VLEUTEN, Cees. Revisiting ‘Assessing professional competence: from methods to programmes. **Medical Education**, v. 50, n. 9, p. 885–888, 2016.

VAN DER VLEUTEN, Cees e FREEMAN, Adrian e COLLARES, Carlos Fernando. Progress Test Utopia. **Perspectives on Medical Education**, v. 7, n. 2, p. 136–138, 1 Abr 2018.

VAN DER VLEUTEN, Cees e SCHUWIRTH, Lambert. Assessing professional competence: From methods to programmes. **Medical Education**, v. 39, n. 3, p. 309–317, 2005.

VAN DER VLEUTEN, Cees e SCHUWIRTH, Lambert. Assessment in the context of problem-based learning. **Advances in Health Sciences Education**, v. 24, n. 5, p. 903–914, 2019. Disponível em: <<https://doi.org/10.1007/s10459-019-09909-1>>.

VINHA, Luís Gustavo do Amaral. **Estudos longitudinais e tratamento de dados ausntes em avaliações educacionais**. 2016. 124 f. Tese (Doutorado em Psicologia Social, do Trabalho e das Organizações) – Instituto de Psicologia, Universidade de Brasília, Brasília, 2016.

VINHA, Luis Gustavo do Amaral e LAROS, Jacob Arie. Dados ausentes em avaliações educacionais: comparação de métodos de tratamento. **Estudos em Avaliação Educacional**, v. 29, n. 70, p. 156–187, 2018.

WALANI, Salimah R. e CLELAND, Charles M. The multiple imputation method: A case study involving secondary data analysis. **Nurse Researcher**, v. 22, n. 5, p. 13–19, 2015.



WALSH, Jason L. e colab. Formative student-authored question bank: Perceptions, question quality and association with summative performance. **Postgraduate Medical Journal**, v. 94, n. 1108, p. 97–103, 1 Fev 2018.

WANG, Ling e colab. Using Markov chain model to evaluate medical students' trajectory on progress tests and predict USMLE step 1 scores---a retrospective cohort study in one medical school. **BMC Medical Education**, v. 21, n. 200, p. 1–9, 2021.

WRIGHT, B.D. IRT in the 1990s: Which Models Work Best? 3PL or Rasch? **Rasch Measurement Transactions**, v. 6, n. 1, p. 196–200, 1992.

WRIGLEY, William e colab. A systemic framework for the progress test: Strengths, constraints and issues: AMEE Guide No. 71. **Medical Teacher**, v. 34, n. 9, p. 683–697, 2012.

ZEGOTA, Simon e colab. Using item response theory to appraise key feature examinations for clinical reasoning. **Medical Teacher (online)**, p. 1–7, 2022. Disponível em: <<https://doi.org/10.1080/0142159X.2022.2077716>>.

**APÊNDICE A – PLANILHA COMPLETA DE POSICIONAMENTO DOS 66 ITENS, DE ÓTIMA E BOA QUALIDADE, ELABORADA A PARTIR DAS RESPOSTA DOS ALUNOS DO 6º ANO AO TESTE DE PROGRESSO 2019, APLICADO NA REGIONAL SUL II DA ABEM (NAPISUL II)**

Posicionamento dos itens Teste de Progresso 2019 - NAPISUL II																	âncoras			
Análise das questões do 6º anos em 2019																	quase-âncoras			
66 itens																	dúvida quanto posicionamento do item			
Fórmula utilizada para -4,0 na célula G7 = $\$"c" + (1 - \$"c") * (1 / (1 + \text{EXP}(-\$"a" * (\text{G}\$6 - \$"b"))))$																				
Parâmetros																				
Questões	"a"	"b"	"c"	-4,0	-3,5	-3,0	-2,5	-2,0	-1,5	-1,0	-0,5	0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,00
Q_3	2,1	-0,11	0,19	0,19	0,19	0,19	0,20	0,21	0,23	0,30	0,44	0,64	0,83	0,93	0,97	0,99	1,00	1,00	1,00	1,00
Q_4	1,4	-1,81	0,26	0,30	0,33	0,39	0,47	0,58	0,71	0,82	0,89	0,94	0,97	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Q_7	0,8	2,09	0,26	0,26	0,26	0,27	0,27	0,28	0,29	0,31	0,33	0,37	0,41	0,47	0,54	0,61	0,69	0,76	0,82	0,87
Q_8	0,8	-0,19	0,24	0,28	0,29	0,32	0,35	0,39	0,44	0,50	0,57	0,65	0,72	0,78	0,84	0,88	0,91	0,94	0,96	0,97
Q_9	1,0	1,91	0,28	0,28	0,28	0,28	0,29	0,29	0,30	0,31	0,34	0,37	0,42	0,48	0,57	0,66	0,75	0,82	0,88	0,92
Q_11	1,4	0,53	0,24	0,24	0,24	0,24	0,25	0,26	0,28	0,31	0,38	0,48	0,61	0,74	0,85	0,92	0,96	0,98	0,99	0,99
Q_12	0,8	-0,07	0,26	0,30	0,31	0,33	0,36	0,40	0,45	0,50	0,57	0,64	0,71	0,77	0,83	0,87	0,91	0,93	0,95	0,97
Q_13	1,3	-0,14	0,31	0,32	0,32	0,33	0,34	0,37	0,41	0,48	0,58	0,69	0,80	0,88	0,93	0,96	0,98	0,99	0,99	1,00
Q_16	0,9	-0,74	0,24	0,27	0,29	0,32	0,36	0,42	0,49	0,57	0,66	0,74	0,81	0,87	0,91	0,94	0,96	0,98	0,98	0,99
Q_17	0,9	-0,98	0,21	0,27	0,30	0,33	0,38	0,45	0,52	0,60	0,69	0,76	0,83	0,88	0,92	0,94	0,96	0,97	0,98	0,99
Q_18	1,1	-0,26	0,30	0,32	0,33	0,34	0,36	0,40	0,45	0,52	0,61	0,70	0,79	0,86	0,91	0,94	0,97	0,98	0,99	0,99
Q_19	1,3	1,33	0,40	0,40	0,40	0,40	0,40	0,41	0,41	0,43	0,45	0,49	0,55	0,64	0,73	0,82	0,89	0,94	0,97	0,98
Q_21	0,7	-0,18	0,31	0,35	0,37	0,39	0,42	0,46	0,50	0,56	0,61	0,67	0,73	0,79	0,84	0,88	0,91	0,93	0,95	0,97
Q_23	0,9	-0,14	0,19	0,22	0,23	0,25	0,28	0,32	0,38	0,45	0,53	0,62	0,71	0,79	0,85	0,90	0,93	0,95	0,97	0,98
Q_25	1,1	1,48	0,48	0,48	0,48	0,48	0,49	0,49	0,50	0,51	0,53	0,57	0,61	0,67	0,74	0,81	0,87	0,92	0,95	0,97
Q_26	0,8	2,07	0,20	0,21	0,22	0,22	0,23	0,24	0,25	0,28	0,31	0,34	0,39	0,45	0,52	0,59	0,67	0,74	0,80	0,85
Q_27	0,8	-1,15	0,26	0,32	0,35	0,39	0,44	0,50	0,57	0,65	0,73	0,80	0,85	0,90	0,93	0,95	0,97	0,98	0,99	0,99
Q_29	1,5	1,31	0,35	0,35	0,35	0,35	0,35	0,36	0,36	0,38	0,42	0,49	0,60	0,72	0,83	0,91	0,96	0,98	0,98	0,99
Q_31	1,1	2,72	0,21	0,21	0,21	0,21	0,21	0,21	0,22	0,22	0,23	0,25	0,28	0,32	0,38	0,46	0,56	0,66	0,76	0,84
Q_33	0,9	-1,47	0,26	0,33	0,36	0,40	0,47	0,54	0,63	0,71	0,79	0,85	0,90	0,94	0,96	0,97	0,98	0,99	0,99	1,00
Q_34	0,9	1,06	0,20	0,21	0,21	0,22	0,23	0,25	0,27	0,31	0,35	0,42	0,50	0,59	0,68	0,76	0,83	0,89	0,92	0,95
Q_35	1,0	-3,34	0,25	0,50	0,59	0,69	0,78	0,85	0,90	0,94	0,96	0,98	0,99	0,99	0,99	1,00	1,00	1,00	1,00	1,00
Q_36	0,8	-3,80	0,24	0,59	0,67	0,74	0,80	0,85	0,89	0,92	0,95	0,96	0,97	0,98	0,99	0,99	0,99	1,00	1,00	1,00
Q_37	0,7	-1,71	0,24	0,36	0,40	0,45	0,51	0,58	0,65	0,72	0,78	0,83	0,88	0,91	0,94	0,96	0,97	0,98	0,98	0,99
Q_38	1,3	0,64	0,23	0,23	0,23	0,24	0,24	0,25	0,28	0,31	0,38	0,47	0,58	0,70	0,80	0,88	0,93	0,96	0,98	0,99
Q_40	1,0	1,00	0,25	0,26	0,26	0,27	0,28	0,29	0,32	0,35	0,40	0,46	0,54	0,63	0,71	0,79	0,86	0,90	0,94	0,96
Q_41	1,8	1,24	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,34	0,35	0,39	0,46	0,59	0,74	0,87	0,94	0,97	0,99	1,00
Q_42	1,7	1,86	0,11	0,11	0,11	0,11	0,11	0,12	0,12	0,12	0,13	0,15	0,19	0,28	0,43	0,61	0,78	0,89	0,95	0,98
Q_43	1,1	-0,75	0,31	0,32	0,34	0,36	0,39	0,44	0,52	0,61	0,70	0,79	0,87	0,92	0,95	0,97	0,98	0,99	0,99	1,00
Q_45	0,9	-1,18	0,28	0,34	0,37	0,40	0,45	0,52	0,59	0,67	0,75	0,81	0,87	0,91	0,94	0,96	0,97	0,98	0,99	0,99
Q_46	0,9	-0,68	0,28	0,31	0,33	0,36	0,39	0,45	0,51	0,59	0,67	0,75	0,82	0,87	0,91	0,94	0,96	0,98	0,98	0,99
Q_47	1,6	0,26	0,16	0,16	0,16	0,16	0,17	0,18	0,20	0,25	0,35	0,49	0,66	0,81	0,90	0,95	0,98	0,99	1,00	1,00
Q_48	0,8	-0,16	0,31	0,34	0,36	0,38	0,41	0,44	0,49	0,55	0,61	0,68	0,74	0,80	0,85	0,89	0,92	0,95	0,96	0,97
Q_50	1,1	-0,13	0,27	0,28	0,29	0,30	0,32	0,36	0,41	0,48	0,56	0,66	0,75	0,83	0,89	0,93	0,96	0,97	0,99	0,99
Q_51	0,9	-2,55	0,24	0,40	0,46	0,54	0,63	0,72	0,80	0,86	0,90	0,94	0,96	0,97	0,98	0,99	0,99	1,00	1,00	1,00
Q_53	0,8	-0,98	0,29	0,35	0,38	0,41	0,45	0,51	0,57	0,64	0,71	0,78	0,83	0,88	0,91	0,94	0,96	0,97	0,98	0,99
Q_54	1,0	0,00	0,30	0,31	0,32	0,33	0,35	0,38	0,43	0,49	0,56	0,65	0,73	0,81	0,87	0,91	0,94	0,97	0,98	0,99
Q_56	1,1	1,71	0,32	0,33	0,33	0,33	0,33	0,34	0,34	0,36	0,38	0,42	0,47	0,54	0,62	0,71	0,80	0,87	0,92	0,95
Q_59	0,8	1,73	0,15	0,16	0,17	0,17	0,18	0,19	0,21	0,24	0,27	0,32	0,38	0,46	0,54	0,62	0,71	0,78	0,84	0,88
Q_60	0,9	1,33	0,46	0,46	0,47	0,47	0,47	0,48	0,50	0,52	0,54	0,58	0,63	0,69	0,75	0,81	0,86	0,90	0,93	0,96
Q_64	0,9	0,19	0,32	0,34	0,35	0,37	0,39	0,41	0,45	0,50	0,56	0,63	0,71	0,77	0,83	0,88	0,92	0,94	0,96	0,98
Q_67	1,7	1,75	0,33	0,33	0,33	0,33	0,33	0,33	0,33	0,34	0,34	0,36	0,40	0,47	0,59	0,74	0,86	0,93	0,97	0,99
Q_68	1,2	1,66	0,28	0,28	0,28	0,28	0,28	0,29	0,29	0,31	0,33	0,36	0,42	0,50	0,61	0,72	0,81	0,89	0,93	0,96
Q_73	1,1	0,30	0,25	0,26	0,26	0,27	0,28	0,31	0,34	0,39	0,47	0,56	0,67	0,77	0,85	0,90	0,94	0,97	0,98	0,99
Q_75	0,7	-0,54	0,26	0,31	0,33	0,36	0,40	0,45	0,50	0,57	0,63	0,70	0,76	0,82	0,86	0,90	0,93	0,95	0,96	0,97
Q_77	1,1	-0,26	0,25	0,27	0,28	0,29	0,32	0,35	0,41	0,49	0,58	0,68	0,77	0,85	0,90	0,94	0,96	0,98	0,99	0,99
Q_78	1,1	0,08	0,20	0,21	0,22	0,23	0,25	0,28	0,32	0,39	0,48	0,58	0,69	0,78	0,86	0,91	0,94	0,97	0,98	0,99
Q_79	1,4	-0,51	0,30	0,30	0,31	0,32	0,34	0,38	0,44	0,54	0,65	0,77	0,86	0,92	0,96	0,98	0,99	0,99	1,00	1,00
Q_80	1,1	-0,77	0,24	0,27	0,28	0,31	0,35	0,41	0,48	0,58	0,68	0,77	0,84	0,90	0,94	0,96	0,98	0,99	0,99	1,00
Q_83	1,0	-1,99	0,22	0,32	0,37	0,43	0,51	0,61	0,70	0,78	0,85	0,90	0,93	0,96	0,97	0,98	0,99	0,99	1,00	1,00
Q_88	1,1	1,79	0,25	0,25	0,26	0,26	0,26	0,26	0,27	0,29	0,31	0,34	0,40	0,47	0,57	0,67	0,77	0,85	0,90	0,94
Q_89	1,4	1,74	0,12	0,12	0,12	0,12	0,12	0,13	0,13	0,14	0,16	0,19	0,25	0,35	0,49	0,64	0,78	0,87	0,93	0,97
Q_91	0,8	-0,47	0,25	0,30	0,32	0,34	0,38	0,43	0,48	0,55	0,62	0,69	0,76	0,81	0,86	0,90	0,93	0,95	0,97	0,98
Q_92	0,9	2,29	0,23	0,23	0,24	0,24	0,24	0,25	0,26	0,27	0,29	0,32	0,36	0,41	0,48	0,56	0,65	0,73	0,81	0,87
Q_93	0,7	-1,50	0,24	0,34	0,38	0,43	0,48	0,55	0,62	0,69	0,75	0,81	0,86	0,89	0,92	0,94	0,96	0,97	0,98	0,99
Q_94	0,8	0,68	0,22	0,24	0,25	0,26	0,28	0,30	0,34	0,39	0,44	0,51	0,58	0,66	0,73	0,79	0,84	0,89	0,92	0,94
Q_97	0,8	2,41	0,36	0,36	0,36	0,37	0,37	0,38	0,39	0,40	0,42	0,45	0,48	0,52	0,57	0,63	0,69	0,75	0,81	0,85
Q_98	1,3	0,98	0,17	0,17	0,18	0,18	0,18	0,19	0,21	0,23	0,28	0,36	0,46	0,59	0,72	0,82	0,90	0,94	0,97	0,98
Q_99	0,9	-0,50	0,27	0,30	0,31	0,34	0,37	0,42	0,48	0,55	0,63	0,71	0,79	0,85	0,90	0,93	0,95	0,97	0,98	0,99
Q_100	0,8	-0,50	0,27	0,32	0,34	0,37	0,40	0,45	0,50	0,57	0,63	0,70	0,77	0,82	0,87	0,90	0,93	0,95	0,97	0,98
Q_104	1,0	-0,02	0,24	0,25	0,26	0,27	0,29	0,33	0,37	0,44	0,53	0,62	0,72	0,80	0,87	0,92	0,95	0,97	0,98	0,99
Q_106	0,9	0,34	0,24	0,25																

**ANEXO A - PRIMEIRO ARTIGO – ENVIADO PARA PUBLICAÇÃO, FOI ACEITO E ESTÁ EM FASE DE EDITORIAÇÃO. SERÁ AQUI DISPONIBILIZADO, APÓS TERMINO DO PROCESSO.**

**ANEXO B – NORMAS DE PUBLICAÇÃO DA REVISTA “ESTUDOS EM  
AVALIAÇÃO EDUCACIONAL”**





## ESTUDOS EM AVALIAÇÃO EDUCACIONAL

### NORMAS PARA APRESENTAÇÃO DOS ORIGINAIS

<i>Nome do autor</i>	O nome do(s) autor(es) deve constar conforme o autor deseja que seja publicado.
<i>ORCID</i>	O número do ORCID de cada autor deve ser informado.
<i>Tamanho do título e subtítulo</i>	Títulos e subtítulos dos artigos não podem exceder 80 caracteres com espaços.
<i>Tamanho do resumo</i>	Toda matéria, à exceção de resenhas, precisa vir acompanhada de resumo em português contendo no máximo 800 caracteres com espaços, com o objetivo, o método, os resultados e as conclusões do artigo. O espaçamento deve ser de 1,5 entre as linhas e sem espaço entre parágrafos, sem conter siglas nem referências, trazendo, em seu início, o título do trabalho.
<i>Palavras-chave</i>	Ao final dos resumos, devem ser indicadas, no máximo, quatro palavras-chave (descritores) do conteúdo do texto, que devem ser buscadas no <a href="#">Thesaurus Brasileiro de Educação</a> , no <a href="#">Thesaurus de Educação da Unesco (IBE Education Thesaurus)</a> ou no <a href="#">Thesaurus Europeu dos Sistemas Educativos</a> .
<i>Tamanho da folha</i>	A4
<i>Margens</i>	3 cm de margem superior, 3 cm de margem inferior, 3 cm de margem esquerda e 2 cm de margem direita.
<i>Tipo de letra</i>	Times New Roman
<i>Tamanho da fonte</i>	Corpo 12
<i>Espaço entre linhas</i>	Espaçamento de 1,5 entre as linhas. Sem espaço (anterior ou posterior) entre os parágrafos.
<i>Número de caracteres</i>	Artigos, ensaios teóricos e relatos de experiência não podem exceder 63.300 caracteres com espaços, incluindo tabelas, gráficos, quadros ou figuras, anexos ou apêndices e referências. A extensão máxima da Resenha é de 14.550 caracteres com espaços.
<i>Siglas e abreviaturas</i>	As siglas devem ser desdobradas quando mencionadas à primeira vez no artigo. Ex.: Ministério da Educação (MEC). As siglas pronunciáveis com mais de três letras devem ser grafadas em caixa alta e baixa. Ex.: Exame Nacional do Ensino Médio (Enem).

<i>Notas de rodapé</i>	<p>As notas de rodapé devem ser evitadas e utilizadas apenas quando for estritamente necessário, preferencialmente sem ultrapassar três linhas. Será usada numeração consecutiva das notas dentro do texto.</p> <p>Não usar <i>idem</i>, <i>ibidem</i> ou <i>op. cit.</i></p>
<i>Tabelas, gráficos, quadros e figuras</i>	<p>Todas as tabelas, gráficos, quadros e figuras devem ser enviados preferencialmente em Excel, em arquivos separados, numerados com algarismos arábicos, com títulos (posicionados acima) padronizados quanto ao formato e termos utilizados. Abaixo destes, deve, obrigatoriamente, ser indicada a fonte dos dados (remetida às referências bibliográficas), com autoria e ano, inclusive se for de elaboração própria dos autores. As figuras (ilustrações, imagens, mapas, fotos, etc.) devem ter resolução de, no mínimo, 300 DPIs.</p> <p>Na versão em PDF, tabelas, gráficos, quadros e figuras devem vir inseridos no lugar em que serão publicados.</p>
<i>Metodologia estatística</i>	<p>A metodologia estatística empregada deve ser descrita com detalhamento que permita que os resultados possam ser replicados por um leitor com experiência no assunto que tenha acesso aos dados originais. Por outro lado, deve-se evitar o uso de linguagem excessivamente técnica, visando a apresentar a descrição metodológica com clareza suficiente para que um leitor não especializado no assunto também possa compreendê-la. Essas recomendações se traduzem, principalmente, em medidas como:</p> <ul style="list-style-type: none"> <li>• fornecer detalhes sobre as unidades experimentais, aleatorização e procedimentos amostrais;</li> <li>• descrever e/ou proporcionar acesso aos instrumentos de coleta da pesquisa;</li> <li>• indicar todas as técnicas descritivas e inferenciais utilizadas, fornecendo referências bibliográficas sempre que necessário;</li> <li>• apresentar os resultados quantitativos com suas respectivas medidas de incerteza (intervalos de confiança, <i>p-values</i>, etc.);</li> <li>• adotar um nível de precisão na apresentação dos dados e resultados (por exemplo, número de casas decimais) padronizado e coerente com as possibilidades da pesquisa;</li> <li>• utilizar recursos gráficos e tabelas sempre que esses elementos possam facilitar a organização e síntese dos dados e a compreensão dos resultados;</li> <li>• discutir a razoabilidade dos resultados obtidos e relatar possíveis limitações dos métodos empregados;</li> <li>• definir os termos técnicos, os símbolos matemáticos e as siglas presentes no artigo;</li> <li>• especificar os <i>softwares</i> estatísticos utilizados.</li> </ul>
<i>Citação no texto</i>	<p>As citações literais curtas são integradas ao texto, entre aspas, seguidas de parênteses com sobrenome do autor, em caixa alta, ano da publicação e número da página. Citações de três ou mais linhas devem ser destacadas no texto com recuo de 4 cm, em corpo 10 e com a indicação: (AUTOR, ano, p.).</p> <p>As remissões sem citação literal são incorporadas ao texto, sem necessidade de informar a página. Ex.: Segundo Fonseca (1986).</p>
<i>Referências</i>	<p>As referências, restritas apenas às obras citadas no texto, devem vir ao final deste e se pautarem pelas normas da ABNT (NBR 6023, 2018), apresentando o título das obras em itálico. Obedecer à ordem alfabética do sobrenome do primeiro autor e, quando possível, fazer constar por extenso o prenome dos autores. Nas referências com mais de dois autores, todos os nomes devem vir separados por ponto e vírgula, mesmo aquelas com mais de três autores, dispensando o uso de <i>et al.</i></p>

<i>Nota sobre autoria</i>	Para textos com quatro ou mais autores, deve-se incluir uma nota no final informando qual a participação de cada um na elaboração do artigo.
---------------------------	--

► O *template* para a formatação dos artigos está disponível em:

<https://www.fcc.org.br/fcc/wp-content/uploads/2022/09/Template-EAE.docx>

► Durante o preenchimento da submissão, é necessária a inclusão de dois arquivos:

*a.* um documento “.doc” ou “.docx”, com identificação de autoria; e

*b.* um documento PDF, sem identificação de autoria no corpo do texto ou em notas de rodapé, nem nas propriedades do arquivo.

► Para consultar o passo a passo da submissão, acesse:

[https://drive.google.com/file/d/1yLFCTF2qFqr4e1C5\\_rJJ-D2tytEsVGuS/view?usp=sharing](https://drive.google.com/file/d/1yLFCTF2qFqr4e1C5_rJJ-D2tytEsVGuS/view?usp=sharing)

Atualizado em: setembro, 2022



**ANEXO C – SEGUNDO ARTIGO – PUBLICADO NA “Revista Brasileira de Educação Médica – RBEM”, DA ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO MÉDICA, v. 46, n. 4, p. e142, 2022. Disponível em: <<https://doi.org/10.1590/1981-5271v46.4-20220145>>**

## Teste de Progresso: ausência de alunos de Medicina é sugestiva de mau desempenho acadêmico

*Progress Test: medical student absenteeism is suggestive of poor academic performance*

Carlos Eduardo Pinheiro<sup>1</sup>  [c.e.a.pinheiro@ufsc.br](mailto:c.e.a.pinheiro@ufsc.br)  
Emil Kupek<sup>1</sup>  [emil.kupek@ufsc.br](mailto:emil.kupek@ufsc.br)  
Olavo Franco Oliveira Filho<sup>2</sup>  [olavofranco@gmail.com](mailto:olavofranco@gmail.com)  
Ademir Reberti<sup>3</sup>  [agrebert@terra.com.br](mailto:agrebert@terra.com.br)  
Diogo Onofre Gomes de Souza<sup>4</sup>  [diogo.bioq@gmail.com](mailto:diogo.bioq@gmail.com)

### RESUMO

**Introdução:** Não se sabe se a ausência de estudantes de Medicina ao Teste de Progresso (TP) se dá de forma aleatória ou por alguma característica sistemática deles, o que poderia influenciar a representatividade dos resultados obtidos pelos participantes.

**Objetivo:** Este estudo teve como objetivos comparar os índices de desempenho acadêmico, no curso de Medicina da UFSC, dos alunos presentes e ausentes ao TP em 2019; propor uma maneira de estimar, a partir desses índices, quais seriam as notas dos faltantes se tivessem participado do TP; e identificar fatores associados à ausência ao TP.

**Método:** Foram comparadas as médias dos índices de desempenho acadêmico, globais e nas diferentes fases (semestres) dos grupos de alunos presentes e ausentes ao TP, utilizando teste t de Student para amostras independentes. Por meio de uma técnica de regressão linear, foram imputadas as prováveis notas no TP ao grupo de alunos ausentes.

**Resultado:** As médias globais dos três indicadores acadêmicos foram significativamente menores nos alunos ausentes ao TP (p variando de  $< 0,03$  a  $< 0,0001$ ); em dez das 11 fases (semestres) analisadas do curso, os indicadores acadêmicos dos faltosos foram piores do que dos presentes. A imputação de notas no TP aos ausentes permitiu verificar que existe correlação ( $R = 0,62$ ) entre a porcentagem destes e a diferença de notas entre os grupos que realizaram e os que faltaram ao TP. Entre os alunos do gênero masculino, 25,8% não fizeram o TP, enquanto no gênero feminino foram 16,6% (diferença com  $p < 0,01$ ).

**Conclusão:** A ausência de alunos ao TP não se dá de forma aleatória. Entre os faltosos, há uma tendência sistemática de existirem alunos com piores índices de desempenho acadêmico. O uso de imputação múltipla de dados evidencia uma correlação entre a porcentagem de faltosos e a diferença na média da nota no TP, desse grupo, comparada à média da nota dos participantes. A proporção de homens que faltaram ao TP foi significativamente maior do que a de mulheres.

**Palavras-chaves:** Educação Médica; Avaliação Educacional; Teste de Progresso; Imputação Múltipla; Ausência de estudantes ao TP.

### ABSTRACT

**Introduction:** It is not known whether the absence of medical students at the Progress Test (PT) is random event or if it due to some systematic characteristic of the students, which could influence the representativeness of the results obtained by the participants.

**Objectives:** 1) to compare the academic performance indexes, in UFSC Medical School, of students who were present and absent from the PT in 2019; 2) to propose a way to estimate, based on these indexes, what the absentee's grades would be if they had participated in the PT; 3) to identify factors associated with absence from the PT.

**Method:** The averages of academic performance indexes, overall and in the different phases (semesters) in the groups of students who were present and absent from the PT, were compared using Student's t test for independent samples. Using a linear regression technique, the probable PT scores were assigned to the group of absent students.

**Results:** The global averages of the three academic indicators were significantly lower in students absent from the PT (p ranging from  $< 0.03$  to  $< 0.0001$ ); in 10 of the 11 analyzed course phases (semesters), the academic indicators of absentees were worse than those present at the test. The attribution of PT grades to the absentees allowed us to verify that there is a correlation ( $R=0.62$ ) between the percentage of these students and the difference in grades between the groups that took and those that did not take the PT. Among male students, 25.8% did not attend the PT, while among female students the number of absentees was 16.6% (difference with  $p < 0.01$ ).

**Conclusions:** The absence of students at the PT does not occur randomly. Among the absentees, there is a systematic tendency to have students with worse academic performance. The use of multiple imputation of data demonstrate a correlation between the percentage of absentees and the difference in the average of grades in the PT of this group, compared to the average of the participants' grades. The proportion of male students who missed the PT was significantly higher than that of female students.

**Keywords:** Medical Education; Educational assessment; Progress Test; Multiple imputation; Absence of students at PT.

<sup>1</sup>Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil.

<sup>2</sup>Universidade Estadual de Londrina, Londrina, Paraná, Brasil.

<sup>3</sup>Universidade da Região de Joinville, Joinville, Santa Catarina, Brasil.

<sup>4</sup>Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil.

Editora-chefe: Rosiane Viana Zuza Diniz.  
Editora associada: Pedro Hamamoto Filho.

Recebido em 19/05/22; Aceito em 21/09/22.

Avaliado pelo processo de *double blind review*.

