

# Avaliação do Teste do Progresso de uma faculdade de Medicina pelos pressupostos da Taxonomia SOLO

*Evaluation of the Progress Test of a Medical school according to the SOLO Taxonomy theoretical framework*

Pedro Paulo Trindade Resende<sup>1</sup>  [pedroresende@ufsj.edu.br](mailto:pedroresende@ufsj.edu.br)  
Alexandre de Araújo Pereira<sup>2</sup>  [alexandre.pereira@unifenas.br](mailto:alexandre.pereira@unifenas.br)  
José Maria Peixoto<sup>2</sup>  [jmpeixoto.prof@gmail.com](mailto:jmpeixoto.prof@gmail.com)

## RESUMO

**Introdução:** A capacitação de futuros egressos das faculdades de Medicina para uma prática assistencial responsável e qualificada de atenção à saúde é um desafio. O objetivo é que estejam aptos para a resolução de problemas que demandem habilidades cognitivas de ordem superior. Sendo assim, a avaliação da aquisição de tais competências se torna muito importante. Um método de avaliação que vem ganhando atenção no ensino médico é o Teste do Progresso (TP). Teorias cognitivas têm trazido progresso em pesquisas educacionais relativas aos processos de avaliação. Utilizamos no nosso estudo a Taxonomia *Structure of Observing Learning Outcome* (SOLO) para avaliar e categorizar os itens do TP aplicado em uma faculdade de Medicina. A Taxonomia SOLO (TS) possibilita a análise cognitiva necessária para a realização de determinadas tarefas, permitindo uma observação integral da compreensão do entendimento do aluno. Utilizamos também no nosso estudo a Teoria Clássica dos Testes (TCT) e calculamos para cada item de múltipla escolha (IME) do TP o índice de dificuldade (IDF) e o índice de discriminação (ID), e os correlacionamos com a classificação SOLO.

**Objetivo:** Este estudo teve como objetivo avaliar características do TP aplicado em uma faculdade privada de Medicina, analisando seus itens pelos pressupostos da TS e correlacionando-os com a TCT.

**Método:** Trata-se de um estudo descritivo, de abordagem quantitativa e qualitativa. De acordo com os princípios da TS, foram realizadas a análise e a caracterização dos itens da prova de um TP aplicado em uma faculdade privada de Medicina e a correlação com o IDF e o ID.

**Resultado:** Verificamos um equilíbrio entre a aprendizagem superficial (AS) e a aprendizagem profunda (AP) no total de itens e uma relação direta entre os níveis de AP e IME compostos por casos clínicos. Não verificamos diferença estatisticamente significativa entre as categorias SOLO quanto às médias do IDF e do ID.

**Conclusão:** A análise das atividades avaliativas não deve ser restringida às propriedades psicométricas. Ferramentas taxonômicas, como a TS, podem auxiliar de maneira significativa a realização dessas atividades, de modo a conciliar as avaliações ao currículo, possibilitar a realização de provas adequadas ao nível de aprendizagem desejável e favorecer a progressividade do ensino.

**Palavras-chave:** Taxonomia; Educação Médica; Desempenho acadêmico; Avaliação Educacional.

## ABSTRACT

**Introduction:** the training of future graduates from medical schools for responsible and qualified health care practice is a significant challenge. The goal is for them to be equipped to solve problems that require higher-order cognitive skills. Thus, evaluating the acquisition of such competencies becomes crucial. One assessment method that has been gaining attention in medical education is the Progress Test (PT). Cognitive theories have advanced educational research related to assessment processes. In our study, we used the *Structure of Observing Learning Outcome* (SOLO) taxonomy to evaluate and categorize the items of the PT applied at a medical school. The SOLO taxonomy (ST) allows for the necessary cognitive analysis required for performing specific tasks, enabling a comprehensive observation of the student's understanding. We also applied the *Classical Test Theory* (CTT) in our study, calculating the difficulty index (DFI) and discrimination index (DI) for each multiple-choice item (MCI) of the PT and correlating them with the SOLO classification.

**Objective:** the objective of this study is to evaluate the characteristics of the PT applied in a private medical school, analyzing its items based on the assumptions of ST and correlating them with CTT.

**Materials and Methods:** this is a descriptive study with a quantitative and qualitative approach. According to the principles of ST, we conducted the analysis and characterization of the items from a PT applied in a private medical school and correlated them with the DFI and DI.

**Results:** we found a balance between surface learning (SL) and deep learning (DL) across the total items, as well as a direct relationship between the levels of DL and MCIs consisting of clinical cases. We did not find statistically significant differences between the SOLO categories regarding the means of DFI and DI.

**Conclusion:** the analysis of assessment activities should not be restricted to psychometric properties. Taxonomic tools, such as the ST, can significantly aid in conducting these activities, aligning assessments with the curriculum and facilitating the creation of tests appropriate for the desired level of learning, thereby promoting effective teaching progression.

**Keywords:** Taxonomy; Education, Medical; Academic Performance; Educational Measurement.

<sup>1</sup> Universidade Federal de São João del-Rei, São João Del Rei, Minas Gerais, Brasil.

<sup>2</sup> Universidade José do Rosário Velano, Belo Horizonte, Minas Gerais, Brasil.

Editora-chefe: Rosiane Viana Zuza Diniz.  
Editora associada: Daniela Chiesa.

Recebido em 09/10/23; Aceito em 17/11/24.

Avaliado pelo processo de double blind review.

## INTRODUÇÃO

O propósito das escolas de Medicina é preparar futuros egressos para uma prática assistencial responsável e qualificada. Nesse sentido, o processo de avaliação da aprendizagem é essencial, pois possibilita observar os resultados das intervenções educacionais<sup>1</sup>. O procedimento de avaliação é uma importante ferramenta no arsenal do ensino, sendo um instrumento impulsionador da aprendizagem<sup>2,4</sup>. No entanto, comumente os métodos de avaliação utilizados não são planejados para identificar a evolução das habilidades cognitivas dos estudantes e, em geral, avaliam a aquisição dos conteúdos estudados, muitas vezes cobrando uma simples memorização.

Para que possa contribuir para o processo de aprendizagem, a avaliação deve ser planejada de forma a espelhar os objetivos a serem alcançados, nos níveis cognitivos, afetivos e psicomotores, e permitir que tanto os docentes como os estudantes possam identificar seus progressos e pontos de melhorias<sup>5</sup>.

Um método de avaliação que vem ganhando atenção nas escolas médicas é o Teste do Progresso (TP), que avalia o desempenho cognitivo longitudinal dos estudantes<sup>6</sup>. O TP é aplicado semestralmente ou anualmente a todos os discentes do curso, de forma simultânea, sendo constituído de itens de múltipla escolha (IME). O conteúdo do TP compreende toda a matriz curricular, sendo relacionado aos seus objetivos finais e baseado nas Diretrizes Curriculares Nacionais do curso de graduação em Medicina<sup>7,8</sup>.

OTP deve exigir dos alunos muito além da memorização. Os IME são elaborados por docentes da própria instituição ou de consórcios formados com esse propósito, e espera-se um ganho cognitivo gradativo com o passar dos períodos<sup>8</sup>.

Pela importância que o TP vem ganhando na educação médica, é relevante avaliar se a qualidade de seus itens atende às premissas de um bom processo avaliativo. Isso inclui verificar se os IME do TP apresentam:

- *Capacidade discriminativa*: Os itens devem ser capazes de distinguir entre estudantes com diferentes níveis de conhecimento e habilidades.
- *Equilíbrio quantitativo em relação à dificuldade dos itens*: Deve haver uma distribuição balanceada de itens com diferentes graus de dificuldade, de modo a garantir que o teste seja desafiador o suficiente para todos os alunos, sem ser excessivamente fácil ou difícil.
- *Complexidade cognitiva apropriada*: Os itens devem exigir não apenas memorização, mas também capacidade de fazer relações, inferências e generalizações, refletindo as demandas cognitivas da prática médica.

É importante distinguir entre dificuldade e complexidade em itens avaliativos. A dificuldade se relaciona a fatores que

oferecem obstáculos para a resolução de um problema, enquanto a complexidade envolve as habilidades cognitivas necessárias para resolvê-lo. A dificuldade é geralmente avaliada estatisticamente, enquanto a complexidade é avaliada qualitativamente, muitas vezes utilizando taxonomias educacionais<sup>9,10</sup>. Portanto, um item pode apresentar elevado índice de dificuldade e baixa complexidade ou vice-versa.

Dentre os métodos estatísticos para a avaliação psicométrica, a Teoria Clássica dos Testes (TCT) permite avaliar o índice de dificuldade (IDF) e o índice de discriminação (ID) dos itens que compõem o teste. O IDF é encontrado calculando a razão entre o número de alunos que responderam corretamente e o número total de alunos submetidos ao item. Esse índice varia de 0 a 1. A Tabela 1 apresenta uma classificação dos itens de uma prova em relação ao IDF e o percentual esperado de distribuição em uma avaliação<sup>11</sup>.

O ID é calculado pela diferença entre o percentual de acerto de um determinado item dos estudantes que tiveram melhor desempenho no teste e o percentual dos que tiveram pior desempenho. Para obter esse cálculo, os candidatos serão alocados em três grupos: o grupo superior (27% das maiores pontuações), o grupo inferior (27% das menores pontuações) e o grupo intermediário com os demais 46% dos candidatos<sup>11</sup>. A Tabela 2 apresenta os critérios dos valores do ID e a classificação do item de acordo com seu poder de discriminação.

**Tabela 1.** Critérios de distribuição e classificação do grau de dificuldade do item pela TCT.

Quantitativo ideal de itens em uma avaliação (% esperado)	Índice de dificuldade do item	Classificação do item em relação ao índice de dificuldade
10%	Superior a 0,9	Muito fáceis
20%	De 0,7 a 0,9	Fáceis
40%	De 0,3 a 0,7	Medianos
20%	De 0,1 a 0,3	Difíceis
10%	Até 0,1	Muito difíceis

Fonte: Vilarinho<sup>11</sup> (p. 27).

**Tabela 2.** Valores do poder de discriminação e classificação do item pela TCT.

Valores	Classificação
Discriminação < 0,20	Item deficiente, deve ser rejeitado
0,20 ≤ Discriminação < 0,30	Item marginal, sujeito a reelaboração
0,30 ≤ Discriminação < 0,40	Item bom, mas sujeito a aprimoramento
Discriminação ≥ 0,40	Item bom

Fonte: Vilarinho<sup>11</sup> (p. 28).

Em relação à complexidade dos itens, eles podem ser avaliados por meio das taxonomias educacionais, que são sistemas de classificação que possibilitam a categorização dos níveis de aprendizagem atingidos pelos estudantes, sendo útil para o sistema de avaliação e planejamento dos objetivos educacionais<sup>12</sup>. Dentre as taxonomias educacionais existentes, a Taxonomia *Structure of Observing Learning Outcome (SOLO)* foi concebida com a ideia de que os indivíduos aprendem diferentes conteúdos em estágios de complexidade ascendentes<sup>13,14</sup>.

Biggs et al.<sup>15</sup> propuseram uma categorização dos estágios de entendimento de conteúdo, denominados “modos de pensamento”, baseados nos estágios piagetianos. Esse sistema identifica diferentes graus de formalização do pensamento, permitindo avaliar a qualidade da aprendizagem. A Taxonomia SOLO (TS), derivada desses conceitos, classifica a estrutura da aprendizagem demonstrada em cinco níveis progressivos de complexidade cognitiva<sup>15</sup> (Figura 1):

- Pré-estrutural (SOLO 1): respostas inadequadas, informações irrelevantes ou incoerentes.
- Uniestrutural (SOLO 2): respostas direcionadas a um único elemento da tarefa e, portanto, inconsistentes.
- Multiestrutural (SOLO 3): respostas identificam mais de um elemento da tarefa, mas não há integração das informações, levando a incoerências.
- Relacional (SOLO 4): diversas informações são identificadas e são estabelecidas relações de forma coerente, havendo um entendimento do todo, não havendo inconsistências.
- Abstrato estendido (SOLO 5): resposta vai além dos elementos do item, partindo para abstração e generalização.

Estudos identificaram duas formas principais de aprendizagem: uma denominada superficial e a outra

profunda. A aprendizagem superficial (AS) é caracterizada pela reprodução do conteúdo sem conexões ou reflexões, enquanto a aprendizagem profunda (AP) envolve um entendimento intrínseco e reflexivo, exigindo processos cognitivos elaborados<sup>15</sup>. A AS é formada pelos níveis SOLO 2 e SOLO 3, e a AP, pelos níveis SOLO 4 e SOLO 5. A AS se baseia na retenção de detalhes concretos por meio de memorização. A AP é mais complexa, pois requerer relações de informações, de modo a qualificar o indivíduo para compreender mecanismos e princípios e realizar generalizações ou teorizações<sup>15,17</sup>.

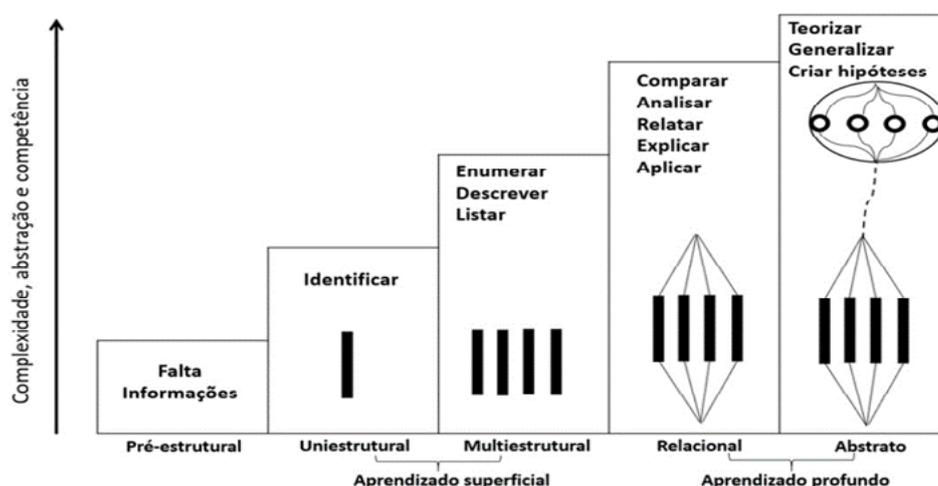
Ao longo da formação, os estudantes progredem de forma ascendente, de uma compreensão concreta para abstrata, desenvolvendo habilidades de estabelecer relações e fazer inferências, refletindo um aumento na capacidade de manejar informações de forma consistente e realizar generalizações. Essa evolução ascendente do processo cognitivo dos estudantes pode ser categorizada como ciclos de aprendizagem, que representam a forma como o discente compreende e opera o conteúdo estudado, do modo mais concreto ao mais abstrato<sup>18</sup>.

Pelos motivos expostos, o objetivo do presente estudo foi analisar as características psicométricas de uma avaliação do TP pela TCT, em relação ao IDF e ao ID obtidos, além de categorizar a complexidade cognitivas de seus itens pelos pressupostos da TS. A TS foi selecionada para este estudo devido à sua capacidade de classificar a complexidade da estrutura da aprendizagem demonstrada pelos estudantes em uma tarefa específica, identificando os processos de pensamento envolvidos, e pela possibilidade de diferenciar a AS da AP<sup>19,20</sup>.

## MATERIAIS E MÉTODO

Trata-se de um estudo descritivo, de abordagem quantitativa e qualitativa. Analisamos o TP aplicado, no

**Figura 1.** Taxonomia SOLO.



Fonte: Elaborado pelos autores a partir de BIGGS, John. Calidad del aprendizaje universitario. Narcea, 2006. p. 71<sup>16</sup>.

**Quadro 1.** Sistema de categorização dos itens do TPU2022-1.

Categoria SOLO da questão	Exemplos de verbos de comando e sua relação com as categorias SOLO	Conhecimento abordado no item		Procedimentos cognitivos envolvidos para resolução do item
		Número de tópicos utilizados	Relação entre os tópicos	
Abstrato (SOLO 5)	Discutir, hipotetizar, avaliar, raciocinar, estimar, criticar, interpretar, prever, refletir, programar, julgar, generalizar, implementar.	Dois ou mais tópicos.	Necessita de associação entre os tópicos.	Indução e/ou dedução; exige identificação de informações relevantes não comumente discutidas no curso médico, elaboração de hipóteses e generalizações.
Relacional (SOLO 4)	Explicar, integrar, referir, analisar, comparar, interpretar, construir, planejar, resumir, relacionar, argumentar.			Indução e/ou dedução; exige identificação de informações relevantes frequentemente discutidas no curso médico.
Multiestrutural (SOLO 3)	Descrever, executar, resolver, aplicar, combinar, completar, classificar, enumerar.	Um único tópico.	Não se aplica.	Indução e/ou dedução; exige identificação de informações relevantes frequentemente discutidas no curso médico.
Uniestrutural (SOLO 2)	Identificar, decidir, organizar, reproduzir, eleger, encontrar, reconhecer, contar, buscar, parafrasear.			Indução e/ou dedução; exige identificação de informação relevante frequentemente discutida no curso médico.

Fonte: Adaptado de Ceia<sup>21</sup> e Pereira<sup>10</sup>.

primeiro semestre do ano de 2022, a todos os alunos, do primeiro ao 12º período, do curso de Medicina da Universidade Professor Edson Antônio Velano (Unifenas) – câmpus de Belo Horizonte (TPU2022-1), de forma simultânea, contendo 120 IME. O conteúdo foi dividido nas áreas do conhecimento: cirurgia, clínica médica, ginecologia-obstetrícia, pediatria e saúde coletiva, com 24 itens para cada área.

A prova do TPU2022-1 foi baseada na matriz de prova do TEP MINAS 2019, que foi elaborada pela equipe do Consórcio Mineiro de Escolas Médicas para o Teste do Progresso (TEP MINAS 1). Fazem parte do TEP MINAS 1 oito escolas médicas do estado de Minas Gerais, entre entidades públicas e privadas.

As encomendas (guia para elaboração das questões) foram enviadas aos professores das instituições de acordo com a sua área de atuação. Quando entregues, as questões foram revisadas, e realizaram-se as correções necessárias.

Os IME do TPU2022-1 foram avaliados por meio da TCT e categorizados pelos critérios da TS. A categorização dos IME pela TS foi realizada por meio de uma adaptação do modelo de classificação da complexidade dos itens de uma avaliação proposto por Mário Ceia<sup>21</sup>. Por esse modelo, o item de uma avaliação é analisado a partir da resposta esperada para a questão, considerando três parâmetros: quantidade de conhecimento necessário para sua resolução, operações cognitivas envolvidas na solução do problema e complexidade da resposta solicitada. O Quadro 1 apresenta o sistema de categorização proposto para este estudo, que foi adaptado dos trabalhos de Ceia<sup>21</sup> e Pereira<sup>10</sup>.

**Quadro 2.** Ficha de Categorização do Item.

A seguir você encontrará a transcrição da questão de nº ..... do Teste do Progresso da Unifenas-BH do ano de 20..... Leia o enunciado da questão e avalie a resposta assinalada como correta. Partindo da resposta correta, analise os conteúdos e os procedimentos cognitivos exigidos para a resolução da questão, considerando o sistema de categorização dos itens apresentado a você. Ao final classifique a complexidade da questão pela Taxonomia SOLO.
Enunciado do item:
Alternativas:
Gabarito da questão: .....
Análise de conteúdo:
Análise de procedimentos:
Categoria SOLO da questão: ( ) SOLO 2 ( ) SOLO 3 ( ) SOLO 4 ( ) SOLO 5

Fonte: Elaborado pelos autores.

A partir das informações do Quadro 1, elaborou-se a Ficha de Categorização do Item – FCI (Quadro 2), em que, para cada item do TPU2022-1, transcreveram-se o enunciado, as alternativas e o gabarito da questão. Ainda constam na FCI espaços para análise do conteúdo do item e dos procedimentos cognitivos necessários para sua resolução, e, por fim, um espaço para sua categorização pela TS. Uma lista de verbos frequentemente associada a cada categoria SOLO foi anexada à FCI para que os avaliadores pudessem consultar e melhor ajustar seus pareceres.

Participaram da categorização dos IME dois professores médicos, doutores, que receberam treinamento sobre os

pressupostos da TS para a categorização de IME, além do pesquisador principal. Foram selecionados 33,33% dos itens do TPU2022-1, que correspondem a 40 itens, utilizando uma amostragem probabilística sistemática para que os três avaliadores realizassem suas análises com o intuito de identificar algum viés de categorização. Cada avaliador recebeu os itens selecionados e, após análise individual, preencheu a FCI. As análises dos avaliadores foram confrontadas, e verificou-se uma alta taxa de concordância: 95%. Ajustes foram realizados por consenso, permitindo uma calibração da categorização por parte do pesquisador principal. Os demais itens foram analisados e classificados somente pelo pesquisador principal.

Todos os itens do TPU2022-1 foram classificados de acordo com os princípios da TS, em quatro níveis de complexidade cognitivas: uniestructural, multiestructural, relacional e abstrato estendido. O nível pré-estrutural não foi incluído, uma vez que o propósito era analisar os IME do TP em relação à complexidade cognitiva exigida para sua resolução; portanto, não eram esperados itens com essa categorização. Após categorização, os itens foram subdivididos em duas categorias de aprendizagem: superficial (SOLO 2 e 3) e profunda (SOLO 4 e 5).

Os IME foram também submetidos à análise psicométrica pela TCT, em que se calcularam para cada item: o IDF e o ID. Como o TP é aplicado a estudantes com diferentes ciclos de aprendizagem, para a análise pela TCT consideraram-se como referência as provas dos alunos do último ano do curso, uma vez que se trata de uma amostra representativa de discentes que integralizaram 83,33% da matriz curricular.

Com o objetivo de investigar se houve diferença significativa nas médias das medidas dos parâmetros IDF e ID quanto aos níveis do SOLO, a análise de variância com um fator (*oneway*) foi aplicada aos dados. O Teste de Igualdade de Variância de Levene foi utilizado para investigar se as variâncias entre as categorias eram estatisticamente diferentes. Além disso, calculou-se o tamanho de efeito (“eta parcial ao quadrado”). Para investigar se ocorreu diferença significativa nas médias das medidas dos parâmetros IDF e ID quanto às categorias SOLO que representam a AS e a AP (2 ou 3 × 4), o teste *t* de Student para amostras independentes foi aplicado aos dados. O Teste de Igualdade de Variância de Levene foi novamente utilizado para investigar se as variâncias entre as categorias eram estatisticamente diferentes. Além disso, calculou-se o tamanho de efeito (*d* de Cohen). Os resultados foram considerados significativos para uma probabilidade de significância inferior a 5%, com pelo menos 95% de confiança nas conclusões apresentadas.

O presente estudo foi aprovado pelo Comitê de Ética em Pesquisa da Unifenas, sob o CAAE nº 56009222.9.0000.5143, Parecer nº 5.379.183, solicitou-se a dispensa do Termo de

Consentimento Livre e Esclarecido, e encaminharam-se o Termo de Compromisso de Utilização de Dados e o Termo de Anuência.

## RESULTADOS

Encontramos um equilíbrio entre a AS e a AP nos itens do TPU2022-1. Cerca de 41% foram classificados como AS, e 59,2% como AP, como mostra a Tabela 3. Quando analisamos por áreas específicas do conhecimento, constatamos esse equilíbrio nas áreas de cirurgia, ginecologia-obstetrícia e pediatria. Nas áreas de clínica médica e saúde coletiva, não verificamos equilíbrio entre a AS e a AP. Em clínica médica, verificamos que 87,5% dos itens analisados estavam relacionados à AP e 12,5% à AS. Já na área de saúde coletiva, constatamos que cerca de 83,3% dos itens relacionavam-se à AS e 16,7% à AP. A Tabela 3 divide os itens do TPU2022-1 entre AS e AP.

Não identificamos nenhum item relacionado ao nível abstrato estendido, nível de maior complexidade cognitiva da TS. A Tabela 4 apresenta o resultado da análise, de acordo com os níveis de complexidade da TS, por áreas do conhecimento: cirurgia, clínica médica, ginecologia-obstetrícia, pediatria e saúde coletiva.

Constatamos que os itens compostos por casos clínicos que exigiam habilidades de resolução de problemas propiciaram maior exploração da AP. A Tabela 5 apresenta a estratégia empregada na formulação dos itens, baseada na presença ou ausência de casos clínicos. A maioria dos itens (84,2%) apresentava a presença de um caso clínico. Verificamos a necessidade da relação de tópicos para a resolução de casos clínicos em grande parte dos itens das áreas de conhecimento, com exceção da área de saúde coletiva, na qual uma considerável parcela dos itens (54,2%) não continha casos clínicos, e, quando um caso clínico estava presente, o que ocorreu em 11 itens (45,8%), somente em dois foi exigida a relação de tópicos para sua resolução.

**Tabela 3.** Distribuição dos itens do TPU2022-1 entre aprendizagem superficial e aprendizagem profunda.

Área/SOLO	Aprendizagem superficial	Aprendizagem profunda
Cirurgia	7 (29,2%)	17 (70,8%)
Clínica médica	3 (12,5%)	21 (87,5%)
Ginecologia-obstetrícia	10 (41,7%)	14 (58,3%)
Pediatria	9 (37,5%)	15 (62,5%)
Saúde coletiva	20 (83,3%)	4 (16,7%)
Total	49 (40,8%)	71 (59,2%)

Fonte: Elaborada pelos autores.

**Tabela 4.** Distribuição dos itens do TPU2022-1 pelos níveis da TS.

Área/SOLO	Uniestrutural	Multiestrutural	Relacional	Abstrato
Cirurgia	3 (12,5%)	4 (16,7%)	17 (70,8%)	0
Clínica	2 (8,3%)	1 (4,2%)	21 (87,5%)	0
Ginecologia-obstetrícia	7 (29,2%)	3 (12,5%)	14 (58,3%)	0
Pediatria	7 (29,2%)	2 (8,3%)	15 (62,5%)	0
Saúde coletiva	18 (75,0%)	2 (8,3%)	4 (16,7%)	0
Total	37 (30,8%)	12 (10%)	71 (59,1%)	0

Fonte: Elaborada pelos autores.

**Tabela 5.** Presença ou ausência de casos clínicos nos itens do TPU2022-1.

Área/item SOLO	Ausência de caso clínico	Presença de caso clínico
Cirurgia	1	23
Clínica médica	0	24
Ginecologia-obstetrícia	3	21
Pediatria	2	22
Saúde coletiva	13	11
Total	19	101

Fonte: Elaborada pelos autores.

**Tabela 7.** Índice de discriminação do TPU2022-1.

Discriminação	Número de questões	Frequência (%)
< 20	95	79,1
0,20 ≤ a < 0,30	15	12,5
0,30 ≤ a < 0,40	8	6,7
≥ 40	2	1,7
Total	120	100,0

Fonte: Elaborada pelos autores.

Quanto à análise psicométrica, verificamos que 10,8% dos itens do TPU2022-1 apresentaram o IDF no nível fácil, 50,9% no nível mediano e 25% no nível difícil. A proporção de itens muito difíceis e muito fáceis foi de 13,3%. A Tabela 6 apresenta os resultados de acordo com o IDF.

Verificamos na nossa análise que 79,2% dos itens apresentaram um ID menor que 20%. Cerca de 2% apresentaram um ID maior que 40%. A Tabela 7 apresenta os resultados de acordo com o ID.

A Tabela 8 demonstra que não foi identificada diferença estatisticamente significativa entre as três categorias SOLO quanto aos parâmetros da TCT. Ressalta-se que o tamanho de efeito calculado é considerado pequeno, o que corrobora a não significância estatística do teste. Cabe ressaltar que não

**Tabela 6.** Índice de dificuldade do TPU2022-1.

% de acerto	Número de questões	Frequência (%)
≤ 10,0% (muito fácil)	1	0,8
De 10,1 a 30,0% (fácil)	13	10,8
De 30,1 a 70,0% (mediana)	61	50,9
De 70,1 a 90,0% (difícil)	30	25,0
> 90,0% (muito difícil)	15	12,5
Total	120	100,0

Fonte: Elaborada pelos autores.

**Tabela 8.** Análises comparativa entre as categorias do SOLO quanto aos parâmetros de avaliação do TCT (dificuldade e discriminação).

Categoria SOLO	Medidas descritivas		Valor de p
	N	Média ± DP	
<i>Índice de dificuldade</i>			
SOLO 2	37	60,9 ± 22,8	0,821 F <sub>2, 117</sub> = 0,197
SOLO 3	12	62,1 ± 19,4	
SOLO 4	71	58,5 ± 24,4	
Geral	120	59,6 ± 23,3	
<i>Índice de discriminação</i>			
SOLO 2	37	15,0 ± 10,6	0,484 F <sub>2, 117</sub> = 0,731
SOLO 3	12	11,3 ± 7,2	
SOLO 4	71	13,4 ± 9,6	
Geral	120	13,7 ± 9,7	

Base de dados: 120 questões (SOLO 2 → 37 casos, SOLO 3 → 12 casos e SOLO 4 → 71 casos)

Nota: d.p. → Desvio-padrão

p → Probabilidade de significância da *Análise de Variância com 1 fator (OneWay)*.

F → Estatística da *Análise de Variância com 1 fator (OneWay)*.

- Teste de *Levene* para igualdade de variância:  $p \gg 0,05$ . Para ambas as variáveis.

- Tamanho de efeito ( $\eta^2$ ):  $\eta^2 \leq 0,01$  → Tamanho de efeito Pequeno (ambas as variáveis).

Fonte: Elaborado pelos autores.

foi observada diferença estatisticamente significativa entre as variâncias (Levene  $\rightarrow p > 0,05$ ). Logo, não há necessidade da aplicação do teste de Welch.

A Tabela 9 demonstra que não foi verificada diferença estatisticamente significativa entre as categorias SOLO que representam a AS e a AP quanto aos parâmetros da TCT.

No final deste artigo, está disposto o *link* de acesso à análise de todos os itens do TPU2022-1 pelos pressupostos da TS.

## DISCUSSÃO

Este estudo teve como objetivo avaliar os itens do TP aplicado em uma instituição de ensino privada, em relação às suas características psicométricas e de complexidade cognitiva pela TS. Nossos resultados constataram no TPU2022-1 um equilíbrio entre a frequência de itens relacionados à AS e à AP, com um predomínio de itens associados à AP. As áreas de cirurgia, ginecologia-obstetrícia e pediatria foram as que mais apresentaram esse equilíbrio.

Os pesquisadores defendem que um equilíbrio na distribuição dos itens de uma avaliação, relacionado aos níveis de complexidade cognitiva, contribui para uma melhor avaliação da aprendizagem. Dessa forma, ganha grande importância uma distribuição balanceada de acordo com as taxonomias de aprendizagem<sup>22</sup>. Apesar da importância atribuída a esse equilíbrio, é de se supor que os egressos do curso médico estejam aptos para a resolução de problemas complexos. A expectativa é que os alunos dos últimos períodos estejam capacitados para a resolução de atividades, como solução de problemas e tomadas de decisões<sup>23</sup>. Dessa forma, acreditamos que no TP deveriam prevalecer atividades destinadas a avaliar a AP. A TS auxilia na construção e seleção de itens de avaliação adequados para a verificação de aquisição de atributos esperados de um graduado e que fomentem a análise e não uma simples memorização<sup>24</sup>.

Os itens classificados como AP do TPU2022-1 continuam, na maioria das vezes, um caso clínico, e o enunciado demandava habilidades de resolução de problemas e integração de conhecimentos, exigindo um raciocínio clínico. O raciocínio clínico exige uma base de conhecimento, possibilitando aos estudantes a geração de hipóteses, o estabelecimento de diagnósticos e o oferecimento de uma conduta para a resolução de problemas clínicos<sup>25,26</sup>. O uso de resoluções de problemas clínicos genuínos é uma medida eficiente de análise do raciocínio clínico<sup>27</sup>.

A construção do pensamento de ordem superior é fundamental na formação e no exercício da medicina<sup>28</sup>. A AP é associada a uma melhor efetividade no ensino médico está mais relacionada à capacidade de os estudantes se atualizarem após a sua formação<sup>17</sup>.

**Tabela 9.** Análises comparativa entre as categorias SOLO (aprendizagem superficial e profunda) quanto aos parâmetros de avaliação do TCT (dificuldade e discriminação).

Categoria SOLO de aprendizagem	Medidas descritivas		Valor de p
	N	Média ± DP	
<i>Índice de dificuldade</i>			
Superficial	49	61,2 ± 21,8	0,544 t118 = 0,609
Profunda	71	58,5 ± 24,4	
Geral	120	59,6 ± 23,3	
<i>Índice de discriminação</i>			
Superficial	49	14,1 ± 9,9	0,684 t118 = 0,408
Profunda	71	13,4 ± 9,6	
Geral	120	13,7 ± 9,7	

Base de dados: 120 questões (SOLO 2  $\rightarrow$  37 casos, SOLO 3  $\rightarrow$  12 casos e SOLO 4  $\rightarrow$  71 casos)

Nota: d.p.  $\rightarrow$  Desvio-padrão

p  $\rightarrow$  Probabilidade de significância da *Análise de Variância com 1 fator (OneWay)*.

F  $\rightarrow$  Estatística da *Análise de Variância com 1 fator (OneWay)*.

- Teste de Levene para igualdade de variância: p  $\gg$  0,05. Para ambas as variáveis.

- Tamanho de efeito (*d de Cohen*):  $d \leq 0,12 \rightarrow$  Tamanho de efeito Pequeno (ambas as variáveis).

Fonte: Elaborado pelos autores.

Outra importante aplicabilidade das taxonomias de aprendizagem, como a TS, é oferecer ao discente dados sobre seu nível de pensamento cognitivo. Isso se torna muito significativo nesse momento de mudança de paradigma do ensino superior, no qual se recomenda uma aprendizagem autônoma, centrada no aluno<sup>29</sup>.

A não identificação de nenhum item do TPU2022-1 relacionado ao nível abstrato estendido pode estar associada ao uso de IME, o que pode trazer uma certa limitação para a avaliação do nível taxonômico mais elevado<sup>30,31</sup>. Outras estratégias de avaliação, como os itens discursivos, são adequadas para demonstrar esse nível, no entanto seria inviável a realização de um exame com esse tipo de item, cujo objetivo é avaliar um número significativo de alunos, com um conteúdo muito extenso<sup>32</sup>. Observa-se, assim, a necessidade da discussão de uma metodologia que propicie a construção de itens em um nível SOLO 5 da TS por meio de IME.

O método de questões de múltipla escolha é muito utilizado nos processos de avaliação das escolas médicas. As avaliações baseadas em IME, quando bem elaboradas, possuem a capacidade de avaliar os discentes nos níveis superiores do conhecimento, tornando essa tarefa desafiadora<sup>33</sup>. A introdução de casos clínicos nos testes de múltipla escolha enriquece a qualidade desse processo de avaliação para a mensuração do raciocínio clínico<sup>34</sup>. Verificamos isso no TPU2022-1, em que

muitos itens exigiram um raciocínio mais complexo por meio da resolução de casos clínicos.

A análise psicométrica do IME de uma avaliação é muito importante para a comprovação de sua qualidade. O IME deve ser avaliado para verificar sua validade e confiabilidade<sup>35</sup>. Torna-se muito relevante examinar a razão da escolha de uma opção de resposta em detrimento de outras e o motivo pelo qual a alternativa mais respondida não é a correta<sup>36</sup>.

Verificamos um satisfatório resultado relacionado ao IDF, com uma porcentagem menor de itens muito fáceis e muito difíceis (13,3%), com a maioria dos itens (86,7%) dentro de uma faixa aceitável para um processo de avaliação. A maioria dos IME encontrou-se em um nível insatisfatório em relação ao ID, não possibilitando uma diferenciação dos alunos de melhor e de pior desempenho, o que sugere a revisão de um número significativo de itens. Isso pode trazer questionamentos quanto a interpretações e conclusões baseadas em uma prova com essa particularidade.

Alguns fatores podem afetar o ID de um item, como ambivalência da pergunta, excessiva dificuldade ou facilidade para a sua solução, presença de tópicos não discutidos no currículo, má preparação dos alunos, motivação dos discentes, número de questões, tempo para a resolução e fatores ambientais, como temperatura, ruído e ventilação<sup>37</sup>.

Não se identificou correlação entre os parâmetros TCT (IDF e ID) e as categorias da TS. Também não se identificou correlação entre as categorias SOLO que representam a AS e a AP aos parâmetros da TCT.

Hamamoto Filho et al.<sup>38</sup> demonstraram, em um Teste de Progresso aplicado a dez escolas médicas brasileiras, que itens relacionados com altos níveis taxonômicos alcançaram uma melhor *performance* em ID.

Quando se analisam as questões pela TS, verifica-se que o grau de complexidade não é determinante do seu nível de dificuldade<sup>30,39</sup>. O grau de dificuldade das questões é avaliado por métodos estatísticos, ao passo que o nível de complexidade pode ser estabelecido por metodologias qualitativas, como a TS<sup>10</sup>. Esse dado demonstra que a qualidade de uma avaliação deve levar em consideração, além dos dados psicométricos, as habilidades cognitivas envolvidas para a resolução dos itens, ou seja, o nível de complexidade cognitiva exigida, dado não avaliado pelas metodologias psicométricas comumente empregadas.

Para um melhor desempenho dos processos de avaliação, como o TP, torna-se muito importante o preparo do corpo docente para a execução dos seus itens. Isso requer um programa de educação continuada docente nas escolas médicas para o desenvolvimento de itens de avaliação de qualidade<sup>40</sup>. As taxonomias de aprendizagem podem colaborar de maneira considerável para a criação de avaliações bem planejadas,

contribuindo para um ambiente eficaz de aprendizagem<sup>24</sup>.

Disponemos de um número grande de técnicas de avaliação. Nenhum método isolado deve ser empregado nos processos de avaliação das escolas médicas<sup>41</sup>. Para que haja uma satisfatória comprovação da *performance* dos alunos, é indispensável uma associação de técnicas<sup>42,43</sup>.

Encontramos na literatura poucos estudos que analisam avaliações externas à TS. Mol et al.<sup>44</sup> fizeram uma revisão de estudos brasileiros que utilizam a TS e encontraram 14 trabalhos, dos quais dez são artigos e quatro são dissertações. Nenhuma tese que abordasse a TS foi encontrada. Somente dois estudos foram relacionados com o ensino superior e nenhum na área do ensino médico. Ferreira et al.<sup>45</sup> fizeram um levantamento das teses e dissertações defendidas no Brasil que utilizaram a TS. Encontraram 12 trabalhos, a maioria na área de ciências exatas e nenhum relacionado ao ensino médico.

Algumas limitações do presente estudo devem ser consideradas. A primeira se refere ao fato de o estudo ter compreendido apenas uma prova do TP, de uma específica escola médica. A análise de um número maior de provas, incluindo outras escolas de medicina, poderá trazer maior subsídio para uma interpretação mais adequada dos dados. O fato de a categorização de todos os itens ter sido feita por um único avaliador pode também ser um fator limitante, no entanto é importante observar a tentativa de padronização prévia dessa análise por três avaliadores com uma elevada taxa de concordância, o que favorece a aplicabilidade dessa metodologia de análise no dia a dia de uma escola, onde o professor muitas vezes não dispõe de uma equipe de avaliadores disponíveis.

No entanto, este estudo apresenta contribuições relevantes para o processo de reflexão e desenvolvimento de avaliações no ensino médico. Demonstra a necessidade de equilibrar as metodologias de análises para além das psicométricas, incluindo aquelas que consideram os processos cognitivos envolvidos para a resolução das questões. Demonstra que a inclusão de casos clínicos favorece a análise da AP, devendo ser considerada a necessidade de buscar estratégias para elaboração de itens com boa capacidade discriminatória e equilíbrio em termos de dificuldade. O estudo apresenta ainda uma metodologia para a categorização da complexidade cognitiva dos IME que deverá ser testada em novos estudos.

Os dados deste estudo podem contribuir para a inclusão de um grupo de análise que considere, além dos dados estatísticos da psicometria, suas características em termos de complexidade cognitiva.

## CONCLUSÃO

A análise do TPU2022-1 nos possibilitou investigar características dessa ferramenta, útil na avaliação do

desempenho cognitivo longitudinal dos estudantes. Verificamos um equilíbrio entre a AS e a AP quando analisados todos os itens. No entanto, não observamos esse equilíbrio nas áreas de clínica médica e saúde coletiva. Verificamos uma relação direta entre os níveis de AP da TS e IME compostos por casos clínicos. Constatamos uma boa relação para o IDF, mas não para o ID, o que sugere a necessidade de uma análise crítica dos itens. Não verificamos diferenças estatisticamente significativa entre as categorias SOLO que representam a AS e a AP quanto às médias do IDF e do ID, indicando que essas metodologias avaliam diferentes particularidades dos IME.

Como o TP é um valioso instrumento de avaliação do ensino, deve ser incentivado nas escolas médicas. Ferramentas taxonômicas, como a TS, podem auxiliar de maneira significativa a realização das atividades avaliativas, de modo a conciliar as avaliações ao currículo, possibilitar a realização de provas adequadas ao nível de aprendizagem desejável e favorecer a progressividade do ensino. A qualificação do corpo docente para a realização de atividades avaliativas é necessária. Oficinas para a orientação na formulação de itens devem ser ofertadas aos docentes.

Estudos complementares são essenciais para crescer a consistência do uso dessa ferramenta taxonômica no ensino médico.

## CONTRIBUIÇÃO DOS AUTORES

Pedro Paulo Trindade Resende participou da elaboração do projeto de pesquisa, da revisão bibliográfica, da análise dos dados, da discussão dos resultados e da produção e revisão do manuscrito. Alexandre de Araújo Pereira participou da elaboração do projeto de pesquisa, da análise dos dados, da discussão dos resultados e da produção e revisão do manuscrito. José Maria Peixoto participou da elaboração do projeto de pesquisa, da discussão dos resultados e da produção e revisão do manuscrito.

## CONFLITO DE INTERESSES

Declaramos não haver conflito de interesses.

## FINANCIAMENTO

Declaramos não haver financiamento.

## LINK DE ACESSO AOS ITENS DO TPU2022-1

[https://docs.google.com/document/d/1V3SM\\_cLmPY-gfWnprkZfRw16QK0C-WjxR/edit?usp=drive\\_link&oui-d=103188252538609047663&rtpof=true&sd=true](https://docs.google.com/document/d/1V3SM_cLmPY-gfWnprkZfRw16QK0C-WjxR/edit?usp=drive_link&oui-d=103188252538609047663&rtpof=true&sd=true)

## REFERÊNCIAS

1. Champlain AFC. Setting and maintaining standards in multiple-choice examinations: guide supplement 37.2 – Viewpoint. *Med Teach*. 2010;32:436-7.
2. O'Shaughnessy SM, Joyce P. Summative and formative assessment in medicine: the experience of an anaesthesia trainee. *Int J Higher Educ*. 2015;4(2):198-206.
3. Ferris H, O'Flynn, D. Assessment in medical education: what are we trying to achieve? *Int J Higher Educ*. 2015;4(2):139-44.
4. Prashanti E, Ramnarayan K. Ten maxims of formative assessment. *Adv Physiol Educ*. 2019;43:99-102.
5. Hadie SNH. The application of learning taxonomy in anatomy assessment in medical school. *Education in Medicine Journal*. 2018;10(1):13-23.
6. Reberti AG, Monfredini NH, Ferreira Filho OF, Andrade DF, Pinheiro CEA, Silva, JC. Teste de Progresso na escola médica: uma revisão sistemática acerca da literatura. *Rev Bras Educ Med*. 2020;44(1):1-9.
7. Pinheiro OL, Spadella MA, Moreira HM, Ribeiro ZMT, Guimarães APC, Almeida Filho OM, et al. Teste de Progresso: uma ferramenta avaliativa para a gestão acadêmica. *Rev Bras Educ Med*. 2015;39(1):68-78.
8. Baldim YL, Vicente C A O, Arcuri MB. O teste de progresso sob a visão do discente. *Revista da Faculdade de Medicina de Teresópolis*. 2018;2(1):41-54.
9. Condé FN. Análise empírica de itens. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais; 2001.
10. Pereira VCAS. Aplicação da Taxonomia SOLO na análise da qualidade da avaliação: validação do método analítico por aplicação aos exames nacionais de Matemática A entre 2006 e 2014. Covilhã; 2019.
11. Vilarinho APL. Uma proposta de análise de desempenho dos estudantes e de valorização da primeira fase da OBMEP [dissertação]. Brasília: Universidade de Brasília; 2015.
12. Aripin MA, Hamzah R, Setya P, Hisham MHM, Mohd Ishar MI. Unveiling a new taxonomy in education field. *International Journal of Evaluation and Research in Education*. 2020;9(3):524-30.
13. Amantes A, Oliveira E. A construção e o uso de sistemas de categorias para avaliar o entendimento dos estudantes. *Revista Ensaio*. 2012;24(2):61-79.
14. Biggs J. Study Process Questionnaire Manual. Student. Approaches to Learning and Studying. Australian Council for Educational Research; 1987.
15. Biggs J, Collis K. Evaluating the quality of learning: the SOLO Taxonomy. New York: Academic Press; 1982.
16. Biggs, J. Calidade del Aprendizage Universitário. Narcea; 2006.
17. Rossi GZ, Fischer JMS, Rocha SR, Casalecchi GA, Avó LSR, Germano CMR. Abordagens de aprendizado e sua correlação com ambiente educacional e características individuais em escola médica. *Rev Bras Educ Med*. 2021;45(3):1-11.
18. Filipe MAER. A Taxonomia SOLO nos Exames Nacionais de Matemática – 9º ano [dissertação]. Lisboa: Universidade Nova de Lisboa; 2011.
19. Yurtyapan MI, Yilmaz GK. An investigation of the geometric thinking levels of middle school mathematics preservice teachers according to SOLO Taxonomy: "Social Distance Problems". *Participatory Educational Research*. 2021;8(3):188-209.
20. Jimoyiannis A. Using SOLO taxonomy to explore students' mental models of the programming variable and the assignment statement. *Themes in Science & Technology Education*. 2011;4(2):53-74.
21. Ceia M. A taxonomia SOLO e os níveis de Van Hiele. Encontro de Investigação em Educação Matemática. Coimbra: Sociedade Portuguesa de Ciências da Educação; 2002.
22. Korkmaz F, Unsal S. Analysis of attainments and evaluation questions in sociology curriculum according to the SOLO Taxonomy. *Eurasian Journal of Educational Research*. 2017;69:75-92.
23. Soobard R, Rannikmäe M. Examining Curriculum Related Progress Using a Context-Based Test Instrument – a comparison of Estonian Grade 10 and 11 Students. *Science Education International*. 2015;26(3):263-83.

24. Hadie SNH. The application of learning taxonomy in anatomy assessment in medical school. *Education in Medicine Journal*. 2018;10(1):13-23
25. Peixoto JM, Santos SME, Faria RMD, Moura AS. Processos de desenvolvimento do raciocínio clínico em estudantes de Medicina. *Rev Bras Educ Med*. 2018;42(1):75-83.
26. Thampy H, Willert E, Ramani S. assessing clinical reasoning: targeting the higher levels of the pyramid. *J Gen Intern Med*. 2019;34(8):1631-6.
27. Van Der Vleuten CPM, Schuwirth LWT. Assessment in the context of problem-based learning. *Adv Health Sci Educ*. 2019;24:903-16.
28. Aragão JCS, Almeida LS. Raciocínio clínico e pensamento crítico: desenvolvimento na educação médica. *Rev Estud Investig Psicol Educ*. 2017;(12):12-20.
29. Jaiswal P. Using constructive alignment to foster teaching learning processes. *English Language Teaching*. 2019;12(6):10-23.
30. Scully D. Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research, and Evaluation*. 2017;17(4):1-12.
31. Sprecher EA. Back to the chalkboard: lessons in scaffolding using SOLO taxonomy from school teachers for university educators. *Psychology Teaching Review*. 2019;25(2):95-102.
32. Kim M, Patel RA, Uchizono JA, Beck L. Incorporation of Bloom's Taxonomy into multiple-choice examination questions for a pharmacotherapeutics course. *Am J Pharma Educ*. 2012;76(6):1-8.
33. Vegi VAK, Sudhakar PV, Bhimarasetty DM, Pamarth K, Edara L, Kutikuppala LVS, et al. Multiple-choice questions in assessment: perceptions of medical students from low-resource setting. *J Educ Health Promot*. 2022;11:1-6
34. Modi JN, Anshu, Gupta P, Singh T. Teaching and assessing clinical reasoning skills. *Indian Pediatr*. 2015;52:787-94.
35. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality Multiple Choice Questions (MCQs) from an Assessment of Medical Students of Ahmedabad, Gujarat. *Indian Journal of Community Medicine*. 2014;39(1):17-20.
36. Piton-Gonçalves J, Almeida AM. Análise da dificuldade e da discriminação de itens de Matemática do Enem. *Revista Eletrônica de Matemática*. 2018;4(2):38-53.
37. Bhattacherjee S, Mukherjee A, Bhandari K, Rout AJ. Evaluation of Multiple-Choice Questions by Item Analysis, from an Online Internal Assessment of 6th Semester Medical Students in a Rural Medical College, West Bengal. *Indian Journal of Community Medicine*. 2022;47(1):92-5.
38. Hamamoto Filho PT, Silva E, Ribeiro ZMT, Hafner MLMB, Cecilio-Fernandes D, Bicudo AM. Relationships between Bloom's taxonomy, judges' estimation of item difficulty and psychometric properties of items from a progress test: a prospective observational study. *Sao Paulo Med J*. 2020;138(1):33-9.
39. Hattie JAC, Brown GTL. Cognitive processes in aSTTle: the SOLO taxonomy. *aSTTle Technical Report*. Wellington: Ministry of Education; 2004.
40. Vanderbilt AA, Feldman M, Wood IK. Assessment in undergraduate medical education: a review of course exams. *Med Educ Online*. 2013;18(1):1-5.
41. Shah SSH, Munir TA, Sabir M, Tipu SA. Psychometric analysis of MCQs used in assessing the students at entrance to a medical college. *Ann King Edw Med Univ*. 2012;18(3):296-9.
42. Fowell SL, Bligh JG. Recent developments in assessing medical students. *Postgrad Med J*. 1998;74:18-24.
43. Khan MUZ, Aljarallah BM. Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for assessing the cognitive skills of undergraduate medical students. *Int J Health Sci*. 2011;5(1):39-43.
44. Mol SM, Matos ASM. Uma análise sobre a taxonomia solo: aplicações na avaliação educacional. *Est Aval Educ*. 2019;30(75):722-47.
45. Ferreira FFG, Rocha MLPC. A Taxonomia SOLO nas teses e dissertações defendidas em programas de pós-graduação no Brasil. *Revista de Matemática, Ensino e Cultura*. 2020;15:32-46.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.