

<https://doi.org/10.18222/ea.v33.7533>

QUALIDADE DOS ITENS DE MÚLTIPLA ESCOLHA UTILIZADOS EM UM TESTE DE PROGRESSO

 ELEN CRISTINA NASCIMENTO DE ARAUJO TOMBI^I

 CRISTINA ZUKOWSKY-TAVARES^{II}

 IRANÍ FERREIRA-GERAB^{III}

^I Universidade Anhembi Morumbi (UAM), São Paulo-SP, Brasil; elentombi@gmail.com

^{II} Centro Universitário Adventista de São Paulo (Unasp), São Paulo-SP, Brasil; cristina.tavares@unasp.edu.br

^{III} Universidade Federal de São Paulo (Unifesp), São Paulo-SP, Brasil; irani.gerab@gmail.com

RESUMO

Este artigo objetiva analisar e discutir a qualidade dos itens de múltipla escolha (IME) elaborados para um Teste de Progresso, por meio da análise quali-quantitativa de 100 itens. Nos resultados há predomínio da dimensão do conhecimento conceitual e o domínio cognitivo sintetizado pelo verbo “entender”, analisado pela Taxonomia de Bloom Revisada, e 44% dos itens estão em acordo com as diretrizes técnicas para elaboração de IME. Pela análise da Teoria Clássica dos Testes, KR20 foi 0,835, dificuldade (P) 0,496 e índice de discriminação 0,240. O teste apresentou bom índice de confiabilidade, com dificuldade dos itens adequada. Deve ser chamada a atenção ao índice de discriminação e à formulação do enunciado no intuito de melhorar esses parâmetros para os itens de forma geral.

PALAVRAS-CHAVE ENSINO SUPERIOR • AVALIAÇÃO DA APRENDIZAGEM • TESTES.

CALIDAD DE LOS ÍTEMS DE ELECCIÓN MÚLTIPLE UTILIZADOS EN UNA PRUEBA DE PROGRESO

RESUMEN

Este artículo tiene el propósito de analizar y discutir la calidad de los ítems de elección múltiple elaborados para una Prueba de Progreso, por medio del análisis cuali-cuantitativo de 100 rubros. En los resultados predomina la dimensión del conocimiento conceptual y el dominio cognitivo sintetizado en el verbo “entender”, analizado por la Taxonomía de Bloom Revisada, y el 44% de los ítems está de acuerdo con las directrices técnicas para elaborar rubros de elección múltiple. Por medio del análisis de la Teoría Clásica de las Pruebas, KR20 fue de 0,835, dificultad (P) 0,496 e índice de discriminación 0,240. El test presentó un buen índice de confiabilidad, con dificultad adecuada de los ítems. Hay que llamar la atención al índice de discriminación y a la formulación del enunciado con el objetivo de mejorar dichos parámetros para los ítems en general.

PALABRAS CLAVE EDUCACIÓN SUPERIOR • EVALUACIÓN DEL APRENDIZAJE • TEST DE ELECCIÓN MÚLTIPLE.

QUALITY OF MULTIPLE-CHOICE ITEMS USED IN A PROGRESS TEST

ABSTRACT

This article aims to analyze and discuss the quality of multiple-choice items prepared for a Progress Test, through the quali-quantitative analysis of 100 items. The results show a predominance of the dimension of conceptual knowledge and the cognitive domain synthesized by the verb “understand”, analyzed by the Revised Bloom Taxonomy, and 44% of the items are in accordance with the technical guidelines for the elaboration of multiple-choice items. According to Classical Test Theory, KR20 was 0.835, difficulty (P) 0.496, and discrimination index 0.240. The test presented a good reliability index, with suitable item difficulty. Special attention should be given to the discrimination index and the writing of the questions in order to improve these parameters for the items in general.

KEYWORDS HIGHER EDUCATION • LEARNING ASSESSMENT • MULTIPLE CHOICE TEST.

INTRODUÇÃO

A avaliação da aprendizagem é parte do processo educacional em que, de acordo com a natureza do conteúdo, diferentes instrumentos podem ser utilizados para subsidiar a tomada de decisão.

Os testes são instrumentos utilizados na avaliação da aprendizagem e podem ser elaborados com itens de múltipla escolha (IME). Vários tipos de testes elaborados com IME são usados em muitos países para diversas finalidades, entre elas, vestibulares, provas de títulos e exames finais (BOLLELA *et al.*, 2018).

Desde as décadas de 1960 e 1970, principalmente nas escolas médicas de formação, o Teste de Progresso (TP) tem servido principalmente como ferramenta de autoavaliação e monitoramento do progresso individual de estudantes de ensino superior nas aprendizagens e domínios curricularmente estabelecidos. O TP é uma prova longitudinal que afere o ganho de conhecimentos ao longo da formação, em que os resultados são acionados na elaboração de *feedback* aos estudantes, aos professores e aos gestores. Permite à instituição de ensino diagnosticar deficiências no processo ensino-aprendizagem de conceitos e habilidades ao longo da estrutura curricular de um curso (HAMAMOTO FILHO; BICUDO, 2020; BICUDO *et al.*, 2019).

De periodicidade semestral ou anual, o TP não é um teste padronizado a ser aplicado a estudantes do ensino superior, é construído de acordo com as especificidades da organização curricular de determinada área de formação ou curso. O teste é aplicado simultaneamente a todos os estudantes do curso e é composto por IME elaborados internamente pela equipe docente do curso ou por especialistas externos. Tem como objetivo medir a aprendizagem teórico-conceitual e aplicada de um concluinte em sua área de formação, independe da organização curricular. Apresenta itens relacionados a temáticas previstas no plano de cada semestre de determinado curso, o que favorece o monitoramento do progresso nas aprendizagens ao longo dos anos. Dessa forma, o esperado seria que graduandos dos últimos semestres apresentassem melhor proficiência que os ingressantes. Com um bom número de respondentes, o teste pode também fornecer informações pertinentes ao currículo de formação e ao planejamento específico de disciplinas ou módulos de ensino.

Nesse sentido, este artigo objetiva analisar e discutir a qualidade dos IME elaborados por docentes de um curso de graduação, com base em resultados de uma pesquisa realizada em uma dissertação de mestrado profissional (TOMBI, 2019). Para o desenvolvimento dessa pesquisa, foram utilizados os IME elaborados para um TP aplicado a estudantes de um curso superior em estética, de uma universidade paulistana, no ano de 2017. Embora tenha sido realizada com itens e resultados obtidos na aplicação de um TP, não é intenção deste artigo discutir o teste, mas a qualidade de IME.

Os IME foram elaborados por docentes do curso, no intuito de mensurar domínios cognitivos conceituais e procedimentais sobre conhecimentos esperados ao longo do curso. O teste foi composto por 100 itens de múltipla escolha, distribuídos em 30 itens referentes aos componentes curriculares do núcleo básico na área da saúde e 70 itens relacionados a temas específicos da área de atuação profissional.

O teste foi respondido por 582 graduandos, correspondendo a 95% dos estudantes matriculados no curso, do primeiro ao último ano. A realização do teste foi obrigatória, e os estudantes realizaram o TP no mesmo dia e horário, em ambiente virtual, cada estudante em um computador.

QUALIDADE DOS ITENS DE MÚLTIPLA ESCOLHA

Os IME podem aferir diferentes níveis do domínio cognitivo e o uso desse tipo de item apresenta vantagens que estão relacionadas à forma prática e objetiva de correção dos testes, possibilitando uma devolutiva rápida do resultado. Ainda, permitem a padronização do instrumento para a aplicação da prova a muitos estudantes (BOLLELA *et al.*, 2018; VANDERBILT; FELDMAN; WOOD, 2013).

Os IME caracterizaram-se por apresentar um enunciado estruturado com uma pergunta ou frase a ser completada seguida de alternativas, que são opções de respostas a serem selecionadas. A resposta correta é considerada o gabarito e as alternativas incorretas são chamadas de distratores (NEDEAU-CAYO *et al.*, 2013).

As instruções do enunciado podem ser antecedidas de um texto-base ou caso clínico, sendo esse último denominado vinheta, na literatura (BOLLELA *et al.*, 2018; PANIAGUA *et al.*, 2016).

A literatura apresenta diretrizes que orientam a redação de IME, relacionadas ao modo de escrever um item, no intuito de direcionar a construção e a formatação para evitar falhas comuns que podem ser cometidas na elaboração de um IME e também ao uso de taxonomia para redigir itens capazes de avaliar diferentes níveis do processo cognitivo (HALADYNA; DOWNING; RODRIGUEZ, 2002; PANIAGUA *et al.*, 2016; TARRANT *et al.*, 2006).

No âmbito educacional, na área da saúde, a taxonomia mais utilizada na prática de pesquisa e na discussão do domínio cognitivo, bem como para direcionar a elaboração e a avaliação dos itens, é a Taxonomia de Bloom (TB) (SCULLY, 2017).

Originalmente, a TB apresentava seis categorias para o domínio cognitivo, estruturadas de maneira hierárquica, com característica linear, e os objetivos de aprendizagem eram compreendidos de maneira crescente, ou seja, do processo cognitivo simples ao complexo (FERRAZ; BELHOT, 2010).

Na revisão da TB, em 2001, a representação do nível cognitivo por substantivos designados de maneira hierárquica – como conhecimento, compreensão, aplicação, análise, síntese e avaliação – foi atualizada. O substantivo “conhecimento”,

que representava o primeiro nível cognitivo, passou a ser uma nova dimensão classificada em quatro tipos: conhecimento efetivo ou factual, conhecimento conceitual, conhecimento procedimental e conhecimento metacognitivo. Os substantivos utilizados para designar os demais aspectos do domínio cognitivo foram atualizados para formas verbais e renomeados ordenadamente: lembrar, entender, aplicar, analisar, avaliar e criar. A revisão e a atualização promoveram uma relação entre o tipo de conhecimento e o domínio cognitivo, configurando um aspecto bidimensional à TB. Isso possibilitou a interação dos níveis cognitivos com os diferentes tipos de conhecimento, sem a obrigatoriedade de elaborar os objetivos, pensando em processos cognitivos de maneira crescente e linear (FERRAZ; BELHOT, 2010; KRATHWOHL, 2002; SCULLY, 2017).

Um teste deve conter itens bem redigidos, que busquem aferir os objetivos de aprendizagem propostos, apresentando resultados fidedignos. Embora essa seja uma premissa, a elaboração de bons IME não é uma tarefa simples. É necessário treinamento, orientação, tempo e disponibilidade dos autores e elaboradores de itens.

Os testes apresentam características que são medidas pela psicometria, como a Teoria Clássica dos Testes (TCT), o que torna possível compreender a confiabilidade do instrumento e a qualidade dos itens que o compõem, conhecendo a discriminação e a dificuldade dos IME.

A TCT vincula os resultados da análise do teste ao grupo que o realizou. Quando um teste composto por um conjunto de itens recebe baixa pontuação, esse resultado pode ser decorrente da dificuldade dos itens ou de baixa proficiência dos respondentes (SCHUWIRTH; VAN DER VLEUTEN, 2011).

Um elemento importante para aferir a qualidade de um teste é a confiabilidade, que indica sua consistência interna, ou seja, quão bom seria o resultado de um teste se repetido (TAVAKOL; DENNICK, 2012, p. 246).

Além da inter-relação entre os itens do teste, o índice de dificuldade (P) pode ser aferido pelo nível de dificuldade do item – quão fácil ou difícil foi para o grupo de respondentes. Esse cálculo é feito com base nos respondentes que acertaram ou erraram o item (PANIAGUA *et al.*, 2016).

Um item pode diferenciar os estudantes mais proficientes, que apresentaram maior pontuação no teste, em relação aos estudantes menos proficientes, que obtiveram menor pontuação, por meio do índice de discriminação (ID) do item (SARTES; SOUZA-FORMIGONI, 2013).

Borgatto e Andrade (2012) relatam que o item capaz de discriminar a proficiência dos estudantes é um bom item. Tavakol e Dennick (2011) corroboram essa ideia afirmando que quando os estudantes com melhor desempenho acertam a resposta do item e os estudantes com proficiência inferior erram o mesmo item, ele pode ser considerado adequado.

O ID oferece informações a respeito de quanto o item pode diferenciar os estudantes que tiveram bons resultados dos estudantes que tiveram pior desempenho no teste. Supõe-se que um bom item seja respondido corretamente pelo grupo de estudantes com melhor desempenho no teste e espera-se que o grupo com pior desempenho acerte pouco o mesmo item. Quanto maior a diferença entre a quantidade de estudantes que acertaram em relação aos estudantes que erraram o item, melhor sua capacidade de discriminação (BORGATTO; ANDRADE, 2012). No entanto, se o item considerado de difícil complexidade é gabaritado por estudantes de baixa e alta proficiência, pode-se inferir que o item não tem poder de discriminação.

Realizar análises após a aplicação dos testes contribui para que a qualidade e a confiabilidade dos instrumentos de avaliação possam ser aperfeiçoados e para que itens sejam identificados quanto a sua discriminação, dificuldade e necessidade de reformulação (TAVAKOL; DENNICK, 2011).

METODOLOGIA

Trata-se de uma pesquisa descritiva com abordagens quantitativa e qualitativa a respeito dos IME. A produção dos dados ocorreu por meio de três etapas: análise taxonômica bidimensional, análise da qualidade técnica na formulação de IME e análise psicométrica (Figura 1).

FIGURA 1 - Produção dos dados da pesquisa



Fonte: Elaboração das autoras.

Análise taxonômica dos IME

A análise taxonômica dos IME foi feita em relação à dimensão do conhecimento, classificado em: factual, conceitual, procedimental ou metacognitivo. E a dimensão do processo cognitivo foi classificada de maneira crescente em: lembrar, entender, aplicar, analisar, avaliar e criar. Para essa etapa, foi elaborada uma tabela bidimensional (Quadro 1), de acordo com a Taxonomia de Bloom Revisada (TBR), abordada

por Ferraz e Belhot (2010), Krathwohl (2002) e Scully (2017), com informações adicionais para auxiliar o processo de análise e classificação de cada item.

Nos espaços em branco foram adicionados os números das questões que se enquadravam nos critérios de classificação.

QUADRO 1 – Análise bidimensional dos itens de múltipla escolha pela Taxonomia de Bloom Revisada

DIMENSÃO DO CONHECIMENTO	DIMENSÃO DO PROCESSO COGNITIVO					
	1. Lembrar	2. Entender	3. Aplicar	4. Analisar	5. Avaliar	6. Criar
	Recuperando conhecimento relevante da memória de longo prazo. 1.1 Reconhecendo 1.2 Recordando	Determinar o significado de mensagens instrucionais, incluindo comunicação oral, escrita e gráfica. 2.1 Interpretação 2.2 Exemplificando 2.3 Classificação 2.4 Resumindo 2.5 Inferenciamento 2.6 Comparando 2.7 Explicando	Executar ou usar um procedimento em uma determinada situação. 3.1 Execução 3.2 Implementando	Partindo o material em suas partes constituintes e detectando como as peças se relacionam umas com as outras e com uma estrutura ou propósito geral. 4.1 Diferenciando 4.2 Organizando 4.3 Atribuindo	Fazendo julgamentos com base em critérios e padrões. 5.1 Verificação 5.2 Criticar	Juntar elementos para formar um todo novo e coerente ou fazer um produto original. 6.1 Gerando 6.2 Planejamento 6.3 Produzindo
A - Factual/ Efetivo						
B - Conceitual						
C - Processual/ Procedimental						
D - Metacognitivo						

Fonte: Adaptado de Krathwohl (2002, p. 214-215).

Nota: **A - Efetivo/Factual:** Os elementos básicos que os estudantes devem conhecer para se familiarizar com uma disciplina ou resolver problemas nela. Conhecimento da terminologia; Conhecimento de detalhes e elementos específicos. **B - Conceitual:** As inter-relações entre os elementos básicos dentro de uma estrutura maior que permitem que eles funcionem juntos. Conhecimento de classificações e categorias; Conhecimento de princípios e generalizações; Conhecimento de teorias, modelos e estruturas. **C - Processual/ Procedimental:** Como fazer alguma coisa; métodos de investigação e critérios para o uso de habilidades, algoritmos, técnicas e métodos. Conhecimento de habilidades e algoritmos específicos do assunto; Conhecimento de técnicas e métodos específicos do assunto; Conhecimento dos critérios para determinar quando usar procedimentos apropriados. **D - Metacognitivo:** Conhecimento da cognição em geral, bem como consciência e conhecimento da própria cognição. Conhecimento estratégico; Conhecimento sobre tarefas cognitivas, incluindo conhecimento contextual e condicional apropriado; Autoconhecimento.

Análise da qualidade técnica na formulação de IME

Para essa fase foi elaborado um formulário com 17 tópicos (Quadro 2) relativos a possíveis falhas na elaboração de IME baseados nos estudos de Haladyna, Downing e Rodriguez (2002, p. 312), Paniagua *et al.* (2016, p. 15-51) e Tarrant *et al.* (2006, p. 670-671).

Os tópicos foram formatados em formulário eletrônico no Google Forms®, em questões de resposta dicotômica “sim ou não”.

Cada avaliador tinha acesso ao *link* para responder individualmente o formulário de maneira independente.

QUADRO 2 – Tópicos para análise da qualidade técnica dos itens do TP 2017

1. Opção longa ou complicada nos distratores.
2. Termos vagos nas opções dos distratores ou no enunciado.
3. A linguagem ou a estrutura das opções é heterogênea.
4. “Nenhuma das anteriores” é usada como uma opção.
5. O enunciado tem frase negativa.
6. Há dicas gramaticais porque um ou mais distratores não acompanha o enunciado gramaticalmente.
7. Uma opção (ou mais de uma) tem dica por estar em par ou ser exaustiva.
8. Alguma opção (ou mais de uma) pode ser eliminada, porque outra cobre todos os resultados possíveis.
9. Alguma opção apresenta termos absolutos, como “sempre” ou “nunca”.
10. A resposta correta é mais longa, mais específica ou mais completa do que as outras opções.
11. Uma palavra ou frase está presente tanto no enunciado como na resposta correta.
12. Apresenta convergência (a resposta correta inclui a maioria dos elementos em comum com as outras opções).
13. A orientação no enunciado é confusa.
14. O conteúdo da questão está relacionado com outra questão.
15. Há erros gramaticais, quanto à pontuação e ortografia.
16. Alguma opção é absurda ou implausível.
17. Há informações desnecessárias no enunciado.

Fonte: Elaboração das autoras.

Tanto a tabela de análise taxonômica quanto o formulário de análise da qualidade técnica dos itens foram refinados a partir de uma análise piloto realizada com uma amostra de 21 IME do TP. Os dois instrumentos foram aprimorados e aplicados aos 100 IME do TP 2017.

Essas etapas foram realizadas por três avaliadores: um avaliador docente do curso de graduação em estética, um avaliador pós-doutor e especialista em avaliação educacional e um avaliador doutor e pesquisador na área de ensino em ciências da saúde.

Inicialmente, cada avaliador analisou os 100 IME de maneira independente. Após essa primeira análise individual, o resultado da análise de cada avaliador foi confrontado em encontro presencial entre os três avaliadores e as discordâncias foram discutidas em forma de painel, onde cada item em divergência era apresentado e discutido para buscar um consenso entre os avaliadores (BOLLELA *et al.*, 2018).

As divergências entre os avaliadores que analisaram individualmente os IME do teste de estética foram de 3,64% (62 tópicos). Esse resultado foi menor do que os resultados apresentados nos trabalhos de Tarrant *et al.* (2006) e Tarrant e Ware (2008). Em ambos os trabalhos, os itens foram analisados por quatro avaliadores e as discordâncias na análise individual foram, respectivamente, de 15 e 13%. O consenso entre os avaliadores após as reuniões de discussão sobre os itens foi comum em todos os trabalhos.

Análise psicométrica do TP

A análise psicométrica por meio da TCT foi feita pelo *software* IteMan®, com os dados do TP 2017. Para a elaboração do banco de dados para as respostas dos estudantes, em cada questão atribuiu-se aos acertos o valor 1, aos erros 0 e à ausência de resposta a letra N. O teste foi analisado quanto aos parâmetros a seguir.

A. Confiabilidade

Diferentes estimativas podem ser aplicadas utilizando a pontuação obtida no teste para determinar sua confiabilidade e calcular a relação entre os itens e a homogeneidade do teste, entre elas, a Confiabilidade Kuder-Richardson (KR-20), aplicada neste estudo (TAVAKOL; DENNICK, 2011).

Valores de KR-20 com escores entre 0,8 e 0,9 indicam boa confiabilidade. Valores acima de 0,9 indicam que o teste foi elaborado com itens homogêneos e valores menores que 0,8 indicam baixa confiabilidade (ABDULGHANI *et al.*, 2017).

B. Índice de dificuldade (P)

O valor do P está relacionado à quantidade de estudantes que responderam corretamente ao item e pode ser “representado como um percentual ou proporção”. Por exemplo, se em uma amostra de 100 estudantes o item foi respondido corretamente por 40 estudantes, pode ser representado tanto por 40% como por 0,40 (PANIAGUA *et al.*, 2016, p. 25).

Para o P do item, as referências foram: abaixo de 0,3, item difícil; entre 0,3 e 0,8, item de dificuldade média; acima de 0,8, item fácil (TAVAKOL; DENNICK, 2011).

C. Índice de discriminação (ID)

A estimativa utilizada para determinar o ID foi o coeficiente de correlação, como a correlação bisserial e a correlação ponto-bisserial, que indica a “[...] relação entre um item específico (correto ou incorreto) em um teste e a pontuação total do teste” (TAVAKOL; DENNICK, 2011, p. 453).

Os seguintes parâmetros foram considerados: valores menores que 0,15, discriminação ruim; entre 0,15 e 0,24, discriminação marginal; entre 0,25 e 0,34, discriminação boa; acima de 0,35, discriminação excelente (HINGORJO; JALEEL, 2012; LINNETTE D’SA; VISBAL-DIONALDO, 2017).

ASPECTOS ÉTICOS

O estudo foi autorizado pela coordenação do curso, cujo documento permanece arquivado com a pesquisadora. O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Paulo/Hospital São Paulo em 14/05/2018, sob o parecer CEP n. 3579190418.

RESULTADOS E DISCUSSÃO

Na análise qualitativa dos itens de múltipla escolha do TP 2017, pela TBR, predominaram IME que solicitavam a dimensão do conhecimento conceitual, caracterizado por itens que solicitavam lembrar, entender ou analisar um conceito. O processo cognitivo mais observado foi representado pela categoria “entender”. Os processos cognitivos “aplicar” e “analisar” não foram expressivos (Tabela 1). Itens relacionados aos processos cognitivos “avaliar” e “criar”, bem como ao conhecimento meta-cognitivo, não foram observados.

TABELA 1 - Análise taxonômica: dimensão do processo cognitivo em relação à dimensão do conhecimento analisados nos Itens de Múltipla Escolha do Teste de Progresso 2017

DIMENSÃO DO PROCESSO COGNITIVO	DIMENSÃO DO CONHECIMENTO			
	FACTUAL	CONCEITUAL	PROCEDIMENTAL	METACOGNITIVO
38% “lembrar”	16%	15%	7%	-
56% “entender”	-	49%	7%	-
2% “aplicar”	-	-	2%	-
4% “analisar”	-	3%	1%	-

Fonte: Elaboração das autoras.

Esses resultados apresentam valores um pouco maiores para a classificação dos itens em relação ao conhecimento conceitual (67%) e processo cognitivo categorizado como “entender” (56%), quando comparados aos resultados do estudo realizado por Silva e Martins (2014), que analisaram 105 itens de física do Exame Nacional do Ensino Médio (Enem), aplicados entre 2009 e 2013, por meio da TBR. Os autores apresentaram em sua pesquisa que 56% dos itens eram referentes ao conhecimento conceitual e 49% referiam-se ao processo cognitivo “entender”.

Outros trabalhos avaliam o domínio cognitivo utilizando a TB original e fazem adaptações à classificação taxonômica para analisar os IME. Os resultados de Tarrant *et al.* (2006) sobre a análise dos níveis cognitivos dos IME, utilizados em instrumentos de avaliação no curso superior em enfermagem, apontaram que mais de 90% dos itens foram escritos solicitando baixos níveis cognitivos. Os autores avaliaram e classificaram os itens em dois níveis cognitivos, o primeiro nível referia-se à recordação de fatos ou compreensão, sendo considerados níveis inferiores, já os itens que solicitavam aplicação e análise foram considerados superiores.

Em outra pesquisa, realizada com IME utilizados em um curso de enfermagem, também se observou que grande quantidade (93%) de itens solicitavam a recordação de fatos (NEDEAU-CAYO *et al.*, 2013).

Tariq *et al.* (2017) avaliaram os níveis cognitivos dos IME utilizados em exames de avaliação em farmacologia. Os autores utilizaram a TB modificada, classificando os domínios cognitivos em três níveis, e identificaram que 65% dos itens estavam avaliando a recordação de informações, enquanto 19% avaliavam a interpretação dos dados e 16% avaliavam as habilidades de resolução de problemas. Considerando que a recordação pode ser relacionada à dimensão do processo cognitivo “lembrar” na TBR, observa-se que a maioria dos itens solicitavam os níveis mais baixos do domínio cognitivo.

Independentemente do tipo de análise, com base na TB Original ou Revisada, observa-se que os trabalhos de Neddeau-Cayo *et al.* (2013), Silva e Martins (2014), Tariq *et al.* (2017) e Tarrant *et al.* (2006) apresentam um grande número de itens que avaliam os domínios cognitivos mais simples. Esse fato foi semelhante ao observado nos IME analisados no presente estudo, em que predominaram as categorias “lembrar” e “entender”.

Para Abdulghani *et al.* (2017, p. 2), a qualidade do item relaciona-se com a capacidade de testar o pensamento crítico e níveis elevados do conhecimento. Segundo os autores, “[...] itens de alta qualidade, no processo de avaliação, interrogam a aprendizagem profunda dos estudantes”, o que pode ser relacionado aos níveis mais altos do domínio cognitivo da TB, como aplicação, análise e síntese do conhecimento adquirido.

Corroborando essa discussão, Scully (2017) relata que a TB, apesar de apresentar uma organização dos domínios cognitivos de maneira crescente, não faz referência quanto ao pensamento de ordem superior ou inferior. O autor ainda infere que pesquisas sobre o tema apresentam diferentes definições para a classificação do pensamento de ordem superior ou inferior e adota em sua pesquisa a classificação dos itens que avaliam o pensamento superior para todos os níveis a partir da compreensão, que na TBR corresponde ao processo cognitivo “entender”.

Javaeed (2018) considera que há uma discussão em relação à qualidade do IME e sua capacidade de avaliar o pensamento de ordem superior, relacionado com os domínios cognitivos superiores da TB. Segundo o autor, um item pode ter essa característica, mas a falta de treinamento dos elaboradores de itens contribui para que haja uma apresentação de grande número de itens com domínios cognitivos inferiores. O autor complementa que para avaliar os estudantes, em especial da área da saúde, os IME devem ser construídos com a intenção de avaliar níveis mais altos do domínio cognitivo e sugere, para esse fim, a elaboração de itens relacionados à atividade profissional para os quais os estudantes estão se preparando.

Na análise dos itens em relação às diretrizes técnicas, para cada IME foi avaliada a possibilidade de 17 falhas ou incorreções técnicas, o que corresponde ao total de 1.700 para os 100 itens investigados. O presente estudo identificou 88 falhas para os

100 itens, o que corresponde a 5,17% de falhas na elaboração técnica (Tabela 2). Dos itens observados, 35% apresentaram uma falha, 21% mais de uma falha e 44% apresentaram-se tecnicamente adequados a partir dos critérios delimitados nessa análise.

TABELA 2 - Falhas identificadas nos Itens de Múltipla Escolha do Teste de Progresso 2017

FALHA	QUANTIDADE	(%)
1. Opção longa ou complicada nos distratores.	5	(5,7)
2. Termos vagos nas opções dos distratores ou no enunciado.	1	(1,1)
3. A linguagem ou a estrutura das opções é heterogênea.	2	(2,3)
4. "Nenhuma das anteriores" é usado como uma opção.	3	(3,4)
5. O enunciado tem frase negativa.	2	(2,3)
6. Há dicas gramaticais porque um ou mais distratores não acompanha o enunciado gramaticalmente.	-	-
7. Uma opção (ou mais de uma) tem dica por estar em par ou ser exaustiva.	8	(9,1)
8. Alguma opção (ou mais de uma) pode ser eliminada porque outra cobre todos os resultados possíveis.	-	-
9. Alguma opção apresenta termos absolutos, como "sempre" ou "nunca".	2	(2,3)
10. A resposta correta é mais longa, mais específica ou mais completa do que as outras opções.	3	(3,4)
11. Uma palavra ou frase está presente tanto no enunciado como na resposta correta.	2	(2,3)
12. Apresenta convergência (a resposta correta inclui a maioria dos elementos em comum com as outras opções).	1	(1,1)
13. A orientação no enunciado é confusa.	3	(3,4)
14. O conteúdo da questão está relacionado com outra questão.	-	-
15. Há erros gramaticais, quanto à pontuação e ortografia.	12	(13,6)
16. Alguma opção é absurda ou implausível.	17	(19,3)
17. Há informações desnecessárias no enunciado.	27	(30,7)

Fonte: Elaboração das autoras com base nos dados da pesquisa.

Embora o número de inadequações técnicas aos critérios delimitados tenha sido menor, algumas falhas ocorreram com maior frequência relacionadas a "erros na gramática, quanto à pontuação e ortografia", "presença de informações desnecessárias no enunciado" e "distratores implausíveis", solicitando atenção e cuidado quanto a esses quesitos.

Tarrant e Ware (2008) analisaram IME de testes aplicados no curso enfermagem. Os autores identificaram falhas em 47,3% itens. As falhas correspondiam a enunciado com frase negativa em 13%, informações desnecessárias no enunciado em 12%, distratores implausíveis em 10%, maior detalhe na opção correta em 8,7% e palavras repetidas no enunciado e na resposta correta em 6% dos itens.

Na investigação de Tariq *et al.* (2017), os tipos mais comuns de falhas foram os distratores implausíveis em 20%, opção correta com mais detalhes em 18%, termos vagos em 9,85% e termos absolutos em 9,09% dos itens.

Algumas das falhas observadas por Tarrant *et al.* (2006) são em relação a termos vagos (1,7%), termos absolutos (1,7%), repetição de palavra no enunciado e na resposta correta (4%), enunciado com frase negativa (6,9%), uso de nenhum dos anteriores (1%) e informação desnecessária no enunciado (6,1%).

É possível notar que o resultado do estudo dos IME do TP 2017 apresenta algumas falhas que estão semelhantes aos resultados da literatura (TARIQ *et al.*, 2017; TARRANT *et al.*, 2006; TARRANT; WARE, 2008). Com relação à falha que caracteriza informações desnecessárias no enunciado, o resultado está acima do observado nesses trabalhos.

No estudo realizado por Khan *et al.* (2013) com itens de provas aplicadas em três anos no curso de medicina, houve diminuição das falhas técnicas nas diretrizes para elaboração de itens. Os autores relatam que o número de falhas diminuiu de 67%, em 2009, para 36,3% em 2010 e para 21% em 2011. Com a apresentação desses resultados, os autores ressaltam a importância da formação docente na elaboração dos itens e sugerem que a elaboração de itens seja programada e durante a preparação das aulas. Consideram que a mesma importância concedida ao preparo da aula deve ser dada à elaboração dos itens que irão avaliar a aprendizagem. Os autores complementam que a elaboração dos IME muito próxima da data de avaliação dificulta a revisão do item e a possível melhora de sua qualidade por falta de tempo.

Em um estudo que analisou falhas técnicas nas diretrizes para redação de itens e a complexidade dos itens do curso de medicina veterinária, observou-se que a maior parte dos itens (61,6%) estava em níveis inferiores do domínio cognitivo. Os autores relatam que esse resultado é acima do ideal e sugerem que itens com níveis mais altos do processo cognitivo compõem 60% da prova em avaliações para treinamento clínico. As falhas mais encontradas foram distratores implausíveis (22,9%), resposta correta mais longa (20,6%) e enunciado com frase negativa (11,8%). Quanto ao resultado da falha “distratores implausíveis”, foi semelhante ao resultado encontrado nos itens do teste de estética (RUSH *et al.*, 2016).

O impacto de incorreções técnicas na elaboração de itens em relação ao desempenho dos estudantes foi discutido em outro trabalho na literatura, em que quase metade dos itens apresentou falhas (47%). A pesquisa comparou os resultados de itens com falhas em relação aos itens sem falhas e mostrou que a dificuldade é 10% maior em itens com falhas. Os autores concluíram que há necessidade de mais pesquisas sobre a influência das falhas no desempenho dos estudantes e que é importante realizar treinamento de docentes para melhorar a qualidade dos itens (TARRANT; WARE, 2008).

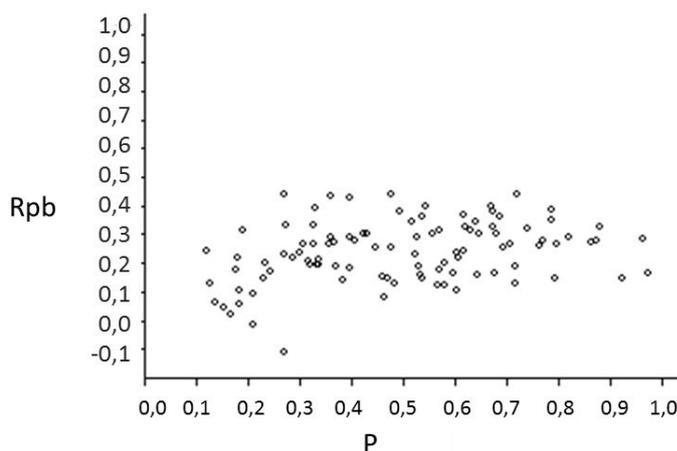
Este trabalho apresentou pequena quantidade de falhas técnicas (5,17%), porém ocorreram em pouco mais da metade dos itens do teste. Isso chama a atenção para

as falhas mais frequentes, que incluem os erros gramaticais, quanto à pontuação e ortografia, presença de informações desnecessárias no enunciado e distratores implausíveis. Essas duas últimas falhas também aparecem repetidamente na literatura, demonstrando maior necessidade de atenção e cuidado quanto a esses quesitos. Nesse sentido, o treinamento docente pode contribuir para melhorar redação e qualidade técnica dos itens e diminuir ainda mais as falhas. O anexo apresenta alguns exemplos de itens utilizados no teste.

A fim de contribuir para a compreensão da qualidade dos itens e complementar os dados da pesquisa de forma quantitativa, a confiabilidade do teste, realizada pela análise psicométrica do TP feita pelo *software* Iteman®, teve como resultado o coeficiente alfa (KR-20) de 0,835, demonstrando que houve boa consistência interna entre os itens do teste e que eles se correlacionam. Tratando do objetivo da pesquisa, que foi avaliar a qualidade dos itens e não a qualidade do TP em si, a análise quantitativa pelo programa de dados foi realizada na busca de informações sobre a qualidade dos itens; dessa forma, não será discutida neste artigo a confiabilidade do teste. Este artigo abordará os índices correlacionados aos IME, o P e ID do item.

Os valores médios para dificuldade dos itens (P) foi 0,496 e para a discriminação (Rpbis – correlação ponto-biserial) foi 0,240. Os resultados caracterizaram o teste com dificuldade média e discriminação positiva, demonstrado na Figura 2 pelo gráfico de dispersão de P (dificuldade) por Rpbis (discriminação) para os itens do TP 2017, produzido a partir da análise gerada pelo *software* Iteman® (ITEMAN, 2019).

FIGURA 2 - Gráfico de dispersão de P (dificuldade) por Rpbis (discriminação) para os itens do TP 2017



Fonte: Dados da pesquisa.

O ID foi bom ou excelente em 50% dos IME do TP aplicado em 2017, porém 29% dos itens apresentaram ID marginal e 21% apresentaram índice ruim, indicando a necessidade de revisão no intuito de melhorar o ID. O teste foi composto por 72% de IME com média dificuldade.

Valores P indicando média dificuldade e ID classificados como bons ou excelentes são considerados características de um bom IME, o cruzamento desses dois índices (P e ID) identificam os itens ideais no teste. O TP 2017 apresentou 42% de itens nessa categoria (Tabela 3).

TABELA 3 - Itens de Múltipla Escolha considerados ideais no Teste de Progresso 2017

REFERÊNCIAS	% DE ITENS COM P ENTRE 0,3 E 0,8	ID MÉDIA E DP	P MÉDIA E DP
ID excelente	14	0,40 ± 0,03	0,58 ± 0,15
ID bom	28	0,29 ± 0,03	0,54 ± 0,15
Total	42		

Fonte: Elaboração das autoras com base nos dados da pesquisa.

Os itens com ID excelente apresentaram a média de ID e desvio padrão (DP) de $0,40 \pm 0,03$, e a média de P e desvio padrão (DP) foi $0,58 \pm 0,15$. Para os itens com ID bom, esses valores foram $0,29 \pm 0,03$ e $0,54 \pm 0,15$, respectivamente.

O resultado da combinação de P e ID no presente trabalho são semelhantes aos encontrados no trabalho de Linnette D'Sa e Visbal-Dionaldo (2017). As autoras relataram que 47% dos itens foram considerados ideais, quando combinados os valores obtidos para os itens de P com média dificuldade com ID bom ou excelente. No trabalho das autoras, a média de P foi 0,67 e a média do ID foi 0,44.

O cruzamento dos dados da análise quantitativa (análise psicométrica) com os dados da análise qualitativa (taxonômica e de critérios técnicos) permitiu relacionar ID, P e dimensão do processo cognitivo dos IME. Essa relação foi feita comparando os itens que não apresentaram falhas técnicas nas diretrizes que orientam a redação de itens (itens sem falha) e tiveram bom desempenho na avaliação psicométrica com bom ID e dificuldade média com os itens que tiveram desempenho ruim no ID pela análise psicométrica e apresentavam incorreções técnicas (Tabela 4).

Observa-se que o processo cognitivo dos itens sem inadequações técnicas e com bom desempenho psicométrico (melhores itens) é semelhante ao dos itens com falhas e ID inferior (piores itens).

Itens elaborados para avaliar domínios cognitivos superiores têm melhores índices de discriminação e avaliam de forma mais abrangente os estudantes (RUSH *et al.*, 2016).

A correção de falhas e a reformulação de IME utilizados em testes para reaplicação em novos testes melhora a qualidade dos itens em relação ao ID (ALI; RUIT, 2015).

TABELA 4 – Comparação das análises qualitativa e quantitativa entre os melhores e os piores Itens de Múltipla Escolha do Teste de Progresso 2017

	ITENS SEM FALHAS E BOM DESEMPENHO PSICOMÉTRICO	ITENS COM FALHAS E ID RUIM
Nº de itens	14	9
Nº de falhas identificadas	-	14 (9,15%)
Falhas observadas (identificadas por seu número)	-	4 5 7 15 16 17
Dimensão do processo cognitivo	Lembrar (28,57%) Entender (64,28%) Aplicar (7,15%)	Lembrar (66,66%) Entender (33,33%)
ID (média e DP)	0,32 ± 0,05	0,08 ± 0,05
P (média e DP)	0,50 ± 0,17	0,33 ± 0,22

Fonte: Elaboração das autoras com base nos dados da pesquisa.

Em suma, os resultados do presente estudo destacam a relevância de uma análise reflexiva dos resultados após a administração de testes, provas e outros instrumentos de avaliação. Conhecendo os resultados da prova e a qualidade dos IME utilizados é possível discutir meios para melhorá-los, revelando preocupação em ter instrumentos de avaliação confiáveis, que apresentem resultados fidedignos e capazes de discriminar a proficiência dos estudantes.

Nesse sentido, o preparo do corpo docente para elaboração de IME, a TB identificando as dimensões do conhecimento e níveis cognitivos, bem como as possíveis falhas que podem ser evitadas na elaboração de itens, podem contribuir para melhorar a redação e qualidade técnica dos itens e produzir resultados ainda mais confiáveis para as reflexões necessárias ao complexo processo de ensino e aprendizagem na educação superior (ZUKOWSKY-TAVARES, 2010; ZUKOWSKY-TAVARES; LIMEIRA; RUIZ-MORENO, 2019).

CONCLUSÃO

A análise taxonômica do teste evidencia predominância de questões que privilegiam domínio cognitivo no âmbito do conhecimento, com destaque para “entender”, embora seja importante considerar domínios mais complexos para melhor compreensão da aprendizagem teórico-conceitual dos estudantes.

É importante oferecer aos docentes ações formativas para que possam aprimorar a formulação de testes de melhor qualidade em seus enunciados e distratores/gabarito.

O TP aplicado em 2017 apresentou boa consistência interna, com bom índice de confiabilidade. O P dos IME foi adequado para a prova. Com relação ao ID dos itens,

é necessário observar a necessidade de melhorar esse parâmetro para os itens de forma geral.

IME são utilizados com frequência como instrumentos de avaliação no ensino superior e a elaboração de itens deve ser um tema abordado e discutido no processo de formação contínua dos docentes. Planejamento, treinamento e preparação dos docentes são quesitos necessários para a elaboração de itens de boa qualidade, bem como uma análise mais bem fundamentada das providências necessárias para a superação de desafios individuais e coletivos, com aprendizagens e competências a serem desenvolvidas na efetivação dos compromissos curriculares e de formação em cada área do conhecimento.

Após a finalização deste estudo, algumas inquietações surgiram e podem direcionar novas pesquisas para compreender a perspectiva dos docentes no processo de elaboração de itens, a fim de entender como são elaborados os IME para compor diferentes instrumentos de avaliação, quais são o conhecimento e a preparação dos docentes em relação às diretrizes técnicas para a elaboração de IME ou que ações de desenvolvimento profissional docente podem contribuir para melhorar a formação docente em avaliação.

REFERÊNCIAS

ABDULGHANI, Hamza Mohammad *et al.* Effectiveness of longitudinal faculty development programs on MCQs items writing skills: A follow-up study. *PLoS One*, v. 12, e0185895, Oct. 2017. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0185895>. Acesso em: 8 jan. 2018.

ALI, Syed Haris; RUIT, Kenneth G. The impact of item flaws, testing at low cognitive level, and low distractor functioning on multiple-choice question quality. *Perspectives on Medical Education*, v. 4, n. 5, p. 244-251, Sept./Oct. 2015. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4602009/?report=reader>. Acesso em: 23 jan. 2022.

BICUDO, Angélica Maria *et al.* Teste de Progresso em Consórcios para Todas as Escolas Médicas do Brasil. *Revista Brasileira de Educação Médica*, Brasília, v. 43, n. 4, p. 151-156, out./dez. 2019.

BOLLELA, Valdes Roberto *et al.* 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*, Brasília, v. 42, n. 4, p. 74-85, out./dez. 2018.

BORGATTO, Adriano Ferreti; ANDRADE, Dalton Francisco de. Análise clássica de testes com diferentes graus de dificuldade. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 52, p. 146-156, maio/ago., 2012.

FERRAZ, Ana Paula do Carmo Marcheti; BELHOT, Renato Vairo. Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gestão & Produção*, São Carlos, SP, v. 17, n. 2, p. 421-431, jun. 2010.

HALADYNA, Thomas M.; DOWNING, Steven M.; RODRIGUEZ, Michael C. A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, Londres, v. 15, n. 3, p. 309-334, jul. 2002.

HAMAMOTO FILHO, Pedro Tadao; 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*, Brasília, v. 44, n. 1, e018, 2020.

HINGORJO, Mozaffer Rahim; JALEEL, Farhan. Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *Journal of Pakistan Medical Association*, Karachi, v. 62, n. 2, p. 142-147, Feb. 2012.

ITEMAN: Software for Classical Analysis Copyright © 2013 – Assessment Systems Corporation. *Relatório Prova de Progresso Estética*, abr. 2019.

JAVAEED, Arslaan. Assessment of higher ordered thinking in medical education: multiple choice questions and modified essay questions. *MedEdPublish*, United Kingdom, v. 7, n. 128, p. 11-18, June 2018. Disponível em: <https://mededpublish.org/articles/7-128>. Acesso em: 23 jan. 2022.

KHAN, Humaira Fayyaz *et al.* Identification of technical item flaws leads to improvement of the quality of single best Multiple-Choice Questions. *Pakistan Journal of Medical Sciences*, Pakistan, v. 29, n. 3, p. 715-718, May/June 2013.

KRATHWOHL, David R. A Revision of Bloom's Taxonomy: an overview. *Theory into Practice*, London, v. 31, n. 4, p. 212-218, Mar./Apr. 2002.

LINNETTE D'SA, Juliana; VISBAL-DIONALDO, Maria Liza. Analysis of multiple-choice questions: item difficulty, discrimination index and distractor efficiency. *International Journal of Nursing Education*, India, v. 9, n. 3, p. 109-114, July 2017.

NEDEAU-CAYO, Rosemarie *et al.* Assessment of item-writing flaws in multiple-choice questions. *Journal for Nurses in Professional Development*, USA, v. 29, n. 2, p. 52-57, Mar./Apr. 2013.

PANIAGUA, Miguel A. *et al.* *Construindo as perguntas do teste escrito para ciências básicas e clínicas*. NBME: Philadelphia, 2016.

RUSH, Bonnie R. *et al.* The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Medical Education*, United Kingdom, v. 16, n. 250, p. 1-10, Sept. 2016.

SARTES, Laisa Marcocela Andreoli; SOUZA-FORMIGONI, Maria Lucia Oliveira de. Avanços na psicometria: da Teoria Clássica dos Testes à Teoria de Resposta ao Item. *Psicologia: Reflexão e Crítica*, Porto Alegre, v. 26, n. 2, p. 241-250, jul. 2013.

SCHUWIRTH, Lambert W. T.; VAN DER VLEUTEN, Cees P. M. General overview of the theories used in assessment: AMEE Guide No. 57. *Medical Teacher*, United Kingdom, v. 33, p. 783-797, Sept. 2011.

SCULLY, Darina. Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment Research & Evaluation*, v. 22, n. 4, p. 1-13, May 2017. Disponível em: <http://pareonline.net/getvn.asp?v=22&n=4>. Acesso em: 24 jan. 2022.

SILVA, Vailton Afonso da; MARTINS, Maria Inês. Análise de questões de física do ENEM pela taxonomia de Bloom revisada. *Ensaio: Pesquisa em Educação e Ciências*, Belo Horizonte, v. 16, n. 3, p. 189-202, set./dez. 2014.

TARIQ, Saba *et al.* Evaluation of cognitive levels and item writing flaws in medical pharmacology internal assessment examinations. *Pakistan Journal of Medical Sciences*, Pakistan, v. 33, n. 4, p. 866-870, July/Aug. 2017.

TARRANT, Marie; WARE, James. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education*, Oxford, v. 42, n. 2, p. 198-206, Feb. 2008.

TARRANT, Marie *et al.* The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education Today*, USA, v. 26, n. 8, p. 662-671, Dec. 2006.

TAVAKOL, Mohsen; DENNICK, Reg. Post-examination analysis of objective tests. *Medical Teacher*, United Kingdom, v. 33, n. 6, p. 447-458, May 2011.

TAVAKOL, Mohsen; DENNICK, Reg. Post-examination interpretation of objective test data: monitoring and improving the quality of high-stakes examinations – a commentary on two AMEE Guides. *Medical Teacher*, United Kingdom, v. 34, n. 3, p. 245-248, Feb. 2012.

TOMBI, Elen Cristina Nascimento de Araújo. *Estudo dos itens de múltipla escolha de um teste de progresso aplicado em um curso de Estética de uma universidade paulistana*. 2019. 89f. Dissertação (Mestrado Profissional) – Programa de Pós-Graduação em Ciências da Saúde, Universidade Federal de São Paulo, São Paulo, 2019.

VANDERBILT, Allison A.; FELDMAN, Moshe; WOOD, Isaac K. Assessment in undergraduate medical education: a review of course exams. *Medical Education Online*, v. 18, n.1, p. 1-5, Mar. 2013.

ZUKOWSKY-TAVARES, Cristina. Formação em avaliação como um caminho para a profissionalização docente. *Revista Lusófona de Educação*, Lisboa, v. 16, n. 16, p. 59-74, ago. 2010.

ZUKOWSKY-TAVARES, Cristina; LIMEIRA, Polyana de Castro; RUIZ-MORENO, Lídia. O portfólio e a construção de saberes docentes na pós-graduação em saúde. *Pro-Posições*, Campinas, SP, v. 30, e20170181, 2019.

ANEXO - ALGUNS ITENS DO TESTE DE PROGRESSO

IME – Morfologia Humana

O Esqueleto é tradicionalmente dividido em esqueleto axial e apendicular. Identifique nas alternativas a seguir o único osso que não faz parte do esqueleto axial.

- a) Rádio.
- b) Crânio.
- c) Coluna vertebral.
- d) Costelas.
- e) Hióideo.

IME – Recursos Manuais em Estética

Para a aplicação da drenagem linfática manual (DLM), são necessários diferentes requisitos, a

integridade do sistema linfático, o correto entendimento entre as diferenças de pressão e um profissional que conheça e realize a técnica de forma correta. Leia as afirmativas sobre a DLM.

- I. Contração muscular, pulsação de artérias vizinhas e movimentos respiratórios e articulares são alguns fatores determinantes para o deslocamento da linfa.
 - II. A DLM proporciona a diminuição da pressão intersticial e do líquido do interstício para o capilar.
 - III. O atrito e pressão corretos promovem o movimento dos filamentos de ancoragem com o consequente afastamento da sobreposição celular, o que promove a passagem de líquidos do interstício para o interior do capilar linfático.
 - IV. O linfangion é o segmento de vasos linfáticos entre dois pares de válvulas. Unidade funcional do sistema linfático, responsável pela propulsão da linfa.
- a) São corretas apenas as afirmativas I, III e IV.
 - b) São corretas apenas as afirmativas I, II e III.
 - c) São corretas apenas as afirmativas II e III.
 - d) São corretas apenas as afirmativas I, II e IV.
 - e) Todas as afirmativas são corretas.

IME - Eletroestética

Em relação à intensidade da corrente elétrica, assinale a alternativa correta:

- a) É a quantidade de carga por unidade de tempo que passa por um condutor e sua intensidade é medida em μA ou mA para os tratamentos estéticos.
- b) É a quantidade de carga por unidade de tempo que passa por um condutor e sua unidade é medida em A para os tratamentos estéticos.
- c) É a quantidade de carga por unidade de tempo que passa por um condutor e sua medida é em KHz para os tratamentos estéticos.
- d) É a quantidade de carga por unidade de tempo que passa por um condutor e sua unidade de medida é em ms para os tratamentos estéticos.
- e) É o número de vezes que um determinado fenômeno se repete durante a passagem do tempo, sejam horas, minutos ou segundos.

COMO CITAR ESTE ARTIGO

TOMBI, Elen Cristina Nascimento de Araujo; ZUKOWSKY-TAVARES, Cristina; FERREIRA-GERAB, Iraní. Qualidade dos itens de múltipla escolha utilizados em um Teste de Progresso. *Estudos em Avaliação Educacional*, São Paulo, v. 33, e07533, 2022. DOI: <https://doi.org/10.18222/ae.v33.7533>

Recebido em: 24 JUNHO 2020

Aprovado para publicação em: 20 OUTUBRO 2021



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.