Int. J. Morphol.,
38(5):1179-1183, 2020.

Development and Validation of a Scale to Measure the Gross Anatomy Preparedness of Medical Graduates

Desarrollo y Validación de una Escala para Medir la Preparación Anatómica General de los Graduados Médicos

Amjad T. Shatarat[1]; Sara S. Elmegarhi[1]; Justin Z. Amarin[2]; Darwish H. Badran[1]; Mouna M. Al-Sa'ad[3]; Nader M. Albsoul[4] & Raed N. Al-Taher[4]

SHATARAT, A. T.; ELMEGARHI, S. S.; AMARIN, J. Z.; BADRAN, D. H.; AL-SA'AD, M. M.; ALBSOUL, N. M. & AL-TAHER, R. N. Development and validation of a scale to measure the gross anatomy preparedness of medical graduates. *Int. J. Morphol., 38(5)*:1179-1183, 2020.

SUMMARY: Postgraduate refresher courses may address deficiencies in the gross anatomy preparedness of medical graduates. However, the literature does not offer a method to identify such deficiencies. Our aim is to develop and validate a scale to measure the gross anatomy preparedness of medical graduates. First, we defined gross anatomy preparedness (the construct) as "the benchmark of personal ability in gross anatomy against the standard required for clinical practice." Next, we conducted a literature search for extant items related to our definition. To develop our scale, we grouped the items under three headings: proficiency, preference, and pertinence. Finally, we constructed item-specific response anchors to "Likertize" the items. We recruited experts to validate the content and conducted cognitive interviews to validate the response process. To evaluate the internal structure and reliability of the scale, we invited a purposive sample of 120 surgery residents to complete the scale and explored the results of the pilot test using data reduction and reliability analysis. A total of 77 surgery residents completed the scale. Varimax-rotated principal components analysis revealed three components with eigenvalues greater than one, and the components explained 64 % of the total variance. The rotated solution was consistent with the original structure of the questionnaire. The components, which represented the proficiency, preference, and pertinence item sets, explained 25 %, 23 %, and 16 %, respectively, of the total variance. Cronbach's α coefficients for the item sets were 0.72, 0.71, and 0.61, respectively. We developed and validated a scale to measure the gross anatomy preparedness of medical graduates. In addition, we offer conceptual guidelines to help users interpret the results of the scale. Outcome data are required to substantiate the predictive validity of the scale.

KEY WORDS Anatomy; Medical Education; Graduate Medical Education; Surveys and Questionnaires; Preliminary Data.

INTRODUCTION

Gross anatomy is a cornerstone of medicine. Medicine comprises a diverse slate of core and elective courses. In medical education, courses are designed to improve clinical performance. Unfortunately, the contribution of a course to clinical performance is difficult to assess. Claims of importance (for example, the opening sentence) are not always substantiated. Intersubjective metrics are a popular and convenient form of evidence. Indeed, many authors have studied the clinical importance of gross anatomy using intersubjective metrics (Pabst & Rothkötter, 1997; Cottam, 1999; Ahmed *et al*., 2010; Smith & Mathias, 2011; Lazarus *et al*., 2012).

Pabst & Rothkötter studied the perceptions—regarding medical education—of 109 physicians who had completed postgraduate training in Germany. The sample rated gross anatomy higher in clinical relevance than 27 out of 28 other courses. More recently, Smith & Mathias surveyed 140 alumni of Southampton Medical School. The majority of the sample had found gross anatomy education "invaluable". In another study, the clinical importance of gross anatomy was stressed across the spectrum of medical education (Ahmed *et al*.). However, gross anatomy is more relevant to some specialties than others. Cottam sampled 801 physicians in leadership positions from four types of

[1] Department of Anatomy and Histology, School of Medicine, The University of Jordan, Amman, Jordan.
[2] School of Medicine, The University of Jordan, Amman, Jordan.
[3] Department of Special Surgery, School of Medicine, The University of Jordan, Amman, Jordan.
[4] Department of General Surgery, School of Medicine, The University of Jordan, Amman, Jordan.

residency programs. Gross anatomy was rated first in importance by 89 % of diagnostic radiologists, 46 % of general surgeons, 16 % of emergency physicians, and 10 % of family physicians. In another study, the importance of gross anatomy was rated highest by surgeons (Lazarus *et al*.).

One theme recurs in the literature of gross anatomy education: the standards are dwindling. According to Waterston & Stewart (2005), the late Professor David Sinclair is the fons et origo of the theme. Many authors have since criticized the gross anatomy preparedness of medical trainees (Waterston & Stewart; Bergman *et al*., 2008).

Postgraduate refresher courses may resolve deficiencies in gross anatomy preparedness (Cottam). However, two challenges hamper the process: deficiencies must be identifiable, and progress must be trackable. A valid scale that measures the salient dimensions of gross anatomy preparedness addresses both challenges. Our aim is to develop and validate a scale to measure the gross anatomy preparedness of medical graduates.

MATERIAL AND METHOD

Setting. The Jordan University Hospital is the inaugural teaching hospital in Jordan (bed capacity of 550–600). The hospital was annexed to The University of Jordan in 1975 and currently employs some 250 residents (120 of whom are surgery residents). The hospital houses 24 operating theatres in which an estimated 25,000 surgeries are performed annually.

Scale development. Unobservable constructs may be measured using self-administered scales. Using a guide by the Association for Medical Education in Europe, we set out to develop a self-administered scale to measure the gross anatomy preparedness of medical graduates (Artino Jr. *et al*., 2014). Three members of our study team completed the development process in collaboration. First, we defined gross anatomy preparedness (the construct) as "the benchmark of personal ability in gross anatomy against the standard required for clinical practice." Next, we conducted a literature search for extant items related to our definition. After collecting the items, we attempted to capture the salient dimensions of the construct by grouping similar items under emergent headings—namely, proficiency, preference, and pertinence. We consolidated redundant items and discarded irrelevant items. Twelve items remained. Finally, we constructed item-specific response anchors to "Likertize" the remaining items.

Scale validation. We conducted content and response process validation based on a guide by the Association for Medical Education in Europe (Artino Jr. *et al*.). To evaluate the content, we recruited three content experts and asked them to comment on the representativeness, clarity, relevance, and distribution of each item. The experts rejected three items and suggested minor modifications to the other items. We discarded the three items and applied the modifications. After refining the scale, we invited a purposive sample of all 120 surgery residents employed by the Jordan University Hospital to participate in our study. To evaluate the response process, we conducted cognitive interviews with 20 participants by retrospective probing after they had completed the scale. We did not identify any respondent errors, so we did not modify the scale further. To evaluate the internal structure and reliability of the scale, we asked the remaining participants to complete the scale. We explored the results of the pilot test using data reduction and reliability analysis.

Data analysis. We manually entered the data into the IBM SPSS Statistics Data Editor and used the software package (version 23.0) to perform data analysis. We pooled the data from cognitive interviews and pilot testing and performed varimax-rotated principal components analysis of the nine items. We generated a correlation matrix to assess the assumption of linearity and set the eligibility criterion for each variable to a minimum of one bivariate correlation ($r \geq 0.3$). We computed the individual and overall Kaiser–Meyer–Olkin measures and ran Bartlett's test of sphericity to assess the assumption of sampling adequacy. We set the Kaiser–Meyer–Olkin measure threshold to 0.5 and the significance threshold for Bartlett's test of sphericity to $P=0.05$. The data met both assumptions. We used the eigenvalue-one criterion to choose the number of components to retain. To measure the internal consistency of the components, we used Cronbach's α. We present numerical data according to the recommendations of Cole (2015). We report component-based scores as medians and ranges (the latter within parentheses).

Ethics. The Institutional Review Board of the Jordan University Hospital (Amman, Jordan) approved the study protocol and we obtained written informed consent from all participants.

RESULTS

A total of 77 surgery residents (response rate, 64 %) completed the nine-item scale. We report a summary of the specialty, residency training year, and alma mater of the residents in Table I. Varimax-rotated principal components

analysis revealed three components with an eigenvalue greater than one, and the components explained 64 % of the total variance. Each of the nine items loaded strongly on only one component and the rotated solution was consistent with the original structure of the questionnaire (Table II). Proficiency items loaded strongly on the first component and explained 25 % of the total variance (2.2 eigenvalues). Preference items loaded strongly on the second component and explained 23 % of the total variance (2.0 eigenvalues). Pertinence items loaded strongly on the third component and explained 16 % of the total variance (1.5 eigenvalues). Cronbach's α coefficients for the item sets were 0.72, 0.71, and 0.61, respectively. The median proficiency, preference, and pertinence scores of the sample were 3.5 (2.0–4.8), 4.3 (3.0–5.0), and 4.0 (2.0–5.0), respectively.

Table I. The specialty, residency training year, and alma mater of N=77 surgery residents.

Characteristic	***n* (%)**
Residency program	
General surgery	54 (70)
Orthopedic surgery	10 (13)
Neurosurgery	6 (8)
Urology	6 (8)
Otorhinolaryngology	1 (1)
Residency training year	
First	35 (46)
Second	13 (17)
Third	17 (22)
Fourth	4 (5)
Fifth	8 (10)
Alma mater	
The University of Jordan	35 (46)
Jordan University of Science and Technology	14 (18)
Mutah University	8 (10)
The Hashemite University	8 (10)
Islamic University of Gaza	5 (7)
Misr University of Science and Technology	3 (4)
October 6 University	1 (1)
Cairo University	1 (1)
Alzaiem Alazhari University	1 (1)
Alfaisal University	1 (1)

DISCUSSION

Gross anatomy preparedness is an undefined construct. Our definition of the construct is "the benchmark of personal ability in gross anatomy against the standard required for clinical practice." We developed a nine-item scale to measure three dimensions of the construct: proficiency, preference, and pertinence. We validated the content and response process, and we evaluated the internal structure and reliability of the scale.

Many authors have advocated for postgraduate refresher courses in gross anatomy (Cottam; Fitzgerald *et al*., 2008; Lazarus *et al*.). However, the decision to provide refresher training carries economic implications and must be balanced against the anticipated benefit. Indeed, Cottam showed that 71 % of residency program representatives do not feel that gross anatomy refresher training is necessary. Currently, the decision process is bereft of objective metrics. In response, we propose below conceptual guidelines to help users calculate a preparedness score and measure, by extension, the anticipated benefit of a refresher course.

Proficiency is a positive indicator of preparedness; a high score indicates a high degree of preparedness and vice versa. However, according to our definition of the construct, proficiency does not equate preparedness because the standard of proficiency required for clinical practice varies from

Table II. Rotated component matrix generated from varimax-rotated principal components analysis of 77 complete responses to the nine-item Gross Anatomy Preparedness of Medical Graduates (GAPMG) scale.

Item	Component		
	1	2	3
How satisfied are you with the gross anatomy training you received during medical school?	**0.7**	-0.2	0.05
How much information have you retained from the gross anatomy training you received during medical school?	**0.8**	-0.2	0.03
How is your ability to recall the gross anatomy training you received during medical school?	**0.7**	0.2	0.1
How difficult is it for you to apply your gross anatomy training from medical school in a clinical setting?	**0.8**	0.2	-0.07
In your opinion, how helpful would further gross anatomy training be?	-0.04	**0.8**	0.2
How interested are you in further gross anatomy training?	0.04	**0.9**	-0.07
In your opinion, to what extent would your clinical performance improve following further gross anatomy training?	-0.003	**0.7**	-0.2
In your opinion, how important is gross anatomy in a clinical setting?	-0.01	-0.02	**0.8**
How valuable is gross anatomy to your clinical performance?	0.1	-0.02	**0.8**

Note: Coefficients greater than 0.3 are emboldened

specialty to specialty. Preference and pertinence complement proficiency because they reflect the standard of proficiency required for clinical practice. Both preference and pertinence are negative indicators of preparedness; high scores indicate a low degree of preparedness and vice versa. For example, a group with a low proficiency score and high preference and pertinence scores reasonably stands to benefit from a refresher course more than a group with a high proficiency score and low preference and pertinence scores. The minimum and maximum dimension scores are 1.00 and 5.00, respectively. To reduce the unweighted parameters into one simple metric, we derived the following equation:, where

$$P_4 = \frac{9 + P_1 - P_2 - P_3}{12} \times 100\%$$

P_1, P_2, P_3 and P_4 represent the proficiency, preference, pertinence, and preparedness scores, respectively. The equation yields a preparedness score ranging from 0 % to 100 %. The preparedness score of our sample, for example, is 35 %. We present a scheme to translate the preparedness score into the anticipated benefit of a refresher course in Table III. Again, the method is conceptual and must be substantiated in future studies.

Table III. A conceptual scheme for the interpretation of the preparedness score.

Preparedness score	Anticipated benefit
<25 %	Major
≥25 % and <50 %	Moderate
≥50 % and <75 %	Minor
≥75 %	Minimal

From an administrative standpoint, the demand for refresher courses may be justified by a heterogeneous collection of evidence. For instance, historical data exhibit a worldwide downtrend in the time allocated to gross anatomy training throughout medical school (Fitzgerald *et al*.; Craig *et al*., 2010). The dwindling standards of gross anatomy education may compromise the safety of clinical practice. Indeed, a number of anatomists have argued that the modern physician is poorly equipped for safe clinical practice (Cahill *et al*., 2000; Older, 2004; Waterston & Stewart). Professor Harold Ellis went a step further by demonstrating an association between poor gross anatomy knowledge and increased medico-legal claims against surgeons in the United Kingdom (Ellis, 2002).

From an academic standpoint, the target population of refresher courses is likely receptive. Smith & Mathias reported that new medical graduates from Southampton Medical School "were concerned that there was still so much anatomy they did not know". In addition,Pabst & Rothkötter reported that a sample of new residency graduates was heavily in favor of postgraduate refresher training. Interestingly, success stories are documented in the literature. For instance, Ergül *et al*. (2002) documented the ability of two junior residents to identify nerves in the inguinal region during inguinal hernia surgery. Following a refresher course, the ability of the junior residents to correctly identify the nerves significantly improved. Cabrera *et al*. (2011) documented another success story: 19 radiation oncology residents from Duke University and the University of North Carolina were highly satisfied with a 20-module refresher course spanning two years. Upon further questioning, a subset of the sample unanimously agreed that the experience had improved their clinical performance. Therefore, evidence suggests that refresher courses are doubly valuable from the perspective of both the administrator and the learner.

The practical value of our study is limited because we did not collect evidence for predictive validity. Our recommendations are conceptual and must be appraised in the future using outcome data. In addition, according to our results, the reliability of the subscales is acceptable. However, the small number of items in each subscale compromises interpretability. For instance, the pertinence subscale consists of two items only. Cronbach's α coefficients may underestimate the true reliability of two-item subscales. Instead, the Spearman–Brown coefficient is more appropriate (Eisinga *et al*., 2013). However, our analysis yielded identical Cronbach's α and Spearman–Brown coefficients. Therefore, we reported the former to simplify our report. All in all, the choice of coefficient is perhaps tangential; the most appropriate approach to two-item subscales is the addition of items (Eisinga *et al*.). Therefore, the subscales may be improved by the addition of items in future studies.

In conclusion, we developed and validated a scale to measure the gross anatomy preparedness of medical graduates. In addition, we offer conceptual guidelines to help users interpret the results of the scale. Outcome data are required to substantiate the predictive validity of the scale.

SHATARAT, A. T.; ELMEGARHI, S. S.; AMARIN, J. Z.; BADRAN, D. H.; AL-SA'AD, M. M.; ALBSOUL, N. M. & AL-TAHER, R. N. Desarrollo y validación de una escala para medir la preparación anatómica general de los graduados médicos. *Int. J. Morphol., 38(5)*:1179-1183, 2020.

RESUMEN: Los cursos de actualización de posgrado pueden abordar las deficiencias en la preparación de la anatomía macroscópica de los graduados médicos. Sin embargo, la literatura no ofrece un método para identificar tales deficiencias. Nuestro objetivo fue desarrollar y validar una escala para medir la preparación anatómica general de los graduados médicos. Primero, definimos la preparación para la anatomía macroscópica (el constructo)

como "el punto de referencia de la capacidad personal en anatomía macroscópica frente al estándar requerido para la práctica clínica". A continuación, realizamos una búsqueda bibliográfica de elementos existentes relacionados con nuestra definición. Para desarrollar nuestra escala, agrupamos los ítems bajo tres encabezados: competencia, preferencia y pertinencia. Finalmente, construimos anclas de respuesta específicas del ítem para "dar me gusta" a los ítems. Reclutamos expertos para validar el contenido y realizamos entrevistas cognitivas para validar el proceso de respuesta. Para evaluar la estructura interna y la confiabilidad de la escala, invitamos a una muestra intencional de 120 residentes de cirugía a completar la escala y exploramos los resultados de la prueba piloto utilizando la reducción de datos y el análisis de confiabilidad. Un total de 77 residentes de cirugía completaron la escala. El análisis de componentes principales rotados con Varimax reveló tres componentes con valores propios mayores que uno, y los componentes explicaron el 64 % de la varianza total. La solución rotada fue consistente con la estructura original del cuestionario. Los componentes, que representaban los conjuntos de ítems de competencia, preferencia y pertinencia, explicaban el 25 %, el 23 % y el 16 %, respectivamente, de la varianza total. Los coeficientes de Cronbach para los conjuntos de elementos fueron 0,72, 0,71 y 0,61, respectivamente. Desarrollamos y validamos una escala para medir la preparación anatómica general de los graduados médicos. Además, ofrecemos pautas conceptuales para ayudar a los usuarios a interpretar los resultados de la escala. Se requieren datos de resultados para corroborar la validez predictiva de la escala.

PALABRAS CLAVE: Anatomía; Educación médica; Educación Médica de Posgrado; Encuestas y cuestionarios; Datos preliminares.

REFERENCES

Ahmed, K.; Rowland, S.; Patel, V.; Khan, R. S.; Ashrafian, H.; Davies, D. C.; Darzi, A.; Athanasiou, T. & Paraskeva, P. A. Is the structure of anatomy curriculum adequate for safe medical practice? *Surgeon, 8(6)*:318-24, 2010.

Artino Jr., A. R.; La Rochelle, J. S.; Dezee, K. J. & Gehlbach, H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med. Teach., 36(6)*:463-74, 2014.

Bergman, E. M.; Prince, K. J.; Drukker, J.; van der Vieuten, C. P. & Scherpbier, A. J. How much anatomy is enough? *Anat. Sci. Educ., 1(4)*:184-8, 2008.

Cabrera, A. R.; Lee, W. R.; Madden, R.; Sims, E.; Hoang, J. K.; White, L. E.; Marks, L. B. & Chino, J. P. Incorporating gross anatomy education into radiation oncology residency: a 2-year curriculum with evaluation of resident satisfaction. *J. Am. Coll. Radiol., 8(5)*:335-40, 2011.

Cahill, D. R.; Leonard, R. J. & Marks, S. C. Jr. Standards in health care and medical education. *Clin. Anat., 13(2)*:150, 2000.

Cole, T. J. Too many digits: the presentation of numerical data. *Arch. Dis. Child., 100(7)*:608-9, 2015.

Cottam, W. W. Adequacy of medical school gross anatomy education as perceived by certain postgraduate residency programs and anatomy course directors. *Clin. Anat., 12(1)*:55-65, 1999.

Craig, S.; Tait, N.; Boers, D. & McAndrew, D. Review of anatomy education in Australian and New Zealand medical schools. *A. N. Z. J. Surg., 80(4)*:212-6, 2010.

Eisinga, R.; te Grotenhuis, M. & Pelzer, B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *Int. J. Public Health, 58(4)*:637-42, 2013.

Ellis, H. Medico-legal litigation and its links with surgical anatomy. *Surgery (Oxford), 20(8)*:1-2, 2002.

Ergül, Z.; Kulaçoglu, H.; Sen, T.; Esmer, A. F.; Güller, M.; Güneri, G. & Elhan, A. A short postgraduate anatomy course may improve the junior surgical residents' anatomy knowledge for the nerves of the inguinal region. *Chirurgia (Bucur), 106(5)*:599-603, 2002.

Fitzgerald, J. E.; White, M. J.; Tang, S. W.; Maxwell-Armstrong, C. A. & James, D. K. Are we teaching sufficient anatomy at medical school? The opinions of newly qualified doctors. *Clin. Anat., 21(7)*:718-24, 2008.

Lazarus, M. D.; Chinchilli, V. M.; Leong, S. L. & Kauffman, G. L. Jr. Perceptions of anatomy: critical components in the clinical setting. *Anat. Sci. Educ., 5(4)*:187-99, 2012.

Older, J. Anatomy: a must for teaching the next generation. *Surgeon, 2(2)*:79-90, 2004.

Pabst, R. & Rothkötter, H. J. Retrospective evaluation of undergraduate medical education by doctors at the end of their residency time in hospitals: consequences for the anatomical curriculum. *Anat. Rec., 249(4)*:431-4, 1997.

Smith, C. F. & Mathias, H. S. What impact does anatomy education have on clinical practice? *Clin. Anat., 24(1)*:113-9, 2011.

Waterston, S. W. & Stewart, I. J. Survey of clinicians' attitudes to the anatomical teaching and knowledge of medical students. *Clin. Anat., 18(5)*:380-4, 2005.

Corresponding author:
Amjad T. Shatarat, MD, PhD
Department of Anatomy and Histology
School of Medicine, The University of Jordan
Queen Rania al-Abdullah Street
Amman 11942
JORDAN

Email: a.shatarat@ju.edu.jo

Received: 03-03-2020
Accepted: 29-04-2020