JOURNAL OF INTERDISCIPLINARY PERSPECTIVES

ISSN Print: 2984-8288, ISSN Online: 2984-8385
Vol. 3 No. 8, 866-873, August 2025

The Development of a Cognitive Test on Exercise Physiology for Exercise and Sports Science College Students

Julia Pearl M. Arroyo
University of the Philippines, Quezon City, Philippines

Author Email: jmarroyo@up.edu.ph

Date received: June 20, 2025
Date revised: July 7, 2025
Date accepted: July 28, 2025

Originality: 99%
Grammarly Score: 99%
Similarity: 1%

Recommended citation:
Arroyo, J.P. (2025). The development of a cognitive test on exercise physiology for exercise and sports science college students. *Journal of Interdisciplinary Perspectives*, 3(8), 866–873. https://doi.org/10.69569/jip.2025.517

Abstract. Future exercise and sports science professionals must master human physiological concepts and their applications to be effective in their fields. The education of aspiring professionals in exercise and sports science requires quality assessment methods, from which results serve as a basis for educational and training programs. To ensure coherence with the program objectives of academic institutions and industry standards, the use of well-developed cognitive tests is imperative. This study aimed to develop a cognitive written test for an introductory course in Exercise Physiology for use in an Exercise and Sports Science program. Eighty sports science students participated in the development of the cognitive test. The constructed forty-item multiple-choice test had a Cronbach's alpha of 0.73, indicating an "acceptable" internal consistency. Item-analysis results showed that 39 out of the 40 items had acceptable item difficulty levels, with a good combination of easy-desirable-difficult items. However, only 15 out of the 40 items were selected for retention, as 25 items were classified as having negative to poor item discrimination. The test development process entailed careful planning, pre-construction conceptualization, and a multi-step process, from identifying the test's purpose and constructing the table of specifications to item analysis and field testing. This study highlights the importance of test improvements and regular evaluation of curriculum assessment tools.

Keywords: Exercise physiology; Cognitive test; Test development; Sports science; Knowledge test.

1.0 Introduction

The human body is an amazing, complex machine. There is intricate communication that allows the coordination of all physiological functions. Cells, tissues, organs, and systems function and work together to keep the body active and effective in daily tasks and emergencies. A solid foundation in human physiology concepts enables the understanding of how the body responds, adjusts, and adapts to the emotional and physical demands of exercise and sport training. Exercise and sports science professionals must master and apply this knowledge to be effective.

The Chartered Association of Sport and Exercise Sciences (CASES) is a professional association that recommends undergraduate programs in exercise and sport science. It requires programs to "demonstrate the development of understanding of the key bodies of knowledge and application of scientific and practical techniques relevant to sport and exercise science in the subdiscipline areas of physiology, psychology, and biomechanics." (Chartered Association of Sport and Exercise Sciences, 2025, p. 2). To ensure coherence with the program objectives of academic institutions and industry standards, the development of well-developed cognitive tests is imperative.

Assessment methods are critical elements of educational programs that equip students to exhibit the fulfillment of program objectives (Khan et al., 2025).

Lane, Raymond, and Haladyna (2016) have outlined 12 components of test development, from an overall test plan to test documentation. A more straightforward process was devised by Knupp & Harris (2012). It includes test design, item development, field testing, item evaluation, and ends with monitoring and improving the test development procedures. The development of test content and evaluation of its psychometric properties are the initial steps for effective educational measures (Knupp & Harris, 2012; Mallaih, Williams, & Allegrante, 2024). Identifying the aim of the test is key to guiding the test content. Tests are constructed to assess mastery of concepts, evaluate program effectiveness, and determine academic scholarships, among other purposes.

For this research, test development involved a classroom summative assessment using a cognitive test. The validity of the scores depends partly on the content representativeness of the test items about the concepts discussed in class. Assessment methods developed in alignment with the course objectives facilitate student engagement and establish a solid foundation of relevant knowledge (Khan et al., 2025). Content specifications of the test, therefore, should reflect the course objectives or outcomes of the class.

Two key elements of test development are content representativeness and relevance, which are determined by both the content and the cognitive level. To ensure that a test will cover all learning that took place across the range of content areas covered in class and the cognitive processes considered essential, a test "blueprint" or what we commonly call a *table of specifications (TOS)* is constructed. These can be formulated by emphasizing the course material taught and represent the curriculum (Nitko, 2001). The cognitive dimension focuses on the thinking processes required for each course objective. Benjamin S. Bloom proposed the most widely used cognitive taxonomy (Krathwohl, 2002), which includes knowledge, comprehension, application, analysis, synthesis, and evaluation. These ensure that the test will have the necessary relevant content and represent different levels of cognitive complexity (Kubiszyn & Borich, 2013).

Selecting the type and design of a test will be contingent on the available time, the student's level or age, and how the scores will be interpreted. One of the most frequently used types of written tests is the multiple-choice question (MCQ) type. Multiple-choice questions (MCQs) can be used to assess a wide range of learning objectives (Khan et al., 2025). It is time- and cost-effective, efficient with a large amount of content, and suitable for assessing a considerable number of students (Sadaf, Khan, & Ali, 2012). MCQ tests are easy to facilitate, grade, and analyze (Kubiszyn & Borich, 2013).

This study aimed to develop a cognitive test in Exercise Physiology for use in an Exercise and Sports Science program. It sought to address the following: (1) develop an exercise physiology written test for exercise and sports science students, (2) determine the reliability and validity of the instrument/ test through item analysis (and distractor analysis) and Cronbach's alpha, (3) by identifying good questions and those needing improvement or removal, for its overall refinement.

This study is relevant to the students' learning and to enhance the exercise and sports science curriculum, particularly in the field of exercise physiology. It will initiate the development of an assessment tool for exercise physiology. This will ultimately enhance the teaching effectiveness of exercise physiology courses and identify factors that may relate to the learning and retention of the concepts. The results and applicability of this study, however, were limited to the development of a cognitive test from a specific sample population of sports science students at a state university in the Philippines. The process and methods were modified to suit the available time frame of a semester, the resources available to the researcher, and the specific contextual conditions of a state university.

2.0 Methodology

2.1 Participants, Sampling Technique, Context Description

The university practices academic freedom, allowing instructors and professors autonomy to create their assessments, based on the course content objectives established by each department. The researcher and study participants were affiliated with the Department of Sports Science. The participants were conveniently sampled from the students enrolled in sports science classes within the college. Participant inclusion criteria were that they should have completed the introductory exercise physiology course, which covered a thorough discussion of muscle physiology and exercise; they should have been enrolled in any sports science program and any of the higher-level sports science classes, including Advanced Exercise Physiology, Ergogenics, and Acquisition of Motor Skills. These were higher-level exercise physiology classes/ courses within the sports science curriculum.

2.2 Instrument Construction

The cognitive test consisted of a 40-item multiple-choice test with four options for each item. It covered concepts in muscle exercise physiology that were discussed in the introductory course in exercise physiology.

Purpose and Domain Identification *(Construction of the Table of Specifications)*. The purpose of the cognitive test was to evaluate the extent to which students had retained the subject matter content from the introductory exercise physiology course. A Table of Specifications was constructed using Bloom's revised taxonomy (Colorado College, 2022) of the levels of cognitive learning and the course content outline objectives of the introductory exercise physiology course. The Department of Sports Science developed the course content objectives.

The test *table of specifications* (TOS), Table 1, provides the specifications, i.e., course content objectives, and cognitive level of thinking. The cognitive test assessed the following course objectives: (1) describe the functions of muscle by providing examples, (2) define the general properties of muscle, (3) identify the gross and micro structure of skeletal muscle and the motor unit, (4) explain the function of major structures of skeletal muscle highly relevant to muscle contraction, (5) enumerate the steps in sarcomere contraction and relaxation, (6) differentiate the muscle fiber types and relevance to different intensity demands of sport activities, and (7) understand the differences of contraction types and the factors that affect force production for each type.

Table 1. *Table of Specifications used for the construction of the cognitive test for this study*

	Categories/ Levels of Thinking (number of items)							
Content Outline Objectives (COO)	**R**	**U**	**Ap**	**An**	**E**	**C**	**Total**	**%**
1. Describe the functions of muscle by giving examples.			1	1			2	5.00
2. Define the general properties of muscle.	1	1	1				3	7.50
3. Identify the gross and micro structures of skeletal muscle and the motor unit.	4						4	10.00
4. Explain the function of the significant structures of skeletal muscle highly relevant to muscle contraction.	2	3		2			7	17.50
5. Enumerate the steps in sarcomere contraction and relaxation.	2	2	1	1			6	15.00
6. Differentiate the muscle fiber types and their relevance to different intensity demands of sports activities.	2	2	2	2	1	1	10	25.00
7. Understand the differences in contraction types and the factors that affect force production for each type.		1	3	3	1		8	20.00
Total number of items	**11**	**10**	**8**	**9**	**2**	**1**	**40**	**100**
Percentage (%)	**27.50**	**25**	**20**	**22.50**	**5**	**2.50**	**100**	

Multiple Choice Questions for the Cognitive Test. The test was constructed in a multiple-choice question (MCQ) format with one best answer. The test consisted of question-type, stem, and lead-in questions, along with four answer options. Three options were considered acceptable, but four options were preferred (Thellesen et al., 2017). The test items were existing multiple-choice questions that the researcher had created and further improved. The researcher had 15 years of teaching experience in exercise physiology. She has a master's degree in Exercise and Sports Science, majoring in exercise physiology. A professor, PhD, of a graduate-level course on Test and Scale Development, reviewed the MCQ test items. The researcher revised the questions following the careful review.

Criterion-based scoring. The established passing score for the cognitive test considered for this study was the standard used by the university, which is 60%. This is a standard passing mark for all tasks and requirements in the university, regardless of subject or course.

2.3 Data Gathering Procedure

Approval to conduct the study was obtained from the chairperson of the Department of Sports Science and the College Dean. The researcher coordinated with the different professors of the higher-level sports science classes. After the initial revisions of the MCQs from the Test and Scale Development expert review, the cognitive test was field-tested with students from higher-level sports science subjects. Participant informed consent was secured before test administration. The researcher administered the test face-to-face during official class hours. Each student was given sufficient time to answer the test, without any time limit. It was emphasized that answering the instrument with utmost honesty, sincerity, and to the best of their ability was crucial and highly encouraged.

2.4 Data Analysis Procedures

Profile of participants. The students' data were coded and analyzed using descriptive statistics, including frequencies, percentages, means, and standard deviations, with both Excel and SPSS v. 25.0. The participant profile data for the cognitive test included course, year level, number of sports science courses completed, and exercise physiology test score.

Data Analysis for the Cognitive Test: Item Analysis. Item analysis evaluates the reliability and validity of test items. It demonstrates the effectiveness of the questions and their distractors, allowing for appropriate revisions or deletions of questions and resulting in high-quality test questions and examinations for future use (Elgadal & Mariod, 2021). Validation of the instrument involved item analysis, which consisted of two elements: item discrimination and item difficulty. Items that were either very easy or very difficult and those with negative or poor discrimination were recommended for removal. Cronbach's alpha estimates the internal consistency of the test. Cronbach's alpha at a p = .050 level of significance was used to determine the reliability of the instrument. Statistical analysis made use of Microsoft Excel software and SPSS version 25.0.

2.5 Ethical Considerations

This research made sure that the rights and protection of the participants were prioritized throughout the study. Informed consent was secured. The researcher ensured that the purpose, procedures, benefits, risks, confidentiality, and anonymity of participation were clearly explained to every participant. It was emphasized that participation was voluntary. The students were informed that their status as university students would not be affected.

3.0 Results and Discussion

This section presents the descriptions of the participants, the results of the item analysis, and the test reliability. It provides a comprehensive discussion of the results and their practical implications.

3.1 Profile of Exercise Science Students

A total of eighty (80) students from higher-level sports science subjects voluntarily participated in the test. All participants passed the introductory exercise physiology course. Thirty-four percent of the participants were third-year students, 19% were in their second year, 17% were in their first year, 10% were in their fourth year, 17% were in their fifth year, and 6% were in their sixth year of college. For the program of study, most participants were pursuing a Bachelor of Sports Science (BSS, 58%), followed by those pursuing a Bachelor of Physical Education (BPE, 15%), and a Master of Science in Human Movement Sciences (MSHMS, 15%). Students taking the Diploma in Exercise and Sports Science (DESS) accounted for 10% of the participants, while the Certificate in Sports Studies had the fewest students enrolled at 3%. The large proportion of students in their 1st and 2nd year, totaling 36%, was due to the presence of graduate students (MSHMS) in the sports science courses. The students accepted conditionally into the MSHMS program were required to complete the exercise physiology class as part of the bridging courses.

3.2 Cognitive Test Scores

The mean total score in the cognitive test of the sample population was 19.59 (± 5.60). Considering the cut-off (passing score set) of 60%, this mean score was considered failing. This may have implications for the retention of exercise physiology concepts as students take higher-level sports science classes.

3.2 Data Analysis: Item Analysis

Test development entails a series of steps that begin with defining item specifications using the Table of Specifications and conclude with validating the assessment method itself through item analysis and a reliability coefficient. Item analysis effectively identifies the target course objectives in a test, including its content, format restrictions, cognitive demands, the plausibility of distractors, item difficulty, malfunctioning items, and test internal consistency reliability (Elgadal & Mariod, 2021; Mallaih, Williams, and Allegrante, 2024).

In this study, 15 out of the original 40 items were considered acceptable or survived after item analysis. The shaded part of the grid in Table 2 indicates the surviving items after analysis. Items excluded were primarily due to the discrimination index falling into the poor and negative classifications. In a similar study, with a 43-item MCQ science literacy test, 12 items were dropped that had negative or poor discrimination, which were labeled " problematic" test items (Mallaih, Williams, and Allegrante, 2024).

Item Discrimination and Difficulty. The 25 excluded items were not appropriately constructed to differentiate between the outstanding students and the unsuccessful ones. Only three items were classified within the category of good and excellent discrimination, and were within the easy-desirable-difficult levels of difficulty. Twelve items fell into the fair discrimination category, and within the easy-desirable-difficult levels of difficulty. Therefore, the 15 items shown within the yellow-shaded portions of the table were considered acceptable.

Focusing on the difficulty level of the items, 39 out of 40 items were very acceptable. These 39 items fell within the range of easy, desirable, and difficult. The difficulty level of multiple-choice questions (MCQs) is a crucial element in determining their quality. Insufficient cognitive stimulation may result from test items that are too easy to answer. This may reflect an inaccurate assessment of learning and the attainment of course competencies (Khan et al., 2025). The profile of the participants in this study, as described in the previous section, revealed that they were from all the sports science degree programs at the college. Classes composed of students from diverse backgrounds and abilities, such as those in this study, benefit from moderately complex test items, as this increases the instrument's reliability (Kubiszyn & Borich, 2013).

Table 2. *Item analysis results showing the surviving number of items in the shaded region, with their corresponding discrimination index interpretation and difficulty index interpretation*

	Difficulty					
Discrimination	**Very easy**	**Easy**	**Desirable**	**Difficult**	**Very difficult**	**TOTAL**
Very good-Excellent						
Good			2	1		3
Fair		2	8	2		12
Poor	1	3	11	6		21
Negative			2	2		4
TOTAL	1	5	23	11		40

Fate of the Fifteen (15) Surviving Items. The three (3) items classified as good-very good discrimination and easy-desirable-difficult difficulty level were recommended to be included in an MCQ item bank of exercise physiology tests. The remaining 12 items required proofreading and rewording. It was recommended that the items be field-tested after improvements have been made. Enhancing the discrimination of individual items improves the quality of the test (Knupp & Harris, 2012).

Analysis of the Surviving Items: Content and Cognitive Level of Thinking. Upon analyzing the content domain and cognitive skills of each item that survived, the following findings were observed. This is shown in Table 3. Content Outline Objective 1 (COO1): Describe the functions of muscles by providing examples, was no longer represented in the surviving items. Content Outline Objective 3 (COO3): Identify the gross and micro structures of skeletal muscle and motor unit, and Content Outline Objective 6 (COO6): Differentiate the muscle fiber types and their relevance to different intensity demands of sports activities, were the most represented in the surviving test items. Remembering and understanding were the dominant cognitive skills represented in the 15 items. Evaluation and creation of cognitive skills were no longer represented. The alignment of course objectives with the cognitive skills, as seen in the table of specifications (TO), provided a sound foundation for constructing the test. Further development of the test should include the items that cover COO1 to improve their discrimination index. All course objectives must be represented in the final version of the test. Misaligned test items to course content objectives may inaccurately gauge actual understanding and the knowledge gained (Mubuuke et al., 2014).

Nine out of the 15 items that survived utilize factual knowledge, combining the cognitive skills of remembering and understanding. The test had a good mix of items with fair to good discrimination and a balance of easy-to-difficult questions. This was similar to Pal et al. (2025), who refer to remembering and understanding questions as recall questions, which had a moderate, balanced level of difficulty. Their higher-order cognitive skills, including analysis, application, and evaluation, were assessed through data interpretation and problem-solving questions, which proved to be more challenging for the students. Exercise physiology is a course that applies physiological concepts to practical exercise situations, such as when appropriate exercise activities are prescribed in training programs. The evaluation and creation cognitive skills need to be incorporated and matched to the respective course objectives, which were COO3 and COO6. This can support sound learning and the effective transfer of these concepts. These findings align with previous research (Pal et al., 2025), which had demonstrated that multiple-choice questions (MCQs) can effectively assess multiple levels of cognitive function, provided they are constructed in alignment with the cognitive skills and course content objectives. Learning is augmented when content and cognitive processes are congruous (Ajjawi et al., 2020). Assessment needs to shift from simply mastering factual knowledge to proficiency in applying knowledge and skills, as well as critical analysis and problem-solving cognitive skills, in practical situations (Balasooriya, et al., 2024).

The use of Bloom's hierarchy of cognitive skills provides a scaffolding progression from the basic understanding of foundational knowledge and skills to the application and critical analysis of more complex concepts in real-world contexts (Exercise and Sports Science Australia, 2024). The use of Bloom's taxonomy in assessments will assist the achievement of the specifications set by Exercise and Sports Science Australia (ESSA) for the accreditation of exercise scientists for the students of the Department of Sports Science in this study. Furthermore, research on the skills and attributes of sports science professionals specified 'critically analyze' and 'evaluate' sport demands and athlete capabilities, as well as 'plan and create evidence-based interventions' (Bruce et al., 2022; Wheeler & Van Mullem, 2021). This emphasizes the need for higher-level cognitive skills to be an essential part of learning and assessment materials for exercise and sport science programs.

Table 3. *Summary of the analysis of items that survived as a result of the item analysis with the Table of Specifications TOS*

Item Number	Item Classification: Discrimination-Difficulty	Course Competency/ Content Outline Objective (COO)	Level of Thinking
33	Good - Desirable	COO6	Remembering
38	Good - Desirable	COO7	Application
7	Good - Difficult	COO3	Remembering
22	Fair - Desirable	COO7	Application
12	Fair - Desirable	COO5	Remembering
32	Fair - Desirable	COO6	Analysis
8	Fair - Desirable	COO3	Remembering
28	Fair - Desirable	COO6	Understanding
29	Fair - Desirable	COO6	Remembering
9	Fair - Desirable	COO3	Remembering
20	Fair - Desirable	COO4	Analysis
23	Fair - Easy	COO7	Analysis-Application
31	Fair- Easy	COO6	Analysis
21	Fair - Difficult	COO5	Remembering
4	Fair - Difficult	COO2	Understanding

Analysis of Surviving Items: Distractor Analysis

The results of the distractor analysis, presented in Table 4, indicated that mis-keying did not occur in any of the surviving items, as more students in the upper level chose the correct option. It did not appear that guessing had occurred; no item had an almost equal number of students choosing all the options. This suggests that the students were engaged with the content (Pal et al., 2025). There is no ambiguity; no distractor was chosen nearly the same number of times as the correct option. This is similar to Pal et al. (2025), who assessed the effectiveness of MCQs in evaluating different levels of cognitive learning in biochemistry; the distractors in the MCQs were effective, as evidenced by a balanced distribution of responses across the answer choices. However, item 20, distractor A, was found to require modification.

Table 4. *Distractor Analysis Table for Surviving Items*

	Test Item Number														
Option	**4**	**7**	**8**	**9**	**12**	**20**	**21**	**22**	**23**	**28**	**29**	**31**	**32**	**33**	**38**
A	good	good	good	good	good	**X**	good	key	good	good	good	key	good	good	good
B	good	good	good	good	good	key	key	good	good	key	good	good	good	good	good
C	good	good	key	good	key	good	good	good	key	good	key	good	key	key	key
D	key	key	good	key	good	good	good	good	good	good	good	good	good	good	good

3.3 Reliability Test: Cronbach's Alpha

This coefficient measures the internal consistency of the test instrument. It estimates the degree to which the MCQs can assess the same cognitive domain features consistently throughout the test (Elgadal & Mariod, 2021; Aligway et al., 2024). The test item difficulty and discrimination affect the reliability of the test. Due to item discrimination issues, only 15 items survived. Revisions were recommended, and the removal of the 25 items recommended for deletion may further enhance the test's reliability. Despite this, based on the computed Cronbach's alpha of 0.73, the internal consistency of the instrument is "acceptable". Values closer to 1 indicate high reliability (Singh, Singh, & Ananthakrishnan, 2024). In a study by Mallaih, Williams, and Allegrante (2024), the internal consistency for a literacy assessment test improved from $\alpha = .46$ to $\alpha = .69$ after discarding flawed items. A Cronbach's alpha greater than .69 was reported to be an acceptable value for research instruments in science education (Mallaih, Williams, and Allegrante, 2024).

Overall, the multiple-choice test development process successfully produced valid and reliable test items that can be used in exercise physiology classes. Further revisions of the entire instrument were required. Course content objectives and statistical specifications, when carefully thought out and applied to assessments, can facilitate fair and consistent assessment measures across multiple classes and teachers (Knupp & Harris, 2012). The regular review of program objectives and assessment tools by a subject specialist, as part of the program department's protocols, can ensure that assessment methods are aligned with program objectives (Elgadal & Mariod, 2021). Multiple-choice items may be challenging and time-consuming to construct. However, they can include essential higher-level thinking skills that meet the needs of future exercise and sports science professionals (Kubiszyn & Borich, 2013).

The challenge for future research is to create more valid and reliable test items and assessment methods in the field of exercise and sports science. This can be achieved through intentional and coordinated efforts to design assessment methods and learning experiences that enhance the cognitive skills, knowledge, and practical skills sports science students require. It is highly recommended to conduct two field tests and periodically review the items to ensure they continue to align consistently with course outcomes and the type of students or test-takers. A more homogenous study sample population should be used for further and future test item improvements and development. This would prevent inconsistencies caused by the mix of graduate students' abilities. Assessment methods should be further studied for specific degree courses and specializations in exercise and sports science, such as sport psychology and biomechanics. Conducting a focus group discussion will be more effective for creating highly valid items that tap into the expertise of subject matter experts (SMEs).

4.0 Conclusion

Recognizing the need for aspiring athletic trainers, coaches, and sports scientists to have a strong understanding of human physiology, the study involved 80 sports science students in developing a 40-item multiple-choice test. The test demonstrated an acceptable internal consistency. The research emphasized the importance of a structured test development process that included conceptual planning, item analysis, and regular evaluation of assessment tools. Validity and reliability are emphasized as crucial factors in developing practical cognitive tests, with a focus on ensuring well-crafted assessment items that align with academic and industry standards in exercise and sports science education. This study validates multiple-choice questions (MCQs) as an effective tool for assessing a range of cognitive skills and course content objectives in exercise and sports science. It provided evidence for the need to reinforce the importance of standardized assessment methods and departmentalized sports science program activities. This may lead to more uniform and consistent learning outcomes for students.

5.0 Contributions of Authors

Sole Author: Julia Pearl M. Arroyo, 100%

6.0 Funding

The author received no funding for the research and publication of this article.

7.0 Conflict of Interests

This research/ author declares that she has no conflict of interest.

8.0 Acknowledgment

The author would like to extend her heartfelt appreciation to the dean and professors of the college and the university.

9.0 References

Ajjawi, R., Tai, J., Huu Nghia, T., Le, Boud D., Johnson, L., & Patrick CJ. (2020). Aligning assessment with the needs of work-integrated learning: the challenges of authentic assessment in a complex context, Assess Eval High Educ. 45:304–16. https://doi.org/10.1080/02602938.2019.1639613

Aligway, G. J., Delos Angeles, J., Collano, A., Barroca, E., Aves, A. C., Catubay, J., & Cortes, S. (2024). Validity and reliability of a concept inventory test in human physiology. JPBI (Jurnal Pendidikan Biologi Indonesia), 10(1), 273-282. https://doi.org/10.22219/jpbi.v10i1

Balasooriya, C., Lyons, K., Tran, M., Pather, N., Chur-Hansen, A., & Steketee, C. (2024). Learning, teaching, and assessment in health professional education and scholarship over the next 50 years. Focus On Health Professional Education, *25*(2), 110–129. http://dx.doi.org/10.11157/fohpe.v25i2.785

Bruce, L., Bellesini, K., Aisbe1, B., Drinkwater, E., & Kremer, P. (2022). A profile of the skills, attributes, development, and employment opportunities for sport scientists in Australia. Journal of Science and Medicine in Sport, *25*, 419-424. https://doi.org/10.1016/j.jsams.2021.12.009

Chartered Association of Sports and Exercise Science. (2025, March). Undergraduate endorsement scheme application guidelines. United Kingdom. Retrieved from https://www.cases.org.uk/imgs/bues_application_guidelines_new___mar_2593.pdf

Colorado College. (2022, April 14). How to assess learning: Bloom's revised taxonomy. Retrieved from Colorado College: https://tinyurl.com/4tptrmsx

Elgadal, A., & Mariod, A. (2021). Item analysis of multiple-choice questions (MCQs): Assessment tool for quality assurance measures. Sudan Journal of Medical Sciences, *16*(3), 334-346. https://doi.org/10.18502/sjms.v16i3.9695.

Exercise and Sports Science Australia. (2024, September). Accredited exercise science professional standards for accreditation. Australia. Retrieved from https://tinyurl.com/mwdkb8ms

Khan, H. F., Qayyum, S., Beenish, H., Khan, R. A., Iltaf, S., & Faysal, L. R. (2025). Determining the alignment of assessment items with curriculum goals through document analysis by addressing identified item flaws. BMC Medical Education, *25*, 1-9 https://doi.org/10.1186/s12909-025-06736-4.

Knupp, T., & Harris, D. (2012). Building content and statistical specifications. In C. S. Denison, Handbook on measurement, Assessment and Evaluation in Higher Education (p. 680). New York: Routledge, Taylor & Francis Group.

Krathwohl, David (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Kubiszyn, T., & Borich, G. (2013). Educational testing & measurement, 10th edition. Jefferson City: John Wiley & Sons, Inc.

Lane, S., Raymond, M., & Haladyna, T. (2016). Handbook of test development. New York: Routledge Taylor & Francis.

Mallaih, J., Williams, O., & Allegrante, J. (2024). Development and validation of a stroke literacy assessment test for community health workers. Health Education & Behavior, *51*(5), 764–774. http://dx.doi.org/10.1177/10901981241245050

Mubuuke, A. G., Mwesigwa, C., Maling, S., Rukundo, G., Kagawa, M., Kitara, D.L., et al. (2014). Standardizing assessment practices of undergraduate medical competencies across medical schools: challenges, opportunities, and lessons learned from a consortium of medical schools in Uganda. Pan Afr Med J, 19:382.. http://www.panafrican-med-journal.com/content/article/19/382/full/

Nitko, A. J. (2001). Educational assessment of students (3rd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Pal, M., Maji, B., Sakar Mahat, J., Haldar, S., Das, P., & Sethuraman, K. (2025). Unlocking cognitive potential: Validating multiple-choice assessments for higher-order cognitive skills in medical education. Journal of Chemical Health Risks, *15*(3), 250–253. https://www.jchr.org/index.php/JCHR/article/view/8183/4692

Sadaf, S., Khan, S., & Ali, S. (2012). Tips for developing a valid and reliable bank of multiple-choice questions (MCQs). Education for Health, *25*(3). https://doi.org/10.4103/1357-6283.109786

Singh, G., Singh, R., & Ananthakrishnan, N. (2024, December). Analyzing the quality of multiple-choice questions in microbiology for second-year MBBS students. *MGM* Journal of Medical Sciences, *11*, 651–658. DOI:10.4103/mgmj.mgmj_333_24. http://dx.doi.org/10.4103/mgmj.mgmj_333_24

Thellesen, L., Bergholt, T., Hedegaard, M., Colov, N., Christensen, K., Andersen, K., & Sorensen, J. (2017). Development of a written assessment for a national interprofessional cardiotopography education program. BMC Medical Education, *17*(88). https://tinyurl.com/y3yeser9

Wheeler, W., & Van Mullem, H. (2021). High-Impact Educational Practices in Kinesiology: Examples of Curricular Advancements to Prepare Students for the Future of Work. Kinesiology Review, *10*, 419-427. https://doi.org/10.1123/kr.2021-0047