

Standardized Assessment Model in Physical Education 1

Asia Pacific Journal of
Educational Perspectives

Vol. 8 No.1, 72-78

May 2021

ISSN: 2782-9332 (Print)

Emily J. Berja (EdD)

Bicol State College of Applied Sciences and Technology, Philippines

ejberja@astean.biscast.edu.ph

Date Received: November 25, 2020; Date Revised: March 7, 2021

Abstract – Student evaluation is one of the four main elements of physical education. Evaluation collects evidence of student success and draws inferences regarding student progress based on that evidence. The research suggested a mix of content-based and performance-based evaluation frameworks for physical education based on five learning outcomes. The proposed assessment model consists of 50 content-based items and six performance-based topics spread across learning domains. Based on the difficulty index, there are 15 simple questions in the proposed content-based performance evaluation models for physical education. Twelve of these items are easy; seven are average in difficulty, ten are complicated, and six are very difficult. The study utilized a mixed method of research design with 108 research participants: 3 experts in Physical Education, 5 PE teachers, and 100 students undertaking PE 1 in Bicol State College of Applied Sciences and Technology. In evaluating the proposed assessment models, the researcher utilized the five-point Likert scale, Cronbach's alpha, and discrimination index. Questions were evaluated through an index of discrimination test and were found to be good items and acceptable models in terms of distinguishing the examinees that are knowledgeable or not. Based on the results, the proposed content-based and performance-based assessment model is a valid and reliable tool to determine the level of competencies in physical education among the college students in Bicol State College of Applied Sciences and Technology. The proposed and revised content and performance-based assessment models in physical education be presented to physical education teachers, coordinators and be recommended for its implementation.

Keywords: Assessment, Physical Education, content-based assessment, performance-based assessment, college student

INTRODUCTION

In studying, individuals accumulate subject awareness, build skills, develop working patterns, and practice their application to real-world circumstances.

Learning assessment is a set of strategies for improving and applying knowledge, skills, and working practices by implementing important and stimulating exercises to students. Evaluation of physical education (P.E.) is critical and challenging to complete. The need for better assessment systems in physical education is a crucial priority in assessing quality achievement and decreasing marginalization [1]. Despite the effectiveness and prevalence of school tests, considering the numerous challenges that remain, such as significant class sizes, scarcity of funding, and record-keeping problems, the implementation of physical education has been daunting [2].

Physical Education is a learning branch that uses learners' physical strength and ability to develop knowledge, emotional stability, body wellness, and skills. These are factors contributory to the holistic lifelong growth and development of an individual. Human beings live in a cycle centered on moving, thinking, improving, and balancing, all present in Physical Education. Developing assessment tool in Physical education presents a constant challenge for P.E. teacher educators.

An international study in Physical education highlighted assessment for learning. For example, the Norwegian study of Physical Education emphasizes that integrating assessment for learning in the teaching practices and curriculum is an integral part of cultivating society [3]. In the same way, Thompson and Penney revealed that assessment is embedded in teachers' practices; however, they are deemed superficial, and the goal of the evaluation is usually not achieved, and the study also highlighted gaps in assessment knowledge, a lack of moderation and student engagement in assessment [4].

Performance-based assessment of physical education enables learners to show particular abilities and competencies when doing or creating something. It will help Physical Education instructors evaluate both what students can do and what they have accomplished in a given teaching program based on the Curriculum expectations. The performance-based evaluation also measures the learning experience of the learner.

The instructor's role is to connect the unit's material to subjects or processes that empower students to learn and to coordinate learning in a way that promotes the discovery of new interests [5][6]. Content refers to education in the emotional, cognitive and psychomotor worlds that student experiences convey in behaviors, knowledge, and skills. In a content-based assessment in education, the teachers assessed the student accumulated learnings in the three domains.

The researcher conducts this study to standardize performance assessment techniques in P.E. This standardization effort in assessing students' performance in Physical Education increases the level of validity, practicality, and reliability of test scores, thus, the reliability of grades given to P.E. students. It will also serve as a control mechanism in ensuring the delivery of quality instruction to P.E. learners. The study can also be a basis for curriculum enhancement in Physical Education.

In college, under the new curriculum, physical education is mandated by the Commission on Higher Education on the latest policies, standards, and guidelines across disciplines through Resolution No. 724-2017 dated October 3, 2017, focuses on the different movement patterns [7]. Students are required to adapt and pass knowledge and agility skills in various situations. Body movement is essential to the progress of day-to-day work and life. Awareness about what, when, and how the body will travel is vital to the standard of human life.

This study is developed to assess the performance of the students in Physical Education objectively. Since implementing the new curricula in the tertiary levels, teachers face unique challenges in developing instructional and assessment tools. All colleges and universities are challenged to address the need for updated tools aligned to the new content standards and competencies required in each course.

It is hoped that these assessment models can help teachers of Physical Education efficiently and objectively assess their students' performance according to the educational objectives. Moreover, may this prototype performance assessment models standardize physical education teachers' practices at the tertiary level and consequently serve as a control mechanism in ensuring the quality of physical education.

OBJECTIVES

This study proposed content-based and performance-based assessment models in Physical

Education 1 anchored on the five learning outcomes at Bicol State College of Applied Sciences and Technology (BISCAST). Specifically, the study develops a performance-based assessment in Physical Education 1; and determine the content validity and reliability of the assessment.

METHODOLOGY

Research Design

The analysis used the research and development approach. The study developed content and performance-based evaluation models that used physical education performance assessment methods. This research mainly looked at the comments and ratings of the experts and the PE teachers. The evaluation models have been used and tested for content validity and reliability. This research used quantitative and qualitative methods to process, evaluate and interpret numerical data and verbal responses.

Research Locale and Participants

The research study was conducted in BISCAST. It is located along Peñafrancia Avenue in Naga City. There was a total of eight respondents in the study who rated the assessment models proposed in the study. The reliability was tested based on the ratings of the proposed content-based questionnaires and performance assessment rubrics. Validity was also measured by the 8 raters using the validated tool determine practicality of the assessment models. They were composed of 5 PE teachers and 3 experts. There were at least 100 students who participated in the pilot-testing of the assessment models.

Table 1. Respondents of the Study

	Frequency	Percentage
Experts	3	2.8
PE teachers	5	4.6
Students	100	92.6
TOTAL	108	100

Research Instruments

This research utilized a metarubric to determine the content validity and also used as an interview guide, proposed assessment models, and rating scale for practicality.

The Metarubric for Content Validity and as Interview Guide. The metarubric covers three criteria to generate information on their suggestions and

observations for improvement of each of the five assessment models [8].

Rubrics. The proposed rubrics for the five learning outcomes Physical Education 1 were developed and later used in rating the performances. Rubrics were varied in criteria, indicators, weight and descriptors.

Data Gathering Technique

Criteria of the rubrics for assessing Performance in writing were gathered from relevant books and literature and inputs from Physical Education 1. Assessment Models were developed based on the cited sources and from the recommendations of the content validators.

The researcher seeks approval from the President of Bicol State College of Applied Science and Technology to conduct this study by administering the five assessment models based on the five-course learning outcomes. The researcher employed Physical Education experts from other colleges and universities, PE teachers in separate sessions to gather their inputs and comments. In conducting and collecting the necessary data for the study, informed consent and confidentiality agreement were secured from the respondents, assuring the respondents any information collected will be only used for the research.

For the assessment models' content validity, validators were gathered to test and determine the extent to which the rubrics' contents measure what they are supposed to measure. It is also employed to establish the relevance, adequacy, and completeness of the indicators and criteria. When the assessment models are completed, expert reviewers participated in the content validation. Content validation is an analytical study based on individual subjective decisions.

Content validation was carried out to evaluate each assessment model discussed the competencies to be assessed and the extent to which the assessment is appropriate and applicable in the outside world. Respondents are provided explicitly with rubric checklists to determine the headings' performance standards, proficiency, and explanation. Modifications to the evaluation models were made based on the assimilation of the comments of the reviewers.

In evaluating scores involving raters, it is essential to know the extent to which different scorers agree (or disagree) on the values assigned by the experts and the teachers. In this study, three experts and five teachers, and 100 students are asked to rate the Performance and determine their ratings.

Statistical Treatment of Data

For evaluating the practicality of the assessment models, raters are provided with a rating scale and the responses are analyzed using weighted means. Aside from quantitative rating, comments of the raters are analyzed using qualitative methods. Furthermore, results of interviews with experts, coaches, and teachers are collated. Qualitative responses were analyzed and integrated in the discussions. Their suggestions were considered in assessment models' revisions.

Content validity of the assessment model is established based on the evaluation ratings of the panel of experts using the Metarubric and will be analyzed using the weighted mean.

Table 2. Likert Scale for Content Validity

Parameter Limits	Response Category	Interpretation
3.26 – 4.00	Strongly Agree	Very Good
2.51 – 3.25	Agree	Good
1.76 – 2.25	Disagree	Fair
1.00 – 1.75	Strongly Disagree	Poor

For the sake of reliability, the Cronbach Alpha was calculated as an index of internal accuracy of the ratings provided to each article by the raters using the corresponding evaluation model. Cronbach alpha internal consistency, as a measure of reliability, was used to calculate the agreement's degree if the raters were to allocate scores to each article using the suggested evaluation models [9] in which the standardized assessment tool collected 0.941 reliability.

Discrimination Index has been used if the test question/item is valid. The Index of Discrimination is a primary measure of the validity of an item. It is a measure of an object's ability to differentiate between those who scored high on the test and those who scored low [7].

Table 3. Interpretation of Discrimination Index

Range	Verbal Interpretation
0.40 - above	Very Good Item/accepted
0.30 - 0.39	Good Item/accepted
0.20 - 0.29	Fair Item/accepted
0.09 - 0.19	Poor item/rejected

RESULTS AND DISCUSSION

Physical Education 1 offers instruction on various types of movement and core involvement under healthy

eating and physical activity. Students would be able to adapt and transfer their mobility skills in multiple situations. The following are the Physical Education Course Learning Results (CLO).

A. Content -Based Assessment Models

The proposed content-based performance assessment models in physical education are anchored on the five learning outcomes. It has 50 items spread across learning domains. In the course learning outcome 1, there are six (6) items each for remembering understanding questions. There are five (5) questions pertaining to applying and three (3) for analyzing. These 20 questions covered the topics: determinants of lifelong health, wellness and fitness; My food plate plan and balancing diet; Behavior and modification; Importance of physical activity; Introduction to fitness; Benefits of fitness Principles of physical activity; FIT formula; Movement pattern; Self-assessment of Physical Activity; Anatomical types of movement; Phases of exercise program; and Contraindicated Exercises.

B. Performance-based Assessment Model

There were two performance-based assessment rubrics designed under CLO 1. Performance-based assessment No. 1 entitled "My food Plate Plan", with three criteria to be assessed, nutritional components, meal planning and variety. Performance-based assessment no. 2, is "The Anatomical movement" covers the criteria, readiness, mastery of movement and participation/coordination. Both performance-based assessment will measure the learnings on CLO 1.

Under course learning outcome number 2, the six-hour concepts have a total of five (5) questions each for remembering an understanding. CLO 2 include the three phases of exercise program, low intensity workout and contraindicated exercises. The performance-based assessment is on the three phases of exercise program. Students will be rated according to the following criteria: organization of fitness program, completeness of the fitness program, participation, timeliness and safety measures/ social responsibility. CLO 3 has a total of 10 questions. It has 5 questions for remembering and two questions each for understanding and apply and 1 for analyzing. Performance-based assessment rubrics under CLO 3 covers the Cardiovascular Fitness. Students are given to criteria on how they will be rated: Organization of the cardiovascular fitness exercise, appropriateness of the

steps and music of the fitness dance and completeness and correctness of the fitness dance exercise.

CLO 4 is involved in fitness practices and sports. There are ten questions assigned to this learning result, which aim to discover, implement, and review various techniques to enhance the efficiency of one's physical activity and those of others. Performance-based assessment rubrics for Fitness activities and games have the following criteria to rate the students' performance: Participation, cooperation/coordination, and safety measures and social responsibility. It is equivalent to 6 hours with ten questions spread across learning outcomes on content based and the performance-based assessment. For content-based assessment, there were four questions assigned to remembering and six questions are assigned to analyzing. Finally, CLO 5 explores wellness and well-being activism to encourage realistic and innovative interventions that can build a community bond and lead to the school's health and well-being and the broader community. It has five questions for each of you to recall and submit. The "Health and Wellness Advocacy Proposal" was also the performance-based appraisal headings under CLO 5. Students will be scored on the following criteria: proper style, health, and wellness content, and timeliness.

Gareis and Grant (2008) describe the Tab of Specification (TOS) as a map or table that details the substance and level of cognitive level tested on the test as well as the styles and emphases of the test items [10]. The specification table is very relevant to discuss the validity and reliability of the test objects. The test's validity ensures that the evaluation can derive the required conclusion from the assessment since the assessment has secured against any systemic error.

The specification table includes a means for the test constructor to ensure that the evaluation is based on the expected learning outcomes. It is also a means of ensuring that the number of questions on the exam is sufficient to guarantee accurate answers that are unlikely to be attributed to chance. It is also a valuable guide when designing a test and deciding the form of test objects you need to construct [11].

The validity of the proposed content-based assessment model in physical education is rated as very good, with a mean of 3.40. and 3.69 for performance-based assessment models as shown in Table 4. Specifically, these assessment models are very good in content and coverage, with a mean of 3.33 for content-based and 3.63 for performance-based assessment.

Table 4. Validity of the Proposed Content and Performance -Based Assessment Model in Physical Education

Metarubric Traits	Content - Based		Performance- Based	
	WM	VI	WM	VI
Content and Coverage				
Performance criteria are linked to standards.	3.67	VG	3.6	VG
There are a manageable number of performance criteria.	3.33	VG	3.8	VG
The performance criteria are measurable and teachable.	3.33	VG	3.8	VG
Performance criteria match the objectives.	3.33	VG	3.4	VG
The performance criteria clearly stated with a meaningful label.	3.67	VG	3.6	VG
The rubric has these elements: Performance Criteria, Scale, Levels of Proficiency, and Descriptions.	3.33	VG	3.6	VG
Mean	3.44	VG	3.63	VG
Technical Quality				
There is an even number of levels, i.e., 4 to avoid middle scoring.	3.33	VG	3.8	VG
The highest level represents exemplary performance.	3.33	VG	4.00	VG
Each level on the scale is meaningful and non-judgmental.	3.33	VG	3.6	VG
Mean	3.33	VG	3.8	VG
Clarity and Detail				
They are written in student language.	3.33	VG	3.4	G
They are positively stated.	3.67	VG	3.8	VG
There are differences in descriptions.	3.33	VG	3.8	VG
There is a progression of differences among the descriptions.	3.33	VG	3.6	VG
Mean	3.42	VG	3.65	VG
GWA	3.40	VG	3.69	VG

Note: *VG – Very Good; G- Good*

Content and performance-based assessment are both very good in technical quality with a mean of 3.4 and 3.63, respectively. Both assessments are very good in clarity and detail based on the mean of 3.42 for content-based and 3.65 for performance-based. The evaluation models are valid because the performance criteria are related to the standards, have a manageable number, observable and legible, aligned with the objectives, explicitly defined with a meaningful label; and the heading has these elements: performance criteria, size, and level of proficiency.

On the other hand, the technical quality is very good because there are many levels, i.e., 4 to avoid middle scoring; the highest level represents exemplary performance. Each level on the scale is meaningful and non-judgmental. Finally, they are very good in clarity and detail because they are positively stated; there are differences in descriptions, and there is a progression of differences among the descriptions. The result ensures that the parameters or metrics are truly indicative of the results of the operation. It requires the comprehensive administration of the evaluation material to decide if it encompasses a representative range of the measurements or competencies to be measured. The competencies to be measured must be

routinely evaluated to ensure that the evaluation items and inaccurate proportions address all relevant factors.

Table 5. Reliability Statistics Content- Based Assessment Model

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
0.941	0.964	50

Table 5 presents the test of the reliability of the content-based assessment models in physical education. In examining reliability, the proposed content-based assessment model in Physical Education has a Cronbach alpha of .941 (excellent). This implies that concordance in their evaluation as to the reliability of items, particularly in accessing, synthesizing, and evaluating information relative to health; applying concepts.

Based on the results of discrimination of the content-based assessment it implies that mostly of the questions are accepted, some need a little region however there are questions that should be rejected due to very low result of the discrimination index as shown in Table 6.

Table 6. Index of Discrimination

Item No.	No of correct answer from top 27%	No of Correct Answer from the low 27%	Discrimination Index	Verbal Interpretation	Action
1	24	15	0.33	Good Item	Accepted
2	27	18	0.33	Good Item	Accepted
3	25	15	0.37	Good Item	Accepted
4	27	18	0.33	Good Item	Accepted
5	24	15	0.33	Good Item	Accepted
6	20	15	0.19	Poor Item	Rejected
7	22	17	0.19	Poor Item	Rejected
8	15	10	0.19	Poor Item	Rejected
9	23	14	0.33	Good Item	Accepted
10	22	2	0.74	Very good item	Accepted
11	10	7	0.11	Poor Item	Rejected
12	11	5	0.22	Fair item	Accepted
13	2	0	0.07	Poor Item	Rejected
14	19	9	0.37	Good Item	Accepted
15	20	11	0.33	Good Item	Accepted
16	22	10	0.44	Very good item	Accepted
17	20	11	0.33	Good Item	Accepted
18	27	27	0.00	Poor Item	Rejected
19	15	5	0.37	Good Item	Accepted
20	10	2	0.30	Good Item	Accepted
21	18	9	0.33	Good Item	Accepted
22	24	20	0.15	Poor Item	Rejected
23	25	20	0.19	Poor Item	Rejected
24	25	20	0.19	Poor Item	Rejected
25	23	17	0.22	Fair Item	Rejected
26	25	18	0.26	Fair Item	Accepted
27	18	5	0.48	Very Good Item	Accepted
28	27	24	0.11	Poor Item	Rejected
29	26	17	0.33	Good Item	Accepted
30	9	0	0.33	Good Item	Accepted
31	24	10	0.52	Very good item	Accepted
32	9	0	0.33	Good Item	Accepted
33	25	15	0.37	Good Item	Accepted
34	27	24	0.11	Poor Item	Rejected
35	25	17	0.3	Good Item	Accepted
36	20	18	0.07	Poor Item	Rejected
37	18	2	0.59	Very good item	Accepted
38	24	7	0.63	Very good item	Accepted
39	12	1	0.41	Very good item	Accepted
40	4	0	0.15	Poor item	Rejected
41	21	8	0.48	Very good item	Accepted
42	18	4	0.52	Very good item	Accepted
43	23	10	0.48	Very good item	Accepted
44	20	11	0.33	Good Item	Accepted
45	24	13	0.41	Very good item	Accepted
46	20	11	0.33	Good Item	Accepted
47	15	0	0.56	Very good item	Accepted
48	20	6	0.52	Very good item	Accepted
49	22	4	0.67	Very good item	Accepted
50	5	0	0.19	Poor item	Rejected

CONCLUSION AND RECOMMENDATION

Based on the findings, it is concluded that the proposed content-based performance assessment models in physical education is anchored on the five learning outcomes. It effectively assessed the competencies of the students in the accessing, synthesizing and evaluating information relative to health; applying concepts. The content-based assessment model is acceptable because of the items are found to be acceptable based on the discrimination index. It is also a valid and reliable content and performance-based assessment tool to determine the level of competencies in physical education among the college students in Bicol State College of Applied Sciences and Technology. The content-based assessment models also look into the relationship of the between an examinee's performance on the given item (correct or incorrect) and the examinee's score on the overall test. Most of the items are found acceptable.

It is recommended that the proposed and revised content and performance - based assessment models in physical education be presented to physical education teachers, coordinators as a basis for curriculum enhancement. Besides, a seminar can be held to enhance the mastery of the physical education teachers in developing test items based on the learning domains and the competencies required in the college physical education course. A Test Bank for content-based physical education test items be developed so that the P.E. teachers can easily retrieved these items according to the concepts and competencies they cover in particular grading period. Teaching materials and instructional technology be provided for teachers of physical education to effectively tackle and insure mastery on the different course learning outcomes.

The study only proposed two assessment models in Physical Education and were conducted in BISCASST thus the study findings are limited to the current curricula of the institution. For future investigation, assessment models in Physical Education of different state universities may be explored to determine and develop an assessment model that is appropriate for any higher education institutions.

REFERENCES

- [1] Collier, D. (2011). Increasing the value of physical education: The role of assessment. *Journal of Physical Education, Recreation & Dance*, 82(7), 38-41.
- [2] Gallo, A. M., Sheehy, D. A., Patton, K., & Griffin, L. (2006). Assessment benefits and barriers: What are you committed to?. *Journal of Physical Education, Recreation & Dance*, 77(8), 46-50.

- [3] Leirhaug, P. E., & MacPhail, A. (2015). 'It's the other assessment that is the key': three Norwegian physical education teachers' engagement (or not) with assessment for learning. *Sport, Education and Society*, 20(5), 624-640.
- [4] Dinan Thompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, 21(4), 485-503.
- [5] Santangelo, T., & Tomlinson, C. A. (2012). Teacher educators' perceptions and use of differentiated instruction practices: *An exploratory investigation. Action in Teacher Education*, 34, 309-327
- [6] Colquitt, G., Pritchard, T., Johnson, C., & McCollum, S. (2017). Differentiating instruction in physical education: personalization of learning. *Journal of Physical Education, Recreation & Dance*, 88(7), 44-50.
- [7] Commission in Higher Education. CMO Resolution No. 724-2017: Policies, Standards and Guidelines for Bachelor of Elementary Education. Philippines.
- [8] Cronbach's Alpha. (2008). In P. Lavrakas, Encyclopedia of Survey Research Methods. Sage Publications, Inc. <https://doi.org/10.4135/9781412963947.n117>
- [9] Bichi, Ado & Embong, Rahimah & Mamat, Mustafa & Abidin, Zainal. (2015). Evaluating the Quality of Islamic Civilization and Asian Civilizations Examination Questions. Proceedings of ICIC2015 – International Conference on Empowering Islamic Civilization in the 21st Century.
- [10] Gareis, C. R. & Grant, L. W. (2008). Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning. New York: Eye on Education.
- [11] Gabuyo, Y. (2014). Assessment of Learning: Textbook and Reviewer I. Manila: REX Book Store.