

QUALITY OF ONLINE BIODIVERSITY SUMMATIVE ASSESSMENTS ADMINISTERED TO STUDENT TEACHERS AT A SOUTH AFRICAN UNIVERSITY

Kholofelo Makgopa, & Monde Kazeni

Department of Science and Technology Education, University of the Witwatersrand (South Africa)

Abstract

Summative assessments are often used to make high stake decisions about students' academic success, which have long-term implications for both students and their respective institutions. Therefore, these assessments need to be of high quality to ensure that they yield accurate, objective and dependable results, upon which valid decisions could be made. The use of online assessments, where most lecturers have little experience and students have control of the assessment environment could compromise the quality of summative assessments. This is particularly true for Biodiversity assessments, which are complex. There is dearth of literature on the quality of online Biodiversity summative assessments in the South African context. This qualitative case study was conducted to determine the quality of online Biodiversity summative assessments administered at a South African university. The LINQED Quality Assurance Framework for Student assessment guided this study, which involved a review of eight online Biodiversity summative assessments administered to student teachers. Quality was measured by determining the content validity, authenticity, alignment with course objectives and the cognitive levels at which the assessments were pitched. The reviewed assessments were found to have high content validity and alignment to course objectives, but had poor coverage of the content and course objectives. In addition, the assessments were pitch at low cognitive levels and they had low to moderate authenticity. While these findings are based on a small sample, they could signify a global phenomenon. We therefore recommend further research, involving a bigger sample, to determine the recurrence of the findings from this study.

Keywords: *Summative, assessments, Biodiversity, online, student teachers.*

1. Introduction

Educational institutions have a mandate to educate, transform and train students in their respective fields of study (Shukla & Dungsungnoen, 2016). In this regard, assessments play a major role of collecting evidence about students' competencies and learning progress. Assessment, in the context of this study refers to the process of tracking students' learning progress, grading and providing students with feedback (Oyinloye & Imenda, 2019). Summative assessments are used to evaluate students' learning progress, skills acquisition, and academic achievement at the end of a defined instructional period. This type of assessment often takes the form of examinations, assignments, projects, or end of block tests, which are often used to make high stakes decisions about students.

Despite the high-stake decisions associated with summative assessments, literature (States, Detrich & Keyworth, 2018) shows that most summative assessments do not provide accurate and dependable results for making important decisions about students' lives, because of the way they are designed. This is particularly true for assessments in the Life Sciences topic of Biodiversity, which is difficult to assess (Sadler et al, 2013). This difficulty was exacerbated in South Africa in 2020 and 2021, when learning institutions were abruptly forced to administer assessments online, due to the COVID-19 pandemic. At this time, online teaching and assessment were relatively new to most educational practitioners in South Africa. Therefore, assessors might have simply developed online summative assessments that were easy to design and grade, such as items that only measure factual knowledge (Junus et al, 2021), hence lowering the quality of assessments.

Low quality online summative assessments are predisposed to cheating by students, who can easily copy answers from the internet, textbooks, class notes, or by consulting knowledgeable others (Junus et al., 2021). This implies that online summative assessments might not assess students'

competencies accurately, and these students would progress to the next educational level or to the corporate world with inadequate knowledge and skills. The purpose of this study was therefore to establish the quality of online Biodiversity summative assessments administered to student teachers at lower (first-year) and higher (Honours) educational levels at a South African university. Student teachers are students who train to become teachers. The study question was: What is the quality of online Biodiversity summative assessments administered to first-year and Honours student teachers at a South African university?

2. Conceptual framework

The quality of assessment is measured using different parameters, but this study focused on the measurements of content validity, authenticity, alignment with course objectives and of the cognitive levels at which assessments are pitched. The selection of these quality measures was guided by the LINQED Quality Assurance Framework for Student assessment (Educational network LINQED, 2011), which include fitness for Purpose (alignment with course objectives and content), authenticity and cognitive complexity. These assessment quality measures have been used in many studies (Gareis & Grant, 2015; Kabombwe, Machila, & Sikayomya, 2021). Measurement of assessment alignment with course objectives involves relating assessment items to the objectives specified for the course under review, while measurement of content validity involves matching assessment items with course content (Gareis & Grant, 2015). The revised Bloom's taxonomy of educational objectives (Anderson & Krathwohl, 2001) has been used extensively to determine the cognitive levels at which assessment items are pitched (Kabombwe, Machila, & Sikayomya, 2021; Anees, 2017). The measurement of assessment authenticity, which denotes an assessment's applicability to real-life situations has often involved the use of Wiggins' framework for authentic assessment (Wiggins, 1990). These measurement methods were employed in this study.

3. Methodology

The study was a qualitative case study involving the review of eight Biodiversity examination papers, comprising of four main and four deferred examination papers, administered to first-year and Honours student teachers in 2020 and 2021, at a South African university. Main and deferred examination papers were reviewed per year and per level of study to triangulate the results of the review. Moreover, assessments from two educational levels (1st year and Honours) were reviewed to determine whether there was a difference in the quality of Biodiversity assessments administered at lower and higher educational levels.

During the review, each assessment item was examined by three raters independently, to determine its alignment with the stated assessment quality measures. Raters then determined the frequency of each assessment quality measure, in each examination paper. To analyse the collected data, average frequencies (from the three raters) of each assessment quality measure per assessment paper were determined. These frequencies were computed into percentages and presented on tables.

Data collection schedules were developed by the researchers and were independently reviewed by two Life Sciences lecturers. The schedules were piloted using summative assessment papers that were not part of the main study. This was to ensure that the schedules captured the intended information. Furthermore, the study was approved by the participating institutions' ethics committee.

4. Results

As alluded to earlier, assessment quality was determined by measuring the assessment papers' content validity, alignment with course objectives, cognitive levels of the items, and their authenticity.

4.1. Content validity of online Biodiversity summative assessments

The content validity of the summative assessments was determined to find out whether the assessments had a good representation of all the Biodiversity constructs prescribed for the courses (Webb, Katz & Decker, 2006). To determine the content validity of the assessments, prescribed topics for both first-year and Honours Biodiversity courses were listed, and items from the respective summative assessment were matched against these topics. The assessment items were then categorized as either 'based on' or 'not based' on the stipulated topics. Percentages of items that were 'based on' the prescribed topics were determined and converted into content validity indices (CVI). The results of these computations are presented on Table 1 below.

Table 1. Content validity of online Biodiversity summative assessments.

Year	Assessment type	First-year				Honours					
		*T 1 (%)	T 2 (%)	*NA (%)	CVI*	T 1 (%)	T 2 (%)	T 3 (%)	T 4 (%)	NA (%)	CVI*
2020	Main	49	48	3	0.97	87	0	0	0	13	0.87
	Deferred	39	47	14	0.86	100	0	0	0	0	1
2021	Main	69	25	6	0.94	75	17	0	8	0	1
	Deferred	69	21	10	0.9	75	0	0	0	25	0.75

*CVI - represents Content Validity Index *T - represents Topic *NA - represents 'Not aligned' to any topic

The results on Table 1 show that all the online Biodiversity summative assessments were content valid, as they had CVIs of more than 0.7, which is considered to be high (Polit & Beck, 2006). However, the Honours assessments did not have a good coverage of the prescribed topics, as the assessment items were mainly based on one topic out of four: Biodiversity and climate change.

4.2. Alignment of online Biodiversity summative assessments with course objectives

The alignment of summative assessment items with the course objectives was determined by listing the Biodiversity course objectives at first-year and Honours levels. Coincidentally, both courses had the same number of course objectives, but differed in focus. Each assessment item was assigned to respective course objectives. However, some assessment items aligned with more than one objective, and such items were counted more than once, when computing alignment percentages. The results are displayed on Table 2 below.

Table 2. Alignment of online Biodiversity summative assessments with course objectives.

Educational level		First-year						Honours						
Year	Examination Type	*O1	O2	O3	O4	O5	O7	O1	O2	O3	O4	O5	O6	O7
2020	Main	60	20	20	0	0	0	56	33	0	0	0	0	0
	Deferred	60	20	20	0	0	0	63	25	12	0	0	0	0
2021	Main	80	16	2	2	0	0	50	17	0	17	0	8	8
	Deferred	83	24	0	0	0	0	25	0	0	50	0	25	0

*O - represents Objective

Table 2 shows that first-year assessments were related to three objectives out of seven. These three objectives required students to recall learnt information. Honours assessments focused mainly on objectives 1 and 2, which required students to describe and to explain the importance of Biodiversity concepts.

4.3. Cognitive levels of online Biodiversity summative assessment items

The revised Bloom's cognitive levels of remember, understand, apply, analyse, evaluate, and create (Anderson & Krathwohl, 2001), were used as the framework for determining the cognitive levels of the reviewed assessments. Demonstrative verbs suggested for each cognitive level by Anderson and Krathwohl (2001) were used to identify the cognitive levels of the assessment items. Items which related to more than one cognitive level were counted more than once, when computing the percentages of items pitched at a particular cognitive level. The results are presented in Table 3.

Table 3. Cognitive levels of online Biodiversity Summative assessments.

Year	Exam type	Remember (%)		Understand (%)		Apply (%)		Analyse (%)		Evaluate (%)		Create (%)	
		Hons	1 st year	Hons	1 st year	Hons	1 st year	Hons	1 st year	Hons	1 st year	Hons	1 st year
2020	Main	57	87	43	13	0	0	0	0	0	0	0	0
	Deferred	29	100	57	0	0	0	14	0	0	0	0	0
2021	Main	61	87	31	13	0	0	8	0	0	0	0	0
	Deferred	76	75	8	25	0	0	8	0	0	0	8	0

The results show an almost consistent pattern of assessing the lower order thinking skills of remembering and understanding, at both first-year and Honours levels of study. Only a few assessment

items were pitched at analysis (8% and 14%) and creation (8%) cognitive levels, in the Honours summative assessments.

4.4. Authenticity of online Biodiversity summative assessment items

Wiggin's 1998 framework for authentic assessments was used to categorize assessment items as either authentic or non-authentic. For an assessment to be classified as authentic, more than 50% of its items had to be authentic (Frey, et al., 2012), based on the authenticity indicators of Wiggin's framework (Wiggins, 1990). Each item in the different assessment papers was analysed to determine whether it had indicators of authenticity or not. The percentages of assessment items classified as authentic were determined and are presented in Table 4.

Table 4. Authenticity of online Biodiversity summative assessments.

Year	2020		2021	
Exam type	Main	Deferred	Main	deferred
1 st year	6	0	0	0
Honours	88	71	75	75

The results of the review show that all first-year assessments had low or no elements of authenticity, while all Honours assessments were moderately authentic (average of 77% authenticity).

5. Discussion

The study reported in this paper investigated the quality of online Biodiversity summative assessments administered to first-year and Honours student teachers at a South African university. The first finding of the study was that all the reviewed summative assessments were of high content validity, as they had CVIs of more than 0.7, which is an indication of high content validity. However, there was poor coverage of all the prescribed topics for Honours course. Similarly, a study conducted by Gareis and Grant (2015) to determine the content validity of educator made assessments in Malaysian found that even though assessments had high CVIs, assessment items were not evenly distributed across the prescribed content. This is particularly common in assessments of a topic such as Biodiversity, which is complex (Sadler et al, 2013). This means that students might pass a course without fully understanding the entire course content. Lack of time and expertise by assessors could account for the uneven distribution of summative assessments items across prescribed content (Webb, Katz & Decker, 2005).

The second finding was that both first-year and Honours summative assessments were highly aligned to course objectives. Nonetheless, assessment items were mostly aligned to three objectives out of seven, for first-year assessments, and to two objectives out of seven, for Honours assessments. A study conducted by Smith (2012) found that most online assessments were not aligned to course objectives. The reason for the misalignment of assessment items with course objectives could be that some objectives require assessment items to be pitched at higher order thinking levels, and it might be difficult for some assessors to design such items. Moreover, some assessors might ignore course objectives when designing assessments because it takes time to align assessment tasks to courses objectives.

The third finding relates to the cognitive levels of the reviewed summative assessments, which showed that the assessments were mostly pitched at lower cognitive levels of blooms' taxonomy; 'remembering' and 'understanding' levels (Anderson & Krathwohl, 2001). Similar findings have been found in several studies (Kabombwe et al, 2021; Shukla & Dungsungnoen, 2016). This finding is concerning because it implies that the quality of assessments is compromised (Anees, 2017), in the sense that students could achieve high scores from summative assessments but lack higher order thinking skills, which are essential for solving problems and making rational decisions in society and in workplace (Junus, et al., 2021). Shukla and Dungsungnoen (2016) suggested that most lecturers understand the importance of higher order thinking skills, but they do not know what they entail or how to integrate them in assessments. Low order thinking assessments administered online could be prone to cheating by students because they do not require individual engagement (Junus, et al., 2021).

The fourth finding revealed varying trends in the authenticity of first-year and the Honours assessments, in that first-year assessments had very low authenticity, while Honours assessments had moderate authenticity. The implication of this finding is that online Biodiversity summative assessments administered to first-year did not prompt students to apply knowledge to real-life situations. This shortcoming could be a result of assessors' lack of understanding of authentic assessment (Frey, et al., 2012), which could prevent them from planning for authentic assessments consistently. Furthermore, authentic assessments are usually open-ended questions that require higher order thinking ability, which some lecturers find difficult to assess (Shukla & Dungsungnoen, 2016).

6. Conclusion and recommendations

The study was set out to answer the question: What is the quality of online Biodiversity summative assessments administered to student teachers at a South African university? The findings showed that the reviewed online summative assessments had high content validity and alignment to course objectives, with poor coverage of the content and course objectives respectively. In addition, the assessments were pitched at low cognitive levels and they had low to moderate authenticity. Based on these findings, it could be concluded that the reviewed online Biodiversity summative assessments were of moderate quality. This finding is concerning because online summative assessments are prone to cheating because they are usually not invigilated and it is easy for students to access sources of information, such as the internet and other learning resources during the assessment (Junus, et al., 2021). Moderate quality Biodiversity assessments could adversely affect the Quality of Life Sciences graduates and subsequently the efficiency of the labor force in life sciences related professions. We recommend further studies to establish the recurrence of the findings from this study. We further recommend that lecturers in higher learning institutions need to revisit their assessments practices to ensure that they develop assessments that have good coverage of prescribed course content, course objectives, are pitched at higher cognitive levels, and that they are authentic.

References

- Anderson, L. W, & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives* (Complete Ed.). New York: Longman.
- Anees, S. (2017). *Analysis of Assessment Levels of Students' Learning according to Cognitive Domain of Bloom's Taxonomy* (ED586762). Institute of Agricultural Extension and Rural Development. ERIC. <https://files.eric.ed.gov/fulltext/ED586762.pdf>
- Frey, B. B., Schmitt, V. L., & Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research, and Evaluation*, 17(2), 1-18.
- Gareis, C. R, & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and Student Learning*. (2nd Ed.). UK: Routledge. <http://dx.doi.org/10.4324/9781315855240>
- Educational network LINQED. (2011). *Quality assurance framework for student assessment – Guidelines for the design and implementation of effective student assessment*. Retrieved, July 13 2022 from: http://www.linqed.net/media/18640/quality-assurance-framework-for-student-assessment-english-proofed-reading_final.pdf
- Junus, K., Santoso, H. B., Putra, P. O. H., Gandhi, A., & Siswantining, T. (2021). Lecturer readiness for online classes during the pandemic: A survey research. *Education sciences*, 11(3), 1-14.
- Kabombwe, Y., Machila, N., & Sikayomya, P. (2021). A Comparative Analysis of the Zambian Senior Secondary History Examination between the Old and Revised Curriculum using Blooms Taxonomy. *Yesterday & Today*, 25, 1-25. DOI: <http://dx.doi.org/10.17159/2223-0386/2021/n25a2>
- Oyinloye, O. M., & Imenda, S. N. (2019). The Impact of Assessment for Learning on Learner Performance in Life Science. *Eurasia Journal of Mathematics, Science and Technology Education*, 15(11), 1-8. <https://doi.org/10.29333/ejmste/108689>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, 29(5), 489-97.
- Sadler, P. M., Coyle, H., Smith, N. C., Miller, J., Mintzes, J., Tanner, K., & Murray, J. (2013). Assessing the life science knowledge of students and teachers represented by the K–8 national science standards. *CBE — Life Sciences Education*, 12(3), 553-575.
- Shukla, D., & Dungsungnoen, A. P. (2016). Students' Perceived Level and Teachers' Teaching Strategies of Higher Order Thinking Skills: A Study on Higher Educational Institutions in Thailand. *Journal of education and Practice*, 7(12), 211-219.
- Smith, R. J. (2012). *Alignment of intended learning outcomes, curriculum and assessment in a middle school science program* [Master's thesis, Edith Cowan University]. <https://ro.ecu.edu.au/theses/489>
- States, J., Detrich, R., & Keyworth, R. (2018). *Overview of Summative Assessment*. Oakland, CA: The Wing Institute.
- Webb, V. J., Katz, C. M., & Decker, S. H. (2006). Assessing the Validity of Self-Reports by Gang Members: Results from the Arrestee Drug Abuse Monitoring Program. *Crime & Delinquency*, 52(2), 232-252. <https://doi.org/10.1177/0011128705277972>
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2), 1-3. <https://doi.org/10.7275/ffb1-mm19>