

ESSAY ITEM CONSTRUCTION IN ODL: CHALLENGES AND FUTURE DIRECTION

DR. KUREBWA MERCY*

MRS NGARA ROSEMARY*

Abstract

Student assessment is critical in Open and Distance learning. Essay items are used in most disciplines at Zimbabwe Open University (ZOU) to assess students formatively and summatively. This study sought to investigate the effectiveness of essay test construction and make recommendations on how to improve essay construction in ODL. A case study was conducted at ZOU Midlands region to determine lecturers' and students' opinions on the process of item essay construction, their quality and challenges faced in whole exercise. Commensurate with the qualitative research approach, focus group interviews and, individual interviews were employed to collect data from the participants. Documents that included examination reports, assignments and examination items were perused to triangulate data from participants. Thematic analysis was used to analyse the data. The study established that the process of constructing essay items was viewed as satisfactory since initial items were moderated at some stage in the process but the exercise did not wholly perfect the items. The participants opined that in some instances principles of item construction were followed, while in others they were not. General weaknesses were attributed to some factors including lack of proper lecturer education in essay item construction at regional and national levels. The study recommends staff development of lecturers in item construction to boost lecturers' competencies, among other things. Furthermore lecturers without an educational base need to be encouraged to do (PGDE) Post Graduate Degree in Education to equip them with the requisite skills.

Key words:-assessment, essay item, item construction, Open and Distance Learning,

* Zimbabwe Open University

Introduction

Assessment is the backbone of teaching and learning. Lecturers and students would want to gauge their students' performance through assessment. At Zimbabwe Open University students are assessed through a variety of means. Formatively they are assessed through coursework. Depending on expectations by different faculties and departments, coursework may be assessed via essays, objective type items, in class tests and practically as is the case with student practicum or Teaching Practice. Examinations are used to assess students summatively. In all the University faculties, the modular approach is used and coursework is done per module, per semester and so are the exams that are written at the end of each semester. Assessment in most cases requires the construction of relevant items. Essays are most commonly used method of assessing student learning at the Zimbabwe Open University. Although essay questions are one of the most commonly used methods for assessing student learning, many are poorly designed and ineffectively used (Reiner et al, 2002). At the Zimbabwe Open University essay items are sometimes not constructed in a scholarly way. It was against this background that this study was carried out.

Conceptual Framework

Essence of an Item

An item colloquially referred to a test question though in several cases it may not be an actual question. So an item is some task based on teaching objectives which is designed to assess learner competence. Items can be essays, objective or practically-oriented tasks.

Steps in developing items

According to General Educational Testing Service (2002), while different testing programs are unique, there are some common steps to be followed among which are the following:

- Deciding how the test should be designed and what should be assessed. This leads to test and item specification – these describe the test as a whole, its purpose and numbers of items, kinds of items and how the test will be scored.
- Creation of items
- Internal review of items

Items are reviewed for their match to standard, content, accuracy, relevance, grade level appropriateness, any bias, sensitivity concerns and editorial considerations (Stone, 2001 and

Haladyna & Downing (1989:2). At this stage use of checklists is critical (Mallis & Martin, 2011). Revision of items includes,

- external review for same concerns as in the internal review;
- Field testing of item. Items are administered to get data on each item to help determine which items worked well and
- Creation of operational items which may be incorporated in the bank (Store, 2001).

Keys to writing winning essay items

According to Case and Swanson (2002) there are principles that should be followed in writing essay items and some of them are:-

- Questions should be framed so that learners' task is explicitly defined. the value and time limit for each question should be specified
- Sample items should be representative of instructional objectives
- Verification of a question's quality by writing a trial response to the question and
- Envisioning the rubric.

Review of Related Literature

A study titled, 'The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education' was conducted by Steven Downing (2005). The purpose of this study was to examine the effects of violations of standard multiple-choice item writing principles on test characteristics, student scores, and pass-fail outcomes. Four basic science examinations were administered to year-one and year-two medical students after the students had been sampled randomly. Test items were grouped as either standard or flawed by three independent raters, blinded to all item performance data. Flawed test questions violated one or more standard principles of effective item writing. Thirty-six to sixty-five percent of the items on the four tests were flawed. Flawed items were 0-15 percentage points more difficult than standard items measuring the same construct. Over all four examinations, 53% of the students passed the standard items while 47% passed the flawed items. The median passing rate difference between flawed and standard items was 3.5 percentage points, ranging from one to 35 percentage points. Findings showed that item flaws had little effect on test score reliability but disadvantaged some medical students. Item flaws introduce the systematic error of construct-irrelevant variance to assessments, thereby reducing the validity evidence for examinations and penalising some examinees.

Although the study involved medical students, it is clearly linked to the current study whose focus is on participants' views on the quality of essay items.

Statement of the problem

There seemed to be some lack of scholarly attention paid to the construction of essay items, both for coursework and exams.

Objectives of the study

To establish how items are generated

To establish perceived weaknesses in item construction

To suggest possible solutions to the problem of item construction

Research questions

The study was guided by the following questions:

- How are item for examinations and assignments constructed?
- What are the perceived weaknesses of the essay items?
- What can be done to improve essay item construction?

Research Methodology

The study was principally qualitative in nature. According to Cohen & Manion, (2007) qualitative approach to research is an approach that advocates the study of direct experience at face value. Qualitative researchers see behaviour as determined by the phenomena of experience rather than by external, objective and physically described reality. For Creswell (2007) qualitative research allows the inquirer to develop a composite description of the essence of the experience for all of the individuals involved. In tandem with the qualitative research approach data were collected through document analysis and in depth face to face, telephone interviews. The documents interrogated were examination and assignment question paper and regional examination reports.

Data collection

Given that more and more researchers seem to come to the conclusion that “it is better to select a variety of research methods when approaching a problem, rather than relying on a single method (Creed et al 2004:60), the study used document analysis, individual and focus group

interviews to collect data. Mixing data sourced through different methods enhances participants' diverse viewpoints that cast light upon a topic (Sayer, 2000) The interviews were conducted with ten students and eight lecturers. These participants were purposefully selected. A student focus group of eight (8) was also conducted.

A total of 40 assignment and examination items were drawn from all the four faculties. These items were for the years 2011 to 2014. Examination reports for the same period were also interrogated.

Data analysis techniques

Data from in-depth interviews were thematically analysed, commensurate with qualitative methods of summarising narrative data. According to Henning et al (2004) data analysis takes place throughout the data collection process, commencing with reading all data and dividing it into smallest meaningful units. Cohen and Manion et .al (2007) sees this as a process of organising of data under headings to conclusive analysis of the ways in which the data contributes to the research question. Data analytic strategies Bogdan and Biklen (2003:145) define qualitative data analysis as “working with data, organizing, breaking it into manageable units, synthesizing it, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others”. Data were coded to come up with categories and themes to organize research findings.

Discussion of findings

The overall goal of the study was to establish the effectiveness of essay construction in ZOU and accordingly make recommendations to improve the quality and nature of item construction. Generally, observations from document analysis and responses from interviews indicated that challenges were being experienced in item construction. The findings were organized using predetermined thematic categories associated with the questions and then re-analysed it using the themes that emerged from the data.

Setting of assignments in ZOU

The study revealed that in the in the Zimbabwe Open University lecturers are allocated different areas in which to set items. Assignment and examination items are then forwarded to the chair person who works with the programme coordinators to select examination and assignment questions from a pool of submitted items. Apparently, during the interviews with lecturers it emerged that most (70%) of the lecturers had not done a course in item writing. When asked on how they learnt how to construct items the following responses were advanced;

“I had limited training while I was in the schools. I attended a one day workshop”

“I have never been trained in item writing but I have learnt how to set items through reviewing past examination papers until I gained the relevant experience.”

The above comments tend to suggest that some of the lecturers were never trained in item writing and might not have the relevant competences. The other lecturer agreed with this view and commented that, *“I did my first degree, master’s degree and currently on doctorate studies but in all these degrees I never met a component of item writing.”* These findings are commensurate with Kurebwa’s (2012) (unpublished thesis) which revealed that, some teachers and college lecturers lacked competences in item writing due to lack of appropriate training in the field. Such lack of competences might lead to poorly constructed items in the Zimbabwe Open University.

Perceived weaknesses of essay items in ZOU

Typographic , grammatical and use of wrong word errors

The responses from participants indicated that assignment and examinations items for each semester have errors as supported by some of the documents that were perused. The errors were said to include typographic errors, grammatical errors, missing information and the use of wrong words. The following errors were noted, the use of *talk-shop instead of talk-shore, ruin instead of run, minutes instead metres, major instead of majority and factor-factor relationship instead of factor–sector relationships*. All (100%) examination reports reported on these errors and recommended the need for their attention in future examinations. In one examination session a total of (10)ten such errors were revealed. Many such errors occurred in examination and assignment questions and distorted the assignment and examination questions. One student lamented on the inconveniences caused by such a scenario and commented, *“I collected assignment questions during registration and never got a chance to have the assignments with errors changed. In rectifying the errors some questions were changed and I was not informed”*. This tends to suggest that errors of that nature distort the context and meaning of questions and fail to solicit for the needed knowledge from students. Cognisant of the tone of the student, such a situation may de-motivate the student leading to dropouts.

Limited sampling

Limited sampling was perceived by lecturers as another challenge in essay item construction. It emerged from the study that in one set of assignments or examination paper items were set from one chapter of the module. Furthermore some essay items were copied from the module to assignments and sometimes the same question is repeated in the examination. This exacerbates

one of the limitations essay items notably, poor content sampling in essay questions. The content needs to be representative of the skills that are supposed to be measured, that is, test objectives, syllabus content, and the test contents (Maizan, 2005). Content covered and the cognitive or skill should conform with the syllabus (Black, 1997:47).

Duplication of assignment questions in examinations

The study also established some assignment questions were duplicated following the same sequence and verbatim in some examination papers. One student commented that, *“It has become very easy to spot examination questions especially after studying the trend of assignment and examination item writing in our institution”* “Therefore if assignments are duplicated, the students are likely to produce what they presented during assignment writing without being challenged. Furthermore content sampling in both assignments and examinations will be further minimized.

It is also disturbing to note that some examination papers are recycled verbatim and in the same sequence with changes in dates only. One student regretted ignoring a previous examination paper which she met in the examination. She commented, *“I think all the students who revised the paper performed very well. I never thought this paper would be brought back into the examination.”*

Questioning Techniques

In addition to duplication of assignments as a challenge in item construction, further challenges were also mentioned. These included poor questioning techniques. The study revealed that in some assignment and examination papers all questions or three (3) out of four (4) questions required pupils to discuss, compare and contrast among other forms of verbs that were repeated. Varying questioning techniques is likely to enable students to view content from a different perspective (Popham, 2004).

The study also revealed that some essay questions were sometimes double barreled and some were asked in very general terms and were prone to many interpretations.

It was also observed that some essay questions demanded simple recall. Such essay questions limited students in acquiring deep and elaborate understanding of the subject matter. In setting such items, consequently lecturers too often omit important curricular outcomes, in their place, test the trivial instead of items that reveal depth and understanding.

Other challenges

Further findings of this study indicated that ,some students had problems with tackling assignments and examination because the literature set books were no longer in circulation and the students had no where to find them. The quoted books included “Waiting for the rain, Why don’t we curve the other animals, Fine madness and Caribbean literature. One student who participated in the interviews had this to say,

“I don’t know why questions are asked from these texts that are no longer in circulation. I went to bookshops and I was told that the books were out of print .I end up ignoring questions from such books”

A lecturer in the department agreed with the student and revealed that some set books were out of supply. The response suggested that some irrelevant essay items were set. Such items, no matter how well they were constructed failed to serve the purposes for which they were constructed.

Another finding was that, some questions did not match the course outlineand that they deviated from the module. These findings emanated from interviews and were also supported by documents. Such items were likely to defeat the whole purpose of assessment since the module contains the course outline. Reliable and timely assessment is key to improving learning and enhancing quality of education (UNESCO, 2005) “Improvement in assessment will depend on how well assessment and curriculum and instruction are aligned (Pellegrino et al 2001:25).”

Conclusion

The study sought to investigate essay construction challenges in ODL. Generally it was observed that essay construction is fairly implemented because of some challenges that were observed. The views from participants indicated that some challenges were encountered during essay construction. It emerged that assignments were presented with;

- typographic , grammatical and use of wrong word errors;
- limited sampling which is exacerbated by setting two or more items from a single section of the module and
- poor questioning techniques which included asking double barreled questions, general questions ,repetitive use of verbs in one set assignments as well as asking simple recall questions which fail to adequately challenge student’s faculties at university level.

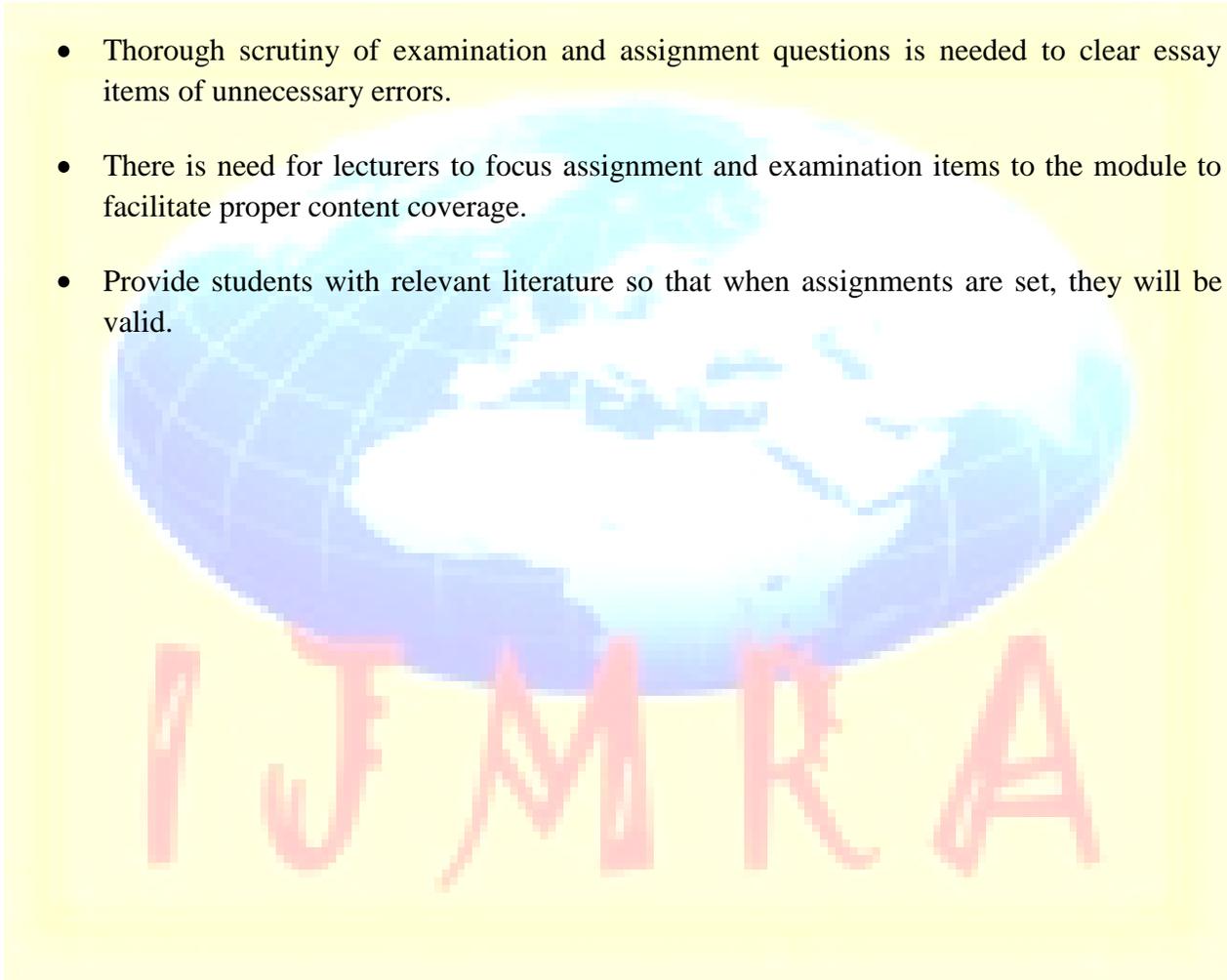
Some questions were also set from unfounded sources which rendered them irrelevant. The study also concluded the errors that were evident in the presented assignments and examination

questions might be because some lecturers lacked relevant competences in item construction. Furthermore, there might be a cursory approach to quality control the set items in the institution.

Recommendations

The following recommendations were made in response to the findings highlighted in the study.

- There is need for the University to have staff development programmes for lecturers on item construction, so that lecturers are equipped with relevant skills.
- Thorough scrutiny of examination and assignment questions is needed to clear essay items of unnecessary errors.
- There is need for lecturers to focus assignment and examination items to the module to facilitate proper content coverage.
- Provide students with relevant literature so that when assignments are set, they will be valid.



References

Bogdan, R. C & Biklen, S. K. (2003). *Qualitative Research for Education: An introduction to Theories and Methods* (4th ed.). New York: Pearson Education group .

Case, S.M and Swanson, D.B .(2002). *Constructing Written Test Questions for the Basic and Clinical Sciences* (3rd .ed) U.S.A NBME.
www.udel.edu/educ/got/ accessed 20/2/2014.

Cohen, L & Manion, L.(2007) *Research Methods in Education* (6th Edition). Amazon.co.uk.

Creed, C .(2004) *Practitioner Research and evaluation skills Training in Open and Distance Learning* .Commonwealth of Learning.

Creswel, J .W(2007) *Qualitative Inquiry & Research Design. Choosing among five approaches*. 2nd Ed. California: Sage.

Dawning , S.M (2005). *The effects of violating Standard Item Writing Principles on tests and students: The consequences of using flawed Test Items on achievement Examination in Medical education in Advances in health Science Education* 10:133-143.

Educational Testing Service K-12 assessments: (2002) Guidelines for item writing. Multiple choice item writing rules, Applied measurement in Education.

Henning, E, Van Rensburg, W & Smit, B. (2004). Finding your way in Qualitative Research. Van Schaik: Pretoria.

Kurebwa, M. (2012) Assessment Problems in the Primary Schools with special reference to Gweru District schools. (Phd Thesis, Unpublished).

Maizan Alias (2005). *Assessment of learning outcome: Validity and reliability of classroom tests.* World Transaction on Engineering and Technology education. Vol.4 No. 2

Pellegrino, J., Chudowsky, N., and R. Glaser (eds.). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment.* Report of the National Research Council, Center for Education, Division on behavioral and Social Sciences and Education. Washington, D. C.: National Academy Press. Can be read on-line at:
<http://www.nap.edu/catalog/10019.html>.

Popham, W. J. (2004). *Classroom assessment: What teachers need to know* (4th ed.). Boston, MA: Allyn & Bacon.

Reiner, C.M., Bothgwe T. W; Sidweeks, R.R. & Wood, B. (2002) *Preparing*

effective essay questions: A self directed workbook for educators .New

Forums Press

Sayer,A.(2000).*Realism and Social Science*.London:Sage

Stone , G.E (2001) *Professional item writing for High stakes Testing*,

Matriks Consulting Ltd U.S.A

Timss (2011) Trends in International Maths and science study. Item

writingguidelining by Mullis, I.V.S and Martin, M.O Timss and

PIvRIS International study centre ,BOSTON College

UNESCO(2005),FAGlobal monitoring

report.[http://portal.unesco.org/education/en/ev.php_URLID=34742&U](http://portal.unesco.org/education/en/ev.php_URLID=34742&URL_DO=DO_TOPIC&URL_SECTION=201.html)

[RL_DO=DO_TOPIC&URL_SECTION=201.html](http://portal.unesco.org/education/en/ev.php_URLID=34742&URL_DO=DO_TOPIC&URL_SECTION=201.html)

[www.assess.com/docs/ASC Item writing pdf](http://www.assess.com/docs/ASC_Item_writing_pdf) accessed 2/01/2014

scoring.msu.edu/ accessed 2/12/2014