

NIGERIAN OPEN AND DISTANCE LEARNING LECTURERS' DIFFICULTY IN CONSTRUCTING STUDENTS' TEST ITEMS

Johnson Ayodele Opatеye

Department of Educational Foundations

Faculty of Education

National Open University of Nigeria, Lagos

jayopateye@yahoo.com

ABSTRACT

This paper examines the level of difficulty lecturers experience in constructing test items for various types of tests based on gender and institutional mode of delivery in an Open and Distance Learning higher education context. Descriptive cross-sectional survey research was adopted as the research paradigm, using lecturers from one single mode and one dual mode ODL institution in South West Nigeria as the study population. Stratified simple random sampling techniques were used to select 240 lecturers as the sample. Three research questions and two hypotheses guided the study's investigation. A questionnaire with a reliability coefficient of 0.78 using Cronbach's alpha value was used as the research instrument. Frequency, percentages, t-test statistics and charts were utilised to analyse the data. The results showed that ODL lecturers exhibited moderate levels of difficulty in test item construction. Female ODL lecturers found case study, multiple choice, matching, essay, and completion items more difficult to generate than their male counterparts did. There was a significant difference in the difficulty male and female ODL lecturers experienced in constructing test items. Significant differences were also found in the difficulty experienced by single and dual mode ODL lecturers in construct test items. The findings of the study have implications for the capacity development of lecturers, ODL institution management, and test experts in order to improve lecturers' capacities in test items construction.

Keywords: Delivery mode; difficulty; gender; lecturers' capacity; Open and Distance Learning; test item construction.



1. INTRODUCTION

Open and Distance Learning (ODL) is increasingly accepted globally, with many learners seeking to obtain diploma and degree certificates via media that connect them to their instructors. The characteristics of flexibility and accessibility have given credence to Open and Distance Learning. The Commonwealth of Learning (2000) describes distance education as a mode of delivering education and instruction, often on an individual basis, to students who are not physically present in a traditional setting such as a classroom. Distance learning provides access to learning when the source of information and the learners are separated by time and distance, or both (Oblinger 2000). The Commonwealth of Learning (2000) identifies four broad rationales for embracing distance education, namely expanding access, alleviating capacity constraints, capitalising on emerging market opportunities, and serving as a catalyst for institutional transformation. A variety of terms describe the type of educational provision that involves some form of an open and distance learning approach and uses open and distance learning techniques to a greater or lesser extent. Single mode institutions are set up to offer programmes of study at a distance, with teaching and learning mediated in some way (for example Indira Ghandi National Open University (IGNOU) and National Open University of Nigeria (NOUN)). Dual mode ODL institutions combine traditional classroom-based (face-to-face) and distance methods to deliver academic programmes to learners (UNESCO 2002). Such universities include the University of Botswana, University of Ibadan and University of Nairobi.

Designing tests is an important part of assessing students' understanding of course content and their level of competency in applying what they are learning. Classroom assessment is an integral part of teaching (Popham 2002; Linn and Miller 2005) and may take up more than one third of a teacher's class time (Ainsworth and Viegut 2006). Most classroom assessment involves tests that teachers have constructed themselves. Barton (2002) reports that teacher-made tests are used in a typical classroom, resulting in perhaps billions of unique assessments yearly worldwide. Teachers regularly use tests they have constructed themselves (Angelo and Cross 2008). Furthermore, teachers put greater store by their own tests in determining grades and student progress than they do by assessments designed by others or by other data sources. Most teachers believe that they need strong testing skills, while some report that they are confident in their ability to produce valid and reliable tests and others report a level of discomfort with the quality of their own tests or believe that their training was inadequate (Kifer 2001).

A manual for testing produced by the University of Washington (2015) gives some guidelines that could aid test-item generators in test construction. Some of these general guidelines are: considering reasons for testing; maintaining consistency between goals for the course, methods of teaching, and the tests used to measure achievement of goals; using testing methods that are appropriate to learning goals; helping students prepare; using consistent language (in stating goals, in talking in class, and in writing test questions) to describe expected outcomes; and designing test items that allow students

to show a range of learning. Devine and Yaghlian (2000) identify the three stages in test construction as planning for the test, preparing the test, and analysing and revising the test. Guskey (2003) also reports that planning for the test involves outlining subject-matter content, identifying learning outcomes, preparing a table of specifications, and choosing appropriate test types. Test preparation activities involve writing the test items according to rules, selecting items to be included, arranging items, and preparing directions for the test. In order to analyse and revise test construction activities, one needs to perform test analysis so as to retain or discard of test items.

Multiple choice questions can be difficult to construct, especially if a teacher wants students to go beyond the mere recall of information, but such examinations are easier to grade than essay or short-answer examinations. On the other hand, multiple choice examinations provide less opportunity than essay or short-answer examination for a teacher to determine how well the students can think about the course content or use the language of the discipline in responding to questions (Ayodele, Adegbile and Adewale 2003). Essay tests ensure that students display their overall understanding of a topic and demonstrate their ability to think critically, to organise their thoughts, and to be creative and original. While essay and short-answer questions are easier to design than multiple-choice tests, they are more difficult and time-consuming to score. Moreover, essay tests can suffer from unreliable grading; that is, grades of the same response may vary from reader to reader or from time to time by the same reader. For this reason, some faculties prefer short-answer items to essay tests. On the other hand, essay tests are the best measure of students' skills in higher-order thinking and written expression.

Test constructors need to have knowledge and understanding of the materials being tested, have continuous awareness of objectives and understanding of the learners for whom the items are intended. Test item writers should possess skills in written communication and the techniques of item construction. Greiff, Holt and Wüstenberg (2013) assert that student interaction with a case study test item could be interactive or non-interactive. A non-interactive case study test item form is one in which the test-taker is required to construct at least one test item to solve the problem and provide a supporting explanation for the proposed problem. An interactive case study test item is one in which the test taker may request additional information or clarification of the information presented in the study. The test taker is asked to construct a solution to the problem and provide a simple explanation of the proposed solution.

The formal assessment training teachers receive often focuses on large-scale test administration and standardised test score interpretation rather than on the test construction strategies or item-writing rules that teachers need (Izard 2005). The current empirical research literature for item-writing guidelines focuses on studies which look at the relationship between a given item format and either test performance or psychometric properties of the test (related to the format choice). There are some guidelines supported by experimental or quasi-experimental designs, but the foundation of best practices in this area remains, essentially, only recommendations of experts.

Ramos-Mattoussi and Milligan (2013) give guiding principles for lecturers' capacity building and emphasise that lecturers' industry engagements must produce benefits for lecturer, college and employer; lecturers are responsible for industry engagements as part of their own professional development; the nature and length of engagements should fit their purpose; industry engagements must not disrupt teaching, and lessons from industry engagements must be systematically integrated into teaching. ODL lecturers should thus be developed by integrating lessons from industry and revising (old) lesson plans to incorporate real-life examples and case studies and also by supplementing the core curriculum with content not prescribed but relevant to the industry, assigning students activities like those in the industry, assessing performance by industry standards and simulating the workplace in the college, e.g. in layout of workshops.

Researchers have discovered gender differences in constructed-response and multiple-choice assessments in mathematics and found that males outperformed females in problem solving (Dayioglu and Turut-Asik 2007). Differences generated by the broad and narrow interpretations of problem solving show some gender differences that might not have been apparent otherwise. Gender disparity in schooling is also observed among the younger population, where female school enrolment in basic and secondary education lags behind that of male children (Tansel 2002). Young and Fisler (2000), examining SAT-M scores of high school seniors, find males to score better than females. However, they note that males generally come from households where the parents' socio-economic status, as measured by examinee-reported educational levels and income, is higher. In contrast, female test-takers are more socio-economically diverse and include more low-income students than the male group.

The effect of gender differences on cognitive abilities has been under continued scrutiny by researchers. Studies have attempted to explain such differences by investigating genetic and hormonal factors, neuroanatomical functions and environmental influences (education and social class). However, the results have been mixed, indicating the need for a psychobiosocial approach to understanding gender differences in intellectual ability (Doherty, Kovas and Plomin 2011; Nisbett, Aronson and Blair 2012). If these differences do exist, then it is unclear when they emerged and how they developed. Confirmation of the existence or absence of gender differences in cognitive abilities is crucial to political decision making and has a wide range of implications for public policy (Halpern and LaMay 2000). If intellectual gender differences are confirmed, then a full array of psychological and educational tests must be revised to avoid bias in specific items or activities to ensure the fair application of test results (Van de Vijver and Leung 2000). As yet, there exists no evidence on gender differences in test item construction among lecturers in higher institutions, which is one of the focuses of this paper.

Koksal (2004) also noted that, in test construction, it is essential that the lecturer asks the following questions: Is the task perfectly clear? Is there more than one possible

correct answer? Can test-takers arrive at the correct response without having the skill supposedly being tested? Do test-takers have enough time to perform the task(s)?

The abilities of test-takers are measured by the type of responses given to these questions. Lecturers in higher education institutions are not all trained to assess students' learning outcomes and may find test item construction difficult. The level of difficulty experienced by lecturers in constructing test items for Open and Distance Learning students should be investigated so as to suggest areas to help them improve their capacity in this regard. Therefore, there is a need to establish the level of difficulty experienced by ODL lecturers in constructing test items for students.

2. OBJECTIVES OF STUDY

The following objectives were formulated for this study:

- To ascertain the level of difficulty experienced by Open and Distance Learning lecturers in constructing test items.
- To determine which types of test items ODL lecturers find difficult to construct.
- To examine the difference in difficulty experienced by male and female ODL lecturers in generating test items.
- To examine the significant difference in difficulty experienced by single and dual mode distance education lecturers in constructing test items.

3. RESEARCH QUESTIONS

The following research questions guided the study:

1. What is the level of difficulty ODL lecturers experience in constructing test items?
2. What types of test items do ODL lecturers find difficult to construct?
3. Which of the test types do ODL lecturers find difficult to construct based on gender?
4. To what extent do lecturers from single and dual mode ODL institutions differ in the level of difficulty experienced in generating test items?

4. HYPOTHESES

Two hypotheses were tested in the study:

1. There is no significant difference in difficulty to construct test items between male and female ODL lecturers.
2. There is no significant difference in difficulty to construct test items between single and dual mode ODL lecturers.

5. METHODOLOGY

A descriptive cross-sectional survey research design was used for the study. The population is comprised of lecturers from distance education institutions in South Western Nigeria. The country is divided into six geopolitical zones. Other zones are: South East, South South, North East; North West and North Central. South West was selected because the only single mode Open and Distance Learning institution is located in this zone. The researcher also resides in this zone, which has six states; the zone also has the highest number of universities in Nigeria. Out of four official distance education institutions in this region, two were used as part of this study. One (single mode) was purposively selected since it is the only single mode ODL institution (NOUN), while the other was selected from the three dual mode institutions by using simple random sampling. The sample consisted of 240 lecturers selected by using stratified and simple random sampling techniques. Data were collected using a questionnaire with two sections. Section A of the questionnaire contains questions related to the bio-data of lecturers like gender, academic rank, mode of institution, and area of discipline. Section B contains questions on the level of difficulty experienced in constructing six test type items and three testing modes. The response is in a modified three Likert scale format of Very Difficult (1), Difficult (2) and Not Difficult (3). As this instrument was constructed by the researcher, it was given to a test and evaluation expert for validation. The comments made by the expert were utilised to improve the quality of the instrument. The corrected instrument was pilot tested on lecturers at the University of Ibadan Distance Learning Centre. The reliability coefficient for its internal consistency, using Cronbach's alpha after pilot testing, was 0.78. The instrument was administered in each of the sampled universities to sampled lecturers. The questionnaires were completed according to the instructions given and returned to the researcher. The research questions were answered using frequency counts, percentages and charts, while hypotheses were tested using t-test at 0.05 level of significance.

6. RESULTS AND DISCUSSION

Research Question 1: What is the level of difficulty in constructing test items experienced by ODL lecturers?

6.1. Research Question 1: What Is the Level of Difficulty in Constructing Test Items Experienced by ODL Lecturers?

Table 1: Description of ODL Lecturers' Level of Difficulty Experienced in Test Item Construction

Test Item Types	VD (%)	D(%)	ND(%)	Mean	S. D.	Std. Error
Essay	32 (13.3)	46 (19.2)	162 (67.5)	2.54	.719	.046
Multiple Choice	157 (65.4)	53 (22.1)	30 (12.5)	1.47	.708	.046

Test Item Types	VD (%)	D(%)	ND(%)	Mean	S. D.	Std. Error
Matching	40 (16.7)	142 (59.2)	58 (24.2)	2.08	.636	.041
Completion	-	70 (29.2)	170 (70.8)	2.71	.455	.029
True / False	17 (7.1)	132 (55.0)	91 (37.9)	2.31	.597	.039
Case Study	199 (82.9)	41 (17.1)	-	1.17	.377	.024
Weighted Mean	2.05					

Key: VD = Very Difficult

D = Difficult

ND = Not Difficult

From Table 1, it is evident that lecturers in Open and Distance Learning had difficulty in setting case study items in that 199 (82.9%) found it very difficult and 41 (17.1%) found it difficult. Case study items also had the lowest mean value (1.17), which implies that they are the most difficult items for ODL lecturers to construct. This is followed by multiple choice items, which 157 (65.4%) of the lecturers found it difficult ($\bar{x} = 1.47$). The test items that lecturers considered to be easiest to construct were essays, which 162 (67.5%) responded it was not difficult to set. In all, since the benchmark of mean difficulty was 2.0 and the weighted mean was 2.05, it then means that ODL lecturers had moderate levels of difficulty in constructing test items for students. That ODL lecturers have moderate difficulty in test construction stems from the fact that most of the ODL lecturers did not undergo teacher training where the intricacies of test item construction are taught. Higher education lecturers lack the requisite knowledge of item construction and therefore lack the skills to set quality test items for students.

6.2. Research Question 2: What Types of Test Items Do ODL Lecturers Find Difficult to Construct?

In Figure 1, it is shown that the most difficult test items to construct were case studies ($\bar{x} = 1.17$), followed by multiple choice items ($\bar{x} = 1.47$), matching ($\bar{x} = 2.08$), true or false ($\bar{x} = 2.31$), essays ($\bar{x} = 2.54$) and completion items ($\bar{x} = 2.71$). It is therefore concluded that the test item ODL lecturers found the least difficult to construct was completion while case studies were the most difficult.

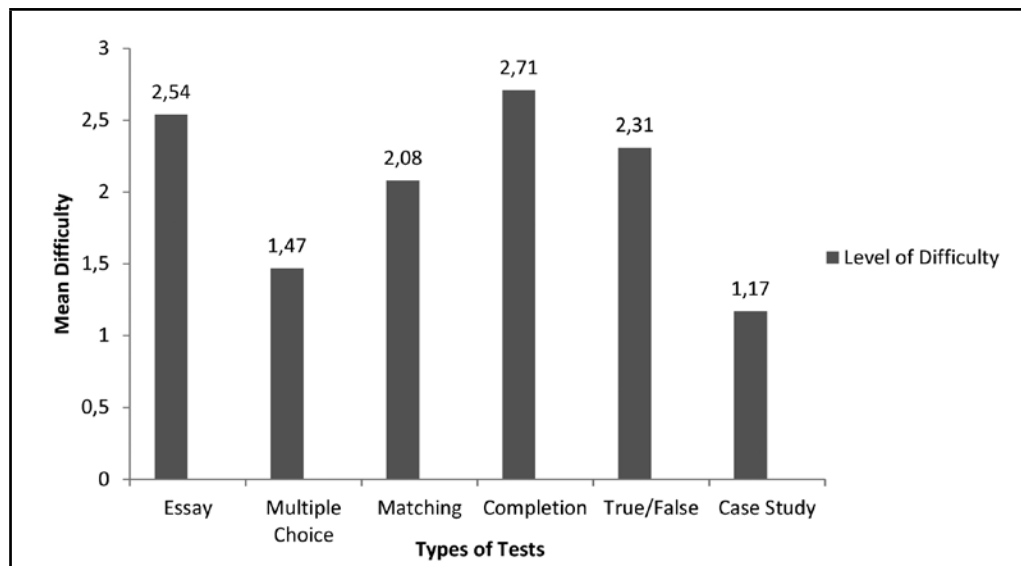


Figure 1: Types of Test Item Construction Difficulty

6.3. Research Question 3: Which Test Type Items Do ODL Lecturers Find Difficult to Construct based on Gender?

Table 2: Level of Difficulty Experienced in Constructing Test Items according to Gender

Test Item Types	Male (n = 156)		Female (n = 84)	
	Mean	S. D.	Mean	S. D.
Essay	2.58	.728	2.48	.702
Multiple Choice	1.49	.714	1.44	.700
Matching	2.17	.570	1.89	.712
Completion	2.72	.451	2.69	.465
True/False	2.29	.653	2.35	.478
Case Study	1.19	.390	1.14	.352

Male lecturers experienced a high level of difficulty in constructing case studies ($\bar{x}=1.19$), multiple choice ($\bar{x}=1.49$) and matching items ($\bar{x}=2.17$), while completion ($\bar{x}=2.72$) and essay ($\bar{x}=2.58$) test items were found less difficult to construct by male lecturers. Female lecturers regarded completion ($\bar{x}=2.69$) and essay ($\bar{x}=2.48$) test items as very easy to set while their male counterparts found case study ($\bar{x}=1.14$) and multiple choice ($\bar{x}=1.44$) items very difficult to construct. Therefore female lecturers found case study, multiple choice, matching, essay and completion test items more difficult to

construct than the male lecturers did. A case study is always a high-order thinking test item that enables those being tested to critically examine the case presented so as to find a solution to the problem. It is a difficult item to set in that the test item constructor has to take the time to read widely so as to find an appropriate case to be presented based on the concept s/he wishes to examine, as observed by Ainsworth and Viegut (2006). Female lecturers may not have enough time to search for an applicable case due to official and domestic commitments.

6.4. Research Question Four: To What Extent Do Lecturers from Single and Dual Mode ODL Institutions Differ in the Difficulty Experienced to Generate Test Items?

Table 3: Difficulty in Constructing Test Items according to Single and Dual Modes of ODL Delivery

Test Item Types	Single (n = 120)		Dual (n = 120)	
	Mean	S. D.	Mean	S. D.
Essay	2.31	.848	2.78	.458
Multiple Choice	1.71	.824	1.23	.463
Matching	1.85	.703	2.30	.460
Completion	2.72	.453	2.70	.460
True / False	2.32	.550	2.30	.643
Case Study	1.05	.219	1.29	.453

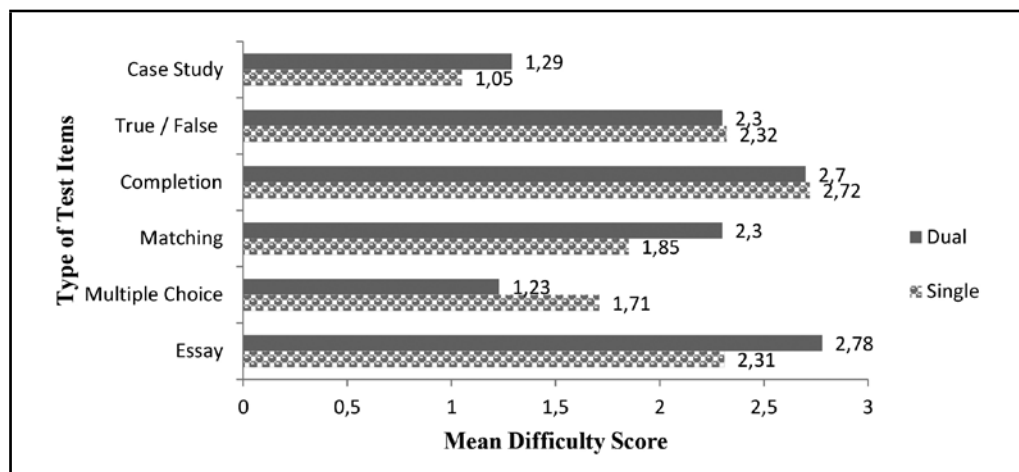


Figure 2: Difficulty in Constructing Test Items by ODL Mode of Delivery

Lecturers from the single mode ODL institution find multiple choice ($\bar{x} = 1.71$), matching ($\bar{x} = 1.85$), and case study ($\bar{x} = 1.05$) type items more difficult to set than their dual mode counterparts do. Multiple Choice ($\bar{x} = 1.23$), completion ($\bar{x} = 2.70$) and true or false ($\bar{x} = 2.30$) test items were more difficult for dual mode ODL lecturers to generate than for single mode lecturers. However, lecturers from the single mode ODL institution found test items more difficult to generate ($\bar{x} = 1.90$) than those from dual mode ODL universities ($\bar{x} = 2.1$). Some factors that could be adduced to dual mode ODL lecturers' greater confidence in constructing test items when compared to their single mode counterparts is that dual mode ODL institutions combine conventional education systems with distance education in which interactions among lecturers are more effective. Such interactions foster the exchange of ideas on test item construction, especially with education lecturers. Also, dual mode lecturers comply with the mandate given to all university lecturers to have a postgraduate diploma in education, and so the majority have obtained professional teaching qualifications and, more specifically, skills in test development.

6.4.1 Hypothesis One: There Is No Significant Difference in the Difficulty Experienced by Male and Female ODL Lecturers in Constructing Test Items

Table 4: T-test of Difficulty Experienced in Constructing Test Items according to Gender

Gender	N	Mean	S. D.	Std. Error	t	df	P	Remark
Male	156	16.33	2.031	.163	3.001	238	.003	Significant
Female	84	15.50	2.220	.242				

The mean difficulty experienced by male lecturers in constructing test items was 16.33 while that of the female lecturers was 15.50 (Table 4). This implies that female lecturers have more difficulty in constructing test items than do male lecturers in ODL institutions. The table also reveals that there exists a significant difference in difficulty experienced by male and female ODL lecturers in constructing test items ($t(238) = 3.001$, $P = 0.003$). The significant difference observed in the difficulty experienced by male and female lecturers in constructing test items was in consonance with Dayiogu and Turut-Asik's (2007) findings that adult males and females differ significantly in terms of human cognition, which is the foundation for constructing quality test items. In contrast, Koksal (2004) observed that female teachers had better test construction skills than male teachers.

6.4.2 Hypothesis Two (Ho2): There Is No Significant Difference in the Difficulty Experienced by Single and Dual Mode ODL Lecturers in Generating Test Items

Table 5: T-test of Difficulty Experienced in Constructing Test Items according to Institutional Mode of Delivery

Mode of ODL Delivery	N	Mean	S. D.	Std. Error	t	df	P	Remark
Single	120	14.35	1.358	.124	20.590	238	.000	Significant
Dual	120	17.76	1.202	.110				

It can be deduced from Table 5 that lecturers from single mode ODL institutions had greater difficulty in constructing test items ($\bar{x} = 14.35$) than lecturers from dual mode institutions with a mean of 17.76. The difference in the level of difficulty experienced between lecturers from single and dual mode ODL institutions was significant: $t(238) = 20.590$ and $p = .000$ which is less than 0.05. Therefore, there is a significant difference in difficulty experienced by single and dual mode ODL lecturers in constructing test items between. This result was at variance with the Commonwealth of Learning's (2000) notion that lecturers in single mode ODL institutions are better at constructing test items than those in dual mode institutions.

7. CONCLUSION AND SUGGESTIONS

As is the case with any institutions of higher learning, assessment is the most important of all institutional responsibilities because the quality of students' achievement is based on their performance in continuous assessments and final examinations. In ODL institutions, learners are assessed mainly through tutor-marked assignments and examinations. These are based on questions generated as supply response (essay) or select response (objectives) or case study. Test items for any of these test types are constructed by the lecturers or tutors. The level of difficulty experienced among lecturers in single and dual mode ODL institutions in setting the test items was considered in the study. Lecturers found test items like case studies, multiple choice and completion difficult to generate. Male and female ODL lecturers also differ significantly in the difficulty they experience in constructing test items but female lecturers have greater difficulty. Also, dual mode ODL lecturers found it easier to develop test items than those in single mode institutions did, with a significant difference in the level of difficulty experienced. In view of the findings of the study, the following suggestions would help to alleviate the difficulties in constructing test items experienced by ODL lecturers:

1. ODL lecturers should show interest in test item construction through self-capacity development by studying test item construction in textual and online materials, irrespective of their areas of specialisation.
2. There is a need for institutional support for female lecturers to improve their skills in test item generation by giving them incentives in the form of sponsorships to attend assessment-based training in test item construction.
3. Lecturers from single mode ODL institutions in Nigeria should liaise with lecturers from dual mode institutions with a view to sharing their experiences in test item development.
4. Further training and workshops should be given to ODL lecturers on aspects of constructing case study, multiple choice and completion test items.
5. Capacity development in test item construction should not be a one-size-fits- all exercise but should be selective, relevant and aligned to the needs of academics within a continuous professional learning approach.
6. ODL institutions should organise frequent test item construction workshops to be sponsored by each institution or collaborating with external bodies like Commonwealth of Learning.

REFERENCES

- Angelo, T. A., and K. P. Cross. 2008. *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass.
- Ainsworth, L., and V. Viegut, D. 2006. *Common formative assessments: How to connect standards-based instruction and assessment*. Thousand Oaks: Corwin Press.
- Ayodele, S. O., J. A. Adegbile, and J. G. Adewale. 2003. *Evaluation studies*. Ibadan: Powerhouse Press and Publishers.
- Barton, P. E. 2002. *Staying on course in education reform*. Princeton: Educational Testing Service.
- Commonwealth of Learning. 2000. *An introduction to open and distance learning*. <http://www.col.org/ODLIntro/introODL.htm> (accessed January 15, 2016).
- Dayioglu, M. and S. Turut-Asik. 2007. Gender differences in academic performance in a large public university in Turkey. *Higher Education* 53(2): 255–277.
- Devine, M., and M. Yaghlian. 2000. *Test construction manual: Construction of objective tests*. Cornell University Center for Teaching Excellence.
- <http://www.clt.cornell.edu/campus/teach/faculty/Materials/TestConstructionManual.pdf> (accessed February 3, 2016)
- Doherty, Y., S. J. Kovas, and R. Plomin, R. 2001. Gene-environment interaction in the etiology of mathematical ability using SNP sets. *Behavior Genetics* 41(1): 141–154.
- Greiff, S., S. Wüstenberg, D. V. Holt, F. Goldhammer, and J. Funke. 2013. Computer-based assessment of complex problem solving: Concept, implementation, and application. *Educational Technology Research and Development* 61(3): 407–421.

- Guskey, T. R. 2003. How classroom assessment can improve learning. *Educational Leadership* 60(5): 7–11.
- Halpern, D. F., and M. L. LaMay. 2000. The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review* 12(2): 229–246.
- Hyde, J. S. 2005. The gender similarities hypothesis. *American Psychologist* 60(6): 581–592.
- Izard, J. 2005. *Overview of test construction*. Paris: International institute for educational planning.
- Kifer, E. 2001. *Large-scale assessment: Dimensions, dilemmas, and policies*. Thousand Oaks: Corwin Press.
- Koksal, D. 2004. Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the Net. *Turkish Online Journal of Distance Education* 5(1): 105–124.
- Linn, R. L., and M. D. Miller. 2005. *Measurement and assessment in teaching*. 9th ed. Upper Saddle River: Prentice Hall.
- Nisbett, R. E., J. Aronson, C. Blair, W. Dickens, J. Flynn, D. F. Halpern, and E. Turkheimer. 2012. Intelligence: New findings and theoretical development. *American Psychologist* 67 (2): 130–159.
- Oblinger, D. G. 2000. The nature and purpose of distance education. Technology Source. technologysource.org/article/nature_and_purpose_of_distance_education/ (accessed June 4, 2015)
- Popham, W. J. 2002. *Classroom assessment: What teachers need to know*. 3rd ed. Boston: Allyn and Bacon.
- Ramos-Mattoussi, F., and Milligan, J. A. 2013. *Building research and teaching capacity in Indonesia through international collaboration*. New York: Institute of International Education.
- Tansel, A. 2002. Determinants of school attainment of boys and girls in Turkey: Individual, household and community factors. *Economics of Education Review* 21(5): 455–470.
- UNESCO. 2002. *Open and Distance Learning trends policy and strategy considerations*. Paris: UNESCO.
- University of Washington 2015. *Manual for testing*. Washington, D.C.: University of Washington Press.
- Van de Vijver, F. J. R., and K. Leung. 2000. Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology* 31(1): 33–51.
- Young, J. W., and J. L. Fisler. 2000. Sex differences on the SAT: An analysis of demographic and educational variables. *Research in Higher Education* 41(3): 401–416.