

Volume 9, Issue 1, 2016

International Journal of Scientific Research in Education



ISSN: 1117-3259

Serving the needs of educators worldwide

The mission of the International Journal of Scientific Research in Education (IJSRE) is to be the premier journal on educational issues. To this end, IJSRE emphasizes quality and relevance in the papers it publishes. In addition, IJSRE recognizes the global influences on the education domain and its development. Therefore, IJSRE seeks international input in all aspects of its endeavour, which includes content, authorship of papers, peer review process, readership, and its International Editorial Board Membership.

**International
Journal of
Scientific Research in
Education
Volume 9(1), MARCH 2016**

ISSN: 1117-3259

IJSRE is Indexed in

EBSCO PUBLISHING. Providers of EBSCOhost



International Scientific Indexing (ISI) (Impact factor 1,413) for the year 2014



Email: editor.ijsre@gmail.com, editor@ijsre.com, submission.ijsre@gmail.com

Copyright © IJSRE

Publisher's website: www.ijsre.com

INSTRUCTION FOR AUTHORS

The **International Journal of Scientific Research in Education (IJSRE)** is a meeting point for the dissemination and explanation of research findings by practitioners, researchers and academicians to the multifaceted problems associated with education. The journal is written and published for an international audience of practitioners, researchers and academicians. IJSRE [invites the submission of manuscripts](#) with a strong applied orientation on any topic in education and non-education domains applicable to education. Manuscripts may be, but are not limited to, empirical articles and theoretical and methodological papers. Applied value and other innovative formats are also considered. IJSRE is published quarterly on the Web. Our goal is not simply to provide an Internet accessible version of a journal that appears in print. Rather the journal intends to exploit the electronic medium to its fullest degree. Authors are highly encouraged to submit their manuscripts as attachments via e-mail to the editor-in-chief editor.ijsre@gmail.com, editor@ijsre.com, submission.ijsre@gmail.com or to the managing editor admin@ijsre.com. The journal's [instructions for authors](#) should be consulted. Paper submissions are accepted but not encouraged. Only in a few, well-justified exceptions will the journal accept submissions in paper form with an enclosed file on a diskette or other digital media.

Originality

Prospective authors sending papers for consideration and possible publication in **IJSRE** should note that only original and previously unpublished manuscripts will be considered. Submission of manuscripts implies a certification that no part of the manuscript is either copyrighted or under review by any other publication. It is the primary responsibility of authors to obtain proper permission for the use of any copyrighted materials in their manuscript prior to submission.

Formatting information

On a separate sheet the author should include the following: a title that indicates exactly but as briefly as possible the subject of the article, the names, present professional affiliations, addresses as well as telephone, fax number and e-mail addresses of authors, an abstract, preferably no longer than 150 words, and a minimum of five key words for referencing.

The main body of the work should begin on a separate sheet than the title, abstract, key words etc. The author's names should not be included anywhere in the manuscript, except on the cover page. Please use Arial 12-point font for the title, which should be in bold upper and lower case.

Use Times New Roman 12-point font for the remainder of your article. Tables and figures should be 10-point font. Tables, drawings, diagrams and charts with a clear title should be presented and numbered with Arabic numerals. These materials should be placed in their exact positions in the manuscript. First level headings should be in bold upper case, and second level headings in bold upper and lower case.

Fully formatted manuscripts should be submitted electronically to editor.ijsre@gmail.com or editor@ijsre.com or to the managing editor admin@ijsre.com. They are to be single-spaced and justified on A4 paper only. Allow adequate "white space" in the margins to allow for reviewing (at least 1" on the four margins). Contributors are asked to use MS Word 5.0 or a later version. The main body of the e-mail message should contain the title of the article and the names, affiliations and addresses of all authors.

Style

Papers should be written in a clear, concise style appropriate to an international readership. Familiar technical terms may be used without explanation. Acronyms and abbreviations are likely to need full presentation at least once. Please write your text in good English (American or British usage is accepted, but not a mixture of the two). Authors should edit and proofread the language of their work before submission. Manuscripts may be between 5000-9500 words in length.

Manuscripts should be prepared in the latest [American Psychological Association \(APA\)](#) editorial style. References

should relate only to material cited within the manuscript and should be listed in alphabetical order by authors' name, year of publication, complete title of the cited work, title of the source, volume, issue and pages cited. Please do not include any abbreviations. See the following examples:

- Single author periodical publication
Nwagwu, N. A. (1997). The environment of crises in the Nigerian education system. *Comparative Education*, 33(1), 87-95.
- Multiple authors periodical publication
Maehr, M. L., & Midgley, C. (1991). Enhancing student motivation: a schoolwide approach. *Educational Psychologist*, 26(3 & 4), 399-427.
- Single author text (book) publication
Naylor, J. (1999). *Management*. Harlow: Prentice Hall.
- Multiple authors text (book) publication
Austin, H., Dwyer, B., & Freebody, P. (2003). *Schooling the child: the making of students in classrooms*. London & New York: RoutledgeFalmer.

Review process

The review process is expected to take approximately 8-12 weeks and the author will be notified of the status of the review as soon as the process is complete. All correspondence will be directed to the lead author of multi-authored manuscripts. It is the responsibility of the lead author to communicate the status of the review process and editorial recommendations to other contributing authors.

To ensure a high quality of published material, the **IJSRE** provides content and writing experts to review submitted manuscripts. Manuscripts are judged by the Editor-in-Chief for focus and format upon receipt. Suitable submissions are assigned at least two reviewers selected from the Review Board based upon the particular content involved. Both the original manuscripts and the reviewers' comments are sent to an associate editor for a final recommendation. Manuscripts with potential for publication will be returned to the author for revision, though requests for re-writes do not guarantee acceptance of the manuscript for final publication. Revised manuscripts will be reviewed again by the original review panel and one additional reviewer. Final recommendations will be based upon the quality of the revision and comments of reviewers and associate editors.

All contributions and correspondence should be addressed to:

Editor-in-Chief of International Journal Scientific Research in Education

Dr. Nwachukwu Prince Ololube
Department of Educational Foundations and Management
Faculty of Education
University of Education
P. M. B. 5047, Port Harcourt
Rivers State
Nigeria.
editor.ijsre@gmail.com, editor@ijsre.com, submission.ijsre@gmail.com

NOTE: IJSRE publication cost is covered by the author's (US\$220 for each paper within 9,500 words and \$50 for additional words/pages). IJSRE may grant discount fees for authors who are student. IJSRE publication fees are extremely competitive and are amongst the lowest of any journal offering open access publication. Fees are paid when your paper has been accepted after the blind review process is complete.

Call for Papers

The **International Journal of Scientific Research in Education (IJSRE)** publishes article that promotes the advancement of education at all levels, encompassing all domains of teaching and learning. The primary mission of the **IJSRE** is to serve as a medium for initiating, collaborating, analyzing, synthesizing and evaluating original and groundbreaking contributions to the theory, practice and research in education applicable to kindergarten, primary, secondary and higher education. The journal serves as a platform for researchers whose work is applicable to education, and for non-education researchers from other disciplines (such as political science, sociology, psychology, accounting, marketing, management, government policy, ICT etc) whose work can be applied to the education context. Thus the objective of this international journal is to disseminate scholarly papers from multiple disciplines as applied to education globally.

The [journal welcomes contributions](#) from an empirical and/or conceptual point of view. If you would like to submit a paper for consideration to **IJSRE**, please send an e-mail with the manuscript as a file attachment in Microsoft Word or RTF to the editor-in-chief editor.ijsre@gmail.com, editor@ijsre.com submission.ijsre@gmail.com or the managing editor admin@ijsre.com. Contributors will receive acknowledgement of the receipt of the manuscript within three working days. The main body of the e-mail message should contain the title of the manuscript and the names, affiliations and addresses of all authors. Manuscripts should be in English. The author's names should not be included anywhere in the manuscript, except on the cover page. Manuscripts must also be accompanied by an abstract of 100-150 words, succinctly summarizing the aims and objectives of the manuscript. Authors should kindly consult the [instructions for authors'](#) page for style and format requirements.

Papers are solicited that address these issues (but are not limited to):

Art of teaching	Instructional methodologies
Case studies from schools around the world	Instructional supervision
Change management and education quality	Knowledge and education
Citizenship education	Leadership in education
Classroom management	Lifelong learning and development of competences
Computers in educational administration	Management in higher education
Differing cultural perceptions of management in education	Management of e-education
Distance education and multimedia environments	Managing the curriculum
Early Childhood Education	Marketing in education
Economics of education	Method courses in education
Educational administration	Organizations as learning communities
Education environments (political, social, legal, cultural)	Policy analysis and evaluation of institutions
Educational leadership	Principles and practices of public administration
Educational policy and management	Principles of development administration
Educational systems planning/strategic planning	Public policies management
E-education / E-learning	Quality improvement
Entrepreneurial development	Research methodology (Quantitative/Qualitative)
Equity and education	School business administration
Finance and accountability in education	School effectiveness
Globalization and education	School plant planning and management
Human resources management	Special Education
Individual professional learning portfolio	Staff professional growth and development
Information Communication Technologies (ICT) in Education	Teacher education
Information systems for education and training support	Teacher effectiveness
Information systems management	Theories of educational management

Editor-in-Chief

Dr. Nwachukwu Prince Ololube
Department of Educational Foundations and Management
Faculty of Education
University of Education, P. M. B. 5047, Port Harcourt, Nigeria.
editor.ijsre@gmail.com or editor@ijsre.com

Associate Editors

Prof. Princely Ifinedo
Cape Breton University, Canada

Dr. Daniel Elemchukwu Egbezor
University of Port Harcourt, Nigeria

Advisory/Editorial Board Members

Dr. Richard Wamai
Harvard University, USA

Prof. Ravinder Rena
University of the Western Cape, South Africa

Prof. David A. Adewuyi
Virginia Union University, Richmond, Virginia, USA

Dr. Chi Kim Cheung
The University of Hong Kong, Hong Kong

Dr. Nneka Nora Osakwe
Albany State University, USA

Dr. Dwedor Morais Ford
Winston-Salem State University, USA

Prof. Reggie Davidrajuh
University of Stavanger, Norway

Dr. Bhattacharya Maumita
Charles Sturt University, Australia.

Dr. Vikram Koundinya
University of Wisconsin, USA

Review Board Members

Dr. Chi Kim Cheung, The University of Hong Kong, Hong Kong

Dr. Dwedor Morais Ford, Winston-Salem State University, USA

Prof. Faith-Michael Uzoka, Mount Royal University, Calgary, Canada

Dr. Nneka Nora Osakwe, Albany State University, USA

Prof. Charles Korede Ayo, Covenant University, Ota, Nigeria

Assoc. Prof. Princely Ifinedo, Cape Breton University, Canada

Prof. Ravinder Rena, University of the Western Cape, South Africa

Dr. M. Azath, VMKV Engineering College, India

Prof. Reggie Davidrajuh, University of Stavanger, Norway

Dr. Richard Wamai, Harvard University, USA

Prof. Charles Uwadia, University of Lagos, Nigeria

Assoc. Prof. Alafuro Epelle, Rivers State University of Education, Nigeria

Dr. Smita Rajpal, ITM College of Engineering and Technology, Gurgaon

Dr. Bhattacharya Maumita, Charles Sturt University, Australia.

Dr. Wichian Sittiprapaporn, Mahidol University, Thailand

Assoc. Prof. Norzaidi Mohd Daud , Universiti Teknologi MARA Malaysia

Dr. Daniel N. Nte, NOVENA University, Nigeria

Assoc. Prof. Alexander E. Sobolev, Tver State Technical University, Russia

Dr. Erlane K Ghani, Universiti Teknologi Mara, Malaysia.

Mohd Nazri Ismail, Universiti Kebangsaan Malaysia

Prof. Arul Lawrence Selvakumar, RCE & Technology, Thrissur, India

Prof. C. Suresh Gnana Dhas, Vel Tech Multitech, India

Dr. Mahdi Salehi, Zanjan University, Iran

Dr. Mohamad Noorman, Masrek Universiti Teknologi, Mara, Malaysia

Dr. Alakanani Alex Nkhwalume, University of Botswana, Botswana.

Dr. Erlane K Ghani, Universiti Teknologi, Mara, Malaysia

Dr. Shamira Malekar, Aruna Manharlal Shah's Institute of Management and Research, India

Dr. Latisha Asmaak Shafie, Universiti Teknologi MARA Perlis, Malaysia

Imran Naz, Kohat University of Science & Technology, Pakistan.

Dr. Vikram Koundinya, University of Wisconsin, USA.

CONTENTS

- 1-6 The Effects of Analogy on Male and Female Chemistry Students' Problem-Solving Ability in Electrolysis.
Gongden, E. J.
- 7-19 Establishing the Assessment Model for English Language Continuous Writing Component
Mogapi, M. O.
- 21-28 Teacher Education and Teacher Efficacy do they Match in Botswana Colleges of Education
Moalosi, W. T. S., & Matsoga, J. T.
- 29-40 Culture and Differential Item Functioning in National Examinations Council Senior School Certification Mathematics Multiple Choice Test in Nigeria
Ebesine, S. S.



The Effects of Analogy on Male and Female Chemistry Students' Problem-Solving Ability in Electrolysis

Ephraim Joel Gongden¹

Department of Science and Technology Education
University of Jos, Nigeria

gongdenej@yahoo.com estephephest@gmail.com

Abstract

Studies have shown that Nigerian secondary school chemistry students have difficulty solving problems involving electrolysis and other concepts. The study was carried out to find out the comparative effects of analogy on male and female students' problem solving ability in electrolysis. The pretest-posttest-control group design was employed. Three instruments, Chemistry Achievement Test (CAT), Mathematical Skill Test (MST) and Chemistry Problem Solving Test (CPST) were developed and used. The CAT and MST were used alongside the pretest and randomization of the sample to ensure equality of the groups. Sixty four students were randomly selected, pre-tested and assigned into control and analogy (experimental) groups. The control group was taught using lecture method while the experimental group was taught using analogy. Analyses of students' posttest mean scores in the CPST using t-test ($\alpha = 0.05$) showed that students taught with analogy performed better than those in the control group in a chemistry problem solving test involving electrolysis. There was also a statistically significant difference between the posttest mean scores of male and female students in a CPST involving electrolysis when they were taught using analogies. The male students performed better than the female students. The study recommended the use of analogies as strategies for teaching problem solving tasks in electrolysis especially to male students.

Keywords: Analogy, Problem-solving, Electrolysis, Chemistry, Strategies.

Reference to this paper should be made as follows:

Gongden, E. J. (2016). The Effects of Analogy on Male and Female Chemistry Students' Problem-Solving Ability in Electrolysis. *International Journal of Scientific Research in Education*, 9(1), 1-6. Retrieved [DATE] from <http://www.ij sre.com>

INTRODUCTION

Science and technology have always been recognized as the basic tool of industrialization and national development and could bring economic and social happiness by providing employment and improving the welfare of the citizenry. As a matter of fact, discoveries in information technology have reduced the world to a global village with the World Wide Web being an enormous information base (Oak, 2011). In Nigeria, some provisions of the National Policy on Education and the change to the 9-3-4 system all aim at providing sufficient opportunities are opened to citizens to get the best scientific education possible.

Chemistry has played a major role in science, technology and society, and it still does so today. There is hardly found anything in nature that chemistry does not have an influence or impact over. No wonder the assertion that without chemistry there will be no life. Unfortunately, the subject has always come under threat with low performance by secondary school students in Nigeria.

A number of studies such as Jimoh (2004) and Njoku (2007) amongst others have reported the poor performance of chemistry students at the secondary and tertiary levels over the years. The National Examination Council (NECO) revealed that the percentage credit pass of students in chemistry during the November/December General Certificate of Education examination for 2011, 2012 and 2013 as 5.32%, 30.17% and 66.41% respectively (Mosadomi, 2014). A source from the West African Examination Council (2012) gives the percentage credit pass in chemistry as follows: 2007 (45.96%), 2008 (44.44%), 2009 (43.70%), 2010 (50.70%) and 2011 (49.54%).

Various reasons have been given for students' poor performance in chemistry amongst which is their inability to solve chemical problems (Dajuma, 2005). Some of the concepts that present such difficulty to students include electrochemistry, chemical equilibrium, redox reactions, mole concept and stoichiometry. However, most researchers attribute the poor performance of students in chemistry and problem solving ability to the pedagogical approaches adopted by teachers in schools (Gabel, 2003; Mtsem, 2011; Owolabi & Oginni, 2013). The poor problem solving ability of students points to a likely deficiency in method of instruction and neglect of students' centered learning strategies. Metacognitive instructional strategies have emerged as a result of various researches and have been found to assist in this direction (Orgil & Thomas, 2007). One of these metacognitive instructional strategies is the use of analogies.

An analogy is a comparison between two domains of knowledge: one that is familiar and another that is not (Orgil & Bofner, 2004). The familiar one is called the analog while the unfamiliar one is the target domain. The target is what needs to be learnt. Researchers in science education have identified analogy as a powerful tool for explaining and facilitating an individual's construction of knowledge (James, Scharmann & Lawrence, 2007). However, despite the effectiveness of these student-centered strategies, little is understood about the effect of analogies on male and female students' performance in problem-solving tasks involving the electrolysis. How will the male and female students differ in their problem-solving ability in electrolysis when taught using analogies?

Objective of the Study

The main purpose of this study was to find out the comparative effects of analogy on male and female students' problem solving ability in electrolysis. Specifically, the study sets to:

- Find out if the male and female students differ in their chemistry problem solving ability in tasks involving electrolysis before being taught the concepts;

- Find out if male and female students taught with analogy teaching strategy differ in their problem solving ability in tasks involving electrolysis.

Two null hypotheses were formulated and tested during the study:

- There is no significant difference between the pretest mean score of male and female students in a chemistry problem solving test involving electrolysis;
- There is no significant difference between the posttest mean scores of male and female students in a chemistry problem solving test when taught electrolysis using analogy instructional strategy.

METHODS

Research Design

The study was a pretest-post-test control group design. The main strength of this design is that the initial random assignment of subjects to the groups and the administration of a pretest to all the groups help to control all threats to internal validity. It also ensures that both groups are equivalent on all important dimensions and that there are no systematic differences between the two groups. The design also controls all the threats to internal validity. The purpose of the experiment was to show that any difference obtained between the initial scores and the final scores in the groups were as a result of the different treatment received by each group.

Sampling/Data analyses

The sample for the study (consisting of sixty four male and female students) was arrived at through random sampling after administering CAT and MST. The data analyzed was obtained through the administration of a chemistry problem solving test, CPST. The reliability coefficient of the CPST was found to be 0.87. Those of the CAT and MST were 0.80 and 0.93 respectively. The t-test for independent sample was used to analyze the data.

The students were grouped into two classes of thirty two students each, one the control and the other the experiment (analogy) class.

RESULTS

The pre-test mean scores of the male and female students were analyzed and the results used to answer research question one and to test hypothesis one as follows:

Table 1a: Group statistics of pretest of male and female students in CPST

Test Group	N	Mean	S.D	S.E Mean
Male	32	27.63	1.786	0.258
Female	32	27.10	2.080	0.300

Research question one: What is the difference between the pretest mean scores of male and that of female students in a chemistry problem solving test?

The results as presented in tables 2a and 2b show that the pretest mean score of the male students was 27.63 while that of the female students was 27.10. The mean difference between the two was 0.53, a negligible figure.

Table 1b: Independent sample test for equality of pretest means of male and female students in CPST

Group	Mean diff.	t	df	S.E diff	P – sig (2-tailed)
Male & Female	0.53	0.421	62	0.396	0.675

Hypothesis one: There is no significant difference between the pre-test mean score of male students and that of female students in a chemistry problem solving test.

The p-significant value was found to be 0.675 ($p > 0.05$). This means that there was no statistically significant difference between the pretest means score of the male students and that of the female students. Therefore the null hypothesis one was accepted. The researcher analyzed the posttest scores of the students and presented it appropriately.

Research question Two: What is the difference between the post test mean scores of male and female students in a chemistry problem solving test when taught using analogy instructional strategy?

Table 2a: Group statistics of posttest of male and female students in CPST

Test Group	N	Mean	S.E	Mean diff
Male	16	66.00	0.593	6.60
Female	16	59.40	0.518	

The results analyzed and presented in tables 3a and 3b showed that the posttest mean scores of male and female students taught with analogy are not the same. While the male students in the analogy group had a mean score of 66.00, the female students in the same analogy group had a mean score of 59.40. The mean difference was 6.60, the mean scores of male students being higher than females’.

Table 2b: Independent sample test for equality of posttest means for male and female students in CPST

	Mean diff.	t	df	S.E diff	P – sig (2-tailed)
Equal variance Assumed	6.60	2.539	30	0.788	0.017

Research hypothesis Three: There is no significant difference between the posttest mean scores of male and female students in a chemistry problem solving test when taught using analogy instructional strategy

The p-value, $0.017 < 0.05$ as presented in table 2b. This showed that the mean score of male students in the CPST when taught with analogy differed significantly from that of the female students. The null hypothesis was rejected and the alternate hypothesis accepted. Therefore, there is a significant difference between the posttest mean scores of male and female students in a chemistry problem solving test when taught using analogy instructional strategy.

Discussion of Results

There was no statistically significant difference between the pretest mean scores of male students and female students in a chemistry problem solving test task involving electrolysis. This was to be expected as both groups were equivalent given the mode of selection of the sample.

The study also revealed the existence of a statistically significant difference in the posttest mean scores of male and female students in a chemistry problem solving test task involving electrolysis when both were taught using analogies. This finding showed that male chemistry students benefitted more in problem solving task involving electrolysis when taught using analogy than female students – indicating gender influence. The reasons for this may include male students' better reasoning ability and high level of abstract thinking, exposure of males than female to the environment, restrictions on females due to customs and religious beliefs which reduce their exposure, etc. The life-world experience of male students helped them visualize abstract ideas than females.

The finding here agrees with (Eribe & Ande, 2006) who found out that there exists gender difference or inequality in science achievement among secondary school science students and that of (Onekutu, 2002) who found out that male students performed better than females with an increasing gap in chemistry examination. Adesoji and Babatunde (2008) reported that female students encountered problem solving difficulties more frequently than their male counterparts. The result is inconsistent with Armagan, Sagir and Celik (2009), who reported that female students did better than male students when they investigated the effect of problem solving skills on the achievement chemistry students. It also disagreed with Olorundare and Aderogba (2009) who reported that there was no significant difference between the academic performance of male and female students exposed to treatment with analogy (but problem solving in electrolysis)

CONCLUSION

Based on the findings of the study, male students benefit more than female students in chemistry problem solving test when analogy is used to teach the concept of electrolysis. This may be due to the ability of male students to think abstractly than the females. The reasoning ability of males is usually higher than that of females hence the males think and relate to the analogy more easily than females.

Recommendations

The result of this study has implication for the teaching and learning of chemistry in secondary schools. The findings show that teaching strategies influence the performance of students in problem solving tasks involving electrolysis. Chemistry teachers' training programs should include a deliberate preparation of teachers for the acquisition of skills in the use of analogy strategies that are useful in this direction. Education authorities and professional bodies should organize seminars, workshops, refresher courses and conferences on the relevant use of analogies as an instructional strategy on regular basis for teachers.

REFERENCES

- Oak, M. (2011). *Importance of science and technology*. Retrieved July 17, 2014 from <http://www.buzzle.com/articles/importance-of-science-and-technology.html>
- Jimoh, A. T. (2004). Influence of gender difference on students' achievement in chemistry. *Nigerian Journal of Educational Studies and Research*, 2(1), 88-94.
- Njoku, Z. C. (2007). Comparison of students' achievement in the three categories of question in SSCE practical chemistry examination. *Journal of Science Teachers Association of Nigeria*, 42(1&2), 67-72.
- Mosadomi, W. (2014). Analysis of the 2012 November/December NECO examination. Retrieved June 4, 2014 from <http://www.vanguardngr.com/2013/03/neco-release>
- Danjuma, I. M. (2005). *An investigation of pre-service NCE teachers' numerical chemical problem solving behaviours*. Unpublished doctoral thesis, A.T.B.U Bauchi.
- Gabel, D. (2003a). *Problem solving in chemistry: Research matters to the science teacher*. Retrieved June, 12 2013 from www.narst.org/publications/research/problem.cfm
- Mtsem, A. A. (2011). Effects of diagnostic and contextual instructional strategy on students' interests and achievement in secondary school Algebra. Unpublished Ph.D thesis, Abia State University, Uturu
- Owolabi, O. T., & Oginni, O. I (2013). Assessing the relative effectiveness of three teaching methods in measuring students' performance in physics. *International Journal of Materials, Methods and Technologies*, 1(8), 116 -125.
- Orgil, M. K., & Thomas, M. (2007). Analogies and the 5E model. *Science Teacher*, 74(1), 40-45.
- Orgil, M., & Bodner, G. (2004). What research tells us about using analogies to teach chemistry. *Chemistry Education: Research and Practice*, 5(1), 15-32.
- James, M. C., Scharmann, L. C., & Lawrence, C. (2007). Using analogies to improve the teaching performance of preservice teachers. *Journal of Research in Science Teaching*, 44(4), 565- 585.
- Eribe, J. O. & Ande, S. (2006). Gender difference in achievement in calculating reacting masses from chemical equations among secondary school students in Makurdi metropolis. *Educational Research and Reviews*, 1(6), 170-173.
- Onekutu, P. O. (2002). Gender difference in achievements in junior secondary school examination in integrated science: Implications for national development. In O. O. Okpoh (Ed), *Review on Gender Studies in Nigeria*.
- Adesoji, F. A., & Babatunde, A. G. (2008). Investigating gender difficulties and misconception in inorganic chemistry at the senior secondary Level. *International Journal of African and African- American Studies*, 7(1). Retrieved May 30, 2015 from www.ojcs.siue.edu
- Armagan, D., Sagir, S. U., & Celik, A.Y. (2009). The effects of students' problem solving skills on their understanding of chemical rates and their achievement on this issue. *Procedia Social and Behavioural Sciences*, 1(1), 2678 - 2684.
- Olorundare, A., & Aderogba, G A. (2009). Comparative effects of concept mapping, analogy and expository strategies on secondary school students' performance in chemistry in Ilesha, *Nigeria Journal of Curriculum and Instruction*, 7(1&2). Retrieved July 25, 2014 from www.unilorin.edu.ngon

ⁱ Ephraim Joel Gongden, is a lecturer in the Department of Science and Technology Education, University of Jos, Nigeria. He can be reached via email at: gongdenej@yahoo.com estephepest@gmail.com



Establishing the Assessment Model for English Language Continuous Writing Component

Molefhe O. Mogapi¹
Department of Educational Foundations
University of Botswana
Gaborone, Botswana
Molefhe.mogapi@mopipi.ub.bw

Abstract

Validation of assessment instruments has often been confined to ensuring that items in the test are a representative sample of the domain. However, structural validation of psychological instruments through multivariate statistical procedures can be used to enhance the construct validity of psychological scales by providing an empirical model of both the dimensionality of the construct itself and fidelity of the scale developed to measure the construct. The current research is an attempt to establish the assessment model of the English Language continuous writing component. The study uses exploratory factor analysis with orthogonal rotation techniques to map-out the dimensional structure the construct. Specifically, final examination composition and letter writing scores were submitted to a principal component analysis extraction method with Varimax rotation technique to arrive at a more parsimonious and theoretically meaningful language proficiency model. Items with significant loadings on the recovered components were subsequently used to name each component in the scale.

Keywords: Structural validity, Principal component analysis, Varimax rotation, Eigenvalue, Psychometric uniqueness.

Reference to this paper should be made as follows:

Mogapi, M. O. (2016). Establishing the Assessment Model for English Language Continuous Writing Component. *International Journal of Scientific Research in Education*, 9(1), 7-19. Retrieved [DATE] from <http://www.ij sre.com>

INTRODUCTION / BACKGROUND

Validation of achievement tests has typically been concerned with issues of content validity. The major objective of content validation of tests is to ensure that items used in the test are a representative sample of the domain. However, construct validation of tests and scales is much more imperative as it is concerned with the extent to which a scale actually conforms to the structure of a construct of interest. The underlying dimension or structure of a psychological construct may be derived from a theory or gleaned from experience. A good example in this case is the construct of intelligence. Psychologists have been fascinated with human intelligence and the best way to measure it. Different theories about the structure of the intelligence construct have been offered and one of the theories suggests that intelligence construct is two dimensional (Spearman, 1904). Therefore, any measurement instrument that purports to measure intelligence with a set of items should conform to the two factor structure. Likewise, achievement instruments designed to assess performance in a given area should exhibit internal structure that conforms to the domain structure.

Messick (1996) defines validity as "...an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment" (p. 6). The author proposed a unified concept of validity that is characterized by six elements to be considered when assessing a measurement scale. According to Messick (1989) quoted in Messick (1996):

In particular, six distinguishable aspects of construct validity are highlighted as a means of addressing central issues implicit in the notion of validity as a unified concept. These are content, substantive, structural, generalizability, external, and consequential aspects of construct validity. In effect, these six aspects function as general validity criteria or standards for all educational and psychological measurement (p. 12).

Specifically, structural validity demands that the internal structure of a measurement instrument be consistent with the theoretical model of the construct of interest. Currently, English Language subject matter at primary school level is assessed by means of two dimensions; namely Comprehension and Language Use. Comprehension assesses the ability of the candidate to extract relevant information from a variety of sources while Language Use dimension focuses on correct application of language devices such as conjunctions, adjectives and the ability to present ideas in a coherent and logical manner. Composition and letter writing falls exclusively under Language Use dimension. Each modality (i.e., Composition and Letter) is assessed by means of 10 criteria as indicated in Table 1.

Table 1: Composition and Letter writing assessment criteria

Composition Writing Criteria	Letter Writing Criteria
Introduction	Correct Address
Descriptive solution	Key Sentence
Sequence of events	Consistency in Content
Realistic composition	Different Adjectives
Correct spelling	Correct Conjunctions
Appropriate Adjectives	Sentence Coherent
Defined Sentence	Varied sentence length
Correct punctuation	Reasons for the job
Assistance required	Consistent Register
Feeling for assistance	Consistent Tense

The categorization of all the 20 criteria under one dimension implies a unidimensional structure for continuous writing. The unidimensional model has not been verified empirically and there is no theoretical justification of its use. The major research objective for the current study, therefore, is to establish the goodness-of-fit of the two factor model for continuous writing scores. The covariance matrix from PSLE composition and letter writing scores were factor analyzed to map out patterns of relationships amongst the items. Multiple model fit criteria were used to extract reliable factors that best represents the language proficiency construct.

LITERATURE REVIEW

The need to consult relevant literature sources on language assessment stems from the controversy surrounding the number of reliable factors that represent English language proficiency construct. According to Oller (1976), English language ability is represented by only one unitary factor; in other words, performance in English language is unidimensional. However, subsequent research work (e.g., Sasaki, 1999) has largely disconfirmed the unitary hypothesis in favour of the multi-componential structure. The general consensus amongst language researchers is that language proficiency is best represented by a high-order secondary factor and several correlated first order factors (Stricker, Rock & Lee, 2005). The controversy relating to the unitary hypothesis claim and multi-componential structure viewpoint can only be resolved by frequent and sustained research work to determine which of the two arguments is supported by the data. One institution that has carried out extensive and intensive research work on the dimensionality of language proficiency construct is the Educational Testing Service (ETS) in the USA. ETS administers a Test of English as a Foreign Language (TOEFL) to candidates all over the world. As a result, ETS has managed to accumulate a wealth of knowledge on language

assessment through its numerous research studies and publications. A few of the relevant studies are reviewed below to shed light on the language dimensionality issue.

TOEFL was developed in 1963 by the National Council on the Testing of English as a Foreign Language (Sawaki, & Orange, 2008). It is specifically designed to test the English language proficiency of non-native speakers of the language applying for admission to institutions in the United States. TOEFL has developed over the years from a paper-based test to an internet-based format (i.e., TOEFL iBT). Extensive research work has been done to assess the relevance and appropriateness of TOEFL test to different populations across the globe. For example, from 1977 to 2005, nearly 100 research and technical reports have been published (ETS, 2008). Several of the TOEFL studies focus on establishing the dimensionality of the test and the invariance of its factors across different subgroups. Some of the key research studies are enumerated below.

Stricker, Rock and Lee (2005) conducted a study aimed at establishing the factor structure of the LanguEdge test and its invariance across various TOEFL populations around the world. The test has four sections; namely, Listening, Reading, Speaking, and Writing Sections. The LanguEdge test was administered to a total of 472 candidates recruited from both domestic and international TOEFL testing centers. The sample was made up of mostly Arabic (N=160), Chinese (N=225), and Spanish (N=114). Maximum likelihood estimation procedures were applied to analyse the covariance matrix for the normalized scores of each subgroup using LISREL version 8.53 (Joreskog & Sorbom, 1996b). The researchers tested the composite hypotheses of the latent factor structure of TOEFL language test and the invariance of the final solution across the three subgroups. Initially four competing models were tested.

The first model tested was a four factor model which proposes a unitary language proficiency trait. In this model, the four language modalities (Reading, Listening, Speaking and Writing) load on a single language ability factor. The second model tested hypothesized existence of two factors defined by a separate factor for Speaking and a second factor created by a fusion of Reading, Writing and Listening. The third model proposed three correlated factors representing Listening and Speaking and a combination of Teaching and Writing components. The fourth model had four distinct factors that corresponded to the four sections of the test (e.g., Reading, Listening, Speaking and Writing). Items in each of the modalities were expected to have significant loadings on their respective sections. Several goodness-of-fit indices were employed to identify the model that has a superior fit to the data

The analysis converged on only two models; the one factor model and two factor model. The fit indexes for the individual subgroups and the overall analysis were also deemed to be satisfactory for the two models. Subsequent detailed comparison of the two competing models indicated a superior fit of the two factor model. Once the researchers established the baseline model; secondary analysis of the data was carried out to determine the extent of the measurement non-invariance of the LanguEdge test. Hierarchically ordered nested models were tested about the factor invariance specifically relating to the invariance of the number of factors; invariance of the factor loading; the invariance of the error variances; and the invariance of the factor intercorrelation. The analysis showed that only the factor correlations were deemed to be invariant across groups. Therefore, the researchers were able to establish that the LanguEdge test is essentially invariant across groups and can best be represented by two factors; these factors were identified as Speaking and a fusion of Reading, Writing, and Listening. However, the two factor solution was at variance with earlier research studied by Carrol (1983) and Bachman et al

(1995) and Kunnan (1995); these studies extracted three correlated factors with Listening defined as a distinct factor.

Another TOEFL instrument that has been thoroughly researched is the TOEFL Internet Based or TOEFL iBT (ETS, 2005). The test was developed in 2005 and consists of five sections; Reading, Writing, Speaking and Listening. Candidates who write this test receive a total of five scores; a separate score for each of the five language modalities and a global language ability score (Sawaki, et al., 2008). The internet based test has a compulsory Speaking section thus making its design to be different from the LanguEdge test. Therefore, it became necessary to investigate the new test's internal structure as well as generate evidence to support the five score reporting policy. Sawaki, Stricker and Orange (2008) conducted a confirmatory factor analysis study to model the test's internal structure. Separate analyses were done for Reading, Listening and a combination of Writing and Listening. Writing and Listening were combined due to limited number of items in these two sections (Sawaki, et al., 2008). The researchers adopted the multitrait-multimethod factor analytic approach to test for the presence of trait and method effects in the correlation matrix. The traits identified were listed as Basic Comprehension, Inferencing, and Reading to learn. On the other hand, the items in the sections were grouped to define item set factors. Four different models that differed according to the number of path coefficients created in each model were sequentially tested. The models were as follows:

- Model A: Correlated traits and correlated item set model. Within this model the three traits and item sets were allowed to correlate among themselves but not between the two factors.
- Model B: Correlated traits and Uncorrelated item set model. This model limited the degree of correlation of the item sets and as such the model is nested within Model A.
- Model C: Correlated trait model. This model imposes maximum restriction on the data as it essentially eliminates item set correlation.
- Model D: Correlated trait and correlated uniqueness model. The model is different from the other models above in the sense that variance within each item is partitioned into specific and error variance to estimate residual variance. However, due to its complexity, the model was not considered as a likely candidate for representing the factor structure of the test but was included to evaluate the appropriateness of other models (p. 16).

The researchers used three criteria to identify the best fitting model to the data. The criteria were: (1) the extent to which the solution was proper, (b) Substantive interpretation, (c) goodness of fit for the model. Also the magnitude of inter-factor correlation cut off point of .90 or more was also used to identify high inter-factor dependency. The pre-determined fit criteria suggested that none of the four models exhibited satisfactory fit to the data. Model A factor loadings were uninterpretable while the remaining three models had high inter-factor loadings exceeding .90. High inter-factor correlation indicates lack of psychometric uniqueness of the factors (Bagozzi & Yi, 1992). The results therefore point to the likelihood that the three traits (e.g., Basic Comprehension, Inferencing, and Reading to Learn) are evidence of a single trait. As a result, the inter-factor correlations were set at 1 essentially creating one factor upon which all items loaded. Further analysis confirmed the superiority of the single trait model and it was subsequently adapted as a parsimonious representation of the trait structure of the reading section. This suggested a unidimensional nature of reading ability (p. 27).

A similar analysis was done for the Listening section and another one for combined Speaking and Writing sections. In the case of listening, the model specified three traits as being Basic Understanding, Pragmatic Understanding, and Connecting Information. The results showed lack of psychometric distinctiveness amongst the three traits due to significant inter-factor correlations that was well above the pre-determine cut off level of .90. A single trait approach was then followed to establish the best fitting model. The analysis indicated that the single trait model and the single trait uniqueness model had equivalent fit to the data. For this reason, the single trait model was adopted on the principle of parsimony.

The analysis of the Speaking and Writing sections was not as thorough as the previous ones for Reading and Listening. This was mainly due to the number of items available for analysis. However, the analysis converged on a correlated trait model that combined Speaking and Writing traits. A chi-square difference test suggested a significantly better fit for a two factor model.

To establish the best fitting model for the entire TOEFL iBT test, the four preceding models were combined. As a result, four nested models were tested at the global level; these are the Bifactor model, the Correlated trait model, the Single trait model, and the Higher-order model. The main difference between the Bifactor model and the High-order model was that the former allowed for direct influence of the second-order factors on the measured variables whereas the latter specified presence of a common underlying dimension that influenced the individual items only through the first order factors (p. 53). Several goodness of fit indices were used to evaluate model fit (e.g., NNFI, CFI, RMSEA, GFI, ECVI). The fit indices converged on two models; the Correlated trait model and the Higher-order model. The fit statistics indicated that the two models had an equivalent fit. The high-order model was considered to be a reasonable representation of the factor structure of the entire TOEFL iBT test, thus, supporting claims that language ability was essentially multi componential in nature.

Another very informative study on TOEFL assessment instruments was done by Shin (2000). This particular study was conducted to provide evidence that will shed some light on the relationship between language proficiency and the structure of language tests. The issue to be addressed was whether the structure of language tests is invariant across different ability levels. Previous research that sought to understand the relationship led to the emergence of two schools of thought. The first school of thought proposes that as language proficiency improves, the factor structure also becomes more differentiated. A study by Swinton and Powers (1980) for example, 'pointed to greatest amount of factor differentiation for the group with the highest proficiency and the least amount of factor differentiation for language group with the lowest proficiency (p. 33). The observed positive relationship was also confirmed by Ginther and Stevens (1980) in their multi-group structural equation modelling research. On the other hand, some research studies have provided evidence showing an inverse relationship between language proficiency and factor differentiation. The studies show that as learners become more proficient in language use, the factor structure of the test becomes less differentiated. Studies by Kunnan (1992) and Oltman, Stricker, and Barrows (1988) demonstrated that the factor loadings for the low ability groups become more salient.

The primary goal of Shin (2000) study was, therefore, to investigate the relationship between language proficiency and the structure TOEFL. The sample of the study comprised 779 candidates who participated in the Cambridge TOEFL Comparability Study (Bachman, 1995). The sample was divided into three groups; low, intermediate, and high language ability groups. The candidates wrote TOEFL test and the scores were used to generate a covariance matrix of

correlation coefficients. A principal axis factoring modeling techniques were used to extract initial components. Amongst the several competing models, the high-order model was selected as best fitting model because it was able to account for the high correlation amongst the three factors identified. Also, the researcher noted that the selected model structure replicated previous research findings demonstrating that language proficiency consists of one general factor and several distinct abilities (e.g., Fouly xx, 1990).

Since the analysis indicated the suitability of the second order model, the model was further analyzed to test its invariance across the three language proficiency groups. The model was subsequently tested for measurement in variance and structural invariance. Measurement invariance holds when the factor loadings of a measure are found to be equal across groups (p.). The results largely indicated that the TOEFL test exhibited measurement and structural invariance; however, the pronunciation and fluency subscales were an exception to the general observation. The evidence lead the researcher to conclude that final model supported neither of the hypotheses of increasing factor differentiation nor that of decreasing factor differentiation as a function of increasing examinee proficiency (p. 53). The results also supported the policy of using a single score for all language groups and attest to the high construct validity of the TOEFL test.

Recently, a study was also undertaken by Gu, Turkan, & Gomez, (2015) to investigate the dimensionality of the Test of English for Teaching or TEFT. The main objective of the study was to build a validity argument in support of the test's internal structure and the relevance of the score reporting practice. Specifically, the researcher noted that 'With regard to score reporting, the most prominent information on the score report is the total score scale and the associated band and band descriptors' (Gu et al., p. 3). The researchers sampled 1,307 participants from a group of students who wrote one form of TEFT administered during a pilot study in 2012. TEFT is designed in such a way that it measures language proficiency in two broad areas. Firstly, the test items are categorized into four primary language skill area; namely, Reading, Writing, Speaking, and Listening. Secondly, the test is organized into three content areas based on the functional use of language in a classroom setting. The functional skill areas are Managing the classroom, Understanding and Communication lesson content, and Providing feedback(Gu et al, 2015).. A candidate receives a separate score for language skill area and another score for the content component. The overall score is obtained by combining the two component scores.

According to Gu et al. (2015), "A multitrait-multimethod confirmatory factor analysis approach was taken to examine the influence of both skill and content on test performance' (p. 3). The analysis first focused on the establishment of a plausible baseline model for skill and content dimensions separately. The best fitting models for the separate dimensions were then combined. The decision to test separate models was motivated by the score reporting policy (i.e., reporting separate score for skill and content area dimensions).

In terms of the skill dimension, three competing models were tested; (a) A correlated four factor model, (b) A higher-order model, and (c) A Bifactor model. The analysis indicated an adequate fit for the four factor skill model and the higher-order model. A parallel analysis was done for the content section; three models were also hypothesized. However, all the three content models produced unsatisfactory fit indices and were subsequently considered to be inadmissible. This scenario compelled the researchers to create a new model where the Managing the classroom and Understanding and Communication lesson content factors were combined to form a new factor. This essentially created a two factor content model.

To test an overall model for the entire test, the best fitting skill model and best fitting content models were combined to create four models (i.e., . High-Order Skill Relationships and Two Content Factors, Four Skill Factors and Two Content Factors, High-Order Skill Relationships and Three Content Factors, Four Skill Factors and Three Content Factors). The analysis showed that only the Higher-Order model converged. The researchers noted that ‘Generally, the results support the current score reporting practice, that is, to report a total scaled score along with score information on skills and language use in specific content areas’ (p. 9). Also, the results supported findings by other researchers suggesting that language proficiency is hierarchical and multi-componential in nature (Fouly, Bachman & Cziko, 1990; Bachman, Davidson, Ryan & Choi, 1995; Sawaki, Stricker & Orange, 2008; Sawaki et al., 2009). Therefore, the literature sources reviewed indicates an apparent consensus amongst language experts that language proficiency is best represented by a multi-componential hierarchical model.

METHODS

The study follows the quantitative approach as correlation coefficients for PSLE English Language continuous writing component scores are used in the analysis. The correlation matrix of the scores is submitted for analysis using CFA procedures to identify reliable dimensions.

Sampling procedures

There are ten educational districts in Botswana; each district is divided into regions. The South Central District has four educational regions with a total of about 30 public primary schools. The current study derived the sample from the four regions within South Central District. Simple random sampling procedures were applied at the level of schools; as a result, 22 schools were sampled and all of the composition and letter scripts in each of the sampled schools were included for analysis making a total sample of 1800.

Instrument

The candidates wrote PSLE English Language composition and letter examination in 2003. An analytic marking scheme comprising twenty criteria was used to determine the language proficiency of each candidate. Though composition and letter modalities are graded separately, the two marks are combined to produce a score for continuous writing.

Analysis

The suitability of the data for factor analysis was assessed by means of Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity. Both tests yielded favourable results as indicated in Table 2.

Table 2: KMO and Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.899
Bartlett's Test of Sphericity	Approx. Chi-Square	8111.610
	df	190
	Sig.	.000

The methodological literature on factor analysis strongly recommends use of multiple goodness-of-fit indices because each criterion has its own advantages and disadvantages (Henson, Capraro & Capraro, 2004; Hu & Bentler, 1999). For the present data, two factor extraction techniques were used; these are the Kaiser criterion (K1 rule) and the Scree Plot. The rationale of the K1 rule is that a factor or dimension with substantive meaning should account for variance in the data that exceeds the variance explained by a single item. Since the maximum variance that an item can explain is 1, the K1 rule leads to the retention of dimensions with eigenvalues greater than 1 (Costello & Osborne, 2005). Table 3 shows that four factors satisfy the KI rule and their cumulative variance explained is 48.684%. It should be noted that the fifth factor missed the cut of point by a very small margin.

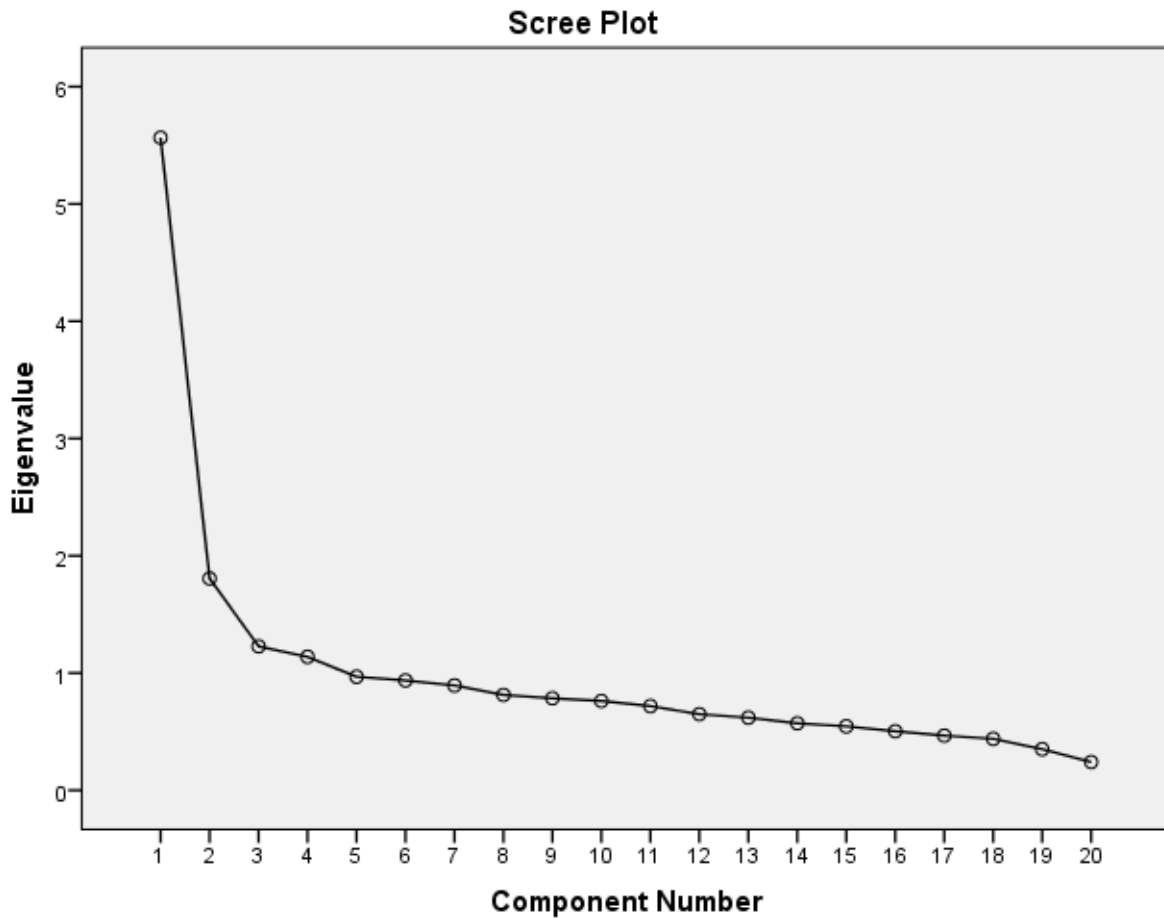
Table 3: Total variance explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.565	27.826	27.826	5.565	27.826	27.826
2	1.806	9.028	36.854	1.806	9.028	36.854
3	1.228	6.139	42.993	1.228	6.139	42.993
4	1.138	5.691	48.684	1.138	5.691	48.684
5	.968	4.841	53.525			

Extraction Method: Principal Component Analysis. Rotation method: Varimax

The four dimensional structure represented by the KI rule is corroborated by the Scree plot (Figure 1). The Scree test arranges dimensions by order of significance with the factor that explains the larger variance listed first. An examination of the plot shows that a bend or 'elbow' occurs after the fourth factor. Factors that occur after the bend lie in an almost horizontal position. However, it should be noted that the point where the lines changes from vertical to horizontal is not so pronounced. This could mean that Dimension 4 is not well represented in the data or Dimension 5, though excluded on the basis of the K1 rule, has some theoretical value. Generally, the K1 rule and Scree plot suggest a four dimensional solution as a more parsimonious representation of the language proficiency construct.

Figure 1: The Scree Plot



Factor Loadings

The rotated component matrix (Table 4) below shows the factor loading of the 20 measured variables used to assess language proficiency. Varimax rotation was used to identify items that have significant loading on one of the four dimensions extracted. This ensures that each item has a higher loading on only one factor to avoid cross loading items and at the same time bring about simple structure. For example item 1 (Introduction) has a higher loading of .715 on Dimension 2 an insignificant loading of .096 and .065 on Dimension 1 and 2 respectively.

Table 4: Variable loading for the rotated component matrix

	Component			
	1	2	3	4
1. Introduction	.096	.715	.065	.030
2. Descriptive solution	.122	.374	.581	-.014
3. Sequence of events	.107	.021	.615	-.112
4. Realistic composition	.253	.428	.292	-.070
5. Correct Spelling	.016	.115	.613	.204
6. Appropriate Adjective	.189	.666	.305	.034
7. Defined Sentence	.057	.204	.531	.111
8. Correct punctuation	.149	.584	.034	.115
9. Assistance required	.238	.525	.226	-.040
10. Feeling for assistance	.175	.693	.085	-.039
11. Correct address	.121	-.091	.254	.633
12. Key Sentence	.632	.158	.151	-.081
13. Consistency in content	.783	.277	.054	.117
14. Different adjectives	.672	.063	.183	.116
15. Correct conjunctions	.375	.208	-.165	.464
16. Sentences coherent	.781	.262	.073	.089
17. Varied Sentence length	.488	-.062	.180	-.542
18. Reasons for the job	.745	.212	.089	.168
19. Consistent Register	.374	-.002	.130	.430
20. Consistent Tense	.667	.294	-.058	.055

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 7 iterations.

Naming of the four dimensions

Since the analysis converged on four dimensions, the next logical stage is to find an appropriate name for each of the dimensions. Naming of a dimension is determined by content of items that have significant loading on the dimension under consideration. Dimension 1 was named ‘Logical Development of Ideas’ because most of the items that load on this dimension dealt with the ability of the candidate to put ideas in a logical and coherent manner. The Dimension 2 was named ‘Communication of Feelings’ as the majority of items associated with the dimension required the candidate to show their feelings and/ or emotions. Dimension 3 and Dimension 4 were labelled as ‘Correct Use of Language Devices’ and ‘Appropriate Register’. Dimension 3 deals with correct use of adjectives and conjunctions while Dimension 4 examines the ability to write an address and salutation correctly when writing a letter.

DISCUSSION

The original 20 metric variables used to measure language proficiency were submitted to PCA with Varimax rotation. The factor analysis applied successfully generated four dimensions which were subsequently named. The components extracted had substantive meaning as the minimum number of items loading on a dimension was three. The four factor solution is in support of the multi-componential nature of language previously established other language researcher (). Most importantly, the four factor solution is contrary to the current PSLE assessment practice that regards continuous writing as being unidimensional. Categorizing all items relating to continuous writing under the Language Use dimension does not seem to be a true reflection the continuous writing construct. The current results identify a four factor model as a more parsimonious structure for continuous writing. Consequently, candidates should be assessed by means of four separate scores; there should be a score for Logical Development of Ideas, Communication of Feelings, Correct Use of Language Devices and Appropriate Register. However, more research work needs to be done to establish the substantive and theoretical relevance of Dimension 4 and Dimension 5. The cu-off point between these two dimensions is currently ambiguous.

CONCLUSION

Research on the dimensionality of language proficiency strongly points to the multidimensional nature of the language proficiency construct. The extraction of four factor model that account for 48.684% of the variance in the matrix provides additional evidence for the four multi-componential hypothesis. Therefore, score reporting practice that utilizes four separate scores for each of the identified dimensions would serve as a better diagnostic tool than the current unidimensional approach.

REFERENCES

- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*, 74-94.
- Carrol, J. B. (1983). Psychometric theory and language testing. In J. W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 80-107). Rowley, MA: Newbury House.
- Costello, Anna B. & Jason Osborne (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, *10*(7). Available online: <http://pareonline.net/getvn.asp?v=10&n=7>
- Fouly, K. A., Bachman, L. F. & Cziko, G. A. (1990). The divisibility of language competence: A confirmatory approach. *Language Learning*, *40*, 1-21
- Gu, L., Turkan, S., & Gomez, P. G. (2015). Examining the internal structure of the test of English for Teaching (TEFT). New Jersey: Educational Testing Service.
- Henson, R. K., Capraro, R. M., & Capraro, M. M. (2004). Reporting practice and use of exploratory factor analysis in educational research journals: Errors and explanations. *Research in Schools*, *11*(2), 61-72.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional versus new alternatives. *Structural Equation Modeling*, *6* (1), 1-55.
- Joreskog, K. G., & Sorbom, D. (1996b). PRELIS 2: User's reference guide [Computer software manual] Chicago: Scientific Software.

- Kunnan, A. J. (1995). Test taker characteristics and test performance: A structural modeling approach. Cambridge, England: Cambridge University Press.
- Messick, S. (1996). Validity and washback in language testing. Princeton, New Jersey: Educational Testing Service.
- Oller, J. W., Jr. (1976). Evidence of a general language proficiency factor: An expectancy grammar. *Die Neuren Sprachen*, 76, 165-174.
- Oltman, P. K., Stricker, L. J., & Barrows, T. (1988). Native Language, English proficiency, and the structure of the TOEFL. TOEFL research report 27. Princeton, NJ: Educational Testing Service.
- Sasaki, M. (1999). Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses. New York: Lang.
- Sawaki, Y., Stricker, L., & Orange, A. (2008). Factor structure of the TOEFL Internet Based Test (iBT): Exploration in the field trial sample. New Jersey: Educational Testing Service.
- Shin, S. (2005). Did they take the same test? Examinee language proficiency and the structure of language tests. *Language Testing*, 22(1), 33-57.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Stricker, J. L., Rock, A. D., & Lee, Y. (2005). Factor structure of the LanguEdge test scores across groups. Monograph Series. New Jersey: Educational Testing Service.
- Swinton, S. S., & Powers, D. F. (1980). Factor analysis of the TOEFL. Research Report 6. New Jersey: Educational Testing Service. Retrieved from: <http://www>.

 ©IJSRE

¹ Mogapi, M. O. is a lecturer in the Department of Educational Foundations, University of Botswana Gaborone, Botswana. He can be via email at: Molefhe.mogapi@mopipi.ub.bw



Teacher Education and Teacher Efficacy do they Match in Botswana Colleges of Education?

Waitshega Tefo Smitta Moalosiⁱ
University of Botswana
smitta.moalosi@mopipi.ub.bw

Joe T. Matsogaⁱⁱ
University of Botswana

Abstract

The training of teachers has been a priority to the government of Botswana since the country's independence in 1966. Education Policies have been revised to cater for quality education for citizens. This paper argues that although it is relevant to upgrade colleges, and offer quality education and assurance processes to monitor training institutions; teacher efficacy remains fundamental among pre-service teachers while they are still at training. The researchers present their arguments in relation to the mentioned above title. Research findings carried in the five colleges of Botswana that train primary and secondary school teachers showed a low teacher efficacy in relation to the three subscales of Teacher's Sense of Efficacy Scales, Efficacy in Student Engagement, Instructional Strategies and Classroom management. The paper further states that, ecological research as it focuses on classroom management and student engagement is essential, as it highlights how teachers should manage classroom behaviors and provides some strategies to be implemented to engage students in learning.

Keywords: Quality Education, Teacher Efficacy, Ecological research, Classroom Management, Student Engagement.

Reference to this paper should be made as follows:

Moalosi, W. T. S., & Matsoga, J. T. (2016). Teacher Education and Teacher Efficacy do they Match in Botswana Colleges of Education. *International Journal of Scientific Research in Education*, 9(1), 21-28. Retrieved [DATE] from <http://www.ij sre.com>

INTRODUCTION

The Botswana government as evidenced in several policy documents such as the Vision 2016, the Revised National Policy on Education (RNPE) of 1994 and many more acknowledges the fact that giving quality education remains an important practice all the time since education is a fundamental human right. It is a driver of “sustainable development, peace and stability within and between countries (Sota, 2013). It is an indispensable tool for active participation in the economy and society of the twenty-first century, which affects the rapid globalization (p. 437).

In Botswana, like in most of her African counterparts, compared to at Independence, education landscape has changed tremendously, across all sectors. It must however, be noted that for all achievements attained in the system it is the teacher who has always been very pivotal for all teaching and learning processes. Emphasizing the significance of the role that teachers play is underscored by the country’s Revised National Policy on Education (1994) which stipulates that “the success of any education system depends largely on teachers, they are the catalysts of the learning process and on them mainly rests the whole system” (p. 4). Sharing similar sentiments about the significance of teachers in the education system is the Botswana Education Statistics (2008) which stipulates that: Teachers are among the most important resources in the education system and the qualifications of teachers, methods of teaching, and resources provided have a significant influence on the overall quality of education children receive” (p. 31).

With teacher education accorded the most significant role expressed in the above it becomes even apparent why teacher education has taken great strides, globally. Further, with the above in mind, the paper argues that the training of teachers is not the only answer to educational challenges facing Botswana today. To put discussions here in context the subsequent section of this paper gives a brief background of teacher education processes in the Botswana colleges of education.

Programme Context in the Colleges of Education

From the above introductory remarks it is clear that education remains a priority area for the Botswana government both in terms of its goals and its budgetary allocation. It is therefore not amazing that over the years government channeled millions of Pula (local currency) into all efforts targeting teacher development. Other factors that prompted government commitment to educational development has been the persistently low performance in the academic achievement of learners forced government to undertake a number of unprecedented initiatives designed to improve the quality of schooling.

The Diploma in Primary Education

The Diploma in Primary Education (DPE) programme was piloted in Tlokweng College of Education in 1985. Specifically, by introducing the diploma; the government was keeping with the socio-economic and educational advancements of the country which made it imperative to upgrade teacher qualifications in the primary education to remain competitive in the global market (Tlokweng College of Education Prospectus, 2004/2005). Reiterating this move by government, a major education commissions undertaken in 1993 and 1994 noted that the Primary Teachers’ Certificate (PTC) was no longer relevant for the socio-economic and educational

system of the country. DPE was then extended to other colleges of education, (Lobatse 1996; Serowe 1997 and Francistown 1998).

The Diploma in Secondary Education

During the years 1985 and 1990 respectively, government established the Molepolole and Tonota Colleges of Education as fully-fledged colleges offering a Diploma in Secondary Education (DSE). Primarily, the DSE is a full time programme extending over three years and its aim is to train teachers for Community Junior Secondary Schools each college approximately producing about 400 graduates (Ministry of Education and Skills Development 2004). With regards to courses offered the programme offers both major and minor teaching subjects for instance, core areas of Communication and Study Skills, Foundations of Education Business studies, Social studies, Home Economics, Physical Education and Library Studies. In addition, the programme has service branches, namely, Educational Technology and Special Needs Education.

QUALITY ASSURANCE PROCESSES IN THE COLLEGES

With regard to all quality control processes in the colleges of education the University of Botswana (UB) has been overseeing this through its Office Affiliated Institutions (OAI). Primarily, the OAI is one of the five units that constitute the Centre for Academic Development (CAD). This unit is responsible for the ensuring the maintenance of quality of teaching and learning at the affiliated institutions. Other than the University, there is the Botswana Qualifications Authority (BQA) formerly the Tertiary Education Council (TEC) whose primary mission is to ensure that institutions offer quality programmes that directly meet all market demands. As noted above as demanded by both the UB and the BQA for quality assurance purposes all colleges of education have been instructed to align their programmes to the national, regional and international best practices in tertiary education. It is in attempts to meet this need that since 2012 all colleges have slowly been introducing semesterization of their programmes as well as the use of credit system (Ministry of Education and Skills Development, 2013).

Though the above-mentioned plans for upgrading programs and producing qualified personnel will help serve the nation at large, it is not known whether the trainee teachers believe that they can impact student learning through their classroom management, instructional strategies, and student engagement. The educational vision advocates for all Botswana's citizens to be offered "equitable lifelong education and training that is relevant and responsive to the rapid technological development and the changing socio-economic environment, and that produces knowledgeable, skilled, enterprising, and independent individuals" (National Development Plan, 2009, p. 268). Teacher efficacy is therefore essential in Botswana education system and practice. Teachers or trainees ought to believe in themselves to have a positive impact on students' learning. The terms teacher efficacy, teachers sense of efficacy and self-efficacy will be used interchangeable in this paper.

TEACHER EFFICACY AN IMPORTANT DIMENSION IN EDUCATION

The concept of teacher efficacy is important; Poulou (2007) suggested that psychology and education researchers have based their ideas about teacher efficacy on Bandura's theory of self-

efficacy. Self-efficacy, as defined by Bandura (1997), involves “beliefs in one’s capabilities to organize and execute the course of action required producing given attainments” (p. 2). The definition is relevant to teachers, because they have to believe that they can influence the learner positively, and organize their instruction effectively so that good results are produced, showing that their students know how to learn.

Teacher efficacy is also defined as the teacher’s “judgment of his/her capabilities to bring about the desired outcomes of student engagement and learning, even among those who may be difficult or unmotivated” (Tschannen- Moran and Woolfolk Hoy, 2001, p. 783). This definition implies that individual teachers are able to evaluate their performance to find out whether it has an impact on students learning or not, if not the teacher is expected to improve their performance in order to assist students. Teacher efficacy research can assist pre- service teachers who have insufficient beliefs about their teaching abilities. Personal teaching efficacy has been thought of as having an impact on the growth of beliefs about being a good teacher according to Ng et al. (2010). Efficacy is the ability to bring into being the desired results (Tschannen- Moran and Hoy, 2007). Therefore, teacher self-efficacy can influence pre-service and in-service teachers to be effective and manage difficult students in their work. Rizvi and Elliot argued that teacher efficacy is "an important dimension of teacher professionalism and, together with other dimensions such as teacher practice, leadership and collaboration" (as cited in Cheung, 2008, p.103). Research on pre-service teachers’ beliefs is important because they can be assisted while training to become effective teachers with a high teacher efficacy.

Mulholland and Wallace (2001) reported greater students’ teacher efficacy beliefs that were related to high pupil achievement and desire for positive teacher characteristics. Poulou, (2007) investigated the students teachers efficacy beliefs and factors influencing their beliefs and reported that students teachers stressed the importance of personal motivation particularly their interest in working with pupils as well as (they) students teachers to perform better in teaching.

The findings of Gowie, (2010) with pre-service teachers reported that "instructional practices are examined as potential influence on development of efficacy beliefs" (p. 1). Thus, while pre-service are training to become teachers, teacher educators should use more of instructional strategies that will have an impact on the development of their high efficacious beliefs that can benefit learners at schools.

Teacher efficacy research can reduce some problems in education, especially if researchers investigate teacher efficacy with other factors. For example, experience in teaching has been associated with positive teacher efficacy. A teacher with positive beliefs about being a teacher is useful because it can bring convictions to teachers with regard to students learning.

Teacher efficacy construct is vital among in- service and pre-service teachers in Botswana. The training of teachers may be a priority of Botswana government and teacher education programs upgraded to improve education in general. But what is a concern to the nation at large is declining performance of students from primary, junior and senior secondary schools.

The discussion of this paper borrows from the research studies of Moalosi, and Forcheh (2015). The study was carried on the two colleges that train secondary school teachers and the other three colleges of education in Botswana.

Moalosi (2015) researched Molepolole and Tonota Colleges of Education pre-service teachers on teacher efficacy. Teachers’ Sense of Efficacy (TSES) instrument comprising 25 items were used to measure the trainee’s teacher efficacy. Pre-service teachers in Molepolole college of Education had a high teacher efficacy in Instructional Strategies, Student Engagement

and Classroom management. While their counterparts reported a low teacher efficacy beliefs. What is a challenge to Botswana education is, these mentioned above teachers trainees are assumed to be given quality education as discussed earlier, and also teachers in Botswana are perceived to be the key leaders in teaching and learning process. Teachers or teacher trainees who have a low teacher efficacy, who do not believe in themselves that they can positively influence to students to learn are a concern to the community at large. Education does seem to be an important human right because teachers cannot deliver effective instruction to students. The study findings of Moalosi and Forchheh (2015) used TSES subscales with the five colleges of Botswana that train teachers reported Serowe College of Education having a higher teacher efficacy in all the three subscales of teacher efficacy, Efficacy in Student Engagement, Instructional Strategies and Classroom management. Interestingly, the mentioned above college has mature in-service teachers than other colleges. Could the teaching experience they obtained prior training influenced their teacher efficacy? Younger males had higher efficacy beliefs in classroom management than females in all the five colleges.

The mentioned above results shows the teaching profession shifting from being a female job to male (see Moalosi and Forchheh, 2015). Research of teacher efficacy is not new in Botswana, Dibapile, (2012) conducted a study on teacher efficacy with Botswana Junior Secondary school teachers in the city of Gaborone and surrounding areas. There were 1, 000 participant's and only teachers with Post-graduate Diploma in Education rated themselves higher in one subscale of Student Engagement. Most teachers did not have high efficacious beliefs in instructional strategies and classroom management. While success in any education system is perceived as depending greatly on the training of teacher Botswana government need to change ways of training teachers.

SUGGESTED APPROACHES TO THE SOLUTION

The discussion in this paper, though argumentative offers some useful solutions that could reduce poor performance in Botswana Education system. Ecological research could offer shed light on what teachers ought to do to engage students in learning as well as reflecting to their practice. Ecological research of Kounin and Gump (as cited in Brophy, 2006) described how teachers can manage classroom behaviors as well as how environmental factors that can affect learning. They focused on different activities; both supported (affordances) and prohibited (constraints). During their classroom investigations classrooms were perceived as ecologies that can be investigated in terms of innovations created to fulfill a specific goal (Brophy, 2006). Kounin and Gump focused on different aspects of classroom environments (e.g., whole class, small groups, individual) and the activities that took place in them (e.g.,) more teacher-student discourse occurred in lesson settings than in seatwork settings' (p. 759).

The above mentioned concepts could be allowed in Botswana education system for a change since the performance of students from primary to senior secondary schools decline each year. These researchers reported the following from ecological research conducted in classrooms:

- *Withitness*, “remaining with it” teachers concentrated much on what was happening in the classroom most of the time; they examined the classroom settings while engaged with individual students or small groups.

- *Overlapping*: Teachers performed more than one task at a time. For example, they remained close to students to capture students' attention and to conduct the lesson without interrupting.
- *Signal continuity and momentum during the lesson*: The teacher is expected to teach well planned lesson, efficient classes centered on capturing students' attention. The teacher should also present the content continually ("the signal") that is more compelling than the noise of competing distractions, and sustaining the momentum of the signal throughout the lesson.
- *Conveying purposefulness*: The teachers who were efficient classroom managers took advantage of the time allocated for teaching and evaluated whether the students were participating and learning. They encouraged students to be responsible for finishing their work on time. Daily revision of work was arranged, and students were give effective evaluations. (Brophy, 2006).
- *Teaching appropriate conduct*. Effective classroom managers are defined as having an understanding about expectations of students as well as what they cannot accept. Teachers focused on students 'work, which was important, and also on teaching them how to do their work effectively (Brophy, 2006)
- *Maintaining attention*. Effective managers pinpointed students who were confused or not paying attention in class. They organized seating arrangements so that students faced the direction in which they could be best concentrate. In addition, effective managers changed the tone of their voices when they spoke with students, they moved around the class or "paced[d] to sustain attention" (Brophy, 2006).

The discussed findings of ecological researchers conducted on observations in the classroom settings sheds light to our understanding of what happened in classes that could bring improvement in engaging students in learning and managing classroom behaviors. The writers have not discussed all the research findings of ecological research, but highlighted some that they think could be more useful to teachers and teacher educators. It should be borne in mind that the writers do not imply that teachers in Botswana are not applying in their practice what is discussed in this paper and do not generalize that all teachers in Botswana have a lower teacher efficacy. It is the research findings of teacher efficacy in Botswana that guides this argument. The researchers are aware that there are some factors affecting students' performance in Botswana. For example, parental involvement is another challenge in Botswana schools particularly at primary, and secondary schools. Some parents do not attend Parents and Teachers Associations meetings where they can work together with teachers to help their students learn. Shortages of classrooms, where students are taught in open spaces affect learning in one way or the other and insufficient teaching materials still hampers Botswana schools.

The other thing that is a major problem is the teacher student- ration. Teachers teach many students that it is not easy for them to manage unruly behaviors. In her research reports, Pheko, (2010) with school visits, primary and junior secondary school asserted " But the reality in most schools the researcher visited that the teacher- student ratio at junior secondary schools is higher as it was 1: 51" (p.218). Contrary to the Botswana Education Statistics that reported that, the teacher ratio at primary and secondary schools is 1: 28 and 1: 40. Such crowding in classes will results in poor performance because the teacher won't be able to attend individual learners effectively as desired. Student engagement is vital in learning, learners need to be task focused

if they have to improve their performance and this not only achieved by offering quality education, but by producing teacher with a high teacher efficacy who will have an impact on students' learning. Daugherty (2005) reported that, efficacy for student engagement was "a significant predictor to three of the five (engaging the struggling learner, motivating students, instruction and assessment." (p.59). The continuation of declining performance of students in Botswana makes an individual to ponder and ask him/herself, are the students engaged in learning effectively?

CONCLUSION

To conclude, much is appreciated about how the government of Botswana has done in previous years to improve the training of teachers. Colleges of Education have been upgraded to make sure teachers obtain higher qualifications to enhance students' learning. However, the construct, teacher efficacy is still essential because the students cannot perform well in their studies unless they interact with teachers who have high efficacious beliefs.

Research studies about teacher efficacy reported in this paper shows most teachers as well as pre-service have a low teacher efficacy; particularly in the three subscales, Efficacy in Student Engagement, Instructional Strategies and Classroom Management. The results causes concern because it shows that, it is not how good the training programme can be, but teachers has to believe in themselves, be confident, and committed to their work so that they can impact students' learning. Ecological research findings reveal that teachers engage students in learning. It is beyond doubt that if some teachers in Botswana schools can use ecological research in their classrooms, they can maintain effective instruction, and manage classroom behaviors well as depicted in ecological research. Thus, poor performance in Botswana schools which has been a challenge for some years could reduce. What happens in the learning environment is vital particularly inside the classroom students should be given the opportunity to know how to learn and be given activities that will motivate them to learn.

REFERENCES

- Bandura, A. (1997). *Self- efficacy: The exercise of control*. New York: Freeman & Company.
- Brophy, J. (2006). Observational research on generic aspects of classroom teaching. In P. A. Alexander s & P. H. Winne (Eds.), *Handbook of Educational Psychology (2nd ed)* (pp.755-737). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Cheung, H.Y. (2008) Teacher efficacy: A comparative study of Hong Kong and shanghai primary in-service teachers. *The Australian Educational Researcher*, 35, (1)103-123.
- Daugherty, S. G. (2005). Teacher efficacy and its relation to teachers' behaviors in the classroom. Unpublished doctoral dissertation, University of Houston, Texas.
- Dibapile, W.T.S. (2012) Teacher efficacy and classroom management among botswana junior secondary school teachers. Published doctoral dissertation, University of Tennessee, Knoxville, USA.
- Diploma in Secondary Education (2010). Revised Academic Regulations
- Government of Botswana (1977). Education for Kagisano. Gaborone: Government Printer.
- Government of Botswana (1997). Long Term Vision for Botswana. Gaborone: Government Printer.
- Government of Botswana (1999) National Development Plan 9, Ministry of Finance and

- Development Planning: Gaborone: Government Printer.
- Government of Botswana (2003) National Development Plan 9, Ministry of Finance and Development Planning: Gaborone: Government Printer.
- Gowie, C. K. (2010). Developing Efficacy Beliefs in Preservice Teachers. *NERA Conference Proceedings*. Paper 11, 2010.
- Makki , Y. Y, Themba, A., & Tshekedi .S. (2002) Tracer Study of Tonota College of Education. Ministry of Education and Skills Development, (2013). Gaborone: Government Printer.
- Moalosi, W. T. S., & Forcheh, N. (2015). Self-efficacy levels and gender differentials among teacher trainees in colleges of education in Botswana. <http://dx.doi.org/10.5539/jel.v4n3p>.
- Mulholland, J., Dorman, J. P., & Odgers, B. M. (2004). Assessment of science teaching efficacy of pre-service teachers in an Australian University. *Journal of Science Teacher Education*, 15, 313-331.
- National Development Plan 9, (2009). Long term vision for Botswana: Ministry of finance development and planning. Gaborone: Government Printer.
- Ng, W., Nicholas, H., & Alan W. (2010). School experience influences on pre-service teachers' evolving beliefs about effective teaching. *Teaching and Teacher Education*, 26, 278-289.
- Pheko, B. C. (2010). A conflict between aims and outcomes: Comparison of one cohort's primary school leaving examinations and junior certificate results of 2003 and 2006. *Journal of Social Science*, 22(3).
- Poulou, M. (2007). Personal teaching efficacy and its sources: Student teachers' perceptions. *Educational Psychology*, 27(2), 191-218.
- The Revised National Policy on Education: Government Paper Number 2 of 1994. Gaborone: Government Printer.
- Tlokweng College of Education Prospectus (20004/2005).
- Tschannen-Moran, M., Hoy, A., & Woolfolk, A. E. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Sota, J. (June, 2013) Globalization and Education for All: The Challenge of the Twenty First Century. *Sociology Study*. Volume 3, (6). Tonota College of Education Improvement Plan 1995-1999.
- University of Botswana (2009). Teaching and Learning Policy, Affiliated Institutions.
- University of Botswana (2005). Semesterisation and Modularization. Unpublished
- University of Botswana (2000). Semesterisation and Modularization. Unpublished

 ©IJSRE

ⁱ Dr. Waitshoga Tefo Smitta Moalosi is a Senior Lecturer in the Department of Educational Foundations, University of Botswana. Smitta's research interests are: Teacher efficacy and classroom management, reasons for choosing teaching as a career, self- concept, adolescent development and trauma in orphans and vulnerable children. He can be reached via email at smitta.moalosi@mopipi.ub.bw

ⁱⁱ Dr. Matsoga is a Senior Lecturer in the University of Botswana, Faculty of Education where he teaches courses in Curriculum Theory and Design. He attained his MEd from the University of Birmingham, UK and a Ph.D. from Ohio University in the United States of America. Dr. Matsoga is a former police officer and a police trainer in the Botswana Police Service. Before joining the police he worked in several high schools where he taught English Language and Geography Education. Since December 2004 till March 2010 he served as the Chairperson of the Botswana Educational Research Association and remains very active in university, research and teaching.



Culture and Differential Item Functioning in National Examinations Council Senior School Certificate Mathematics Multiple Choice Test in Nigeria

Sele Sylvester Ebesine¹
Department of mathematics,
College of Education, Warri, Delta State, Nigeria
ebesineleesylvester@yahoo.com

Abstract

The present study investigated the influence of culture on DIF in National Examination Council Senior School Certificate Mathematics multiple choice test in Nigeria. The study ascertained the instances of test items in NECO SSCE mathematics multiple choice questions that function differently for Ibo and Yoruba testees. To guide the study, three hypotheses were formulated and tested. An ex-post facto research design was employed in the study using 1,200 SSS3 students to obtain data on students' differential performance in mathematics across the different cultural environments. The instrument for data collection was General Mathematics paper I multiple choice test used by NECO in June/July, 2012 that was made up of 60 items. Chi-square goodness-of-fit test statistics was used to test the hypotheses at 0.05 levels of significance. The study revealed that the test item with DIF in NECO SSCE mathematics multiple choice questions significantly differ for testees from Ibo and Yoruba cultural environments and that the test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for males and females testees from Ibo cultural environment. It equally revealed that the test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for male and female testees from Yoruba cultural environment. Based on these findings, it was recommended that mathematics teachers should adopt the use of culturally responsive instructional techniques to reduce students' differential performance in mathematics irrespective of gender.

Keywords: Culture, DIF, Multiple Choice test, Items bias, Gender

Reference to this paper should be made as follows:

Ebesine, S. S. (2016). Culture and Differential Item Functioning in National Examinations Council Senior School Certification Mathematics Multiple Choice Test in Nigeria. *International Journal of Scientific Research in Education*, 9(1), 29-40. Retrieved [DATE] from <http://www.ij sre.com>

INTRODUCTION

Mathematics is a compulsory subject in Nigeria education system because of its importance as emphasized in the National Policy on Education. Mathematics has been found to be very important because it is needed for all scientific, technological research and technical training. It is a fundamental science subject, which acts as a base index for understanding the complexities for most other fields. Mathematics is also rightly described as the language of science, as well as the central intellectual discipline of the technological societies. Its contribution to mankind has given the subject the prominent it enjoys among other school subjects. This had earned it the status of core subject in the school curriculum at both primary and secondary schools levels. Its status has also made it compulsory for any student seeking admission to read any science or science related course in any tertiary institution to secure a credit pass in it. It is in view of this that Adamu (2007) described mathematics as key to scientific and technological advancement in this age and as the lifeblood of modern science.

Despite the importance of mathematics, there have been reports of students' poor performance in the subject at the secondary school level. To demonstrate the trends of mathematics failure in Nigeria secondary schools, the West African Examination Council (WAEC) revealed consistent declining performance in Senior School Certificate Examination (SSCE) results for the past nine years. While the percentage success of the result in 2006 was 39.92% in 2009, it declined to 31.50%. Reasons given by WAEC report for the low-level performance were dearth of qualified teachers, inadequate and dilapidated classrooms as well as other poor conditions in secondary schools across the country (Ojo, 1993). Similarly, Adamu (2007) identified poor quality of instructional techniques employed by teachers, overcrowded nature of the classes and teachers' poor method as some of the factors responsible for students' poor performance in mathematics.

Several research efforts have been made to solve the problem of poor performance in mathematics. Some of these efforts are in the improvement of method of teaching, quality of instructional technique employed by teachers, and method of presentation before the students in order to improve mathematics results in our secondary schools. These efforts have also resulted in the orientation programme for mathematics improvement projects intends to improve mathematics teachers in primary, secondary and tertiary level of education. Despite these efforts, available statistics show that students' performance in the subject is still very poor. It is interesting to note that none of these efforts looked at the inherent unfairness in test item used in determining of learning outcome.

The results of researches in the field of evaluation, Flores (2000), and Lam (1995) showed that it is evident that a test may be bias, which may unfairly influence examinee's scores due to the presence of irrelevant, non-target constructs which are related to gender; ethnicity, race, linguistic background, socio-economic status or handicapping conditions (Flores, 2000; Lam, 1995), differences in upbringing environment, cultures (Fortunes, 1985). The general concept is that as much as possible test constructed should represent the best test one could obtain. But if a test so constructed is unfair or yields different scores for subgroups for instance (gender, location, religion, socio-economic background and culture) with the same ability as evident from their scores, then such a test is said to be biased or function differentially (Zumbo, 1999). This phenomenon is termed Differential Item Functioning (DIF). Differential Item Functioning (DIF) occurs when examines from different groups shows differing probabilities of success on the item after matching on the underlying ability that the item is interested to measure

(Zumbo, 2009). In essence, DIF occurs when people from different group with the same ability have systematically different responses to test items. DIF of an item can therefore be understood as a lack of conditional independence between an item response and group membership (often gender or ethnicity) giving the same latent ability or trait (Clickman, Seal & Eisen, 2009).

Admittedly, one of the explanation for different responses to test items for people with the same ability is culture. For instance, Price-William, Gordon and Ramirez in Eibisine (2014) reported that the rate of development of universal types of mathematics abilities varied among culture. Lave (1988) and Scribner (1986) averred that cultural variables such as economic specialization, schooling, and number word system have influence on mathematical knowledge. According to them, mathematical thinking used by students to solve problem, the intellectual tools students acquire from their culture and other support parents provide help them develop mathematical knowledge. Therefore, socio-cultural influences such as values, beliefs, communication patterns, and socio-economic conditions prevailing in students cultural groups may influence the way in which they make sense of mathematics terms and the way in which they solve them (Guillerman & Sheron, 2001). As such, students' performance in a test is influenced by cultural activities students engaged in for a long period of time (Lipka, 1991). This suggests that cultural influences that shape the ways in which students interpret and solve mathematics problems affect their performance in the subject.

Significantly, Ercikan (1998) reported that different groups of examinees may have different multidimensional ability distribution due to language and culture. From this perspective, two examinees from different cultural environment may be functioning differentially in each of the components of mathematics. For example, mathematics has these components viz arithmetic and geometry. Arithmetic is related to gain, loss discount, interest charges and cheques among other things which is different from geometry. Geometry has to do with the study of properties of shapes. A student from Ibo cultural environment whose culture favour trades and spatial components of mathematics as depicted in apprenticeship to various trade is reflected in such topics as interest rate, simple interest, profit and loss fractions, percentages and exchange rates. The trade and spatial aspects of mathematics are specifically reflected in Ibo culture is in their construction of houses especially the thatching of the roof which obeys a certain order that suggests the criss-crossing relationship of line, parallels, vertically and horizontally such that the general outlay is graphical and has some rectangle or square in between. Thatching also exposes ideas of inclination, made so to allow rain drop to slide off the rooftops (Odili & Okpobiri, 2011). Similarly, a student from Yoruba whose cultural activities involves weaving by intertwining of ropes to make a perfect cloth materials and measurement of items of clothing with the use of arm lengths, will understand computation, measurement, patterns and relations, combinatory and probability and areas and volumes. This shows that the average Nigerian student in secondary school already has some mathematics inside of him which is used in lives quantifications, measurement, estimations, calculations of periods, recurrence of peculiar events and age even before coming to school depending on his cultural environment.

Therefore, the nature of mathematics is related to culture as such those whose cultural activities are related to a given mathematical components are likely to do better than individuals whose activities are not related to that mathematics component when test items reflect them. Test items in this aspect of mathematics may be simple to an individual from such cultural environment. Thus, there is the need to investigate through research the influence of culture on

DIF in mathematics in National Examination Council Senior School Certificate Examination, given the cross cultural groups that take the examinations.

Purpose of the Study

The study specially focused on:

- Investigating the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that functioned differently for testees from Ibo and Yoruba cultural environments.
- Determining the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that functioned differently for male and female testees from Ibo cultural environment.
- Ascertaining the instances, of test items with DIF in NECO SSCE mathematics multiple choice questions that functioned differently for male and female testees from Yoruba cultural environment.

Hypotheses

The following hypotheses were formulated to guide the study at 0.05 level of significance:

- The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for testees from Ibo and Yoruba cultural environments.
- The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Ibo cultural environment.
- The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Yoruba cultural environment.

Theoretical Framework

The theoretical framework of this study is anchored on the theory of multiple intelligence. This theory was propounded by Gardner (1985) and it posits that individuals have different intelligence and that the mind is not a holistic entity, but instead consists of distinct, independent modules. According to Gardner (1985) intelligence is an aspect of human activities, that is made up of many independent components. The theory listed seven independent components of intelligence such as (1) linguistic intelligence used in musical appreciation, composition and performance, (2) mathematical-logical intelligence used in arithmetic, numerical calculation and logical reasoning, (3) spatial intelligence used in arranging objects spatially, as well as visual art and finding one's way around, (4) bodily-fkinesthetic intelligence use in sport, dancing or simple everybody movement and dexterity, mechanical component which have to do with ability to work with objects, (5) interpersonal intelligence used in relating to others, interpreting social signal outcomes and (6) intrapersonal intelligence used in understanding and prediating one's own behavior and in indentifying aspects of the self and one's own personality. These

components are in agreement with Piaget and Inhelder (1976) assumption that intelligence development is influenced by activities. The position is also supported by Gardner (1985) that depending on what an individual is doing the various independent components of intelligence can be differentially influenced by culture. Thus, the fact that our activities are inherent in our culture, it stands to show that the different aspects of intelligence can develop based on the activities in our culture.

Essentially, Gardner (1985) theory of multiple intelligence appears to be a reasonable and sound explanation of influence of culture on students' performance in the various aspects of mathematics viz: algebra, arithmetic, statistics, geometry and trigonometry. Drawing from Gardner's proposition the aspect of intelligence that has to do with spatial will be more related to a cultural activity that has to do with ability to visualize a space. Whereas trading abilities may align with Gardner's (1985) mathematical-logical intelligence used in arithmetic, numerical calculation and logical reasoning. Based on this premise, differential item functioning in mathematics should be said to have its root in the theory of multiple intelligence, since culture independently determine what individuals sense and perceive within their environment, predict behavior, as well as, performance difference between two groups of comparable ability or performance.

Concept of Differential Functioning (DIF)

A test that exhibits DIF is one that is unfair to a subgroup of the general population in which it is being used. DIF occurs when two groups (reference group and focal group) that are matched in terms of their relevant knowledge and skills perform differently in an item (Umoinyang, 2011). DIF is a threat to test validity and invalidates interpretation of the test results for some groups of the same population (Pido, 2012). DIF occurs when examinees of the same ability do not have the equal probability of getting an item correct (Perrone, 2006; Roever, 2005). This arises mainly due to the sex, cultural, ethnic, religious, or class background of the examinees. Item bias manifest itself in context, language and item structure and format bias (Rovwer, 2005). Content bias refers to a situation where knowledge and or skills tested are not part of the educational background of the examinees. Lack of familiarity with content in test items disadvantages individuals in their performance. The individual's responses to items are not based on other irrelevant abilities. Language bias occurs where words in items have different or unfamiliar meanings for different examinee subgroups. The item has difficult vocabulary, group specific language, and vocabulary and reference pronouns. Item structure and format bias occurs where there is ambiguity in the instructions, items stem or options. The content or clues and explanations given to successfully complete the task provided disadvantage individuals in some subgroups (Perrone, 2006).

Item bias can occur when a characteristic of the item that is not relevant to the test purpose differently influences responses of examinee groups (Ercikan & Lyong-Thomas, 2013). There is an expectation that if an item on a test is not biased, then examinees from two groups who have equal overall ability ought to have the same probability of correctly responding to it. When examinees from different groups that have comparable ability levels have different probabilities of getting on item correct, differential item functioning (DIF) is said to occur (Hambleton, Swaminathan & Rogers, 19991).

It is important to analyze whether items have DIF for at least two reasons. The presence of DIF signals potential bias and bias has an impact on validity of inference drawn from group

comparison. Therefore DIF items, if confirmed to represent underlying bias, are often removed from future administration of a test. Second, items that exhibit DIF may have implications for curriculum and instruction (Lane, Wang & Magone, 1996), particularly if no reason for bias can be found. For example, test items that are presented in a multiple choice item format may consistently exhibit DIF favouring one group, whereas items presented in constructed response item format may favour another group. In this case, if no bias is established, it may be desirable to ensure that all groups receive adequate instruction in completing all types of test items, and that test contain balanced proportions of various item types. DIF should be distinguished from differences in overall group ability.

METHODS

Design

This study adopted the ex-post facto research design. This design was deemed suitable because it enables the researcher to collect data that can reveal DIF in mathematics among students from Ibo and Yoruba cultural environments.

Population of the study

The population of the study was senior secondary school students in class three who are preparing for the senior school certificate examination among the Ibo and Yoruba in Owerri North and West Local Government Area of Imo State and Ibadan North-East and South-West Local Government Areas of Oyo State. The choice of SSS3 students was informed by the fact that they have covered the NECO mathematics syllabus.

Sample and Sampling Technique

The sample comprised one thousand two hundred (1,200) SSS3 students made up of six hundred (600) students each from Ibo and Yoruba cultural environments using proportionate random sampling. The multi-stage sampling technique was adopted to select the sample for the study. In the first stage, one state each was purposively drawn from the South-East and South-West. Second, a cluster of inhabitants who are predominantly Ibo and Yoruba respectively was sampled. This gave rise to Owerri North and West Local in Imo state and Ibadan North-East and South-West Local Government Area in Oyo State which are predominantly inhabited by the Ibo's and Yoruba's. Third, the researcher used the random sampling technique to select six (6) senior secondary schools from each of the two chosen Local Government Areas in Imo and Oyo States respectively. In all, twenty-four (24) secondary schools were sampled for the study. Having done that, fifty (50) SSS3 students were drawn from each of the schools thereby giving a total sample size of 1,200.

Instrument for Data Collection

The instrument for data collection was General Mathematics Paper I Multiple choice Test used by NECO in the June/July 2012 SSCE. It has sixty items in the multiple choice format. The instrument was a standardized test developed by experts in NECO as such there was no need for validation and reliability because the test is already valid and reliable.

Method of Data Collection

The mathematics teachers in the various secondary schools administered the test to the SSS3 students. The students were given the test after receiving instrument for the test. The students completed the test for two hours under the supervision of their teachers.

Method of Data Analysis

The data collected using the instrument were analyzed so as to enable the researcher to test the hypotheses. The chi-square goodness-of-fit test statistics was used to test the hypotheses at 0.05 level of significance.

RESULTS

Hypothesis 1

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for testees from Ibo and Yoruba cultural environment.

Table 1: Chi-square goodness-of-fit test on significant difference in the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that differ for Ibo and Yoruba cultural environments.

Group	Observed frequency	Expected frequency	DF	Level of significance	Decision
Ibo	27	30	1	0.05	Rejected
Yoruba	20	30			

$$x^2 \text{ Cal.} = 3.63; \text{ Crit. } x^2 = 3.84$$

Table 1 presents chi-square goodness-of-fit test on significant difference in the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that differ for testees from Ibo and Yoruba cultural environments. From the table, the calculated x^2 value is greater than the critical x^2 value, therefore the hypothesis is rejected. This suggests that the test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for testees from Ibo and Yoruba cultural environments.

Hypothesis 2

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Ibo cultural environment.

Table 2: Chi-square goodness-of-fit test on significant difference in the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that differ for male and female testees from Ibo cultural environment.

Group	Observed frequency	Expected frequency	DF	Level of significance	Decision
Male	18	30	1	0.05	Rejected
Female	12	30			

$$x^2 \text{ cal.} = 15.6; \text{ Crit. } x^2 = 3.84$$

Table 2 presents chi-square goodness-of-fit test on significant difference in the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that differ for male and female testees from Ibo cultural environment. From the table, the calculated x^2 value is greater than the critical x^2 value, therefore the hypothesis is rejected. This suggests that the test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for male and female testees from Ibo cultural environment.

Hypothesis 3

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Yoruba cultural environment.

Table 3: Chi-square goodness-of-fit test on significant difference in the instances of test items with DIF in NECO SSCE mathematics multiple choice questions that differ for male and female testees from Yoruba cultural environment.

Group	Observed frequency	Expected frequency	DF	Level of significance	Decision
Male	8	30	1	0.05	Rejected
Female	5	30			

$$x^2 \text{ cal.} = 36.97; x^2 = 3.84$$

Table 3 presents chi-square goodness-of-fit test on significant difference in the instances of test with DIF in NECO SSCE mathematics multiple choice questions that differ for male and female testees from Yoruba cultural environment. From the table, the calculated x^2 value is greater than the critical x^2 value, therefore the hypothesis is rejected. This implies that the test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for male and female testees from Yoruba cultural environment.

DISCUSSION

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for testees from Ibo and Yoruba cultural environments.

Analysis of differential item functioning (DIF) of responses by students in NECO SSCE mathematics multiple choice questions in June/July 2012, showed that the instances of test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for testees from Ibo and Yoruba cultural environment as presented in table 1. The result indicated that the test contain items that function differently for testees from Ibo and Yoruba cultural environment.

In effect the instances of test items with DIF in NECO SSCE mathematics multiple choice questions depend on those experiences and activities of the cultural groups. This finding is consistent with that of Greenfield (1997) that culture and society shape mental functioning, individual have predisposed notions of how to respond to questions, solve problems among others. The predisposition may influence the way in which students interpret materials presented in test and the ways in which students respond to test items. This result further support that of Schliemann, Carraher and Ceci (1997) that the quality of culture in which one lives plays a dominant role in performance and development of mathematics knowledge. According to them, there are mathematical thinking used by students to solve problems that arise in customary activity such as play, trade, farm activities and work. Also, Williams. Gordon and Ramirez (1969) stated that the rate of development of universal types of mathematical abilities varied among culture. This finding further buttressed the finding of Lave (1988) and Scribner (1986) that cultural variables such as economic specialization, schooling and number—word system have influence on mathematical knowledge. They further reported that mathematical thinking used by students to solve problems, the intellectual tools students acquire from their culture and other support parents provide help them (students) develop mathematical knowledge.

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Ibo Cultural environment

Analysis of differential item functioning (DIF) of responses by students in NECO SSCE mathematics multiple choice questions in June/July 2012, showed that the instances of test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for male and female testees from Ibo cultural environment as presented in Table 2. In effect the test had items that measured different things for male and female testees from Ibo cultural environment with the same latent ability in mathematics. Therefore, the instances of test items with DIF in NECO SSCE mathematics multiple choice questions had items that are indepth in male activities and experiences in Ibo cultural environment. This result supported that of Doo Little and Cleary (1987) that males are better on word problem and algebra items favoured females and Beller and Gafni (1996) that measurement items and terms involving problem solving favoured males.

They also considered the age levels and concluded that gender-related DIF becomes larger as age increases.

The instances of test items with DIF in NECO SSCE mathematics multiple choice questions does not significantly differ for male and female testees from Yoruba Cultural environment

Analysis of differential item functioning (DIF) of responses by students in NECO SSCE mathematics multiple choice questions in June/July 2012, showed that the instances of test items with DIF in NECO SSCE mathematics multiple choice questions significantly differ for male and female testees from Yoruba cultural environment as presented in Table 3. In effect the test had items that measured different things for male and female testees from Yoruba cultural environment with the same latent ability in mathematics. Thus, the instances of test items with DIF in NECO SSCE mathematics multiple choice questions had item concepts that are in-depth in male activities and experiences in Yoruba cultural environment. This result agreed with that of Abedalaziz (2010a) that gender-related DIF is a constant concern on larger scale standardized achievement test in mathematics because differences between female and male are often found,

The result also tallies with that of Escorial (2004) that there are evidences of gender DIF in Advance Progressive Matrices Test among applicants to a private university. It was attributed to visuo-spatial nature of the test. Males were performing better. Furthermore, another study by Abedalaziz (2010b) found that numerical ability math items favoured female examinees, whereas items involving special and deductive abilities favoured males.

CONCLUSION AND RECOMMENDATION

This study aimed at ascertaining the influence of culture on DIF in mathematics in National Examination Council Senior School Certificate Examination, It specifically determined how the index of DIF varies with different cultures in the different areas of mathematics. The following conclusions are drawn on the basis of data analyzed in the study. First, test items in NECO SSCE mathematics multiple choice questions contained items that function differentially for testees from the different cultural groups such as Yoruba and Ibo respectively. This means that such items measured different things for testees of the same latent ability across the different cultural groups. Second, the analysis of data showed that NECO SSCE General Mathematics Paper I multiple choice test in June/July, 2012 had items that function differently for male and female testees with the same latent ability across the different cultural groups. This is because culture influences the ways in which people construct knowledge and create meaning from experience and in mathematics, the individual construction of meanings takes place in interaction with the culture of environment. Based on the findings of the study, the following recommendations were made:

- Culture has been demonstrated to influence DIF in NECO mathematics test items based on gender in different cultural environments. As such mathematics teachers need to adopt the use of culturally responsive instructional technique to reduce students' differential performance in mathematics irrespective of gender. This is because the students culture can be used to teach mathematics and this may make mathematics teaching and learning meaningful. This will afford the teachers and students to gain insight into the various components of mathematics.
- Mathematics teachers should motivate their students to engage in mathematics thinking in their local parlance, using local examples of things that they can see, recognize in their quantities, shapes and cultural application.

REFERENCES

- Abedalaziz, N. (2010a). A Gender-Related Differential Item Functioning of Mathematics Test Items. *The International Journal of Educational and Psychological Assessments*, 5, 10-19.
- Abedalaziz, N. (2010b). Detecting Gender Related DIF Using Logistic Regression and Mantel-Haenszel Approaches. *Proedia-Social and Behavioural Science*, 7(c), 406-413.
- Adamu, H. A. (2007). State of Learning Science and Mathematics in Katsina State Secondary Schools. A Report Submitted to the Department of Research statistics, Katsina State Ministry of Education, 11th July (8-10).
- Beller, M., & Gafni, N. (1996). The 1991 International Assessment of Educational Progress in Mathematics and Sciences. The Gender Differences Perspective. *Journal of Educational Psychology*, 88(2), 365-377.

- Doolittle, A. E., & Cleary, T. A. (1987). Gender Based Differential Item Performance in Mathematics Achievement Test Items. *Journal of Educational Measurement* 24(2), 157-166.
- Ebisine, S. S. (2014). Influence of Culture on Differential Item Functioning in National Examinations Council Senior School Certificate Mathematics Multiple Choice Test in Nigeria. Unpublished Ph.D. Thesis, Faculty of Education, Delta State University, Abraka.
- Erchiken, K., & Lyons-Thomas, J. (2013). Adopting Tests for Use in other Languages and Cultures. In K. Geisinger (Ed) *APA Handbooks of Testing and Assessment in Psychology*, Vol. 3. Washington DC: American Psychological Association.
- Ercikan, K. (1998). Translation Effects in International Assessment. *International Journal of Testing*, 2(384), 199-215.
- Gardner, H. (1985). *Frames of Mind; Theory of Multiple Intelligence*. London: Paladin
- Glichman, M. E., Seal, P., & Eisen, S. B. (2009). A Non-Parametre Bayesian-Diagnostic for Detecting Differential Item Functioning in IRT Models. *Health Service Outcome Research Methods*.
- Greenfield, P. M. (1997). Culture as Process: Empirical Method of Cultural Psychology. In J. W. Berry, Y. N. Poortinga and J. Pandey (Eds). *Handbook of cross-cultural Psychology* (2nd Ed). Needham Height, M.A: Allyn and Bacon.
- Guillerman, S., & Sharon, H. (2001). On the Cultural Validity of Science Assessment. *Journal of Research in Science Teaching*, 6(1), 37-49
- Hambleton, R. K., Swaminathan, H., & Roger, J. J. (1991). *Fundamentals of Item Response Theory* Clifornia: Sage Publications.
- Lane, S., Wang, N. and Magone, M. (1996). Gender-Related Differential Item Functioning on a Middle-school Mathematics Performance Assessment. *Educational Measurement: Issues and Practice*, 15(4), 21-27.
- Laue, J. (1988). *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
- Lilepa, J. (1991). Experimental, Cognitive and Anthropological Perspectives in Mathematics. Retrieved August 27, 2010 from <http://www.cvic.edu/ylikpa/dl.html>.
- Odili, G. A. & Okpobiri, N. R. (2011). Elechi Amadi “The Concubine”. *Ethnomathematics Resources for Teaching Mathematics in Ikwerre Primary Schools*. *African Journal of Education and Technology*, 2, 84-92.
- Ojo, M. O. (1998). Maximizing Students’ Learning of Mathematics Through Students’ Interaction Patterns. A Paper Presented at the 28th Annual Conference of MAN held at Abuboka Tafawa Belewa University, Bauchi, September.
- Perrone, M. (2006). Differential Item Functioning and Item Bias; Critical Consideration in Test Fairness. *Applied Linguistics*, 6(2), 1-3.
- Piaget, J. and Inhaber, B. (1976). *The Growth of Logical Thinking from Childhood to Adolescent*. New York: Basic Book Publishers.
- Pido, S. (2012). Comparism of Item Analysis Results Obtained Using Item Response Theory and Classical Test Theory Approaches. *Journal of Educational Assessment in Africa*, 7, 192-207.
- Roeber, C. (2005). *“That’s Not Fair” Fairness, Bias and Differential Item Functioning in Language Testing*. Melbourne: SLS Brownbeg.

- Schleiman, A. D., Carraher, D. W., & Ceci, S. J. (1997). Everyday Cognition. In J.W. Berry, P.R. Dasen and T.S. Saeswathi (Eds) *Handbook of Cross-Cultural Psychological*, 2, 177-216.
- Scribner, L. (1986). Beyond the Methods Fetish. Towards a Humanising Pedagogy. *Harvard Educational Review*, 65, 173-194.
- Umoinyang, I. E. (2011). The Challenging of Removing Consistent Errors in Achievement Test Using Differential Item Functioning (DIF) Detection Methods. Paper Presented at the 30th Conference of the Association of Educational Assessment in Africa, Nairobi, Kenya.
- Zumbo, B. D. (2009). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF) Scores* Ottawa: ON: Directorate of Human Resources Research and Education, Department of National Defense.

 ©IJSRE

¹ Dr. Sele Sylvester Ebesine is a lecturer in the Department of mathematics, College of Education, Warri, Delta State, Nigeria. He can be reached via email at ebesineselesylvester@yahoo.com.