

MEASUREMENT OF ACADEMIC ACHIEVEMENT BY CBT

BY DR MAMMAN MUSA

DEPARTMENT OF SCIENCE EDUCATION

AHMADU BELLO UNIVERSITY, ZARIA

A DISCUSSION PAPER PRESENTED AT THE TRAINING OF ACADEMIC STAFF ON SETTING OF MCTI FOR CBT PLATFORM HELD AT NEW CBT CENTRE ABU ZARIA

23TH JUNE 2016

INTRODUCTION

Admissions into Nigerian universities have dramatically increased in recent times from few hundreds to thousands. In this university (Ahmadu Bello University, Zaria) for example, over 7,000 one-hundred level students alone are admitted across the faculties every year. In Faculty of Science alone, there are courses with over four thousand registered students.

But there is no proportional increase in staffing hence lecture halls are overcrowded. Lecturing students under this type of circumstances is very stressful. But most stressful is the setting and marking of examination papers.

Undoubtedly, it is time consuming exercise to read and grade such large number scripts especially with written essay items. Overwhelming evidence indicated that the grades assigned for essay items are of very low reliability. These grades were found to vary from lecturer to lecturer, from time to time and from question to question (Brown, 1970).

It is in the light of these problems that the University Management set up a committee on the Review of the Teaching of 100 level courses in the Faculty of Science. Having successfully started implementation of the recommendations of this Report, one important recommendation was the adoption of CBT as a platform for assessing 100 level students.

Therefore the goal of this paper is shed some lights on the major issues involved in CBT and provides some guidelines that might lead to best practices in the use of CBT as medium of assessment.

Summary of Some Findings

Earlier studies indicated that paper-based test materials were superior to computer based. For example:

1. Gould (1981) found that even experts spent about 50% more time to compose task with computer than on paper.
2. Dillon (1994) also discovered that reading textual materials on computer was about 20%-30% slower than on paper based.
3. Ziefle (1998) found out that these differences were due to the type of screen display because eyes get tired more quickly on computer screen than on paper.
4. Mayes et al (2001) found computer based reading to be significantly slower than on paper-based.
5. In terms of learning, Mayes et al (2004) observed that no significant differences were found between the performances of learners whose medium of instruction was computer-based and those based on paper.
6. Westland et al (2005) further found out that comprehension of materials from paper was superior to those through computer screen.
7. Bodmann and Robinson (2004) presented 30 multiple test choice items to two experiment groups : computer-based and paper-based. They found no significant difference between the performances of the two groups. However, completion time was longer with paper-based group than with computer - based.
8. Vansickle and Kepler (1993) and Carlbring et al (2007) provided evidence that supported the use of CBT.
9. Other studies such as (1) George et al (1992) (2) Lakford et al (1994), (3) van de Vijver and Harsveld (1994),(4) Russel 1999(5), and Mc Coy et al, (2004) supported the use of PBT rather than CBT.
10. However many studies found no significant differences between the scores of tests through CBT and those through PBT. These include:
 - a. Rosenfeld et al (1992)
 - b. Kobak et al (1993)
 - c. Steer et al (1994)

- d. King and Miles (1995)
- e. Di Lalla (1994)
- f. Ford et al (1996)
- g. Mertin and Ruch (1996)
- h. Pinsoneault (1996)
- i. Vispoel (2000)
- j. Cronk and West (2002)
- k. Fauladi et al (2002)
- l. Fox and Schwartz (2002)
- m. Puhan and Boughton (2004)
- n. Willams and Mc Cord (2006)

Prometric (2006) compared CBT and PBT and came up with the following conclusions:

1. Effect sizes of examination performances of both CBT and PBT were found to be statistically insignificant ($P > 0.05$).
2. The Differential Item Functionality (DIF) on individual test items in both modes of delivery (i.e. CBT and PBT) were fairly equal.
3. Good preparation, practice and drills help to remarkably reduce test anxiety for both CBT and PBT groups.

MERITS AND DEMERITS OF CBT

Noyes and Garland (2007) that advanced the merits and the demerits of CBT. These include:

- **MERITS OF CBT**

1. **Richness of interface:** With CBT, presentations can be enriched by graphics, animation and others.
2. **Large number of candidates :** With CBT, large number of candidates can be admitted at examination venues and sometimes candidates can sit for their tests at home.

3. **Standardization of test:** With CBT, tests are presented in the same way and in the same format and for same time limit. Subjective elements such as handwriting and spelling attached to PBT are removed with CBT.
4. **Scoring.** With CBT scoring is fast and accurate. This reduces human errors.
5. **Increase in Quality and Quality Tip.** With essay-based compositions CBT are more in length and quality because computers can correct wrongly spelt words or poorly constructed sentences.
6. **Low Cost:** With CBT, the cost of running tests is generally lower than PBT in the long run. This cost is in terms of human and materials.
7. **Exam Malpractice:** With CBT examination malpractice is reduced to the minimum especially in randomizing the test items.

DEMERITS OF CBT

1. **Computer Hard and Software:** With CBT computers could crash or programme could free cleaning the examination.

PLANNING THE CBT: some considerations

THE BLOOM'S TAXONOMY

1. KNOWLEDGE
2. COMPREHENSION
3. APPLICATION
4. ANALYSIS
5. SYNTHESIS
6. EVALUATION

- **PLANNING THE CBT:** some considerations

2. TEST BLUE-PRINT

- **MULTIPLE CHOICE TEST ITEM AS PLATFORM FOR CBT**
- ❖ MCTI Consists Of A Stem And A Number Of Possible Responses.

- ❖ Most Experts Agree That MCTI Is The Most Versatile, Valid, Reliable And Widely Of All Types Of Tests.
- ❖ It Was Argued That With Well Designed MCTI, Most Educational Objectives Can Be Assessed

(STOREY, 1970)

- **RULES FOR SETTING MCTI**

- KEEP MCTI VOCABULARY LEVEL SO THAT THE ITEM IS FULLY UNDERSTOOD BY ALL
- EACH ITEM SHOULD BE INDEPENDENT OF ALL OTHER ITEMS.
- AVOID TRICK QUESTIONS
- RANDOMIZE THE POSITIONS OF THE CORRECT RESPONSES.
- THERE SHOULD BE ONE AND ONLY ONE CORRECT RESPONSE.

- **REQUIREMENTS FOR A GOOD TEST**

- A Good MCTI Must Be Valid: Content.
- A Good MCTI Must Be Reliable.
- A Good MCTI Must Have Uniform And Conducive Conditions During Administration.
- A Good MCTI Must Be Of Average Difficulty Level.
- A Good MCTI Must Be Of Average Discrimination Level

- **CONCLUSION/RECOMMENDATIONS**

- From the findings of the literature cited above, it could be deduced that CBT should be adopted by any institution confronted with similar challenges.
- Orientation and demonstration for students writing CBT.
- Extra time between 5 to 10 minutes should added to all CBT candidates.
- Stable power supply should be provided.
- Technical staff should be at the CBT venues.

- **REFERENCES**

- Bodmann, S.M. and Robinson, D.H., 2004. Speed and performance differences among computer- based and paper-pencil tests. *Journal of Educational Computing Research*, 31, 51-60.
- Brown, F. G. (1970): *Principles of Educational and Psychological Testing*: Illinois The Dryden Press Inc.
- Carlbring, P., et al., 2007. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior*, 23, 1421-1434.
- Cronk, B.C. and West, J.L., 2002. Personality research on the Internet: A comparison of Web-based and traditional instruments in take-home and in-class settings. *Behavior Research Methods, Instruments and Computers*, 34 (2), 177-180.
- DiLalla, D.L., 1996. Computerized administration of the Multidimensional Personality Questionnaire. *Assessment*, 3, 365-374.
- Dillon, A., 1994. *Designing usable electronic text: Ergonomic aspects of human information usage*. London: Taylor & Francis.
- Ford, B.D., Vitelli, R., and Stuckless, N., 1996. The effects of computer versus paper-and-pencil administration on measures of anger and revenge with an inmate population. *Computers in Human Behavior*, 12, 159-166.
- Fouladi, R.T., McCarthy, C.J., and Moller, N.P., 2002. Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment*, 9, 204-215
- Fox, S. and Schwartz, D., 2002. Social desirability and controllability in computerized and paper- and-pencil personality questionnaires. *Computers in Human Behavior*, 18, 389-410.
- George, C.E., Lankford, J.S., and Wilson, S.E., 1992. The effects of computerized versus paper-and- pencil administration on negative affect. *Computers in Human Behavior*, 8, 203-209.
- Goldberg, A., Russell, M., and Cook, A., 2003. The effect of computers on student writing: A meta- analysis of studies from 1992 to 2002. *Journal of Technology, Learning, and Assessment*, 2, 1-51. [online]. Available from: <http://www.jtla.org> [Accessed 15 June 2006].

- Gould, J.D., 1981. Composing letters with computer-based text editors. *Human Factors*, 23, 593- 606.
- King, W.C. and Miles, E.W., 1995. A quasi-experimental assessment of the effect of computerizing noncognitive paper-and-pencil measurements: A test of measurement equivalence. *Journal of Applied Psychology*, 80, 643-651.
- Kobak, K.A., Reynolds, W.M., and Greist, J.H., 1993. Development and validation of a computer- administered version of the
Hamilton Anxiety Scale. *Psychological Assessment*, 5, 487-492.
- Lankford, J.S., Bell, R.W., and Elias, J.W., 1994. Computerized versus standard personality measures: Equivalency, computer anxiety, and gender differences. *Computers in Human Behavior*, 10, 497-510.
- McCoy, S., et al., 2004. . Electronic versus paper surveys: Analysis of potential psychometric biases. In: *Proceedings of the IEEE 37th Hawaii International Conference on System Science*. Washington, DC: IEEE Computer Society, 80265C.
- Mayes, D.K., Sims, V.K., and Koonce, J.M., 2001. Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics*, 28, 367-378.
- Merten, T. and Ruch, W., 1996. A comparison of computerized and conventional administration of the German versions of the Eysenck Personality Questionnaire and the Carroll Rating Scale for Depression. *Personality and Individual Differences*, 20, 281-291.
- Noyes, J.M., Garland, K.J., and Robbins, E.L., 2004. Paper-based versus computer-based assessment: Is workload another test mode effect? *British Journal of Educational Technology*, 35, 111-113.
- Pineseault, T.B., 1996. Equivalency of computer-assisted and paper-and-pencil administered versions of the Minnesota Multiphasic Personality Inventory-2. *Computers in Human Behavior*, 12, 291-300.
- Puhan, G. and Boughton, K.A., 2004. Evaluating the comparability of paper and pencil versus computerized version sofa large-scale certification test. In: *Proceedings of the annual meeting of the American Educational Research Association(AERA).SanDiego,CA: Educational Testing Service.*

- Rosenfeld, R., et al., 1992. A computer-administered version of the Yale-Brown Obsessive- Compulsive Scale. *Psychological Assessment*, 4, 329-332.
- Russell, M., 1999. Testing on computers: A follow-up study comparing performance on computer and paper. *Education Policy Analysis Archives*, 7, 20.
- Steer, R.A., et al., 1994. Use of the computer-administered Beck Depression Inventory and Hopelessness Scale with psychiatric patients. *Computers in Human Behavior*, 10, 223-229.
- van De Vijver, F.J.R. and Harsveld, M., 1994. The incomplete equivalence of the paper-and-pencil and computerized versions of the General Aptitude Test Battery. *Journal of Applied Psychology*, 79 852-859
- Vansickle, T.R. and Kapes, J.T., 1993. Comparing paper-pencil and computer-based versions of the Strong-Campbell Interest Inventory. *Computers in Human Behavior*, 9, 441-449.
- Vispoel, W.P., 2000. Computerized versus paper-and-pencil assessment of self-concept: Score comparability and respondent preferences. *Measurement and Evaluation in Counseling and Development*, 33, 130-143.
- Wastlund, E., et al., 2005. Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior*, 21, 377-394.
- Williams, J.E. and McCord, D.M., 2006. Equivalence of standard and computerized versions of the Raven Progressive Matrices test. *Computers in Human Behavior*, 22, 791-800.
- Ziefle, M., 1998. Effects of display resolution on visual performance. *Human Factors*, 40, 554-568.