

**COMPARING THE RELIABILITY OF THE CONVEYOR BELT MARKING
SYSTEM WITH THE TRADITIONAL MARKING SYSTEM**

**MATTHEW BUKENYA
EXECUTIVE SECRETARY
UGANDA NATIONAL EXAMINATIONS BOARD**

A PAPER PRESENTED AT THE 32ND IAEA ANNUAL CONFERENCE

SINGAPORE

22nd – 26TH MAY 2006

Introduction

In Uganda, the Uganda National Examinations Board (UNEB) conducts Primary Leaving Examinations (PLE), which consists of four papers: English, Social studies (SST) Science and Mathematics. The Ministry of Education and Sports uses the results of (PLE) to select pupils for post primary education. As such, it is imperative that UNEB maintains high standards of reliability and validity of PLE. Since its inception in 1980, marking of the UNEB examination papers have been done using the traditional marking system (TMS). A description of TMS is given below.

Traditional Marking System

In this marking system, an envelope containing answer scripts of candidates in a paper from a particular school is given to one examiner to mark. The examiner marks the answers to all the questions attempted by the candidates. Once all the scripts in that envelope are marked another envelope of scripts is given to the examiner to mark. The process is continued until all the scripts are marked. After this checkers who are not examiners are employed to check through the marked answer scripts to detect any errors in marking. Such errors are corrected by a UNEB officer in charge of checking. In case a portion of a script is not marked, the examiner or team leader is requested to mark the unmarked portion. A team leader normally coordinates ten percent of the scripts marked in each envelope to ascertain the consistency in marking. The deviations in marks between the examiner's mark and that of the team leader also act as a check on the consistency in the marking of a particular examiner. Sometimes these deviations are more than the allowed level of ± 2 marks.

There were concerns about the reliability of marks obtained by TMS method as inconsistencies in marking were sometimes discovered in the scripts. Okelowange (2004a) investigated mark-remark reliability of History paper 3 at 'A' level. He experimented marking the History dummy scripts using another system of marking he called the conveyor belt marking system (CBS) and TMS as the control.

Conveyor Belt System of Marking (CBS)

In the CBS, examiners are organized in groups. Each group is composed of a team leader, a starter, markers, and checkers. Each marker marks only a set of questions and passes the candidate's answer script to the next marker who will also just mark the set of questions allocated to him/her. The marked scripts are passed over to the checkers, who are also examiners, to check through the script for any errors. Any errors detected are referred immediately to the markers to correct. Finally the team leader samples ten percent of the scripts in an envelope and remarks to assess the consistency in marking and interpretation of the marking scheme. Based on Okelowange's (2004a) experimental results, CBS was piloted in PLE marking of the year 2003. The examiners were happy with CBS and recommended that it should be adopted in marking all PLE scripts with

effect from year 2004 (Okelowange, 2004b). Following this, a study was conducted to estimate the reliability of CBS and TMS with respect to PLE.

The procedures and results of the study, and the advantages, challenges as well as recommendations made by the examiners are presented in this paper

Sampling Procedures

For each of the four examination papers, ten districts whose scripts were marked using TMS were selected using simple random sampling procedure. From each district five schools were selected using stratified systematic sampling procedure. From each school four scripts were selected for each of the nine mark range categories to cover grades 1 - 9. A total of 1800 scripts were selected from the ten districts. A similar procedure was used to select the scripts marked using CBS in the year same year. In addition, examiner-team leader deviations were computed for 3000 scripts marked using TMS and another 3000 marked using CBS.

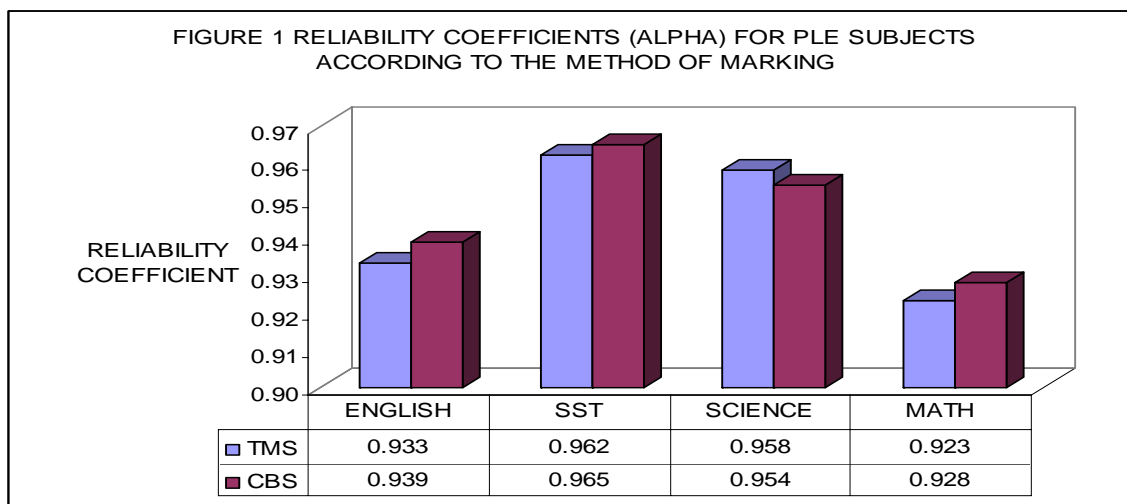
Data Analysis

SPSS program was used to analyze the data. This means the internal consistency reliability was computed using Cronbach coefficient (alpha). The significance of the differences in the reliability coefficients were tested using z-test. The mean, SD, and variance of the deviations were computed for each set of scores. The significance of the difference in the mean deviations was tested using one way analysis of variance at level of significance of 0.05, two tailed.

Results and Interpretations

Reliability Coefficients

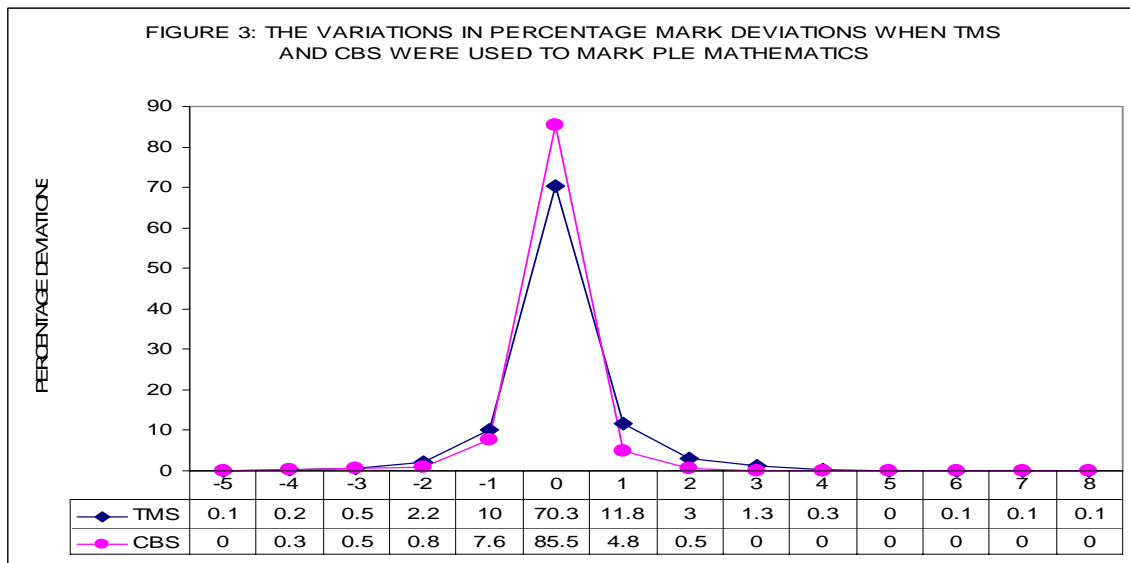
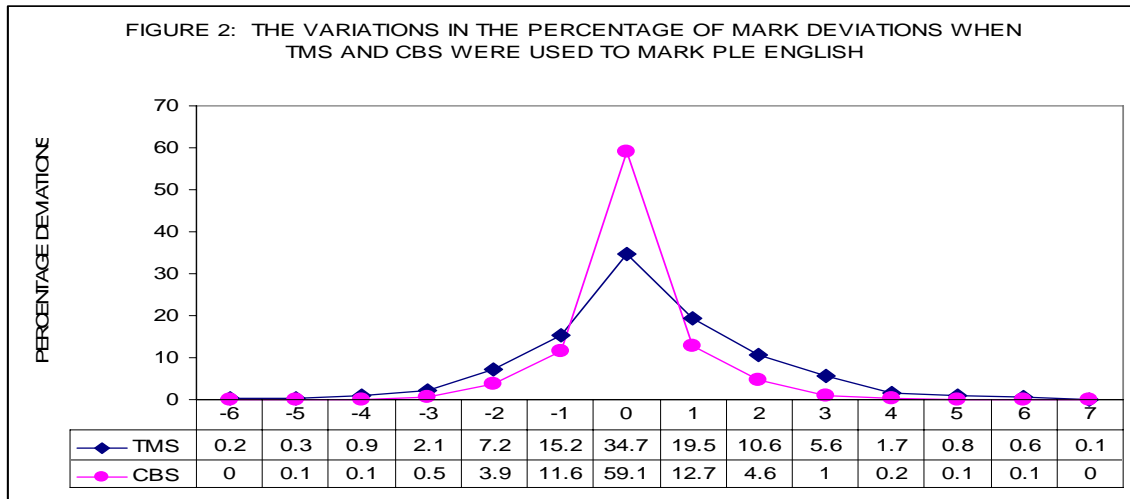
Figure 1 presents the internal consistency reliability coefficients (alpha) for PLE subjects marked using TMS and CBS.



The reliability coefficients were high, above 0.92, for both methods of marking. This indicates that the marking in both systems in the year 2003 was reliable. For both methods SST and Science produced higher consistency reliability coefficients than English and Mathematics. Apart from Science, the reliability coefficients were higher in CBS than in TMS, and these differences were statistically implying that CBS is a more reliable method of marking.

Analysis of Mark Deviations

Figure 2 shows the percentage distribution of deviations of marks between examiners and the team leader in English.



The percentage of scripts with zero deviation increased when CBS was used in marking English, from 34.7% for TMS to 59.1% for CBS. If more scripts have zero deviations it means that examiners are marking more consistently. At a deviation range of -1 to +1,

there were 69.7% and 83.4% of the English scripts marked using TMS and CBS respectively. This implies that using CBS reduced large deviations in the marking of English scripts. In Mathematics, while 70.3% of the scripts marked using TMS had zero deviation, 85.5% of those from CBS had similar deviation. The deviation range of -1 to +1, had 93.1% of the Mathematics scripts marked using TMS and 97.1% of the scripts marked using CBS.

Advantages of CBS

When examiners were asked about the advantages of CBS, they gave the following;

a) Team Spirit

CBS encourages examiners to work as a team
CBS builds more cooperation among examiners
An examiner's commitment to marking is higher in CBS

b) Security of scripts

There is less risk of scripts getting lost in CBS than in the TMS
In CBS individual examiners have less control over a script than in TMS

c) Marking speed and transparency

CBS controls the very fast examiners more than TMS
There is less need for a rush to mark more scripts in CBS than in TMS
CBS is a more transparent system of marking than the TMS

d) Fairness to candidates

A script of a candidate is marked by more than one examiner. This reduces the chances of bias and over-marking or under-marking.

e) Maximization of examiners concentration

Examiners are kept busy. This reduces redundancy and laziness on the part of the examiners.

f) Minimising Malpractice

Cases of malpractice during marking are minimized. Since an examiner marks only a few questions. Moreover movement of examiners from one room to the other is reduced.

Challenges of CBS

a) Demands on examiners: CBS;

- Tedious, examiners tend to breakdown more frequently than in the traditional marking system
- is time consuming
- leaves no room for relaxation for examiners
- confines examiners unnecessarily

b) Reducing marking speed: CBS

- cheats the fast examiners financially
- does not cater for individual differences among the examiners
- leads to clogging of work when a member of the group is slow, lazy, absent, or sick and hence delays the speed of marking of the group.
- assumes that no participant falls sick or comes late or be absent
- it is difficult to balance the good examiners evenly within the teams
- weakens the sense of responsibility among examiners
- encourages dodging of work by examiners
- reduces (positive) competition among the examiners

Discussions

The reliability coefficients were found to be high (above 0.92) for both CBS and TMS. However, in CBS the reliability coefficients were generally higher than those of TMS and the differences were statistically significant. The examiners also expressed the views that CBS was more reliable than TMS in marking PLE. For high stake examinations, Ebel (1972) stated that a well constructed test should yield a reliability coefficient of 0.90 or higher. PLE of 2003 satisfied this condition and the use of CBS improved the reliability coefficient of the PLE.

The finding that CBS enhances team spirit, offers greater security of script, reduces chances of malpractice, and is fair and transparent agrees with the earlier findings by Okelowange (2004a).

The findings that CBS is tedious, exhaustive and slows down speed of marking disagree with the findings reported by Okelowange (2004a, 2004b) that CBS speeds up the marking and is less tiring to the examiners than TMS. These challenges hinge on the organization of marking which can be reviewed and improved upon. Another worrying finding is that CBS reduces the sense of responsibility of the examiners. The examiners dodge work and have the feeling that after all they will all be paid the same amount of money. Attention should be paid to such examiners and weed them out from marking. This measure should also apply to lazy examiners and those who frequently absent themselves from marking room.

Recommendations

The majority (96.4%) of the examiners recommended that marking should be fully residential. To reduce the number of days of marking and also reduce chances of malpractice in the examinations at the marking stage UNEB should consider this recommendation and implement it for all examinations.

On payment for the scripts marked, the recommendation that examiners should be paid according to the scripts marked by the team is an appropriate one. The members of the team should include the team leader, checkers, starter and markers. The team leader could be given a top up of 20% of the group marking fee. 20% may be changed to fit with the available funds. This will ensure that the team leader gets more than the assistant examiners. The value may not be the same for all teams and this will mean that even the TLs may not get uniform rate. The pay should reflect the amount of work the TL has put in. For CE and ACE flat fixed rate should be maintained.

Examiners also made varied recommendations on how to pay an examiner who drops out of making in CBS. The recommendation that he/she should be paid only for the scripts marked is logical. However such examiners should be followed and considered for retirement from marking depending on the reason for dropping out.

Conclusions

Internal consistency reliability coefficients were found to be high for all subjects when TMS and CBS were used. However CBS produced higher internal consistency coefficients than TMS. The percentages of scripts with zero deviations were greater when CBS was used than in TMS. All findings show that marking PLE using CBS produced more reliable results than in TMS. CBS was found to have more advantages than TMS. The advantages include building team spirit, offering more security for scripts and control of scripts and marking, increasing examiners concentration and reducing chances of malpractice during marking. Nevertheless it was also found that CBS is tedious and straining on the examiners, slows down marking and reduces examiners' sense of responsibility.

There were strong recommendations that CBS be adopted by UNEB to mark all examinations, and that the management be reviewed with a view to enhancing efficiency.

REFERENCES

Okelowange, P. J. (2004a). Report of the pilot study on the implementation of conveyor belt system of marking in UNEB (2003 PLE & UACE)

Okelowange, P. J. (2004b). Marking reliability in Uganda Advanced Certificate of Education (UACE) examination with reference to history paper 3. A research report published by Uganda National Examinations Board, Kampala

Pido, S. (1998). Reliability and validity of the 1997 Uganda Advanced Certificate of Education examinations. A research report published by Uganda National Examination Board, Kampala