



COMPARISON OF THE RELATIONSHIP BETWEEN EXAMINEES' HANDWRITING RATING AND THE AVERAGE MARK CHANGES IN ESSAY TEST SCORES IN WHOLE SCRIPT AND SEGMENTED MARKING METHODS.

AGBONKPOLO, MONDAY UGIAGBE

DEPARTMENT OF EDUCATION,

BENSON IDAHOSA UNIVERSITY, BENIN CITY. EDO STATE. NIGERIA

Abstract: The main purpose of the study was to find out experimentally whether there is difference between the influence of examinees' handwriting clarity on marking error in whole script marking method (WSMM) and segmented marking method (SMM) in essay tests with regards to the systematic and random errors. Thus the average mark changes (AMC) was used to measure the effectiveness of the two methods with regards to the minimization of marking errors arising from examinees' hand writing clarity. Two research questions and one hypothesis were raised for the study. The post test – only control group design, in which two equivalent groups are subjected to different treatments, was used in the study. The two methods (treatments) were explained and randomly given to two equivalent groups, each consisting of 24 NECO markers (one group to one method), to mark thirty two examination scripts in a conference marking. Clean photocopies of the thirty two scripts were also assessed for examinees' hand writing clarity by twenty independent research assistants. Thereafter the AMC in the scores awarded by the twenty four (24) markers for each of the thirty two (32) examination scripts and the mean hand writing clarity rating for the same candidates were used for correlation analyses and testing the hypothesis. The statistics used were the Pearson correlation, Fishers' Z transformation test and regression statistics. Findings showed that poor hand writing clarity of some of the examinees tends to increase AMC in SMM more than in WSMM, which implies that poor handwriting clarity was more influential in the SMM which employs direct comparative scoring technique, than in whole script scoring and as examinees' handwriting clarity increases, AMC becomes smaller in SMM than in WSMM although the difference in the Pearson r was not significant. These findings showed that WSMM is more effective than SMM in minimizing marking errors arising from poor examinees' handwriting clarity. The researchers thus concluded that the WSMM may be better than the SMM for marking examinees' scripts with poor hand writing clarity.

Keywords: Examinees, Handwriting, Segment, Marking Methods.

Introduction

In the light of the crucial role of accurate evaluation, a considerable effort is being made to ensure objectivity in scoring test. This is most pertinent in the scoring of essay tests. In external examinations, for instance, several steps are taken to create the enabling conditions for objectivity in scoring essay test. Despite the painstaking steps, it is a common fact that unacceptable variation between the original scores and vetted scores still exist.

Many measures to improve the marking reliability of essay tests have been suggested. In line with this desire many scoring methods are currently being used by markers. The two most popular ones are the whole script marking and the segmented marking methods. Whole script marking is the marking of all the responses in one script before going to another script (Bakker & Van Lent, 2003; Meadows & Billington, 2005). The reliability of this method of marking is constrained by the problem of halo effect. Halo effect is the bias, in the marking of a particular item due to an earlier impression from the marking of previous items, the marker had of the script owner. Hence measurement error tends to be higher among the scores of upper achievers and that of the lower achievers. The marking of all the items before going to another script increases marker's familiarity with examinee's handwriting and writing style. This means that the reliability of the marks awarded to items towards the end of the script would be higher than the reliability of marks awarded to items preceding the former. This is most likely to be true with scripts with low handwriting quality.

Whole script marking allows for the general impression marking which tends to increase measurement errors. Segmented Marking Method refers to the marking of one item across the scripts before engaging in the marking of another item. (Bakker & Van Lent, 2003; Meadows & Billington, 2005). Segmentation or part making allows the marker the opportunity to compare item response of each candidate to other candidates' responses to the same item with the aim of minimising the level of inconsistency arising from non-detailed marking scheme, halo effect and the effect of poor concentration which inexperienced markers are likely to face when marking essay scripts.

There is greater reliance on comparative scoring compared to whole script marking hence non-cognitive elements such as handwriting, may have greater influence on students' scores. Briggs, (1970, 1980); Bull and Stevens, (1979) provided evidence that good handwriting enables the teacher to make out easily the intent of pupil's responses, poor handwriting on the other hand makes the reading of their response difficult (Bull and Stevens, 1979).

Briggs (1970) observed the same phenomenon. Briggs (1980), conducted an experiment using practicing teachers of English to assess copies of 16+ external examination scripts rewritten in five different handwriting styles. On the strength of this experiment he expressed "that handwriting may significantly affect the chances of some 16-year-olds passing or failing the 16+. Markham (1976) carried out an experiment to examine the influence of handwriting quality on evaluation of written work by markers. The study used 45 teachers and 36 student teachers to rate descriptive paragraphs with varying quality in contests and handwriting style. Multiple Classification Analysis revealed that the teacher characteristics of experience, level taught, degrees held, age, and the student teacher characteristic of level taught do not have a significant influence on the score given to a paper. In any case the analysis of variance showed that the variation in scores due to handwriting was significant. Papers with better handwriting consistently received higher scores than did those with poor handwriting regardless of content.

It seems that the handwriting factor is not straightforward, as other variables, such as gender and attractive photo of the student tend to interact with it. Bull and Stevens (1979), conducted a study in which an identical essay in content was marked by 72 raters who were mostly school teachers, with few students. Some of these raters have their essay scripts in typed form, some in good handwriting and some have scripts with poor handwriting. "A photograph of the supposed author of the essay was attached to the essay. This photograph was of a male or a female who was either highly physically attractive or rather unattractive" (Meadows & Billington, 2005).

The result shows that when the authors were female the awards given to the essays were influenced by handwriting and attractiveness. But where the authors were male the influence of handwriting was absent. On this phenomenon Bull & Stevens say that it is possible that society expects females to have better handwriting than males and so when a female has poor handwriting the resulting variable impression created is poor. Similarly perhaps women are judged more on attractiveness than are men. (p. 58 cited by Meadows & Billington, 2005). Truly, good handwriting seems to favor pupils in the marks they get for their written work. Handwriting bias is magnified by gender and attractiveness.

Purpose of the Study

The main purpose of the study therefore was to experimentally determine the effectiveness of the segmented marking method relative to the whole script method of scoring essay test with regards to the minimization of marking error arising from poor examinees' handwriting.

Research Questions

The following research questions guided the study

1. What is the relationship between examinees' handwriting and the average mark changes in the marks awarded by the segmented marking group.
2. What is the relationship between examinees' handwriting and the average mark changes in the marks awarded by whole script marking group.
3. How does the relationship between examinees' handwriting clarity and the average mark changes in the marks awarded through the SMM compare with that of the WSMM?

Hypothesis

The following null hypothesis was tested by the researcher at 0.05 alpha level.

1. The segmented making and whole script making methods are not significantly different in respect of the relationship between examinees' handwriting clarity and the average mark changes in the marks awarded through the two methods

Method

The data for the study were obtained through conference marking in which two groups of twenty four markers, marked the same 32 scripts. The 32 marks awarded by the 30 markers constituted the data for the study. The marks were used to compare the two groups.

30 NECO markers were invited for one day conference marking through correspondence by mail and telephone. 24 of them were selected through stratified random sampling technique to ensure that the two groups were equivalent in respect of qualification and experience. Two centers namely Benson Idahosa University and Our Lady's of Fatima at Auchi were used for the exercise. Upon arrival they were given attendance sheet to fill, each of them was given a parcel containing 32 clean photocopies of the students' answer scripts in Economics essay test, mark sheet and the question paper. The two groups were not separated, they all sat together to avoid environmental differences.

Before the commencement of marking, the researcher spelt out the financial terms. The marking guide was then discussed, item by item. Corrections that were accepted by all were effected in the guide and the corrected version was used for the marking. The marking exercise was co – ordinated by the researcher and two assistants. Furthermore all the markers were instructed to write their names on the parcel that was used for packaging of the scripts.

The average mark change (AMC) in the scores awarded by 24 NECO markers for each examinee was computed. with the following formula for AMC:

$$AMC = \sum (X - X_i - j) / N$$

The thirty two scripts clean copies were also given to twenty independent assessors to rate the hand writing clarity of each script and then the mean rating (MREH) for each script was computed.

To solve research questions 1 and 2 the thirty two scores awarded by the two groups NECO examiners were correlated with the MREH of the thirty two scripts using the Pearson Product Moment Correlation Statistics. To answer the research question 3, the relationship between AMC and the MREH of the two methods were compared using the Micro Soft Regression line since both groups have a common independent variable, MREH. To test the hypothesis, the two Pearson r of the two groups were tested for significant difference using the Fisher Z transformation Statistics.

Results and Discussion

Research Question 1:

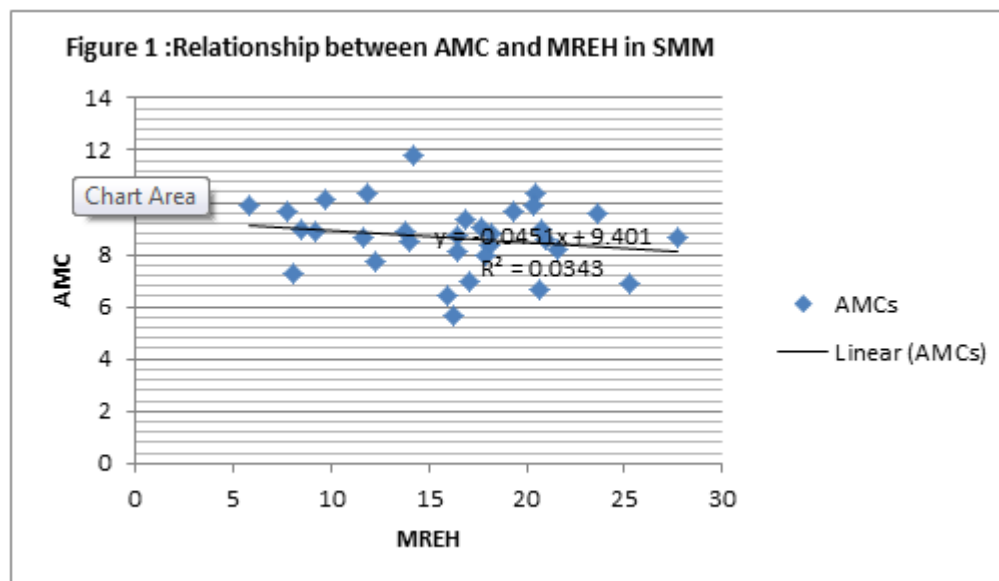
This question was aimed at determining the nature of the relationship existing between students' handwriting and the AMC in the scores awarded by the segmentation marking group. The study reveals that

the relationship between the mean rating of examinees' handwriting and the AMC in the scores awarded by the segmentation marking group has low negative correlation coefficients. The summary of the computation of the Pearson correlation is shown in table 1.

Table 1: Pearson Correlation between the Average Mark Changes and the Mean Rating of Examinees' Handwriting in the Scores Awarded by the Segmentation Marking Group.

No	Sum of Squares and products	Covariance	r
32	-40.063	-1.292	-0.185

The computation of the coefficient of relationship between the average mark changes in 32 scripts and the mean rating of examinees' handwriting as shown in table 4 reveals a low negative correlation coefficient of **-0.185**. A graphic illustration of the linear relationship is shown in figure 1.



The regression equation is: $y = -0.045x + 9.401$ Y and X represent the mean rating of examinees' handwriting and the average mark change.

Research Question 2.

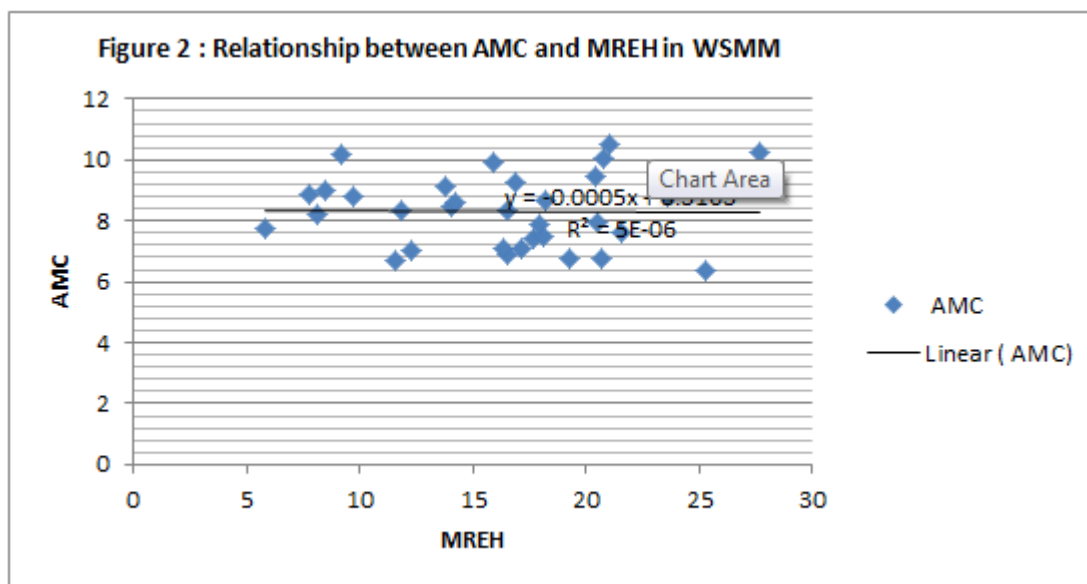
The research question on the relationship between students' handwriting and the average mark changes in the scores awarded by the whole script marking group was aimed to establish the nature of interaction

between the two variables. The study shows that the relationship between the mean rating of examinees' handwriting and the AMC in the scores awarded by the whole script marking group has a low negative correlation as shown in table 2.

Table 2: Pearson Correlation between AMC and MREH in the Scripts Marked by the Whole Script Marking Group.

No	Sum of Squares and products	Cross-Covariance	r
32	-0.432	-0.014	-0.002

The coefficient of relationship between the average mark changes in 32 scripts and the mean rating of examinees' handwriting is **-0.002**, a low negative coefficient of relationship as table 7 shows. A graphic representation of the linear relationship between AMC and MREH is shown in figure 2



The coefficient of the linear relationship is depicted by the slight gradient of the slope of the regression line. The equation of the line is: $y = -0.000x + 8.316$. Y and X represent the average mark changes and the mean rating of examinees' handwriting respectively.

Research Question 3

The quest to find out how the relationship between the average mark changes in 32 scripts and the mean rating of examinees' handwriting in the SMM compare with that of WSMM is the task of this research question. The answer to the question is shown in table 8.

Table 3: Comparison of the Relationship between AMC and MREH in SMM and WSMM

Group	N	Pearson r	Sum of Squares and Cross Product	Covar	r Diff
SMM	32	-0.185	-40.062	-1.292	- 0.05
WSMM	32	-0.002	-0.432	-0.014	

Table 3 shows that the coefficients of relationship between the average mark changes and the mean rating of examinees' handwriting in WSMM and SMM are **-0.002** and **-0.185** respectively with a difference of **-0.05**. The linear relationships in the two methods are shown in figure 3

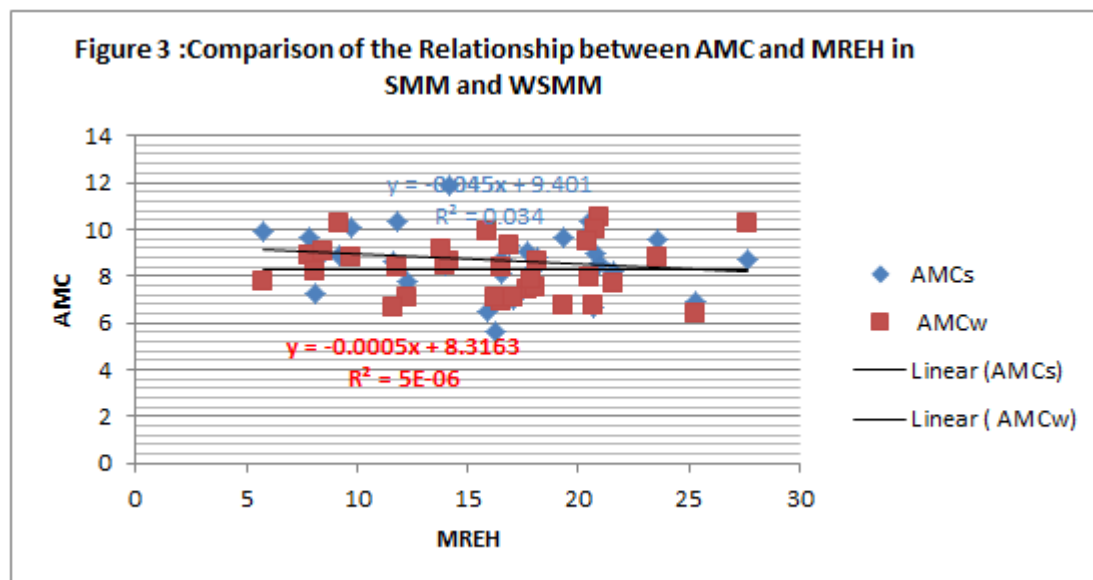


Figure 3 shows that the equations of linear relationship between AMC and

MREH in SMM and WSMM are **Y= -0.045x +9.401** and **Y= -0.000x + 8.316** respectively. The reduction of the gap between the lines towards the right hand side shows that the effect of handwriting is higher in SMM.

Hypothesis

There is no significant difference between the coefficients of correlation of examinees handwriting and the average mark changes in the marks awarded by the segmentation (SMM) and the whole script (WSMM)

marking groups. This hypothesis was tested using the Fisher's Z Transformation test. The result is presented in table 4.

Table 4 –Summary of the Fisher's Z Transformation Test of Significance of the Difference between the Coefficients of Correlation of Examinees' Handwriting and the AMC in the Marks Awarded by the SMM and WSMM groups.

Group	No of Scripts	Pearson r	Sig. Level (2 – tailed)	Fisher's Z Z(r)	Z_{Cal}	Z_{Crit} (two- tailed)
SMM	32	-0.185	0.310	-0.187	1.00	1.96
WSMM	32	-0.002	0.990	-0.002		

Table 4 shows that the SMM group with a coefficient of **-0.185** was more influenced by examinees' handwriting than WSMM group. The Pearson correlation coefficient (r) of the relationship between the clarity of examinees' handwriting and the average mark changes in the marks awarded by the whole script marking group is **-0.002**. r is not significant at **0.05** alpha level since p- level (**0.99**) is greater than the alpha level. The Pearson correlation coefficient (r) of the relationship between the clarity of examinees' handwriting and the average mark changes in the marks awarded by the segmentation marking group to the same scripts in table 11, is **-0.185**. r is not significant at **0.05** alpha level since p- level (**0.310**) is greater than the alpha level. The difference is not significant at alpha-level of **0.05** because the Z calculated (**1.00**) is less than the critical value (**1.96**). Thus the Hypothesis is not rejected.

Discussion of Result

The effect of examinees' handwriting on marking was not significant in the two groups. This finding did not agree with previous findings that handwriting significantly affects the reliability of essay marking (Markham, 1976; Briggs, 1980). This is most probably because the responses on the scripts were very scanty, which, by the estimation of the researchers, was about an average of two and a half pages, the highest being about four pages on A4 paper. The segmentation marking group was more susceptible to handwriting influence with Pearson coefficient of -0.187 as against -0.002 for the whole script marking group. This may be due to the fact that markers while marking whole script become used to the handwriting of the examinee before they finish marking the script. Thus the influence of handwriting was less in the whole script marking method. The comparative approach in the SMM may also have contributed to the higher influence of handwriting in the SMM group.

Recommendation

The researchers recommend that the SMM should be used in marking the scripts of examinees with poor hand writing and where examiners have decided to use the SMM because of the advantages of the SMM as shown in the work of Agbonkpolo 2015, scripts with poor hand writing should first be selected out for WSMM.

REFERENCES

- Agbonkpolo, U, M (2015) Effectiveness of whole script marking and segmented marking methods in the reliability of essay tests.
- Bakker, S. & Van Lent, L. G. (2003). National testing on - line: How far can we go? Paper presented at the IAEA Conference, Manchester. Retrieved 5 January 2005 from <http://www.aqa.org.uk/support/iaea/papers.html>.
- Briggs, D. (1980). A study of the influence of handwriting upon grades using examination scripts. *Educational Review*, 32 (2), 185-193
- Briggs, D. (1970). The influence of handwriting on assessment. *Educational Research*, 13, 50-55.
- Bull, R. & Stevens, J. (1979). The effects of attractiveness of writer and penmanship on essay grades. *Journal of Occupational Psychology*, 52 ,53-59.
- Markham, L. R. (1976). Influence of handwriting quality on teacher evaluation of written work. *American Educational Research Journal*, 13, 277-283.
- Meadows, M & Billington, L (2005). A review of the literature on marking reliability. Retrieved from W:\Michelle\Michelle Meadows\Quality of Marking\Lit Review\Review Sections\A Review of the Literature on Marking Reliability.doc