

Use of Student Achievement Scores as Basis for Assessing Teachers' Instructional Effectiveness: Issues and Research Results

Monday T. Joshua¹, Akon M. Joshua² and Anges Ebi Maliki³

1. Faculty of Education, University of Calabar, P.M.B.1115, Calabar, 540004,
Cross River State, Nigeria

E-mail: mtjoshua2004@yahoo.com

2. Faculty of Education, Cross River University of Technology, Akamkpa Campus,
Cross River State, Nigeria

E-mail: konjoshua@yahoo.com

3. Faculty of Education, Niger Delta University, Wilberforce Island, 560001,
Bayelsa State, Nigeria

E-mail: agnesmaliki@yahoo.com

KEYWORDS Achievement. Assessment. Effectiveness. Performance

ABSTRACT Assessment is very important in the learning and instructional processes. Equally important is the use to which assessment results are put. Use of student achievement scores (SAS) as a basis for assessing the teacher's instructional competence or effectiveness is one of the controversial approaches to teacher evaluation. This paper examined the issues characterizing the use of SAS as a basis of teacher assessment, and presents research results from Nigeria and abroad on the attitude of teachers to this approach of assessing teachers. The research findings from abroad were extracted from the literature. The research in Nigeria sampled 480 secondary school teachers in Akwa Ibom State using stratified random sampling technique. The instrument for data collection was a questionnaire with a reliability estimate of .94. Three hypotheses were tested. The results showed that secondary school teachers displayed a significantly negative attitude to this teacher assessment approach; that this attitude was not significantly influenced by the professional status of the teachers; and that this attitude, when the purpose of teacher assessment is summative, is significantly more negative than the attitude when the purpose of teacher assessment is formative. It was concluded that Nigerian teachers are not different from their counterparts abroad in their showing disdain or condemnation to the use of student achievement scores as indicators of teacher's competence, performance or effectiveness. Therefore, great care is required in the use of assessment results in our schools, especially in attempts to blame teachers for poor academic performance of learners.

INTRODUCTION

Assessment is one of the processes/activities that characterize the school system; and many things/components and persons in the system can be, and have actually been, assessed. Usually, the learners in the school system are the principal focus in assessment; and various tools have been utilized for this purpose. One major tool for assessment has been the test. Testing or measurement has been a process of gathering quantitative estimates of the amount of knowledge, skills, traits or characteristics possessed or acquired by learners in the school system. Tests are administered on learners at various stages of their learning/educational experience, and testing produces scores or data on the basis of which some crucial decisions are taken. So, the use of results of assessment becomes a crucial issue in the school system, and in the entire educational enterprise.

All correspondence to: E-mail: agnesmaliki@yahoo.com

During the process of human development, assessment information/data are generated and used in a variety of ways to improve administration, teaching and learning; and therefore, to enhance the likelihood of success by both the learner and the teacher. Thus, assessment (i) provides useful information, as well as formative or regenerative feedback to the learner to ensure his/her progress towards success; (ii) provides feedback to the teacher with which to effect improvement in instruction, set realistic objectives, evaluate the degree to which course objectives have been met, provide opportunity to improve his/her professional skills, and provide counseling and guidance to the learner; and (iii) enables constant monitoring of the instructional and learning processes, through which the administrator keeps the quality of human development process on track (Nenty, 1997). Thus, the results of school assessment are relevant not only to the learners, but also to the teachers.

The quality of human development process refers essentially to the quality of education, and the quality of education is largely operationalised as the quality of teaching that goes on in our schools. Over the years, the public has become increasingly inquisitive and bothered about the activities going on in schools, and also about the results schools are producing in their products (graduates). Governments, communities, proprietors, employers, parents and learners themselves have had reasons to worry about the results and products schools are producing. The decline and deteriorating results from schools in terms of academic achievement, attitudes, values, intelligence, psychomotor skills and other affective measures in their graduates have been matters of concern across the entire universe. The returns from schools, vis-a-vis the huge investments in education are quite disturbing, and have made some stake holders to associate quality of school products (in terms of achievement scores/grades) with quality of school personnel, who are largely teachers. Some have wondered whether achievement scores of learners in/from schools do not actually reflect the quality of teaching, and therefore, the quality of teachers (Heyneman, 1983; Tsang, 1988). Thus, the use of student achievement scores as the basis, or one of the bases, of assessing the quality of teaching/teachers has become a dominant issue, albeit a controversial one, in education and in educational research community.

Use of Student Achievement Scores to Assess Teachers – Issues

Use of student achievement scores as the basis to assess or evaluate teachers is one of the many approaches of teacher evaluation identified by Darling-Hammond, Wise and Pease (1983) and Joshua (1998). Other approaches include classroom observation, student ratings, peer ratings, principal/HOD/administrator ratings, self-rating, teacher interview, parent rating, competency tests, and other indirect measures. In this approach (use of student achievement scores), test scores of students are used as a measure of not only student achievement, but also of teacher achievement, performance and effectiveness. In education, the ultimate concern is student learning. Because of this, some authors (such as Ebel, 1980; Harris, 1979; McDonald, 1980) maintain that evaluating or assessing teachers

through the use of student achievement data is the most legitimate procedure as it is the only true indicator of teacher effectiveness. Student achievement can be measured in many ways: like comparing student test scores to a State/National norm, net gains over time (like pretest to posttest scores), and so forth (Haefele, 1980).

Use of student achievement scores as bases of teacher assessment is the epitomized approach when teaching is conceived as ‘the production of a product’ in the different conceptions of teaching outlined by Stark and Lowther (1984) and Joshua and Joshua (2000). Also, in their discussion on competing rationales for teacher evaluation systems, Stufflebeam (1992) and Joshua (1998) have located use of student achievement scores to evaluate teachers under ‘consumer protection and community responsiveness’ and ‘merit pay’ rationales. The consumer protection and community responsiveness rationale views teaching as a vital public service, and advocates the delivery of effective teaching to students, the protection of their welfare and responsiveness to community and societal needs. Teachers should be evaluated by student advocates and protectors, and are expected to show evidence of effective execution of teaching responsibilities. Of prime importance is the assessment of student achievement to ensure that graduates will be highly employable, competitive in the world economy, and well qualified for further education. On the other hand, ‘merit pay’ refers to the award a teacher receives for excellent performance. It could be one-time monetary bonus, pay rise, or promotion to a higher rank (which will also result in pay rise). Often, the merit assessments are based on student achievement, student ratings of instruction and supervisor/HOD judgments and recommendations. Thus, the results of student assessment, which usually show up in scores constitute ‘a testimonial’ about the teacher’s performance and effectiveness.

Use of achievement scores as a basis of assessing teachers’ effectiveness is one of the controversial approaches of teacher evaluation (another main controversial one is student rating of the teacher). While some authors maintain that student achievement data constitute the most legitimate procedure for evaluating teachers (Ebel, 1980; McDonald, 1980), some others show concern about the validity and reliability of student achievement scores derived from

standardized or teacher-made tests. Are the teachers in complete control of the teaching - learning situation in order to be considered the sole determinant of learning outcomes? Are student achievement scores really the only learning outcomes? What of the moderating effects of student's I.Q., home background, peer group, interest, aptitude and other variables on student achievement? Can one really be sure of what student achievement score indicates or measures in terms of student knowledge, skill and affective outcomes, before we can understand what such score indicates or measures in terms of teacher cognitive, psychomotor or affective outcomes? These are some of the concerns and issues expressed about using achievement scores of students solely for teacher assessment or evaluation, particularly for some summative decisions based on such data (Soar et al., 1983).

Use of SAS to Assess Teachers – Research Results from Abroad

In the study conducted and reported by Kauchak Peterson and Driscoll (1985), high school teachers in their sample were asked about the use of achievement scores to evaluate teachers. Eighty (80) percent of the teachers were opposed to such practice, and this approach of teacher evaluation was ranked 10th of the 10 approaches considered. The authors reported that opposition to the practice centered around two positions; one questioning the validity of the test to assess student performance; and two, the validity of such tests for measuring teacher ability or performance. The teachers noted that so many factors, internal and external to the student and the school, combine to affect student achievement scores. They, thus, argued that these scores should not be used as indices of teaching/ teacher effectiveness. In this study, teachers gave average rating of 3.95 out of 7 to student achievement scores, and this was the lowest of all the ten evaluation approaches rated by them. In a similar study conducted by Stark and Lowther (1984), use of student achievement scores as a teacher evaluation approach was ranked 6th of the six approaches considered in the study. These research findings tend to agree with those Glasman and Pauline (1982) who found that teachers' receptivity to evaluation practices was dependent upon their perceived control of the component being evaluated. If teachers do

not feel they have control over a part of their teaching, such as student performance on an achievement test, then they do not want to be evaluated on/by this measure. Such is the perception of teachers abroad on using student achievement scores as an approach to teacher evaluation.

Research Results from Nigeria

A study was carried out in Akwa Ibom State, one of the 36 States in Nigeria, to determine the attitude of Nigerian secondary school teachers to the use of student achievement scores as a basis of teacher evaluation. The purpose of this study was to determine the nature of the attitude of secondary school teachers toward the use of student achievement scores (SAS) as the basis to assess teachers; and how this attitude measure was influenced by their professional status, and the purpose(s) of the assessment exercise (i.e. the uses to which the results of the assessment will be put). Thus, the study was designed to test the following three null hypotheses:

- (i) the attitude of secondary school teachers to the use of student achievement scores as a basis for assessment of teachers is not significantly positive;
- (ii) the attitude of professionally trained secondary school teachers to the use of SAS as basis for teacher assessment is not significantly different from the attitude of those teachers who are not professionally trained
- (iii) the attitude of secondary school teachers to the use of student achievement scores as basis for teacher assessment is not significantly influenced by the purposes to be served by that assessment (whether, formative or summative).

The study was a survey. The population of the study consisted of all the secondary school teachers in Akwa Ibom State, numbering about 3000 in the year of the study. The sample consisted of 480 teachers randomly selected from the population using stratified random sampling technique, with gender and school geographical location as the bases of stratification. Twenty local government areas (LGAs) were randomly selected from the 31 LGAs in the State. In each of the selected 20 LGAs, one secondary school was randomly selected in such a way that out of the 20 schools selected, 10 were at urban areas, and the other 10 were at the rural areas. In each selected

school, 24 teachers (12 males and 12 females) were randomly selected for the sample, giving a total of 480 teachers (240 males and 240 females; 240 from urban areas and 240 from rural areas).

The instrument for data collection was a questionnaire constructed by the researchers and vetted by four professionals in educational research, measurement and evaluation. The instrument consisted of six items, each with four Likert-type response categories of 'strongly agree', 'agree', 'disagree', and 'strongly disagree' (scored 4, 3, 2, and 1 respectively). The reliability of the instrument was ascertained using the split-half correlation analysis method with the associated Spearman-Brown Prophecy formula. The reliability estimate was 0.94. The 480 copies of questionnaire were administered on the selected teachers personally by the researchers, with the assistance of some teachers in the respective schools. This strategy yielded 100% return rate. The three hypotheses were tested with population t-test, independent t-test and dependent t-test statistical analysis techniques at .05 level of significance.

DATA ANALYSIS AND RESULTS

Hypothesis One

The attitude of secondary school teachers to the use of student achievement scores (SAS) as a basis for assessment of teachers is not significantly positive.

Here, the researchers reasoned that for the attitude of the teachers to be judged significantly positive, the score made by the teachers should be significantly greater than 15.0. This reference mean score (or population mean) was obtained by multiplying the mid-point between 'agree' and 'disagree' (which is 2.5) by the number of items measuring the attitude (which is 6). Thus, the null form of Hypothesis One was that the mean score representing secondary school teachers' attitude

to the use of SAS as basis of teacher assessment is not significantly greater than 15.0 ($H_0: m = 15.0$; $H_1: m > 15.0$). This hypothesis was tested with a t-test of one sample mean (also known as population t-test), and for the separate instances of formative and summative purposes of evaluation. The results of these analyses are presented in Table 1.

The results in Table I have shown that the calculated t-values of -11.57 (for instances of formative purposes), and -39.91 (for instances of summative purposes) in their absolute forms are each greater than the critical t-value of 1.65 at .05 level of significance. Although the null hypothesis is rejected in these two instances, the interpretation of the negative t-values is that the observed attitude measures are significantly lower than the reference mean score. This implies that the attitude of Nigeria teachers to SAS as a basis for teacher assessment is significantly negative, both when the teachers know that the results of the evaluation are to be used for formative purposes, and when they know that the results are to be used for summative purposes.

Hypothesis Two

The attitude of professionally trained secondary school teachers to the use of SAS as basis for teacher assessment is not significantly different from the attitude of those teachers who are not professionally trained ($H_0: m_t = m_u$; $H_1: m_t \neq m_u$) (t=trained; u=untrained).

Teachers in the sample were categorized into two groups of professionally trained, that is, those who have some form of academic/professional qualification in Education ($N_1 = 400$), and professionally untrained ($N_2 = 80$). An independent t-test statistics was applied on the data; and the results of the analysis are as shown in Table 2. Also, the hypothesis was tested for instances of formative and summative purposes of teacher assessment.

Table 1: A population t-test analysis of whether teachers' attitude to use of student achievement scores a basis for teacher assessment is significantly positive

Purpose of assessment	Variable	Mean $\pm$ SD	t-value	df
Formative	Teachers' attitude	12.78 $\pm$ 4.21	-11.57*	479
Reference Mean Score		15.00	0.00	
Summative	Teachers' attitude	9.52 $\pm$ 3.01	-39.91*	479
Reference Mean Score		15.00	0.00	

* Significant ($p < .05$); Critical t = 1.65; N = 480

Table 2: Independent t- test analysis of the influence of teachers' professional status on their attitude to use of SAS as basis for teacher assessment

<i>Purpose</i>		<i>Group (Prof status)</i>		<i>N</i>	<i>Mean ± SD</i>	<i>t-value</i>	<i>df</i>
Formative	(1)	Professionally-trained	400	12.65 ± 4.18	-1.48	478	
	(2)	Professionally-untrained	80	13.41 ± 4.29			
Summative	(1)	Professionally-trained	400	9.46 ± 2.98	-1.00	478	
	(2)	Professionally-untrained	80	9.82 ± 3.12			

Critical t = 1.96; N=480

The results presented in Table 2 indicate that the calculated t-values in both instances of formative and summative purposes of teacher evaluation are each lower than the critical t-value. Thus, the null hypothesis is not rejected. Therefore, the professionally trained teachers in Nigeria are not significantly different from their counterparts who are untrained in their (negative) attitude to use of student achievement scores as basis of teacher assessment, whether the evaluation results are to serve formative purposes or summative purposes.

Hypothesis Three

The attitude of secondary school teachers to the use of SAS as the basis of teacher assessment is not significantly influenced by the purposes to be served by that assessment (whether formative or summative).

Teachers in the sample were required to express their opinions if they knew that the results of such assessment were to be used for formative purposes (geared towards instructional/professional improvement), and also if they knew that the results of the assessment were to be used for summative purposes (like promotion, award and dismissal, etc.). Thus, each teacher in the sample had two scores, and therefore the sample had two mean scores representing the attitudes of the teachers to the use of SAS as a basis for teacher assessment when the purposes of the assessment are formative, and when the purposes are summative. A dependent t-test analysis was applied on the data to test the hypothesis that these two mean scores (being the measures of attitudes of the teachers) were not significantly different. The result of the analysis is presented in Table 3.

The result presented in Table 3 has shown that the calculated t-value of 22.03 is higher than the critical t-value of 1.96. With this result, the null hypothesis is rejected. This means that the measure of teachers' attitude to the use of SAS

to assess teachers when the purpose of assessment is formative, is significantly higher than the measure of attitude when the purpose of assessment is summative. Thus, this result implies that the attitude of secondary school teachers to the use of SAS as a basis of assessing teachers' instructional effectiveness when the purpose of such assessment is summative is significantly more negative than the attitude of those teachers when the purpose of the assessment is formative.

Table 3: Dependent t-test analysis of the difference in teachers' attitude to use of SAS as basis for teacher assessment under formative and summative purposes

<i>Purpose of Assessment</i>	<i>Mean ± SD</i>	<i>t-value</i>	<i>df</i>
Formative	12.78 ± 4.21	22.03*	479
Summative	9.52 ± 3.01		

*Significant (p < .05); Critical t=1.96; N = 480

DISCUSSION

The major finding of this study is that the secondary school teachers in Nigeria sampled for this study have shown a significantly negative attitude to the use of student achievement scores as a basis of assessing the instructional effectiveness, irrespective of the purposes to be served by such assessment. The other findings are that professionally trained teachers and their untrained counterparts are not significantly different; and that the attitude of teachers in the sample to this approach of assessing teachers' performance when the results of the assessment are to serve summative purposes is significantly more negative than their attitude when the results are to serve formative purposes. These findings are very similar to those of Kauchak and others (1985), Stark and Lowther (1984) and Glasman and Pauline (1982). In these studies, use of student achievement scores as a means of assessing teachers' job performance was not only ranked

the last, but was also discountenanced, rejected and condemned as an evaluation or assessment approach, especially if some serious decisions are to be taken on the teacher.

Really, use of SAS as a basis for teacher evaluation is a controversial issue. Some authors support it, while some repudiate the idea, raising questions about its validity and reliability. Student achievement scores (SAS) actually represent many things, and they require many assumptions to link them to teacher competence, performance or effectiveness. Besides, there are so many factors (some internal and some external to the learners and the school) that combine to affect and impinge on student achievement scores. Many of the factors that affect or determine learning are outside the control of the teacher. These serve to threaten the validity, reliability, interpretability and generalizability of these achievement scores as bases for determining teachers' effectiveness, competence or efficiency. Thus, the system will be unfair to teachers if it applies SAS as indicators of teachers' ability, performance or effectiveness. Although analyses of test scores and test items (in item analysis, for example) can, and does reveal issues that could ultimately improve the quality of instruction (and teachers are enjoined to do so), use of student test scores to assess the teacher, and take some important administrative decisions about him/her will continue to be subject of age-long intense controversy; and, therefore, should be handled with utmost care.

CONCLUSION AND RECOMMENDATIONS

From the findings of this study, it is concluded that Nigerian secondary school teachers are not too different from their counterparts abroad in rejecting the use of SAS as basis for assessing teachers' instructional effectiveness/performance, especially when the results of such assessment are to be used for personnel decisions affecting the teachers directly. It is recommended that Nigerian educational policy makers and school administrators should exercise a great deal of caution in using examination results of students as absolute indicators of teachers' ability, performance and effectiveness, especially in matters that bear on personnel decisions affecting teachers directly. However, teachers should

continue to critically examine test scripts of their students to discover the diagnostic information they could use to improve instructional delivery and learning in the school system.

REFERENCES

- Darling-Hammond, L., Wise, A. E., and Pease, S. R.: Teacher evaluation in the organizational context: A review of literature. *Review of Educational Research*, **53**: 285-328 (1983).
- Ebel, R. L. Evaluation of students: Implications for effective teaching. *Educational Evaluation and Policy Analysis*, **2(1)**: 47-51 (1980).
- Glasman, N. S., and Paulin, P. J.: Possible determinants of teacher receptivity to evaluation. *The Journal of Educational Administration*, **XX(2)**: 149-171 (1982).
- Haefele, D. L.: How to evaluate thee, teacher - Let me count the ways, *Phi Delta Kappan*, **61**: 349-352 (1980).
- Harris, A. J.: The effective teacher of reading revisited. *The Reading Teacher*, **33(2)**: 135-140 (1979).
- Heyneman, S. Education during a period of austerity in Uganda, 1971-1981. *Comparative Education Review*, **27(3)**: 403-406 (1983).
- Joshua, M. T.: Teacher evaluation: Different approaches and competing rationales. *Nigerian Journal of Educational Foundations*, **2(2)**: 92-106 (1998).
- Joshua, M. T. and Joshua, A. M.: Different conceptions of teaching: Implications for learning and evaluation of teaching in Universal Basic Education (UBE). *International Journal of Research in Basic and Lifelong Education*, **1**: 366-373 (2000).
- Kauchak, D., Peterson, K. and Driscoll, A.: An interview study of teachers' attitudes toward teacher evaluation practices. *Journal of Research and Development in Education*, **19(1)**: 32-37 (1985).
- McDonald, F.: Assessing teacher effects on student learning. pp. 71-87. In: *Due Process In Teacher Evaluation*. D. Peterson and A. Ward (Eds.). University Press of America, Lanham, Maryland (1980).
- Nenty, H. J.: Basic assessment skills for classroom teachers. *Unpublished Manuscript*. Institute of Education, National University of Lesotho, Lesotho (1997).
- Soar, R. S. Medley, D. M., and Coker, H.: Teacher evaluation: A critique of currently used methods. *Phi Delta Kappan*, **64(4)**: 239-246 (1983).
- Stark, J. S., and Lowther, M. A.: Predictors of teachers' preferences concerning their evaluation. *Educational Administration Quarterly*, **20(4)**: 76-106 (1984).
- Stufflebeam, D. L.: Competing rationales and associated models and approaches for evaluating the performance of teachers. *Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California* (1992).
- Tsang, M. C.: Cost analysis for educational policymaking: A review of cost studies in education in developing countries. *Review of Educational Research*, **57**: 181-230 (1988).