USING FORMATIVE ASSESSMENT RESULTS TO PREDICT STUDENTS' ACADEMIC ACHIEVEMENT ON HIGH STAKES TESTS IN MATHEMATICS

Basil C.E. Oguguo, Ph.D
Department Of Physical Science Education
Imo State University, Owerri,
Imo State.

And

Nkeiruka P. Francis, Ph.D
Department Of Curriculum Studies
Imo State University, Owerri,
Imo State

Abstract

The purpose of this study is to examine whether formative assessments can accurately predict student academic achievement on Yearly Progress (YP) indicators as measured by standardized criterion-referenced mathematics tests. The population for this study comprised JSSI, JSSII, and JSSIII students enrolled in Ohaji High School Mgberichi from 2004-2007. Over 2,900 student assessments were used to conduct the research. The data sources included the first term, second term, and third term post formative assessments which are administered every seven-week grading period in the school system. The findings indicated that various grade levels exhibited a higher predictability factor with certain quarterly assessments than others. Likewise, unit gains on post assessments demonstrated a statistically significant indicator for academic achievement on high stake standardized assessments. Based on the results, it was recommended amongst others that the school system strive to ensure that its pre and post formative assessment exhibit a high degree of test validity and maintain a strong correlation to the Ministry of Education Performance Standards.

Public education in Nigeria has become an assessment-driven system that focuses on accountability towards both school and teacher effectiveness along with students academic achievement. At the heart of student achievement tests the relationship between teacher and Student. Often times, very early in a school year,

teachers attempt to evaluate their individual students' needs. Formative assessments are the instruments used by teachers and school authorities to help ascertain the current performance levels for students. These assessments can be in written form or often, they are simply demonstrations of student abilities informally observed by an instructor when a student attempts a task that would demonstrate content or objective mastery (Cogshall, 2004). With formalized summative assessments garnering most of the focus of the private and public sectors, formative assessments appear to be a vital link to student achievement that is often overlooked. They are often overlooked because their diagnostic characteristics are not used to assess for instruction vis-à-vis curriculum mapping, pacing, and scope and sequencing but rather, they have been relegated to an assessment of post-instruction.

Another key factor in student achievement revolves around tailoring instruction to meet the specific needs of each learner. This differentiation of instruction has occurred for many years in special education classrooms. Teachers were trained in pedagogy that addressed a variety of learning modalities and their instructional program was, and still is, driven by individualized learning objectives. Now, educators are being required to employ these best-practices towards all students often in mixed-ability classrooms (Gregory and Chapman, 2002).

This research used the Junior Secondary School Certificate Examination (JSSCE) as one assessment measure for student achievement. Imo State, along with most other states, has adopted a standards-based curriculum for educational zones to follow. Simply stated, a standards approach means that centralized themes of learning or topics are subdivided into more concise learning objectives or elements and the curriculum attempts to delve deep within a content area rather than cursorily skim the surface. Advocates of the standards-based curriculum state that this centralization of learning objectives provides a coherent and assessable level of achievement for all learners. Opponents of a standards-based curriculum argue that it is dummied down and does not provide for enrichment opportunities nor does it allow higher learning levels of learning to occur. Regardless of where someone stands on the issue of standards-based curricula, the fact remains that students in most states will attempt to master the learning objectives offered on a high-stakes, cumulative, criterion-referenced assessment. Reeves (2001) states that "balance is notably lacking in high-stakes state tests that claim to be based on standards but that are, in fact, limited by budget and time constraints to providing a brief, basic snapshot of student proficiency on a few standards." Moreover, this brief snapshot assessment will attempt to evaluate not only the student's own levels of academic achievement but also, these scores will then attempt to evaluate the effectiveness of a teacher, a school, a school system,

and in reality—the effectiveness of a state's educational system.

Formative assessment has been championed by assessment specialists (e.g., McTighe and O'Connor, 2005; Stiggins and Chappuis, 2005), and it is increasingly endorsed by professional organizations (e.g., Joint Committee on Standards for Educational Evaluation, 2003; Miller, 2005). Formative assessment is thought to have intrinsic acceptability (Black and Wiliam, 2003) to teachers, but system wide implementation has met with some resistance at the secondary level (Hayward and Hedge, 2005; Smith and Gorard, 2005). The Center for Educational Research and

Innovation (CERI, 2005) notes that powerful bureaucratic constraints limit the implementation of formative assessment in secondary schools, even though its use with adolescent students is particularly justified. In their international case studies, CERI (2005) concludes that the benefits of formative assessment outweigh the barriers to its implementation. Although they are negative, these barriers are also a natural part of the transformation currently underway in education. Black and Harrison (2001) point out that "the development of formative assessment has led to more radical changes in the ways of working of many of the teachers involved" and that "it takes time and patience to achieve changes of this type."

The concept of formative assessment evolved from the early definitions by Bloom, Hastings and Madaus (1971). The term formative assessment has been used in conjunction with the popular term, assessment-for-learning (Black, 2003; Earl, 2003), and the two are now considered conceptually identical (Threlfall, 2005). However, considerable confusion remains regarding the nature of the concept (Black, 2003). As Yorke (2003: 478) observes, "formative assessment is a concept that is more complex that it might appear at first sight". It is sometimes described as a linear sequence, involving teacher-directed instruction, feedback and correctives (Guskey, 2005). This definition fits well with behaviorist or early cognitive theories of learning (Allal and Ducrey, 2000; Yorke, 2003), but its limitations are seen in classroom-based research. Extending beyond prescribed instruction, formative assessment is fundamentally a collaborative act (Yorke, 2003: 496) that necessitates interaction between teachers and students. Several recent definitions detail the characteristics and elements of formative assessment (ARG, 2002, CERI, 2005; Crooks, 2001; Leahy, Lyon, Thompson and Wiliam, 2005; Shepard, 2005; Stiggins, 2002; Torrance and Pryor, 2001). In analyzing these sources, formative assessment is described as a composite practice, involving:

a) clearly communicated learning goals and evaluative criteria,

b) varied approaches to elicit information about learning, including questioning and observation,

c) balanced and descriptive feedback in varied forms,

d) the adjustment of teaching and learning as a result of the assessment, and (e) the active involvement of students. First, assessment, teaching, and learning are ideally integrated within a safe learning environment. Second, a multitude of internal and external factors, including teachers' knowledge and beliefs, impact the elements of formative assessment in practice. This model is presented, not as a static or permanent definition of formative assessment, but as a framework for further analysis and discussion around the concept, as it is presently understood (Roos and Hamilton, 2005).

Problem Statement

There are many factors that now challenge the classroom teacher, the public school administrator, and the public school policy maker to create an ideal teaching and learning environment for every student, with individualization of instruction to each learner's needs. Furthermore, the emphasis on accountability as expressed by standardized assessment results has created an educational climate reflective of most business models (Marzano, 2001). In essence, students have not only remained as the primary customer of the educational process, they have now also taken on the role as the primary producer" of achievement data. With this emphasis now on standardized assessment results serving as the qualifier for what an effective education is, education agencies are being forced to reexamine their curricular standards and objectives, their classroom pedagogy, and their ability to accurately formatively assess students prior to completing a comprehensive, high-stakes summative assessment. The present study therefore is an attempt to investigate whether formative assessments can accurately predict student academic achievement in mathematics.

Research questions

1 Do the post test scores in each of three quarters predict JSSCE (pass vs. fail) in Yr1?

2 Do the post test scores in each of three quarters predict JSSCE (pass vs. fail) in Yr2?

3 Do the post test scores in each of three quarters predict JSSCE (pass vs. fail) in Yr3?

Hypotheses

Ho1 The post test scores in each of three quarters do not predict JSSCE (pass vs. fail) in Yr1.

Ho2 The post test scores in each of three quarters do not predict JSSCE (pass vs. fail) in Yr2.

Ho3 The post test scores in each of three quarters do not predict JSSCE (pass vs. fail) in Yr3.

Research Design

A three year period of formative assessments data were collected for over 2,900 school students. The sample for the study consisted of 1801 students (939 males and 862 females) drawn from secondary schools in Ohaji/Egbema Local Government Area of Imo State. These schools were selected using purposive sampling technique. The junior secondary

school students who were administered the Imo state mandated annual assessment, Junior Secondary School Certificate Examination (JSSCE), and the quarterly post assessments at Ohaji High School Mgberichi from 2004-2007 were the subjects of this study.

This study utilized the Junior Secondary School Certificate Examination (JSSCE) required by the Imo State Ministry of Education and the quarterly post assessments given by the School. Since only the mathematics and English Language content domains are used by the Ministry of Education as the YP indicators to evaluate school system effectiveness, the researchers examined only the mathematics scores on the JSSCE. The JSSCE is designed to measure how well students acquire the skills and knowledge described in the State Performance Standards. Reliability is evaluated by statistical methods. For the 2004 JSSCE, total test reliabilities ranged from 0.79 to 0.86 for Igbo Language, 0.85 to 0.89 for English Language, 0.87 to 0.91 for Mathematics, 0.89 to 0.90 for Integrated Science, and 0.88 to 0.91 for Social Studies.

The researchers identified the JSSI, JSSII, and JSSIII students that were enrolled in the School from 2004-2007. Using the student information database, the researchers filtered the data in order to extrapolate only students who had valid test scores for the academic years pertinent to the study. The students' JSSCE scores in mathematics as well as their three quarterly mathematics post assessment scores were collected and entered into an excel spreadsheet. The students' personal data was protected by deleting all identifying test identification numbers, names, and classroom assignments.

In this research, the Statistical Package for the Social Sciences (SPSS) 15.0 for windows was used for data entry as well as for examining the data. Descriptive statistics were conducted on the data which included frequency and percentages for nominal data and means/standard deviations for continuous data. If the data points are all close to the mean, then the standard deviation is close to zero. To examine the hypothesis, nine binary logistic regressions for JSSI, JSSII and JSSII grades were conducted to assess if the school years 2004-2005, 2005-2006, and 2006-2007 post test scores can serve as independent predictor variables in relation to JSSCE (pass vs. fail) as the dependant criterion variable. Binary logistic regressions were chosen as the statistical tool of analysis because of the categorical, dichotomous nature of the criterion variable. In binary logistic regression we are testing the odds in which the predictors can predict the probability of the criterion variable (1=Pass, 0 =Fail).

Results

Means, standard deviations and ranges for JSSCE scores by grade in school year 2004-2005 are presented in Table 1. Table 2 presents the frequencies and percents for pass vs. fail by grade in school year 2004-2005. Mean, standard deviations and ranges for JSSCE scores by grade 2005-

2006 are presented in Table 3. Table 4 presents the frequencies and percents for pass vs. fail by grade in 2005-2006. Means, standard deviations and ranges for JSSCE scores by grade in school year 2006-2007 are presented in Table 5. Table 6 presents the frequencies and percents for pass vs. fail by grade in school year 2006-2007.

Research Question 1: Do the post test scores in each of three quarters predict JSSCE (pass vs. fail) in year 1?

Table 1: Means, Standard Deviations and Ranges for JSSCE Scores by Grade in 2004-2005

2004-2005	Min	Max	M	SD
JSI JSSCE	234	450	326.49	39.45
JSII JSSCE	261	411	317.92	23.53
JSII JSSCE	236	373	300.80	24.19

Table 2: Frequencies and Percents for Pass vs. Fail by Grade in 2004-2005

Pass			Fail	
2004-2005	Frequencies	Percents	Frequencies	Percents
JSSI	232	73.2	85	26.8
JSSII	182	80.9	43	19.1
JSSIII	97	52.7	87	47.3

Table 3: Means, Standard Deviations and Ranges for JSSCE Scores by Grade in 2005-2006

2005-2006	Min	Max	M	SD
JSI JSSCE	332	920.00	806.50	46.17
JSII JSSCE	263	824.00	325.24	47.19
JSIII JSSCE	233	407.00	315.09	33.05

Table 4: Frequencies and Percents for Pass vs. Fail by Grade in 2005-2006

Pass			Fail	
2005-2006	Frequencies	Percents	Frequencies	Percents
JSSI	192	29.0	127	39.8
JSSII	292	76.0	92	24.0
JSSIII	236	66.8	112	33.2

Table 5: Means, Standard Deviations and Ranges for JSSCE Scores by Grade in 2006-2007

2006-2007	Min	Max	M	SD
JSI JSSCE	748	898	806.69	28.69
JSII JSSCE	278	950	800.10	118.45
JSII JSSCE	258	437	327.25	30.88

Table 6: Frequencies and Percents for Pass vs. Fail by Grade in 2006-2007

Pass			Fail	
2006-2007	Frequencies	Percents	Frequencies	Percents
JSSI	235	55.0	192	45.0
JSSII	251	69.9	108	30.1
JSSII	282	83.7	55	16.3

Hypothesis 1a The post test scores in the 1ST quarter do not predict JSSCE (pass vs. fail) in year 1.

Table 7: Logistic Regression on JSSI Posttest Scores for 2004-2005 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	4.914	1.639	.003	136.242
Second Quarter	4.134	1.658	.013	62.407
Third Quarter	1.944	1.296	.134	6.987

A binary logistic regression was conducted to examine the three JSSI post test quarter scores of 2004-2005 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 83.70, $p < .001$. The model (i.e., the three predictors) fit the

data well, shown by the Nagelkerke R2= 45.0% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 84.8% of the participant outcomes in the sample. Table 7 presents the beta coefficients where the first and second quarter predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 136.242 and 62.4 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 1b The post test scores in the 2nd quarter do not predict JSSCE (pass vs. fail) in yr. 1.

Table 8: Logistic Regression on JSSII Posttest Scores for 2004-2005 predicting JSSCE
(Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	6.696	2.224	.003	808.945
Second Quarter	2.170	1.950	.266	8.754
Third Quarter	2.869	1.739	.099	17.613

A binary logistic regression was conducted to examine the three JSSII post test quarter scores of 2004-2005 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x2 (3) = 41.44, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 42.3% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 89.5% of the participant outcomes in the sample. Table 8 presents the beta coefficients where the first quarter predicted pass-fail outcomes: for every one unit increase in the first quarter scores, participants were 808.95 times more likely to pass than fail the JSSCE.

Hypothesis 1c The post test scores in the 3rd quarter do not predict JSSCE (pass vs. fail) in yr 1.

Table 9: Logistic Regression on JSSIII Posttest Scores for 2004-2005 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	5.440	2.170	.012	230.340
Second Quarter	5.231	2.285	.022	187.033
Third Quarter	7.166	1.913	.000	1295.102

A binary logistic regression was conducted to examine the three JSSIII post test quarter scores of 2004-2005 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 80.48, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 44.9% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 81.5% of the participant outcomes in the sample. Table 9 presents the beta coefficients where the first, second, and third quarters predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 1230.34, 187.03 and 1295.10 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 2a The post test scores in the 1st quarter do not predict JSSCE (pass vs. fail) in yr 1.

Table 10: Logistic Regression on JSSI Posttest Scores for 2005-2006 predicting JSSCE
(Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	7.544	1.218	.000	1890.147
Second Quarter	1.068	1.275	.402	2.909
Third Quarter	3.638	1.528	.017	38.032

A binary logistic regression was conducted to examine the three JSSI post test quarter scores of 2005-2006 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 115.86, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 36.5% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 81.2% of the participant outcomes in the sample. Table 10 presents the beta coefficients where the first and third quarter predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 1890.90 and 38.03 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 2b The post test scores in the 2nd quarters do not predict JSSCE (pass vs. fail) in year 1.

Table 11: Logistic Regression on JSSII Posttest Scores for 2005-2006 predicting JSSCE
(Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	6.572	1.736	.000	714.534
Second Quarter	1.460	1.380	.290	4.306
Third Quarter	5.441	1.716	.002	230.561

A binary logistic regression was conducted to examine the three JSSII post test quarter scores of 2005-2006 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 69.60, p < .001. This is in line with the findings of Ajuonuma and Mkpa (2009) that WASSCE do predict performances in first and second years in the University. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 24.8% of JSSCE (Pass vs. Fail) variance. Overall, the model correctly predicts 88.5% of the participant outcomes in the sample. Table 11 presents the beta coefficients where the first and third quarters predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 714.53 and 230.56 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 2c The post test scores in the 3rd quarter do not predict JSSCE (pass vs. fail) in year 1.

Table 12: Logistic Regression on JSSIII Posttest Scores for 2005-2006 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	2.210	1.468	.132	9.119
Second Quarter	1.518	1.309	.246	4.564
Third Quarter	3.432	1.357	.011	30.936

A binary logistic regression was conducted to examine the three JSSIII post test quarter scores of 2005-2006 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 42.00, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 19.0% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 84.4% of the participant outcomes in the sample. Table 12 presents the beta coefficients where the third quarter predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 30.93 times more likely to pass than fail the JSSCE.

Hypothesis 3a The post test scores in the 1st quarter do not predict JSSCE (pass vs. fail) in yr 1.

Table 13: Logistic Regression on JSSI Posttest Scores for 2006-2007 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	6.549	1.440	.000	698.286
Second Quarter	4.636	.990	.000	103.120
Third Quarter	2.224	.990	.025	9.240

A binary logistic regression was conducted to examine the three JSSI post test quarter scores of 2006-2007 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 178.15, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 41.0% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 82.0% of the participant outcomes in the sample. Table 13 presents the beta coefficients where the first, second, and third quarters predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 698.27, 103.12, and 9.24 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 3b The post test scores in the 2nd quarter do not predict JSSCE (pass vs. fail) in yr 1.

Table 14: Logistic Regression on JSSII Posttest Scores for 2006-2007 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	7.108	1.821	.000	1221.504
Second Quarter	1.541	1.859	.407	4.669
Third Quarter	9.895	2.198	.000	19833.766

A binary logistic regression was conducted to examine the three JSSII post test quarter scores of 2006-2007 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass

vs. fail), x 2 (3) = 122.96, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 64.2% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 87.5% of the participant outcomes in the sample. Table 14 presents the beta coefficients where the first and second quarter predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 1221.50 and 19833.77 times more likely to pass than fail the JSSCE, respectively.

Hypothesis 3c The post test scores in the 3rd quarter do not predict JSSCE (pass vs. fail) in yr 1.

Table 15: Logistic Regression on JSSIII Posttest Scores for 2006-2007 predicting JSSCE (Pass vs. Fail)

Predictor	β	SE	Sig.	Exp(B)
First Quarter	2.045	2.500	.413	7.729
Second Quarter	8.340	3.482	.017	4188.377
Third Quarter	9.156	3.177	.004	9470.640

A binary logistic regression was conducted to examine the three JSSIII post test quarter scores of 2006-2007 predicted JSSCE (pass vs. fail). Results of the regression indicate a significant model such that scores do predict JSSCE (pass vs. fail), x 2 (3) = 42.00, p < .001. The model (i.e., the three predictors) fit the data well, shown by the Nagelkerke R2= 29.0% of JSSCE (pass vs. fail) variance. Overall, the model correctly predicts 84.4% of the participant outcomes in the sample. Table 15 presents the beta coefficients where the first quarter predicted pass-fail outcomes: for every one unit increase in the quarter scores, participants were 30.93 times more likely to pass than fail the JSSCE.

Conclusions

The purpose of this research study was to determine the predictability of formative assessment data as an indicator for student academic achievement on standardized summative assessments. Specifically, this study examined frequent formative assessment measures that followed a pre-assessment instrument, instruction, remediation or enrichment as needed, and subsequent post-assessment data. The Junior Secondary school grade levels (JSSI, JSSII and JSSIII), along with multiple assessment years (2004-2005, 2005-2006, and 2006-2007) were studied. Also evaluated were three academic quarters assessment for every academic year that formative assessments were utilized. Nine research questions examined the predictability of data spanning three grade levels, three assessment periods or quarters, and three academic years and examined if available. Likewise, there was an equivalence of nine null hypotheses that attempted to negate a significant predictability of formative assessment data in relation to summative assessment achievement.

Recommendations

1. The school authority should continue to utilize the system of pre and post test assessment analysis in order to plan their

curricular content mapping and pacing sequences for instruction.

2. The researchers also recommends that the school system strive to ensure that its pre and post formative assessments exhibit a high degree of test validity and maintain a strong correlation to the Ministry of Education Performance Standards (MEPS). The vendor that provides the formative assessment software has implemented a team of professionals to ensure item validity on the assessments. Test constructors should use items only proven to be valid during test composition.

3. The school system should monitor and ensure that instructors are utilizing research-proven instructional best-practices in their classrooms in order to limit the amount of variance that students may receive between different instructors.

References

Allal, L. & Ducrey, G. P. (2000). Assessment of - or in- the zone of proximal development. *Learning and Instruction*, 10, 137-152.

Black, P. (2003). *The nature and value of formative assessment for learning.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Black, P. & Harrison, C. (2001). Feedback in questioning and marking: The science teacher's role in formative assessment. *School Science Review*, 82 (301), 55-61.

Black, P. & Harrison, C. (2001). Self and peer-assessment and taking responsibility: The science teachers' role in formative assessment. *School Science Review*, 83 (302), 43-49.

Black, P. & Wiliam, D. (2003). 'In praise of educational research': Formative Assessment 81 assessment. *British Educational Research Journal*, 29 (5), 623-637.

Bloom, B.S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning.* New York; McGraw-Hill.

Center for Educational Research and Innovation. (2005). *Formative assessment: Improving learning in secondary classrooms.* Retrieved March 2014 from the Organization for Economic Co-operation and Development Web site: http://www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=962005021P

Coggshall, J. G. (2004). *Educational administration, policy and reform: Research and measurement.* Alexandria, VA: Age Publishing.

Crooks, T. (2001, September). *The validity of formative assessments.* Paper presented at the annual meeting of the British Educational Research Association, Leeds, England.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press, Inc. *Ed.gov* (n.d.). Retrieved March 11, 2014, from http://www.ed.gov/NCLB

Gregory, G. H., & Chapman, C. (2002). *Differentiated instructional strategies*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2005). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Hayward, L. & Hedge, N. (2005). Travelling towards change in assessment: Policy, practice and research in education. *Assessment in Education,* 12 (1), 55-75.

Joint Committee on Standards for Educational Evaluation. (2003). *The Student Evaluation Standards.* Thousand Oaks, CA: Corwin Press.

Leahy, S., Lyon, C., Thompson, M. & Wiliam, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership,* 63 (3), 18-24.

Marzano, R. (2001). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.

Reeves, D. B. (2001). Standards make a difference: The influence of standards on Classroom Assessment. *Bulletin*

Roos, B. & Hamilton, D. (2005). Formative assessment: A cybernetic viewpoint. *Assessment in Education,* 12 (1), 7-20.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, (3), 66-71

Smith, E. & Gorard, S. (2005). "They don't give us our marks": The role of formative feedback in student progress. *Assessment in Education,* 12 (1), 21-38.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*,3 (10), 75-78.

Stiggins, R. J. & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice,* 44 (1), 11-18.

Threlfall, J. (2005). The Formative Use of Assessment Information in Planning – the Notion Of Contingent Planning. British Journal of Educational Studies, 53 (1), 54-65.

Wiliam, D., Lee, C., Harrison, C., Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education,* 11 (1), 49-65.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education,* 45, 477-501.