

Stewart Dugger Assessment for Learning 4.22.2020

Sarah Dugger

Doctoral Candidate

Austin Peay State University

Clarksville, TN 37044

United States

Dr. J. Gary Stewart

Professor

Austin Peay State University

Clarksville, TN 37044

United States

Abstract

Formative and summative assessment scores from high school students in a rural high school in Middle Tennessee were examined to determine if a correlation existed between said scores. This study also examined whether summative scores increased or decreased after using a series of formative assessments. Students were given a pre-test, ten formative assessments, and a post-test at specified times over a ten-week period. Literature reviewed in this study reveals that more study should be done regarding this potential relationship. Data gathered indicated a statistically significant correlation between formative assessment scores and summative assessment scores. The data also indicated a statistically significant increase in summative assessment scores after using the series of ten formative assessments. While much study has been done on the subject of different types of assessment and student achievement, little has been done to actually correlate formative and summative assessment scores. With the proclivity of emphasis on data-driven instruction, both classroom teachers and administrators need to be able to effectively interpret available formative assessment data as it relates to summative assessment data. By doing this, gaps in learning as well as in effort may be preemptively identified and closed. The use of formative assessments is an effective method of increasing student achievement on summative assessments.

Assessment for Learning

This study examined the correlation between formative assessment and summative assessment scores. One problem with traditional assessment practices is that a gap is created between student learning and summative assessment scores. According to Stiggins and DuFour (2009), “Formative assessment, done well, represents one of the most powerful instructional tools available to a teacher or a school for promoting student achievement” (p. 640). “The gap, of course, varies from student to student, with very real consequences for differentiating instruction” (Roskos & Newman, 2012, p. 535). Teachers and administrators historically have used past standardized assessment scores (summative assessments) as predictors of future scores. Due to the great differences in test-taking abilities and learning styles of students, these scores may not be an accurate predictor of future achievement, thus creating inaccurate data upon which instruction is based.

Data-driven instruction has changed the impetus of instructional practices. Effective educators understand how to use available data to adjust their scope and sequence to best fit the needs of their students. If this data is inaccurate, it presents a faulty picture of student potential and instruction cannot be aligned accurately. Furthermore, if teachers do not use available data in meaningful ways which influence student learning, then students become apathetic toward assessment. According to Cody (2010), “Test scores do not affect teens' lives in any tangible way and therefore students do not care about the tests” (p. 1). Moreover, researchers concluded that the standardized testing practices being used may do more harm than good to students and have called for an overhaul of the entire system (Volante & Ben Jaafar, 2010). Assessment-driven instruction and learning can be valid only if both teachers and students accept responsibility for using the available data efficiently. In his 2012 study, *Formative Assessment: Assessment is for Self-regulated Learning*, Clark asserted that there is enormous potential for formative assessment practices to influence and increase self-regulation of learning. Clark further described and supported the idea of using formative assessments as a part of inquiry-based instruction in order to increase student achievement (Clark, 2012). Implementing consistent, regular formative assessments that are authentic and accurate could provide both instructors and students usable data from which areas of strength and weakness can be identified and addressed efficiently and effectively.

In order to examine the potential relationship between formative assessment and summative assessment, this researcher conducted a ten-week study in a public high school in Middle Tennessee. The study was based on a 2010 study conducted by Bakula and reported in his paper, *The Benefits of Formative Assessments for Teaching and Learning*, Bakula reported that he observed a strong correlation between the use of formative assessment for learning and achievement on summative assessment (Bakula, 2010). Building upon the findings of Bakula, this current study examined formative assessment scores derived from a series of formative assessments delivered through a standardized formative assessment system as compared to summative assessment scores.

The population from which the research sample was chosen was ninth, tenth, and eleventh grade students. Evidence gathered from this study exhibited the potential for the procurement of valid and accurate assessment data to drive instruction and learning as well as predicting student achievement.

As educators are called to implement best practices in instruction, it behooves them to insure they not only are teaching toward areas of identified student weakness but also are challenging in areas of identified student strength. Revamping the way both educators and students view assessment data will result in greater achievement in the long run. By using formative assessment scores to predict summative assessment scores, teachers will be able to encourage students to set attainable, yet challenging, achievement goals and will be able to differentiate instruction more efficiently. “Formative assessment contributes to student ownership of learning more than any other classroom-based practice” (Brookhart, Moss, & Long, 2008, p. 57).

Limitations

The time-frame was a major limitation. This study was conducted over a period of ten weeks. It would be preferable to observe the relationship between formative and summative assessments over an entire school year.

Another limitation encountered was student apathy. Students were subjected to high-stakes standardized testing numerous times during the school year. School administrators required data-driven instruction so tested standards prevailed as the focus of instruction in tested high school subjects. Students received information about testing from the first day of the school year. In the high school from which the research sample was chosen, ninth grade students lost ten instructional days to standardized testing; tenth grade students lost 11 instructional days to standardized testing; and eleventh grade students lost 18 instructional days to standardized testing. Each testing session lasted a minimum of 90 minutes, with two testing sessions required for most of the testing days. By the time students reached the second round of testing days, they were burned out and tired about testing in general.

Review of the Literature

Discriminating Between Types of Assessments

Definition of assessment. Assessment is a term common to education; however, it is a term sans a precise definition. According to Dorn (2010):

The literature base on using formative assessment for instructional and intervention decisions is formidable, but the history of the practice of formative assessment is spotty. Even with the pressures of high-stakes accountability, its definition is fuzzy, its adoption is inconsistent, and the prognosis for future use is questionable. (p. 325)

Migrating to using the term assessment in lieu of the term testing, connotes a changing view of testing, but the variance in definitions demonstrates there is more to assessment than just testing. According to Huba and Freed (2000):

Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning. (p. 8)

Delving further into different definitions of assessment, Palomba and Banta (1999) stated: “Assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development” (p. 4). Common to both of these definitions is that assessment is to be used to improve learning; however, the first definition involves the collection of student-centered data while the second definition involves the collection of program-centered data. Another facet of assessment is defined by Allen (2004) as, “Assessment involves the use of empirical data on student learning to refine programs and improve student learning” (p. 2). By specifying the use of empirical data, Allen narrows the scope of assessment and virtually eliminates the use of observational assessment practices to improve student learning. No matter the type of data collected in assessment, the goal of assessment in education ultimately is improving student learning and refining educational programs. Above all, “assessment goes beyond grading by systematically examining patterns of student learning across courses and programs and using this information to improve educational practices” (www.cmu.edu, 2008).

Types of assessment. Under the umbrella of assessment, one discovers multiple types of assessment such as diagnostic assessment, formative assessment, interim assessment, and summative assessment. Attempting to obtain a universal definition of each type of assessment is as difficult as obtaining a universally agreed upon definition of assessment in general.

Diagnostic assessment. Diagnostic assessment is the most straightforward type of assessment to define. As the name implies, diagnostic assessment is used to diagnose, or quantify, the amount of prior knowledge students are bringing into the school year or into a new unit. Diagnostic assessments may be administered as formal pre-tests or as more informal assessments such as a journal writing assignment or mind map. The information gleaned from a diagnostic assessment is critical for an educator to use in planning a course of study for a class. For example, a high school biology teacher is ready to begin a unit about experimental design. Data collected from a diagnostic pre-test indicates that 95% of the students in the class do not understand the basic steps in the scientific method. The teacher then uses this information and plans an introductory lesson reviewing the basics of the scientific method. By taking the time to establish this foundational skill, the teacher is giving students a tool that will be used throughout the new unit of study. Diagnostic assessment, if used properly, allows educators to establish a baseline for their students against which achievement can be measured.

Formative assessment. According to the Glossary of Educational Reform (2014), “Formative assessment refers to a wide variety of methods that teachers use to conduct in-process evaluations of student comprehension, learning needs, and academic progress during a lesson, unit, or course” (edglossary.org). Garrison and Ehringhaus (2007) described formative assessment as follows: “Formative Assessment is part of the instructional process. When incorporated into classroom practice, it provides the information needed to adjust teaching and learning while they are happening” (p. 1). These definitions lack specificity with regards to the methodology used within formative assessment. Researchers concurred that formative assessment is a low-stakes form of assessment through which educators use collected information to prescribe adjustments to educational programs, but the actual methods used to gather data are diverse and numerous (Allen, 2004; Dorn, 2010; Garrison & Ehringhaus, 2007).

Summative assessment. Conversely, summative assessment is viewed as a high-stakes form of assessment with scores used in an evaluative sense rather than in a prescriptive sense. “Summative Assessments are given periodically to determine at a particular point in time what students know and do not know” (Garrison & Ehringhaus, 2007, p. 1). Summative assessments typically include testing instruments such as end of unit exams, cumulative mid-term exams, cumulative final exams, and standardized tests such as end of course exams. The standardized assessments such as end of course exams and TNReady are criterion-referenced summative assessments. Student performance is dependent upon mastery of predetermined standards. Other summative assessments, such as college entrance exams, are norm-referenced and are used to rank students. All of these are deemed high-stakes because the scores greatly effect overall student grades for a course or college admissions. Summative assessments are given infrequently, relative to formative assessments; therefore, they do not provide teachers with data on a consistent basis. Student performance blossoms when educators

implement data-driven instructional practices. This data primarily comes from formative assessments rather than summative assessments.

Interim assessments. Assessments which may be classified as interim assessments and benchmark assessments are those given between other assessments. The conundrum presented by interim assessments is whether they should be formative in nature or summative in nature. Typically, these assessments are formative because they are given at different specified intervals throughout the school year and the data (theoretically) are used to refine educational practices in order to address deficits identified in the assessments. Interim assessments are standards-based, not content-based. Benchmark assessments, a sub-group of interim assessments, are typically informal and are used to identify areas of strength, weakness, and growth. Benchmark assessments are similar to summative assessments in their formal design, but the information provided is used for instructional adjustment similar to interim assessment.

For the purposes of this research study, formative assessment was defined as a process involving the following formative assessment cycle:

- a. Establish a baseline (pre-test)
- b. Energize students (activate prior knowledge and create interest)
- c. Enlighten students (present lesson content)
- d. Engage students (individual and group activities)
- e. Examine student learning (formative assessment)
- f. Explain (reteach identified areas of weakness) and repeat steps c-f until mastery of each standard is evident
- g. Evaluate (summative assessment)

Summative assessment was defined as a cumulative standards-based formal assessment used to determine what students have and have not mastered throughout the course. No matter which type of assessment was given, the underlying purpose for assessment was improvement and evaluation of student learning.

The Relationship between Learning and Accountability

Achieving a balance between instruction and assessment remains one of the most difficult tasks facing educators. As districts and schools lean more heavily towards data-driven instruction, more assessment is necessary to gauge instructional effectiveness by student learning. It is incumbent upon educators to hold students accountable for learning while using data obtained through assessment to keep instruction student-centered. Data-driven instruction innately is student-centered if it is implemented efficaciously. At the most basic level, effective educators move through the cycle of formative assessment multiple times during each lesson they teach, each time adjusting, remediating, and challenging students based on the feedback received through the formative assessment.

Learning through assessment. Bloom's (1968) model of mastery learning demonstrated the connection between learning and accountability. Guskey (2005) stated:

In other words, instead of using assessments only as evaluation devices that mark the end of each unit, Bloom recommended using them as part of the instructional process to diagnose individual learning difficulties (feedback) and to prescribe remediation procedures (correctives). (p. 3)

These targeted assessments are used to identify strengths and weaknesses thereby allowing the teacher to adjust instruction as needed. Students are offered remediation on concepts where they have not exhibited mastery and then are re-assessed. If a student does not exhibit the target score for mastery, that student repeats the remediation and re-assessment cycle until mastery is reached. This type of system allows for maximum learning and accountability from both teachers and students. Boykin (2014) contended that "more focus should be put on assessment for learning rather than assessment of learning; and on assessment of the learning context and not just assessment of students" (p. 499). Essentially, Boykin advocated the implementation of formative assessment as a tool to adjust and drive instruction. Classroom-based formative assessment, when used consistently, encourages student learning through accountability. It gives educators insight into the abilities their students have and the new concepts they have learned. It is ironic that standardized summative tests instead of formative assessment systems, which have been proven to result in significant increases in student achievement, are used as a means to evaluate school improvement (Black & William, Assessment, 1998).

Accountability through summative assessment. Summative assessment has become the gold standard for evaluating student learning even though formative assessment provides more real-time data for educators. Both students and teachers must be held accountable for the amount of and quality of learning that takes place in the

classroom; however, the method of accountability is widely impugned. Federal accountability requirements play a major role in state accountability requirements, and all of these requirements involve summative assessment of some sort. The debate arises in the lack of immediate feedback from summative assessments. Being able to offer virtually immediate and continuous feedback to students lends to the attractiveness of formative assessment (Clark, 2012). In this fashion, formative assessment provides for both accountability of teachers and students. This prospect of immediate feedback grants teachers the opportunity to remediate or accelerate lessons based on student strengths and weaknesses. Students are held accountable for their learning and are able to monitor their own progress. This relationship between learning and accountability results in higher levels of student achievement.

Assessment of Learning versus Assessment for Learning

Remarkably, the one-word shift from *oftofo* creates a complex paradigm in the realm of assessment. If the underpinnings of assessment relate to the assessment of learning, then the focus of classroom instruction revolves around the assessment. In this situation, educators have limited freedom to allow students to explore beyond the tested standards. On the other hand, if the basis for assessment relate to assessment for learning, then the focus of classroom instruction revolves around learning. Teachers are able to allow students to apply learning to real world situations which demonstrate progress in learning through depth and complexity of application. The US, over the past forty years, has become more focused on assessment of learning rather than assessment for learning. From the beginnings of summative assessment in instruments such as the *Iowa Test of Basic Skills*, to the more recent PARCC (Partnership for Assessment of Readiness for College and Careers), assessment of students has centered on performance for ranking and placement purposes. This type of assessment may do more harm than good for students. In a report issued by the Committee on Incentives and Test-Based Accountability in Public Education, Editors Michael Hout and Stuart Elliott (2000) present findings from a nine-year study examining test-based incentive practices. This study “concluded that the emphasis on testing yielded little learning progress but caused significant harm” (fairtest.org, 2011). These incentives were examined from both an economic perspective and a psychological perspective. From an economic standpoint, an irregularity comes about when “the use of incentives with performance measures that do not adequately reflect the true value of the work that is being done” (fairtest.org, 2011, p. 2). For example, consider the educational incentives attached to students labeled as *on the bubble* between basic and proficient for standardized test scores. These students are identified as needing an extra push to advance to the next level of test scores and may be placed in some type of intervention program to help them move up the ladder of proficiency, often to the neglect of other students. With the current teacher evaluation models, it is detrimental to teachers to have many *bubble kids* in their classrooms because of the potential for negative progress or no progress. Therefore, the more stable and consistent students who score in the proficient to advanced range are aggressively sought out by teachers. In this example, the incentive of receiving special attention and help may cause the bubble students to regress in their testing achievement so they can continue to receive intervention efforts. It is evident through this example that “the incentives placed on the performance measures lead [students] to perform actions that increase the performance measures but not the underlying value of their work” (Hout & Elliott, 2011, pp. 2-3). The predicament stems from the narrow band of constructs that are evaluated through current summative, standardized assessment practices. Unfortunately, these assessments, whether national or state-based, typically assess a limited number of skills based on a pre-determined set of criteria. The scope of these assessments cannot feasibly be large enough to assess all areas of achievement influencing student growth in learning. Negative ramifications of high-stakes summative assessment “include narrowing the curriculum, teaching to the test, pushing students out of school, driving teachers out of the profession, and undermining student engagement and school climate” (fairtest.org, 2011, p. 2). High-stakes summative assessment, such as TNReady and end of course exams, in order to be equitable to students and provide a clear picture of learning, needs to be just one of many measures of student achievement.

Formative assessment provides for ongoing methods of measuring student learning by utilizing the mantra *assessment for learning*. The many configurations of formative assessment have the potential to morph into instruments through which both teachers and students can discern growth and decline and can work together to turn decline into growth. In *The Impact of Formative Assessment and Learning Intentions on Student Achievement* (2014), Hanover Research summarized multiple seminal works and meta-analyses conducted over formative assessment. Two key findings were identified:

1. Existing research yields promising conclusions regarding the impact of formative assessment on student academic achievement. Though the focus of most research is not on summative test scores, findings indicate that students who receive formative assessment perform better on a variety of achievement indicators than their peers do.
2. The literature supports the efficacy of explicitly stated learning intentions and assessment criteria in improving student learning outcomes. (Hanover, 2014, p.3)

Similarly, the National Council for Teachers of Mathematics has drawn the following conclusions (among many):

1. Existing research yields promising conclusions regarding the impact of formative assessment on student academic achievement. Though the focus of most research is not on summative test scores, findings indicate that students who receive formative assessment perform better on a variety of achievement indicators than their peers do. While many of these studies exhibit methodological limitations, experts agree that the practice of assessment for learning shows promise in its ability to improve student performance.
2. The literature supports the efficacy of explicitly stated learning intentions and assessment criteria in improving student learning outcomes.
(William, 2007, pp. 1-4)

Further supporting the benefits of using formative assessment for learning, Karee E. Dunn and Sean W. Mulvenon (2009), from The University of Arkansas, assert “The use of formative assessments, or other diagnostic efforts within classrooms, provides information that should help facilitate improved pedagogical practices and instructional outcomes” (p. 1). Yet another, seldom mentioned, benefit of assessment for learning is moving students into self-regulated learning. Through the use of a formative assessment system, students begin to take ownership in their learning. They transition from being dependent students who are spoon-fed information to be regurgitated on an arbitrary assessment into being independent learners who understand that they have the power to expand their learning by mastering skills. Instead of viewing assessment as a threat to their academic progress, even the most unskilled test-takers will shift their focus to mastery for learning. This construct shift results in true learning which carries over to any type of assessment.

Formal versus Informal Formative Assessment

Formal formative assessment. Formative assessments can take many forms. Any type of assessment that is used to evaluate student learning, especially understanding, may generate data that can be interpreted formatively. Formal formative assessment is a more structured variety of formative assessment. In this variety, teachers may use published assessments or teacher-created assessments. With either type of assessment, the formal aspect comes into play because students actually complete some type of written assessment. Test items on formal assessments may include subjective and objective questions, depending on the standards being evaluated. If possible, coding test items according to which standards are being evaluated assists teachers and students in knowing specifically which strengths and weaknesses students have. This type of formative assessment lends itself to student tracking of data, which has been proven to be a tremendous motivator to self-regulated learning and improvement. According to Cauley and McMillan (2010), “Formative assessment can have a powerful impact on student motivation and achievement” (p. 1). One motivating factor of formal formative assessment comes through emphasizing mastery goals rather than performance goals. By using a mastery approach to formal formative assessment, students are empowered to accept failure as a stepping stone to progress instead of an obstacle to success. Formal formative assessments may be administered as often as necessary to create a picture of student progress over time. The frequency of these assessments is dictated by the needs of the students.

Informal formative assessment. In contrast to formal formative assessment, informal formative assessment may take place many times over the course of each and every lesson. High-quality educators implement techniques such as checks for understanding (CFU), hand signal responses, quick-writes, exit tickets, observations, student self-assessments, student conferences, Socratic discussions, and think-pair-share activities to quickly assess student understanding. One issue that arises with informal formative assessment is record keeping. It is not always feasible for a teacher to record student data based on informal assessments multiple times during a lesson; therefore, it is necessary for a teacher to be discriminate and record enough data to evaluate the progress of students without interrupting the flow of a lesson. Informal formative assessment serves a meaningful purpose within daily lessons in that it can give a teacher immediate information about student understanding of the concepts being presented. This allows teachers to reteach small portions of a lesson as

soon as a low level of understanding is identified. It also gives teachers multiple opportunities to provide positive feedback and remedial feedback to students, which further motivates students to improve. According to Cauley and McMillan (2010), “When formative assessment is focused on progress of learning as students are instructed and process feedback, it is clearly most powerful” (p. 3). No matter which type of formative assessment is being used, the most effective impact on student learning comes when it is coupled with frequent and immediate feedback from teachers.

Formative assessment scores, when used in a summative fashion, provide more information to educators, students, and parents because the information is not limited to a single performance. In view of the notion that instruction is data-driven, it is fitting to insure that the data driving instruction is the most comprehensive and accurate information available. Summative evaluation of formative assessments along with summative assessments appears to be the most beneficial method of using assessment data for ranking and placement purposes. Equally important, formative evaluation of both formative and summative assessment data is will provide all pertinent entities with information which can be used to formulate instructional practices and policies (Dunn & Mulvenon, 2009).

Gaps in the Literature

Although a copious amount of research exists regarding the use of formative assessment to improve learning, little is found that actually correlates formative assessment performance to summative assessment performance. A consensus is lacking with regard to defining formative assessment in practice and policy. According to Dunn and Mulvenon (2009), “the vagueness of the constitutive and operational definitions directly contributes to the weaknesses found in the related research and dearth of empirical evidence identifying best practices related to formative assessment” (p. 2). For this reason, a comprehensive working definition of formative assessment is needed. One of the most basic questions that must be addressed by this definition is: “Is formative assessment an instrument or a process?” Alongside the answer to this question, research needs to be conducted to demonstrate the validity of formative assessment for evaluating student learning progress. “Without a clear understanding of what is being studied, empirical evidence supporting formative evidence will more than likely remain in short supply” (Dunn & Mulvenon, 2009, p. 2). Winger (2005) used a summative assessment as a formative assessment by providing both quantitative and qualitative feedback about the results of the exam. He called this “formative summative assessment” (p. 164). Ambiguity in terminology as well as operationalization of formative assessment contributes to vague contrived labels such as *formative summative assessment*.

Furthermore, while a plethora of literature is found regarding the use of summative assessment determining AYP, there is a gap in the available literature regarding the incorporation of formative assessment data into growth models used to calculate AYP. Although Black, William, and Harrison (2004) present multiple findings regarding the effect of formative assessment on overall classroom achievement with positive effect sizes ranging from .03 - .07, little research exists to demonstrate the effect on summative assessment. Educators and students alike need a clear purpose for assessment and need a clarity of understanding the results of assessment and the use of those results. One of the drawbacks to standardized summative assessments is the lag time between testing and score reporting. After all, if scores are to be used to evaluate progress and rank and place students for the next school year, scores should be published in a timely manner. Current literature does not address the issue of late scores and the effect on students; similarly, current literature does not address the use of formative assessment scores as correlating to summative assessment scores in order to rank and place students. This study examined the relationship between formative assessment scores and summative assessment scores in order to fill this gap.

Research Questions

This study was conducted to determine if the use of standards-based formative assessments to drive instruction had a significant effect on student achievement.

1. Is there a relationship between formative assessment scores and summative assessment scores?
2. Are post-test scores (summative) significantly increased by using a series of formative assessments?

Methodology and Data Collection

In December 2015, English teachers at the high school selected for this study identified a need to more thoroughly utilize data for students in their classes. Specifically, these teachers wanted standards-based data that could be used to differentiate instruction in preparation for summative assessment. Through professional learning community meetings, these teachers collaboratively designed ten, standards-based formative assessments. These assessments consisted of ten questions which were patterned after TNReady questions. The

assessments and answer keys were evaluated for validity by exchanging assessments among the teachers. The assessments were typographically formatted so that they all had the same appearance with regards to page layout, font type, font size, and headings.

Once the formative assessments were designed, the pre-test / post-test instrument was chosen to be the TNReady practice test. This practice test was specific to grades nine through eleven and contained 42 items. During professional learning community meetings in January and February of 2016, the teachers decided to use a uniform implementation procedure for the formative assessments. This procedure was labeled *the formative assessment cycle* and included the following steps:

- a. Establish a baseline (pretest)
- b. Energize students (activate prior knowledge and create interest)
- c. Enlighten students (present lesson content)
- d. Engage students (individual and group activities)
- e. Examine student learning (formative assessment)
- f. Explain (reteach identified areas of weakness) and repeat steps c-f until mastery of each standard is evident
- g. Evaluate (summative post-test)

The pre-test was administered in March 2016, the formative assessments were given twice a week over five weeks, and the summative post-test assessment was given in May 2016. The post-test administered was the same testing instrument as the pre-test.

The experimental group of 271 students came from English 1, 2, and 3 classes at the high school. Because of the different teachers involved in the study, it was imperative that assessments were given using uniform assessments, procedures, and scoring to insure validity.

Teachers collected scores for the pre-test, formative assessments, and post-test scores on a spreadsheet. Data collected from all of the assessments were evaluated through descriptive and inferential statistical methods using SPSS to determine if a correlation existed between the formative assessments and the summative assessments. The academic coach at the high school provided data regarding ethnicity and socio-economic status. Then, all data were analyzed using SPSS software to calculate descriptive statistics, Pearson Product Moment Correlation, Coefficient of Determination (r^2), Cohen's d , and a two-tailed t -test.

Results

A Pearson Product-Moment Correlation was run to determine the relationship between formative test scores and post-test scores (summative assessment). All assumptions were met prior to running the correlation. There was a strong, positive correlation between formative test scores and post-test scores, which was statistically significant ($r = .470$, $n = 271$, $p < .01$). A significance of .000 resulted indicating a statistically significant relationship.

The resulting r of .470 indicated a moderate correlation between formative assessment scores and summative assessment scores (Correlation is significant at the 0.01 level). An r^2 value of .229 confirms that 22.9% of the variance in the dependent variable is related to the variance in the independent variable change in summative assessment scores may be attributed to the formative assessments. While correlation does not imply causality, this analysis does confirm a strong relationship between the variables.

Table 3

Paired samples statistics

	Mean	N	St. Deviation	Std. Error Mean
Pre-Assessment	18.41	271	7.842	.476
Post-Assessment	24.159	271	8.9959	.5465

$\alpha = .05$

Table 4

Paired Samples t-test

	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pre-Assessment - Post-Assessment	-5.7454	6.2312	.3785	-6.4906	-5.0002	-15.179	270	.000

$\alpha = .05$

There was a statistically significant difference in the scores for post-assessment (M=-5.7454, SD = 6.23) after the administration of the formative assessments, $t = -15.179, p < .05$. Specifically, this test indicated that summative assessment scores are significantly higher than pre-test scores after using a series of formative assessments. Calculation of effect size using Cohen’s *d* indicated an effect size of 0.68, which falls between moderate and large (Cohen, 1988). These results led the researcher to reject the null hypothesis.

The calculation of effect size was further broken down by grade level and individual students and is summarized in Table 5.

Table 5

Individual effect sizes by grade level

Grade Level	Effect Size	Percentage above .4	Percentage between 0 and .4	Percentage at 0	Percentage below 0
9	1.09	90	5	2	1.5
10	.5	55	19	2.7	23
11	.28	51	0	0	49

Computation of individual effect sizes, which tends to be more accurate with larger sample sizes, yielded data which can be used by teachers to assess the effectiveness of the formative assessments on individual student achievement.

Discussion

It was concluded from the Pearson Product Moment Correlation that a moderate relationship exists between formative assessment scores and summative assessment scores. The *r* value of .470 is extremely close to the value for a strong correlation. With a larger group of participants and targeted training of teachers in the use of both formative assessments and the data resulting from these assessments, it is the opinion of this researcher that the correlation would be much higher.

It was concluded from the calculation of r^2 that 22.9% of the variance in the summative assessment scores was related to the formative assessment scores.

When examining effect size by using the *r* value of .470 as the effect size, a moderate effect was noted. This demonstrates not only the value of using a larger sample size but also the innate effectiveness of formative assessment. Moreover, by calculating individual effect sizes and evaluating effect sizes by grade level, it was determined that the younger students, in this case the ninth grade students, exhibited a much greater positive effect from the use of the formative assessments than did the older students (11th grade students). This difference could be due to differences in teaching strategies, classroom atmosphere, student attitude about the testing, or true learning deficits. By discussing reasons for the lower effect size for the eleventh grade students and discussing reasons for the higher effect size for the ninth grade and moderate effect size for the tenth grade, teachers can begin to hone strategies and use what is most effective for their students.

It can be concluded from the two-tailed *t*-test results that summative assessment scores are significantly higher after the implementation of a series of formative assessments. This study was limited by the school year coming to a close. A more significant increase in post-test scores could be observed with a greater amount of time available for remediation and acceleration following formative assessment.

Recommendations

This study was conducted with the participation of most of the English 1, 2, and 3 teachers at a rural, Title 1 high school. One teacher was not able to submit enough data to be included in this study. The collected data could be more accurate if all teachers were trained specifically in the use of formative assessments and the data resulting from these assessments. Also, collaboration between teachers when analyzing and interpreting the data obtained would make the data much more meaningful. Teachers need to look at the overall school data, the grade level data, classroom data, and individual student data to appreciate a complete picture of student achievement. This type of data analysis will allow teachers to gauge how their students are performing as compared to other students in their class, their grade level, and their school. Most importantly, it will provide teachers and students both a way to track student achievement and identify deficits in learning. This type of formative assessment system creates an opportunity for teachers and students to prepare for summative assessment throughout the year instead of only preparing when summative assessment was imminent, both groups will have a feel for which standards have been mastered. Students will also have a better feel for the types of questions they will experience on summative assessment, thus assuaging nervousness and uncertainty. It would prove beneficial to the teachers in this school, as well as other teachers in the district, to offer professional development in order to use formative assessments to make data-driven decisions and increase student achievement.

A slightly negative attitude toward using the formative assessments was observed in the participants. By providing more teacher training as to the benefits of using formative assessments, it is assumed that teacher attitudes would improve. As for student attitude, it is recommended that students track their own data, thus allowing students taking ownership for their achievement. This would benefit students because they would be able to monitor their own areas of strength and weakness. It would also benefit teachers because it would provide specific data that could be used to differentiate instruction in order to meet the individual learning needs of each student. Therefore, the specific recommendations resulting from this study are:

1. Increase the use of formative assessments to span the entire school year, instead of just ten weeks.
2. Through in-service, professional development, or professional learning communities, train teachers in the most effective methods for promoting, implementing, interpreting, and applying the formative assessments. The emphasis for these sessions will be using formative assessments as teaching tools. A second emphasis will be on how to incorporate data tracking and identify areas of student achievement and weakness from the data.
3. Implement student data tracking of formative assessments. This practice will increase self-efficacy and self-regulation of learning among students and will encourage students to continually improve.
4. Repeat this study with a longer series of formative assessments and a larger group of participants across multiple subject areas in order to determine if formative assessment significantly impacts summative assessment scores across multiple disciplines.
5. Train teachers at the school level and the district level in the differences between the different types of assessments and the purposes for each. Educators have been exposed to so many different terms for assessments that the terms formative assessment and summative assessment have become blurred. If teachers can become clear about what assessment is as well as why assessment is done, then they can present assessment to students with clarity of purpose.
6. For teachers in tested high school subjects, begin the conversation about implementing year-long weekly formative assessments with the goal of collaboratively writing standards-based formative assessments and incorporating them into the curriculum. Through professional learning community meetings, teachers should create a feasible timeline for the creation of and implementation of said assessments. Teachers should also determine a timeline for collecting and analyzing the data from the assessments.
7. If students are going to be expected to participate in extensive formative assessments for learning, students need to understand the express purpose of these assessments and the potential for the formative assessments to increase their learning achievement. Training students in the basics of growth mindset versus fixed mindset will benefit students by helping them learn to strive to improve their learning.

This research study is generalizable to students in the state of Tennessee because the demographic makeup of the participant pool closely reflects the general demographic makeup of the state. Implications for policy at the school and district level include requiring teachers to incorporate this type of formative assessment system in their classrooms in order to increase student achievement with regards to mastering the standards set for their subjects. At the state level, it would be beneficial to teachers, administrators, and students if state level educational professionals could devise formative assessments based on the tested standards and question types found on the end of year summative assessments. If these were made available and implemented with fidelity in all schools across the state, all students would have the opportunity to increase their achievement on the summative assessments. The tantamount goal of education is to effectively increase student learning and achievement. From the state level down to the local teacher level, implementation of frequent, formal formative assessments will foster these increases.

References

- Allen, M. J. (2004). *Assessing academic programs in higher education*. Bolton, Mass: Anker Pub. Co.
- Bakula, N. (2010). The benefits of formative assessment for teaching and learning. *Science Scope*, 37-43.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74. doi.org/10.1080/0969595980050102.
- Bloom, B. S. (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation comment*, 1(2), n2.
- Brookhart, S., Moss, C., & Long, B. (2008). Formative assessment. *Exceptional Leadership*, 52-57.
- Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1-6. doi.org/10.1080/00098650903267784.
- Clark, I. (2012). Formative Assessment: Assessment Is for Self-regulated Learning. *Educational Psychology Review Educ Psychol Rev*, 24(2), 205-249. doi:10.1007/s10648-011-9191-6
- Cody, A. (2010, April 21). *Should We Try to Make Our Students Care About Test Scores?* Retrieved from http://blogs.edweek.org/teachers/living-in-dialogue/2010/04/should_we_make_our_students_ca.html.
- Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition*. Hillsdale: Lawrence Erlbaum.
- Dorn, S. (2010). The political dilemmas of formative assessment. *Exceptional Children*, 76(3), 325-337. doi.org/10.1177/001440291007600305
- Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical Assessment Research & Evaluation*, 14(7).
- Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2014). Selecting growth measures for use in school evaluation systems: Should proportionality matter? *Educational Policy*, 30(3), 465-500. doi:10.1177/0895904814557593
- Formative Assessment Definition. (2014). Retrieved February 6, 2016, from <http://edglossary.org/formative-assessment/>
- Garrison, C., & Ehringhaus, M. (2007). "Formative and Summative Assessments in the Classroom." *Association for Middle Level Education*. Website available at <http://www.amle.org/Publications/WebExclusive/Assessment/tabid/1120/Default.aspx> (Accessed June 5, 2012).
- Guskey, T. R. (2005). Formative Classroom Assessment and Benjamin S. Bloom: Theory, Research, and Implications.
- Hout, M., Elliott, S. W., National Research Council (U.S.), & National Research Council (U.S.) (2011). *Incentives and test-based accountability in education*. Washington, D.C: National Academies Press.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston: Allyn and Bacon.
- The impact of formative assessment and learning intentions on student achievement. (2014). Retrieved March 10, 2016, from <http://www.hanoverresearch.com/media/The-Impact-of-Formative-Assessment-and-Learning-Intentions-on-Student-Achievement.pdf>, 1-16.
- The Limits of Standardized Tests for Diagnosing and Assisting Student Learning. (n.d.). Retrieved February 29, 2016, from <http://www.fairtest.org/limits-standardized-tests-diagnosing-and-assisting>

- McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation*, 18(2).
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass Publishers.
- Roskos, K., & Neuman, S. B. (2012). Formative assessment: Simply, no additives. *The reading teacher*, 65(8), 534-538.
- Stiggins, R., & DuFour, R. (2009). Maximizing the power of formative assessments. *Phi Delta Kappan*, 90(9), 640.
- The Value of Formative Assessment. (2007). Retrieved March 6, 2016, from <http://fairtest.org/examarts/winter99/k-forma3.html>
- Volante, L., & Ben Jaafar, S. (2010). Assessment reform and the case for learning-focused accountability. *The Journal of Educational Thought*, 44(2), 167-188.
- Wininger, R. S. (2005). Using your tests to teach: Formative summative assessment. *Teaching Psychology*, 32(2), 164-166. doi.org/10.1207/s15328023top3203_7_
- Wiliam, D. (2007). Benefits of formative assessment. Retrieved February 1, 2016, from <http://www.nctm.org/Research-and-Advocacy/Research-Brief-and-Clips/Benefits-of-Formative-Assessment/>
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education*, 11, 49-65. doi.org/10.1080/0969594042000208994_