Exploring differences and relationships between online formative and summative assessments in Mississippi career and technical education

Ben Alexander, Sean Owen and Cliff B. Thames
*Research and Curriculum Unit, Mississippi State University,
Starkville, Mississippi, USA*

Abstract

Purpose – This study, a *post hoc* observational one, attempted to determine if career and technical education (CTE) students in the state of Mississippi would academically benefit from taking multiple formative assessments in an online format prior to completing their summative exams. Most CTE students in the state of Mississippi are required to take an end-of-course exam cataloged as the Mississippi Career and Planning Assessment System (MS-CPAS). Previously, MS-CPAS test score results did not impact school-wide accountability scores, but in recent years, some of the guidelines were changed so that these summative test scores now play a vital role in school accountability and rankings.

Design/methodology/approach – This study examines both formative and summative online exam scores for more than 13,000 students who have taken an MS-CPAS assessment in the 2018 and 2019 school years.

Findings – The results of this study revealed that there were significant differences in summative exam scores for students who took two online formative practice tests when compared to groups of students who did not take any formative practice tests. This study also illustrated a positive correlation between those students' final online practice test scores and their summative exam scores.

Originality/value – These results would prove very beneficial to both CTE teachers and directors in helping them understand the benefits of introducing formative practice tests into their programs to boost student understanding.

Keywords Statewide assessment, Career and technical education, Practice tests, High school students, Formative assessment

Paper type Research paper

Introduction

Much like the rest of the world, the USA is constantly attempting to ensure students are progressing and mastering their course content, especially in environments where technology and blended learning methods can be implemented. Although these challenges existed long before Covid-19 swept across the globe, the pandemic has made many traditional face-to-face institutions understand the complex task that blended learning environments and open and distance learning (ODL) institutions have been operating within for years. In the state of Mississippi, all career and technical education (CTE) courses have been structured for years with an optional hybrid online course management system. The use of this interface can allow teachers to use it for both online enrichment and remediation purposes. However, to

accurately measure student learning and progress, a separate online formative assessment system was developed in the hopes of not only increasing student mastery but also increasing student summative exam scores.

Ever since the passage of sweeping new laws in the USA that were designed to hold schools accountable for their performance based on student assessment scores, many educators have faced the intimidating prospect of having their worked judged by a single metric in the form of a required summative assessment. Due to the limited number of courses that required these assessments to measure both school and student progress, school districts and schools often focused their efforts on a few courses that counted toward accountability (Hunt, 2008). In some states, CTE courses and their instructors were initially spared from these new levels of scrutiny to which their academic counterparts were being subjected. However, as other laws have increased focus on the importance of CTE in recent years, some of these teachers have joined their colleagues in having state and federal accountability standards imposed on their work (Imperatore and Hyslop, 2017). Many of these accountability standards in the form of assessments have given instructors in CTE courses something to gauge their level of progress.

However, the increased accountability also meant something not quite as pleasant: more oversight and higher expectations. For many teachers, these expectations and pressure have meant a renewed focus on not only their course's curriculum but also of their own efficiency and instructional techniques (Cizek and Burg, 2006; Sadler, 1998; Young, 2006; Zimmerman and Dibenedetto, 2008). Inevitably, it has also led to a strong desire by politicians and educational leaders to prove that renewed focus, improved techniques and more money has led to an increase in results that could be easily quantitatively measured. Indeed, the spark of optimism that resulted in increasing funds and expectations has also resulted in higher expectations for students on standardized CTE assessments.

In the state of Mississippi, these standardized CTE assessments are called the Mississippi Career Planning and Assessment System (MS-CPAS). These assessments are delivered exclusively in an online format to tens of thousands of students in the state annually. These MS-CPAS tests have become far more than just a way to ensure that students are achieving measurable goals in their career or vocational coursework. As higher expectations and goals for graduation rates have increased in the state, a decision was made to allow schools to substitute some two-year CTE high school program credits in place of traditional classroom credits. However, in many schools across Mississippi, students who do not perform well on their MS-CPAS test in the first year of their course are often not allowed to advance to the second year of that pathway's course for credit. In this way, the new spotlight on MS-CPAS scores has become a potential high-stakes exit ticket for students to use for possible graduation. Now, the tests are not simply a measure of mastery on certain skills taught in a career pathway program, but they are a tool that can be used to ultimately assist in meeting certain graduation requirements (The Mississippi Department of Education, 2020).

This added weight has not only affected student participants who are required to take the tests, but also has greatly impacted and placed heavy burdens and stresses on the educators who teach these subjects. In short, the high-stakes testing game that many K-12 educators have grown accustomed to and other states have dealt with has now arrived in the laps of CTE teachers. Teachers of courses measured by these type of accountability standards are having their own success measured, fairly or not, by proxy through their students' performances on these summative exams (von der Embse et al., 2016).

With the stakes for CTE testing at an all-time high, Mississippi educators have wanted a tool or program that would allow them to better prepare their own students for these online assessment requirements and help increase learning mastery. This desire has led to the development of new, online formative practice tests designed to hopefully better prepare students for their summative exams.

The Mississippi State University Research and Curriculum Unit (RCU) has been given the responsibility of creating, implementing and assessing curriculum for all CTE courses for nearly 40 years. As a contractor for the Mississippi Department of Education (MDE), the organization is partially funded by MDE to manage these activities in conjunction with both state and school district leaders. When the clamor for more assessment resources (such as formative practice assessments) were first heard, the burden of fulfilling this need fell to the RCU. To accurately promote and implement online formative assessments, a deeper knowledge of these tests and their history needs to be understood.

Theoretical framework

At some point in the latter half of the 20th century, advancements in technology and academic resources had evolved far beyond the one-room schoolhouse environment to allow teachers more flexibility in assessment. Prior to this period, the idea of easily creating, administering and grading a test was a daunting task. Assessments were created by hand in a time-devouring, tedious process of writing out tests to measure student knowledge on any given subject. It is postulated that the advent of standardized testing in the USA occurred during the First World War when large numbers of individuals were likely subjected to standardized testing as a means of evaluating their potential for military service (McGuire, 1994). However, the technological advancements of a post-Second World War society led to the creation of things like better typewriters and first-generation copiers. Armed with these technological marvels, assessment was not quite as dreadful a task, and it began to flourish and evolve in secondary schools across the country. These innovations eventually led to the idea of using tests to enhance instruction and student understanding, instead of simply measuring student mastery. This was the birth of formative assessments in the modern age (Bell and Cowie, 2001; Crawford *et al.*, 2017; Zimmerman and Dibenedetto, 2008).

The idea of teachers administering a test to students in hopes of gauging their mastery on a subject has been around since before John Dewey's educational revolution. When and where educators first began using assessments to improve their own curriculum or teaching practices remains the subject of much speculation. It was likely Scriven (1966) who first coined the term "formative evaluation" as one modern educators may recognize. In fact, he proposed that there were two distinctive types of assessment – formative and summative. Accordingly, formative assessments were the ones utilized to build or evaluate the merits of an educational program while it was taking place, and summative assessments were used to evaluate whether or not the targeted goals of a program were met (Schildkamp, 2019; Scriven, 1966).

What is a formative assessment?

Although Scriven's definitions seemed rather straightforward at the time, the idea of what constitutes formative assessment has become much more complicated in the past few decades. Peering across the landscape of literature for an up-close examination of formative assessments today would likely leave the surveyor bewildered and confused. Educators and researchers have seemingly defined and redefined what exactly formative evaluations and assessments are numerous times in the past few decades (Guskey, 1987; Harlen and James, 1997; Sadler, 1998). This meandering understanding of what comprises a formative assessment essentially ensures that any current literature review on the topic delves into what exactly is the modern-day definition of these tests (Guskey, 2010; Shepard *et al.*, 2018).

Any serious discussion or exposition on the modern definition and use of formative assessments must likely start with Black and Wiliam's 1998 pioneering study on the positive effects of utilizing a formative assessment. The study pointed out that teachers who utilized

formative assessment could expect more gains in student understanding compared to other methods that were readily available to them. Their study of more than 250 sources laid the modern groundwork for not only the increased use of these tests but for more study of this type of assessment.

Discussions of what constitutes formative and summative tests are essentially based on a debate between what better defines an assessment: its design or its eventual purpose. The generally accepted idea of a formative assessment is that it is a test designed to help increase learning by perhaps altering instruction, curriculum or some other internal mechanism. This definition itself creates another question, however. Can an assessment instrument designed for one purpose be effectively used for an alternate purpose? A current example of this altered use could be a school administrator evaluating the effectiveness of a course, curriculum or teacher by examination of a summative end-of-course exam required for accountability models and then altering one of these components to achieve a desired result the next year (Booher-Jennings, 2005; King, 2016).

This blending of distinctive roles between formative and summative assessments persists today. As late as 2008, *Education Week* published an article about the battle inside of academia and assessment companies over the simple usage of the word "formative assessment." The importance of the word parsing reached such a fever pitch that some assessment gurus chose to stop using the word because of its divisiveness (Cech, 2008). However, the term has found greater acceptance in academic circles as the frequency of assessments have increased. In fact, some researchers argue that classroom activities can be used for either formative or summative purposes depending on how a teacher chooses to interpret these activities (Harlen and James, 1997). Black again revisited this topic and stressed that any feedback provided to a student following an assessment makes the test formative in nature despite its potential use of another design (Black *et al.*, 2003). Bell and Cowie (2001) state these findings as well in reasserting that a summative test can be used as a formative assessment in certain circumstances.

Feedback in formative assessment

An examination on the history and effectiveness of formative assessments would be incredibly hard to explore without some discussion as to the role that feedback plays in defining and separating these tests from summative ones. As stated earlier, it is the belief of many experts that giving feedback following an assessment places it into the formative category. Not only may the very use of feedback provide a clear definition for these assessments, but it also likely establishes the level of effectiveness for these tests (Guskey, 1987, 2010; Stiggins, 2018).

One aspect that many researchers have agreed on when it comes to formative assessments is that they are not nearly as effective if proper feedback is not given to the students who undertake these measures (Sadler, 1998). Researchers in many different academic subject areas have reported on the positive impact that feedback has on formative assessment (Nahadi *et al.*, 2015). In and of itself a formative exam is simply a break in the learning cycle to see what has been properly processed and understood by the student. The student answers questions about materials or subjects that have been covered by the teacher or is asked to show some form of mastery on these subjects through a test. For the full potential of that assessment to be realized, two separate factors must take place. First, the teacher must set an objective for the learner and understand the significance the test data represents in relation to that goal. Only then can the teacher offer effective, timely feedback as remediation to address any deficiencies in the subject knowledge (Hattie and Timperley, 2007; Stiggins and DuFour, 2009).

Cautionary tales about effective use of data for feedback in formative tests illustrate that when formative testing does not work as intended, many times, the reasons can be easily

traced back to teacher aptitude in these and other areas (Sadler, 1989). The prevailing thought that most teachers who administer formative assessments understand the data the tests generate and how to properly give feedback to influence better outcomes may be faulty. Elmore (2002) specifically stated that all teachers may not possess those skills, especially not in schools that are historically low-achieving. Other inquiries into factors that may affect teachers' use of data show a myriad of other factors that play a role in a teacher's effective use of data, such as grade-level and team norms, district expectations and school leadership (Young, 2006).

Quality feedback is the cornerstone on which a successful formative assessment is built. In a follow-up to their seminal work, Black and Wiliam state unequivocally that the benefits students receive from formative tests are bound closely to the quality of feedback that students receive from teachers (2009). Instructors have multiple approaches to how they may administer feedback to students; some methods are likely far more effective than others. As Wong states, feedback remains an important part of effective instruction even in these various non-traditional classes like massive open online courses (2016). Spirited debate about which types of feedback produce better results is the subject of numerous research studies. Many researchers have debated whether feedback in the form of oral, written, indirect or direct feedback offers the greatest gains for student understanding (Almasi and Tabrizi, 2016; Cepni, 2016). Other researchers contest that effective feedback can take on many forms, from overarching actions such as rewarding students for achievement results, to effective goal-setting following formative assessments (Tunstall and Gipps, 1996). An instructor may choose to have students take a very simplistic action such as allowing them to correct their mistakes on their tests. Another instructor may decide to have peer-to-peer discussions about misunderstandings they have about previous content covered on the assessment or even decide to reteach the entire unit if formative assessment outcomes did not meet their expectations. The type of feedback the instructor prescribes to students following a task or assessment can have a strong influence on student growth in the future on similar tasks (Butler and Nisan, 1986).

Benefits and limitations of formative assessment

The recent focus on new high-stakes accountability tests in the past two decades has spurred far more research on formative assessments across multiple continents and cultures. The implementation of the No Child Left Behind Law in the USA in 2001 placed heavier emphasis on schools performing well or, in many cases, being subjected to negative consequences. As these high-stakes testing models have been implemented, so too have increased use of formative assessments (Dixson and Worrell, 2016). In many ways teachers, administrators and school district officials have placed their trust and hope that formative assessments would bolster scores on these high-stakes summative exams. Those hopes have been illustrated by numerous organizations and authors who have published both qualitative and quantitative studies illustrating that the use of formative assessments is one of the promising interventions for promoting higher performance in students (OECD, 2011). The cost–benefit analysis of this testing culture gets more complicated when considering the implications of such a testing culture on individual students. In her detailed examination of the Texas accountability model, Booher-Jennings' study (2005) shows the serious ramifications for students who are deemed unlikely to pass these high-stakes exams. In this study of Texas high schools, Booher-Jennings noted that students who were less likely to pass accountability test were, in some cases, ignored or even placed in special education classrooms.

The beginning of the most recent cascade of formative assessment research began with Black and Wiliam's 1998 review of the existing data on the topic. In their follow up research, they advocated for more studies on the topic by stating that the average effect size from these

examined studies showed at least 0.4 significance levels, which were typically larger than other types of remediation that teachers employed (Black and Wiliam, 1998). Other educators also soon began to espouse the merits of formative assessment and its impact on greater student understanding (Smith and Gorard, 2005).

While studies have shown positive impacts that can occur with localized formative assessments originating in a classroom setting, there may be greater promise in developing testing systems that encompass the cohesive additions of summative exams as well. One of the benefits of a more complete, robust testing system that bridges the gap between formative and summative assessments is that better tests not only increase student learning, but also help students understand the expectations of the high-stakes tests utilized in accountability models (Klein *et al.*, 2005). As these accountability tests have proliferated in recent years, various experts have urged for increased consistency of alignment between these formative tests and the end-of-the-year state assessments (Dolin *et al.*, 2018).

Although the positive results of using well-planned formative assessments have been widely recognized in recent years, some unexpected outcomes of this practice have brought criticism as well. This renewed attention on the potential benefits of formative tests to ensure greater success in accountability may also contribute to a culture that is often criticized for placing far too much emphasis on testing. Also, the positive aspects of using formative assessments during the past few decades have been somewhat limited because the current testing environment has encouraged a teaching-to-the-test mentality.

The caution that some researchers have raised about possible overuse of formative assessments is not simply limited to a fear of creating a repetitive testing environment. In several studies that delve into the effects of formative assessment and its accompanying feedback, there is some ancillary concern about how these testing tools impact student motivation levels. In her examination of assessment practices in some UK vocational programs, Dr. Kathryn Ecclestone points to the potentially negative effects of formative assessments' raising of achievement scores at the cost of robbing students of intrinsic motivation and true learning mastery (Ecclesstone, 2007).

Purpose
Using this theoretical lens, the purpose of this investigation was to describe the differences in statewide CTE post-test scores among groups that take a different number of online CTE formative statewide assessments in Mississippi. The researchers also sought to examine the relationship between the number of online formative statewide assessment attempts and post-test CTE statewide assessment scores of secondary Mississippi students.

Objectives
The following objectives were developed to guide this study:

(1) Examine the differences in statewide online CTE assessment post-test scores among groups that take multiple online formative statewide CTE assessments over those who did not take a single formative statewide CTE assessment in Mississippi.

(2) Examine the relationship between online formative statewide CTE assessment scores and online summative statewide CTE assessment scores in Mississippi.

Method
The methodology for this observational study was in many ways non-existent. As has been stated earlier in this paper, the RCU was presented with several concerns leaders of career and

technical centers in Mississippi had about achieving better outcomes on state-required online assessments in CTE courses. As the organization in the state of Mississippi responsible for crafting and validating these end-of-the-year assessments, the RCU was not guided by conducting research, but rather by providing a practical online learning tool in a relatively short amount of time for both student and course improvement. To provide a tool that could quickly help diagnose potential issues students were having with certain aspects of the CTE course content, the idea of building online practice tests for some of the pathways was discussed internally and quickly began. The decision to use an online format would allow quick and even distribution of these assessments to the hundreds of schools located around the state. In the 2017–2018 school year, the RCU built online formative practice tests for 41 of the 57 CTE courses offered in the state.

Members of the RCU assessment team felt as though these online tests would allow teachers an opportunity to evaluate their teaching content in relation to student results on these assessments. Most of the end-of-the-year assessments students are required to take for CTE in Mississippi are comprised of 100 multiple choice questions that cover material from the entire course. These summative assessments have also been taken in an online format. These tests closely follow the curriculum map that is created by a separate team at the RCU. Courses that were selected to participate in the first year of the practice tests were determined based primarily on the number and quality of assessment items the RCU had in its internal test banks that could rapidly be made into adequate online formative practice tests. Following this first year's implementation of the practice test, the RCU constructed more assessments, resulting in 54 courses each having a viable practice test in the 2018–2019 school year.

Very quickly in the construction process of these online assessments, it was determined that these formative tests would reflect the scope of the entire coursework for each class. Another determination made was that these tests would be limited to half the number of questions students generally are responsible for answering on the end-of-the-year assessment. To create these formative tests more quickly, a decision was also made to utilize previous end-of-the-year test questions that were retired after being in service for several years. These test questions had already been checked for validity, so the assessment team was confident in the acumen of these online formative tests.

Members of the RCU selected questions for these formative assessments after carefully examining the required curriculum for each course and ensuring that these former test questions were still relevant considering possibly updated curriculum. End-of-the-year MS-CPAS tests are administered by RCU in an online format, so similarly, a decision was reached to deliver the practice test by the same method. Having students use a similar technological interface associated with their learning should lead to improved views of the technology tools and better student outcomes (Eskil *et al.*, 2010). A decision was made to release the formative assessments six weeks prior to the opening of the end-of-the-year testing window. The online testing platform would allow for each students' personal login to grant them two attempts with this same test within this six-week period.

Members of the RCU team encouraged both career and technical center directors and teachers in the state to utilize the online practice test for their class, but they were not required to do so. These CTE directors and teachers were also heavily encouraged to have students take their practice test at the beginning of the six-week timeframe. RCU staff members had hoped that once students' first practice tests results were available, teachers would then diagnose the data and address deficiencies through quality feedback.

Following two years of administering and scoring these online practice tests, members of the RCU staff wanted to gain an understanding of how helpful these formative tests were. The RCU ran an analysis of the collected data to examine the impact the practice assessments may or may not have had. The statistical analysis included the past two years of assessment data

that covered more than 50 courses and a total of 28,015 student summative exams. Each of these students had online practice tests available to them each year to take before their end-of-course summative exams. Several questions guided this analysis.

Do students who take multiple online formative CTE assessments score at higher levels on their summative assessment compared to peers who did not take any formative tests?

Do these online formative assessments serve as a valid predicator of the summative MS-CPAS test scores?

Findings
Group studies
When practice tests were released online for students to take, teachers were encouraged to administer one practice test, evaluate their data for student strengths and weaknesses, remediate their pupils' lowest areas of knowledge and then to finally take the second practice test to evaluate growth. As the assessment team's intentions with the construction of these tests were for teachers to utilize two practice tests with feedback given between them, the students were divided into two distinct groups. Group 1 was comprised of 12,956 students who had taken both online practice tests that were available in their career pathway course. Group 2 students were made up of 483 students who had never taken a practice test at all. In total 15,059 students in the state of Mississippi took less than two practice tests during the period we examined. However, only 483 of that number had failed to take any practice tests. That left 14,576 students who had taken one practice test during the years that were examined. RCU staff determined that because these 14,000 students obviously did not take both online practice tests, they likely would not be acceptable for a comparison as a "control" group against Group 1. Group 1 students were the group staff members believed to be the most likely students to have had the benefit of formative assessments and feedback.

To conduct a simple comparison between our two groups that were similar in raw participants, a random sample of 483 students was extracted from the 12,956 students in Group 1. This chosen sample size of Group 1 would give us a simple comparison of groups that were both comprised of 483 individuals. For practical reasons, RCU staff believed an "apples to apples" comparison of equal group numbers, as seen in Table 1, would make it easier to quickly interpret our results for those local district decision makers and individual teachers who may not possess as robust of a statistical background as others.

Demographic comparison of groups
The 483 students who were randomly selected from the pool of nearly 13,000 students for group 1 was comprised of 74.6% of 10th and 11th graders. Group 1 students were ethnically made up of 49.9% Caucasian and 42.9% African American students. Group 1 had slightly more females (61.5%) in their group than males (39.5%).

Group 2 were the 483 students had who not taken any practice tests prior to taking their summative assessment. Much like group 1, group 2 was made up overwhelmingly of students

	Total	Practice tests taken
Group 1	12,956	P1 + P2
	483 (random sample)	P1 + P2
Group 2	14,576	Took P1 or P2
	483	None
Note(s): P1 represents first practice test attempt. P2 represents second practice test attempt		

Table 1.
Number of practice assessments delivered by group

who were in 10th or 11th grade (68.9%). Of the 483 students in group 2, 47% were Caucasian and 44.9% were African American. Group 2 also had most of its members be females (56.5%), while males comprised 42.7% of its numbers.

Comparison of summative exam scores by group
A statistical comparison of the mean summative exam MS-CPAS scores was conducted for both groups 1 and 2 using an independent samples *t*-test. The independent samples *t*-test showed a statistically significant effect between the groups, $t(964) = 8.35$, $p = 0.000$. The mean summative exam score for all students who took both online practice tests for the two years we examined (group 1) was 67.11 ± 14.497. The mean summative exam scores for students who did not take any practice tests was significantly lower (59.08 ± 15.386) than those students who took both. The difference in overall mean summative exam scores is shown in Table 2. A careful examination of the effect size for the group analysis was conducted and found to be $d = 0.53$. According to Cohen (1992), this effect size would be considered a medium size effect. Although that number does not reach the high effect threshold of 0.08, it seems to be significant from a practical point of view and clearly falls in line with what Wiliam and Black described as the overall effect size of the studies they examined, which averaged a 0.4.

Correlation comparison
Nearly 13,000 students out of more than 28,000 students whose scores the RCU examined took both available online practice assessments. These students ideally should have been the ones most likely to receive remediation between practice tests 1 and 2. Another question staff wondered was whether there would exist a correlation between these students' final online practice test score and their summative MS-CPAS scores that they would have taken following the last practice test. A Pearson product-moment correlation was conducted to determine the nature of the relationship existing between students' final practice test score and their MS-CPAS scores. There was an overall fairly strong correlation between students' final practice test scores and their MS-CPAS scores for students who took both practice tests ($r = 0.524$, $P < 0.05$). The relationship between final practice test scores and MS-CPAS scores is shown in Table 3.

Another lingering question formed as RCU staff examined overall correlation results for these nearly 13,000 students who took both online practice tests. In total, 13 career pathways had the option of taking the two practice tests. Staff began to question what the nature of the relationships between the final practice test scores and the MS-CPAS scores were within each

	Total		Practice tests taken
Group 1	12,956		P1 + P2
	483 (random sample)		P1 + P2
Group 2	14,576		Took P1 or P2
	483		None

	N	M	SD
Group 1	483	67.11	14.50
Group 2	483	59.08	15.39

Note(s): Group 1 took a practice assessment before the normal test administration. Group 2 did not take a practice assessment before the normal administration

Table 2.
Means and standard deviations for practice assessment groups

of the 13 pathways. Another Pearson product-moment correlation was used to highlight these relationships between the last online practice test and the MS-CPAS exam for each of the 13 pathways. A moderate, significant, positive correlation was found in nine pathways, and a strong, significant, positive correlation was found in three pathways. One pathway was determined to have a weak, significant, positive correlation. These results suggest that students who scored higher scores on their final practice test tended to score higher on the MS-CPAS than other students. Correlations between students' second practice test scores and final MS-CPAS scores can be seen in Table 4.

A positive mean difference between MS-CPAS scores and the final online practice test score was found in seven of the 13 career pathways in which there was a significant, positive correlation between the two assessments. These students scored higher on average on the MS-CPAS than on the final practice test. In six of the 13 career pathways, there was a negative mean difference between students' MS-CPAS scores and their final practice test scores in which there was a significant, positive correlation between the two assessments. These students scored lower on the MS-CPAS than they had on their final practice test.

The fact that six of the career pathways had a lower mean difference on their summative scores than those students made on their final formative assessment was surprising; however, it was not totally unexpected, given how loose parameters were for taking the practice tests. Teachers and leaders in schools were given complete autonomy to administer the online practice tests however they desired. It is certainly possible that some teachers administered the two separate practice tests in the few days leading up to the summative exam. This scenario would have likely eliminated the main benefit of a formative assessment, which is high-quality feedback given to the students. Regardless of the possible positive influence formative assessments can have on student learning and summative assessment outcomes, studies have shown that formative assessments that are implemented in the

Table 3.
Correlations between students' second practice test scores and final MS-CPAS scores (group 1 only)

	n	M	SD	R
MS-CPAS	12,956	66.89	14.911	
Practice test 2	12,956	65.74	22.556	0.524***
Note(s): ***$p < 0.001$				

Table 4.
Correlations between students' second practice test scores and final MS-CPAS scores by cluster

Cluster	r	n	Practice test M	SD	MS-CPAS M	SD
Agriculture, food and natural resources	0.481***	1,589	57.37	22.00	60.53	13.95
Architecture and construction	0.420***	109	75.78	15.60	71.91	11.98
Arts, A/V technology and communication	0.583***	184	67.07	22.77	67.21	16.22
Business management and administration	0.555***	1,204	65.43	24.02	64.56	14.02
Education	0.550***	463	64.87	18.66	75.42	12.21
Health science	0.381***	3,336	72.58	23.42	73.08	12.95
Hospitality and tourism	0.611***	1,137	66.27	20.07	70.45	14.69
Human services	0.446***	874	70.75	19.87	70.10	12.44
Law, public safety, corrections and security	0.434***	758	63.36	20.81	66.89	12.00
Manufacturing	0.628***	266	65.94	20.22	64.41	14.23
Marketing	0.639***	448	64.88	24.08	61.71	16.04
Science, technology, engineering and mathematics	0.573***	1,433	62.13	20.70	65.36	14.77
Transportation, distribution and logistics	0.555***	1,155	59.01	21.42	54.79	14.07
Note(s): $N = 12,956$; ***$p < 0.001$						

absence of appropriate feedback from teachers were less effective. Successfully implementing formative assessments in online environments can increase the immediacy and quality of personalized feedback to students (Bhagat and Spector, 2017). Moreover, Robertson and Steele (2019) found that undergraduate students that took online formative assessments using Web 2.0 tools paired with quality feedback were more prepared for summative assessments in their coursework than students that were assessed using traditional methodologies.

Conclusions, recommendations and implications

The catalyst for deciding to take a closer examination into this assessment data was to evaluate whether these practice tests could help better prepare Mississippi's career and technical students for their summative exams and increase student learning. Our primary hypothesis of whether students who took multiple online practice tests would outperform peers who did not take any practice assessments was confirmed by these results. In fact, students who took both online practice tests statistically outperformed their peers on the summative exam compared to their peers who did not take any formative tests.

Future research recommendations

Considering recent events surrounding the Covid-19 pandemic and the reliance it has placed on educators to use online tools for the delivery of instruction, there also needs to be more research conducted about using formative assessments in an online format. When this study was conducted, students took these assessments in secure locations on their school campuses. The uncertainty surrounding assessment issues during this pandemic makes it critical that educators react quickly to this shifting, new reality. We would suggest quickly identifying protocols in which educators can quickly produce online practice tests and ensure those tests can be considered valid and reliable when taken in a less secure testing environment such as a student's home.

Another consideration that needs to be addressed is better teacher training in utilizing various course management systems to help more quickly identify potential student engagement and mastery problems. All CTE courses in Mississippi have premade companion Canvas course management shells built for the classroom teachers to utilize. New teachers are especially encouraged to share lessons and course content using the prebuilt course shells. Perhaps, teachers could be better trained in the ways of predicting student success and challenges through these data interface reports. As both Estacio and Raga pointed out, the use of data from an electronic interface could help direct instructors into a targeted type of intervention or feedback for students to ensure online activity matches desired learning outcomes for classes (2017). If teachers better understood the online learning habits of their students, it could influence instructor behaviors and feedback patterns, which are necessary for increasing the effectiveness of formative assessments. Some studies have shown that not having enough feedback is one of the challenges specifically mentioned by students in some open and distance learning (ODL) environments as well (Au et al., 2018).

As the quality of feedback plays such an important role in formative assessments, we would also suggest that more work is needed to specifically understand what types of feedback work best to increase student understanding. Perhaps, a possible survey of feedback methodology done by teachers should be conducted to determine which types of feedback show the largest gains in student understanding. Districts were not required to administer either of the online practice tests, so the decision to administer any of the assessments was left entirely up to individual districts or teachers. Another possible area of interest for future study could be an examination of regions in the state who chose to utilize these formative assessments compared to those who did not have as high a participation

level. Finally, these practice tests are not released by the RCU to the districts until six weeks prior to when the summative tests are available. This limited amount of time to test, remediate and test again may need to be further explored as to its impact on summative exam results (Darling-Hammond and Rustique-Forrester, 2005; Shepard *et al.*, 2018; What Works Clearinghouse (ED), 2012).

Relationship between final practice test and summative test scores

This observational study also queried whether group 1 students' final practice tests would serve as a valid predictor of similar results on their end-of-course MS-CPAS exams when data were disaggregated by career pathway. The results of this second question was also confirmed by this study. We examined the relationship between these students' final practice test scores and their summative exam score by career pathways. These observations confirmed a significant, positive relationship between students' final practice test score and their MS-CPAS scores. These results coincide with numerous other studies that demonstrate formative assessments can serve as valid predicators of student summative exam performance as well (Harlen and James, 1997; Zhang and Henderson, 2015).

There is a public perception that the quality of academic instruction in CTE centers and high schools in the USA needs to be substantially improved. To meet this challenge, the federal government has provided more guidance and flexibility for states around statewide testing and measures of program quality in CTE because of the Every Student Succeeds Act (ESSA) and Perkins V legislation (Imperatore, 2020; Perry, 2019). However, the findings of this study have revealed that the use of formative assessments in the appropriate manner, coupled with remediation strategies, can assuage many concerns about the use of tests for CTE program quality purposes if a mastery learning formative assessment and adaptive instruction model are adopted. It echoes the recommendations from past formative assessment research as it relates to large-scale assessment systems such as the MS-CPAS used with Perkins IV technical skills attainment measures (Cizek and Burg, 2006; Sadler, 1998; Shepard *et al.*, 2018; Zimmerman and Dibenedetto, 2008).

This is an important finding for CTE teachers, directors and other support staff, such as student service coordinators or school counselors, who are attempting to improve program instruction under the framework of a statewide testing environment. This study's findings provide educators with (1) guidance on the importance of formative assessments in the CTE classroom to assist in meeting student performance objectives and (2) increased understanding of the relationship between formative assessment frequency and summative assessment scores, returning the focus back to classroom instruction to help all students to achieve mastery. This type of instructional model not only proved to be highly effective, but also enhanced CTE students' learning. Demonstrating mastery in CTE career pathways for high-school students increases the prospects of American high-school students by preparing them to succeed in post-secondary education and to enter the workforce successfully.

References

Almasi, E. and Tabrizi, A.R.N. (2016), "The effects of direct vs. indirect corrective feedback on Iranian EFL learners' writing accuracy", *Journal of Applied Linguistics and Language Research*, Vol. 3 No. 1, pp. 74-85.

Au, O.T.S., Li, K. and Wong, T.M. (2018), "Student persistence in open and distance learning success factors and challenges", *Asian Association of Open Universities Journal*, Vol. 13 No. 2, pp. 191-202.

Bell, B. and Cowie, B. (2001), "The characteristics of formative assessment in science education", *Science Education*, Vol. 85 No. 5, pp. 536-553, doi: 10.1002/sce.1022, available at: https://onlinelibrary.wiley.com.

Bhagat, K.K. and Spector, J.M. (2017), "Formative assessment in complex problem-solving domains: the emerging role of assessment technologies", *Educational Technology and Society*, Vol. 20 No. 4, pp. 312-317, available at: https://www.jstor.org/stable/26229226.

Black, P. and Wiliam, D. (1998), "Assessment and classroom learning. Assessment in education: Principles", *Policy and Practice*, Vol. 5 No. 1, pp. 7-74, doi: 10.1080/0969595980050102.

Black, P. and Wiliam, D. (2009), "Developing the theory of formative assessment", in *Educational Assessment, Evaluation and Accountability*, Vol. 21 No. 1, pp. 5-31, doi: 10.1007/s11092-008-9068-5.

Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003), *Assessment for Learning: Putting it into Practice*, available at: http://www.mcgraw-hill.co.uk/html/0335212972.html.

Booher-Jennings, J. (2005), "Below the bubble: 'Educational triage' and the Texas accountability system", *American Educational Research Journal*, Vol. 42 No. 2, pp. 231-268, available at: http://www.aera.net/publications/?id=315.

Butler, R. and Nisan, M. (1986), "Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance", *Journal of Educational Psychology*, Vol. 78 No. 3, pp. 210-216, doi: 10.1037/0022-0663.78.3.210.

Cech, S.J. (2008), "Test industry split over 'formative' assessment", *Education Week*, Vol. 28 No. 4, pp. 1-17, available at: http://www.edweek.org/ew/toc/2008/09/17/index.html.

Cepni, S.B. (2016), "A replication study: oral corrective feedback on L2 writing; two approaches compared", *Procedia Social and Behavioral Sciences*, Vol. 232, pp. 520-528.

Cizek, G.J. and Burg, S.S. (2006), *Addressing Test Anxiety in a High-Stakes Environment: Strategies for Classroom and Schools*, Corwin Press, available at: https://psycnet.apa.org/record/2005-11324-000.

Cohen, J. (1992), "A power primer", *Psychological Bulletin*, Vol. 112 No. 1, pp. 155-159, doi: 10.1037/0033-2909.112.1.155.

Crawford, A., Zucker, T., Van Horne, B. and Landry, S. (2017), "Integrating professional development content and formative assessment with the coaching process: the Texas school ready model", *Theory into Practice*, Vol. 56 No. 1, pp. 56-65, doi: 10.1080/00405841.2016.1241945.

Darling-Hammond, L. and Rustique-Forrester, E. (2005), "The consequences of student testing for teaching and teacher quality", *Yearbook of the National Society for the Study of Education*, Vol. 104 No. 2, pp. 289-319, doi: 10.1111/j.1744-7984.2005.00034.x.

Dixson, D. and Worrell, F. (2016), "Formative and summative assessment in the classroom", *Theory Into Practice*, Vol. 55 No. 2, pp. 153-159, doi: 10.1080/00405841.2016.1148989.

Dolin, J., Black, P., Harlen, W. and Tiberghien, A. (2018), "Exploring relations between formative and summative assessment", *Transforming Assessment*, Springer, Cham, pp. 53-80.

Ecclestone, K. (2007), "Lost and found in transition: the implications of 'identity', 'agency' and 'structure' for educational goals and practices", *4th CRLL International Conference: The Times They Are A-Changing: Researching Transitions in Lifelong Learning*, Vol. 15.

Elmore, R.F. (2002), "Hard questions about practice", *Educational Leadership*, Vol. 59 No. 8, pp. 22-25, available at: https://eric.ed.gov/?id=EJ644976.

Eskil, M., Özgan, H. and Balkar, B. (2010), "Students' opinions on using classroom technology in science and technology lessons–a case study for Turkey (Kilis city)", *Turkish Online Journal of Educational Technology*, Vol. 9 No. 1, pp. 165-175.

Estacio, R.R. and Raga, R.C. Jr (2017), "Analyzing students online learning behavior in blended courses using Moodle", *Asian Association of Open Universities Journal*.

Guskey, T.R. (1987), "The essential elements of mastery learning", *Journal of Classroom Interaction*, Vol. 22 No. 2, pp. 19-22, available at: https://www.jstor.org/stable/23869735.

Guskey, T.R. (2010), "Lessons of mastery learning", *Educational Leadership*, Vol. 68 No. 2, pp. 52-57, available at: http://www.ascd.org/publications/educational-leadership/oct10/vol68/num02/abstract.aspx.

Harlen, W. and James, M. (1997), "Assessment and learning: differences and relationships between formative and summative assessment", *Assessment in Education: Principles, Policy & Practice*, Vol. 4 No. 3, p. 365, available at: http://10.0.4.56/0969594970040304.

Hattie, J. and Timperley, H. (2007), "The power of feedback", *Review of Educational Research*, Vol. 77 No. 1, pp. 81-112, doi: 10.3102/003465430298487.

Hunt, J.W. (2008), "A nation at risk and no child left behind: Déjà vu for administrators?", *Phi Delta Kappan*, Vol. 89 No. 8, pp. 580-585.

Imperatore, C. (2020), "Perkins V and high-quality CTE", *Techniques: Connecting Education and Careers*, Vol. 95 No. 2, pp. 12-13.

Imperatore, C. and Hyslop, A. (2017), "CTE policy past, present, and future: driving forces behind the evolution of federal priorities", *Peabody Journal of Education*, Vol. 92 No. 2, pp. 275-289.

King, J.B. (2016), *Secretary Letter to CSSO to Testing Action Plan*, United States Department of Education, p. 4, available at: https://www2.ed.gov/admins/lead/account/saa/16-0002signedcsso222016ltr.pdf.

Klein, S.P., Kuh, G., Chun, M., Hamilton, L. and Shavelson, R. (2005), "An approach to measuring cognitive outcomes across higher education institutions", *Research in Higher Education*, Vol. 46 No. 3, pp. 251-276, doi: 10.1007/s11162-004-1640-3.

McGuire, F. (1994), "Army alpha and beta tests of intelligence", *Encyclopedia of Intelligence*, Vol. 1, pp. 125-129.

Nahadi, N., Firman, H. and Farina, J. (2015), "Effect of feedback in formative assessment in the student learning activities on chemical course to the formation of habits of mind", *Jurnal Pendidikan IPA Indonesia*, Vol. 4 No. 1, pp. 36-42.

OECD (2011), *Education at a Glance: OECD Indicators*, OECD Publishing. doi: 10.1787/eag-2011-en.

Perry, A. (2019), "Making the most of Perkins V", *State Education Standard*, Vol. 19 No. 3, pp. 15-17, available at: https://eric.ed.gov/?id=EJ1229625.

Robertson, S.N. and Steele, J.P. (2019), "Using technology tools for formative assessments", *Journal of Educators Online*, Vol. 16 No. 2, available at: http://www.thejeo.com.

Sadler, D.R. (1989), "Formative assessment and the design of instructional systems", *Instructional Science*, Vol. 18, pp. 119-144, doi: 10.1007/BF00117714.

Sadler, D.R. (1998), "Formative assessment: revisiting the territory", *Assessment in Education: Principles, Policy and Practice*, Vol. 5 No. 1, pp. 77-84, doi: 10.1080/0969595980050104.

Schildkamp, K. (2019), "Data-based decision-making for school improvement: research insights and gaps", *Educational Research*, Vol. 61 No. 3, pp. 257-273, doi: 10.1080/00131881.2019.1625716.

Scriven, M. (1966), *The Methodology of Evaluation. Social Science Education Consortium*, Publication 110, p. 61, available at: https://eric.ed.gov/?id=ED014001.

Shepard, L.A., Penuel, W.R. and Pellegrino, J.W. (2018), "Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment", *Educational Measurement: Issues and Practice*, Vol. 37 No. 1, pp. 21-34, doi: 10.1111/emip.12189, available at: https://onlinelibrary.wiley.com.

Smith, E. and Gorard, S. (2005), "'They don't give us our marks': the role of formative feedback in student progress", *Assessment in Education Principles Policy and Practice*, Vol. 12 No. 1, pp. 21-38, doi: 10.1080/0969594042000333896.

Stiggins, R. (2018), "Better assessments require better assessment literacy", *Educational Leadership*, Vol. 75 No. 5, pp. 18-19, available at: https://eric.ed.gov/?id=EJ1170073.

Stiggins, R. and DuFour, R. (2009), "Maximizing the power of formative assessments", *Phi Delta Kappan*, Vol. 90 No. 9, pp. 640-644, doi: 10.1177/003172170909000907.

The Mississippi Department of Education (2020), *Traditional Diploma with Endorsements*, available at: https://www.mdek12.org/ESE/diploma#.

Tunstall, P. and Gipps, C. (1996), "Teacher feedback to young children in formative assessment: a typology", *British Educational Research Journal*, Vol. 22 No. 4, p. 389, doi: 10.1080/0141192960220402.

von der Embse, N.P., Pendergast, L.L., Segool, N., Saeki, E. and Ryan, S. (2016), "The influence of test-based accountability policies on school climate and teacher stress across four states", *Teaching and Teacher Education*, Vol. 59, pp. 492-502, doi: 10.1016/j.tate.2016.07.013.

What Works Clearinghouse (ED) (2012), *What Works Clearinghouse Quick Review: "An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP)*, available at: https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED530900.

Wong, B.T.M. (2016), "Factors leading to effective teaching of MOOCs", *Asian Association of Open Universities Journal*.

Young, V.M. (2006), "Teachers' use of data: loose coupling, agenda setting, and team norms", *American Journal of Education*, Vol. 112 No. 4, pp. 521-548, doi: 10.1086/505058.

Zhang, N. and Henderson, C.N. (2015), "Can formative quizzes predict or improve summative exam performance?", *Journal of Chiropractic Education*, Vol. 29 No. 1, pp. 16-21.

Zimmerman, B.J. and Dibenedetto, M.K. (2008), "Mastery learning and assessment: implications for students and teachers in an era of high-stakes testing", *Psychology in the Schools*, Vol. 45 No. 3, pp. 206-216, doi: 10.1002/pits.20291.

Corresponding author
Ben Alexander can be contacted at: ben.alexander@rcu.msstate.edu