ORIGINAL PAPER

Leveraging the Power of Common Formative Assessments in Mathematics

Authors: Joshua Griffin, Brittany Hurley, Drew Polly

Abstract: Mathematics learning on the North Carolina End of Grade tests and on national assessments continues to show opportunities for growth in the middle grades (Grades 4-8). This article describes how a rural school district in North Carolina created and used common formative assessments in mathematics to strengthen planning, teaching, and learning. The article provides individuals with free, unlimited access to these assessments, as well as guidance on how they have been used most effectively.

Keywords: *assessment, formative assessment, mathematics, test development*

Background

The Benefit of Common Assessments

Have you ever had your car break down? When your car breaks down, you don't just change a spark plug and hope that fixes the issue. Metaphorically, the same goes with teaching and assessing students. When a student appears to not understand a concept, looks puzzled, or does not provide an appropriate strategy and answer, it is important to understand exactly what their strengths are, where and what their misconceptions are, so you can specifically address what the student doesn't understand. Teachers must be able to act like mechanics to diagnose a student's knowledge base. That is where the power of common formative assessments (CFAs) is.

Common assessments, both formative and summative, are important for multiple reasons (DuFour & DuFour, 2013; Schoenfeld, 2015). Common assessments ensure that students are held to the same expectations and learning goals no matter what school or teacher they have (Schoenfeld, 2015). Since students all learn the same standards, it makes sense that the ways that they are assessed are similar as well. If students are assessed on the exact same items, student results can be meaningfully compared. As a result, shared strengths and misconceptions can be identified and discussed collaboratively, allowing educators to plan targeted additional learning opportunities. If a common assessment is designed correctly, the data can paint a picture of how a student thinks mathematically, which will give teachers the ability to fine-tune a student's mathematical understanding.

When all students are assessed using a common set of items, comparisons extend beyond just the results. Teaching practices that were used to gain certain measures can be discussed and shared, which act as a starting point for teachers to tailor to their own students' needs. While obviously there are differences between populations, even within the same school, meeting to talk about data allows teachers to share experiences of what worked and did not work for their students and is a great source of teachers' professional learning. Any time a collaborative professional practice can be

combined with a teacher's growth mindset, then the students benefit.

The Benefit of CFAs

Using Data from CFAs

In our work, we focus on common formative assessments (CFAs) since we were interested in assessments that allowed teachers to respond to the data to plan and provide additional learning opportunities for students. Unlike the End of Grade and/or End of Course (EoG/EoC) tests at the end of the year, where responding with additional learning opportunities is not feasible, CFAs provide opportunities for teachers and school leaders to analyze and plan what will happen as a result of the data (Maskos et al., 2025).

In mathematics, researchers have found that the use of CFAs and responding to the data has led to long-term increases in student learning (Polly et al., 2017; Veldhuis & Van den Heuvel-Panhuizen, 2020). Teachers, when they have access to CFA data that is aligned to Standards, have been able to successfully look at data and determine how to best meet their students' needs (Nortvedt & Buchholtz, 2018; Polly et al., 2017). Beesley and colleagues (2018) found that middle school teachers, when given valid assessments and support with the analysis of CFA data, were able to appropriately plan and teach additional learning opportunities regardless of their knowledge of mathematics teaching strategies.

Using CFAs As a Planning Resource before Teaching

In addition to using CFAs to provide data about their students, teachers have also benefited from looking at the problems on these CFAs ahead of time as a planning resource. The use of items on a CFA to support unit planning helps to ensure that the initial teaching of concepts aligns to Standards and the rigor that is included on these CFAs (DuFour & DuFour, 2013). This consistent alignment between planning, teaching, and assessment around the Standards helps to increase the likelihood that students will learn and retain important concepts that will lead to their academic success (Martin et al., 2016; DuFour & DuFour, 2013).

In mathematics in North Carolina, many Standards include multiple concepts that overlap and intersect with each other. As a result, quality assessment items provide teachers and educational leaders with the entire breadth, depth, and rigor of how the Standards could be assessed. While commercially published curricula and instructional resources can help teachers teach content, many of these do not have direct alignment to the Standards and how the Standards are assessed. By looking at quality assessment items while planning instruction, teachers are able to keep the end in mind and determine how to support their students' learning en route to being able to succeed on the assessment (Wiggins & McTighe, 2005). In the next section of this article, we describe our journey designing, using, and revising CFAs and how it has positively impacted teachers' planning and their teaching, and most importantly students' learning.

Our Journey with CFAs

Context

Josh Griffin is the Director of Testing and Accountability and the Secondary Mathematics Specialist in Franklin County, NC while Drew Polly is a university-based faculty member who supports mathematics teaching and learning in multiple districts and schools around the state. The school district where Griffin works is a rural district approximately 30 minutes outside of the

Raleigh-Durham Research Triangle area. The district includes 16 schools (8 elementary, 4 middle, and 4 high) and about 8,000 students. As a district, approximately 48% of the students have been identified as economically disadvantaged. A majority of these students are concentrated in specific attendance zones. In terms of demographics, the district's student population is 35.8% White, 30.9% Black or African-American, 25.9% Hispanic/Latinx, and the rest identify with two or more races.

As a small rural district, we were limited in the amount of funding and resources available to teachers, schools, and the district office. Since school populations are small in each school, teachers have 1 or 2 teachers per subject in middle and high school and 3 or fewer teachers per subject in elementary school. These logistics limit the amount of in-school collaborations that are possible, which led us to think holistically about whole district efforts where teachers and district leaders could collaborate and grow together to support teaching and learning. The focus on mathematics CFAs was purposeful since that was the content expertise of the authors.

Process of Creating CFAs

The process of creating the CFAs included two main areas: a) ensuring alignment to the Standards, and b) purposeful format of the items.

Alignment to the Standards

In many districts, it is found that quality teaching is not always the issue; sometimes it is the lack of focus or understanding of the content standards. Some of these high-quality teachers hold the misconception that questions aligned to certain standards can only be asked in a limited number of ways, when that is not the case. Due to this, the need to have questions that display the breadth and depth of the standards, and also show the range of rigor, was displayed.

Districts using data and common assessments is not a new idea and has been studied in the past (Blanc et al., 2010; Datnow & Hubbard, 2015; Datnow & Park, 2014). However, we knew anecdotally that we needed a process that focused on Standards and the rigor of how the Standards are assessed in the development of the items on the assessments. Using the same protocol that state test makers use to create interim assessments seemed to be the obvious path for creation. To create the state-made interim assessments, the entire item pool is explored, and items are chosen that are uniquely different to show the variety of the standard. Since we in the district do not have access to the state's item pool, we simply looked at all of the interim questions, state EOG released questions, then filled any gaps (if any) with the unpacking document to ensure the breadth and depth of the standard was reflected. The questions from interim assessments and released tests were cloned, which means the numbers, context, and names were changed, but the skills needed to solve the questions remained the same in order to keep what is important and also not show students questions they may see later in the course.

Purposeful format of the items

While the NC End of Grade and End of Course (EOG/EOC) tests included both multiple choice and constructed response tests, we intentionally decided to limit the use of multiple choice items. The rationale behind this choice was to ensure that students showed their work and provided evidence of the strategies that they used instead of allowing them to just select an option on a multiple-choice item.

Another intentional decision that was made was to provide multi-part scaffolded items. As you can see in Figure 1, in the

example on the left, students must find the linear best-fit model in Part 1 before using it to find the answers to Parts 2 and 3. In the image on the right, sixth-grade students are asked to find the area of the different parts of the shape before finding the area of the entire shape.

Figure 1 *Samples of Multi-part, Scaffolded Items*

Sample from Math 1

8. (4 pts) The table below displays the walking heart rate and running heart rate of eight girls in beats per minute (bpm).

Walking Heart Rate	Running Heart Rate
66	128
72	136
74	134
78	138
80	142
84	146
86	148
88	152

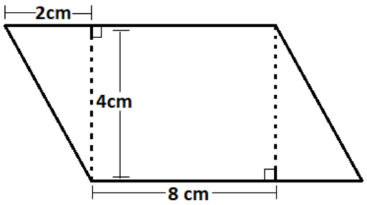
Part A. Determine the linear best-fit model for the data. (*Round your answer to the hundredths place.*)

Part B. Using the linear best-fit model for the data, what is the predicted running heart rate of a girl whose walking heart rate is 100 bpm? (*Round your answer to the hundredths place.*)

Part C. Using the linear best-fit model for the data, what is the predicted walking heart rate of a girl whose running heart rate is 130 bpm? (*Round your answer to the hundredths place.*)

Sample from Sixth Grade

7. **(9 pts) A tile is shaped like a parallelogram.**



I. **(3 pts)** What is the area of **one** of the small triangles? (*HINT: both triangles have the same area*)

II. **(3 pts)** What is the area of the center rectangle?

III. **(3 pts)**What is the area of the entire parallelogram?

When students have to answer multi-part, scaffolded items, teachers can see what part of the Standard students can do successfully and where there are opportunities for growth. Teachers have repeatedly commented about students' specific misconceptions by looking at their work on these types of items.

Additionally, while teachers typically enjoy the convenience of having students take online assessments since they can immediately see tables and graphs of students' data, we strongly encourage teachers to give these assessments on paper without an online option. This decision has not been made lightly as the EoG/EoC and NC Check-ins are 100% online assessments. However, we have found that students are more apt to show their work and strategies if they are required to write things on paper. For those teachers who do give it to students online, they require students to turn in their work on notebook paper to provide evidence of their strategies.

Providing Collaborative Opportunities Around Data

One of the most powerful pieces of common assessments is bringing teachers who don't normally collaborate with one another together to talk about their data and teaching practices. Just simply giving students assessments and collecting data does not improve students' data. The benefit of CFAs is the analysis of data and planning additional learning opportunities for students based on the data. Studies have shown that teachers benefit when they are supported during the analysis of data and planning of future learning opportunities (Polly et al., 2017; Beesley et al., 2018).

As a result, throughout our work with CFAs, we have been committed to holding regular meetings over Google Meet for teachers to collaborate with one another to discuss data, teaching strategies that they used, and plan for additional learning opportunities. These meetings also provided opportunities for teachers to provide us with feedback on the items so that our process of continually refining them included teacher input. In the next section, we share the journey around the Math 1 assessments, which is a course that is taken in our district by both eighth graders and freshmen in high school.

The Beginning of the Process

Creation. The first common assessments created were done with NC Math 1, taken by both eighth graders and high school students. Historically, high school scores across the state are low due to the elimination of double testing and the "advanced" students taking Math1 in middle school. Although our Math1 teachers are veterans and high-quality teachers, our proficiency rate for high school prior to common assessments was slightly below the state average. Through our work with high school teachers, we knew we had a strong staff; however, beyond pacing, there were significant differences in what was taught, the depth of the standards addressed, and how students were assessed. Assessments built at the school were very inconsistent both in content and in rigor, and oddly enough, it was not the same school that had the "weaker" or "easier" test. It varied from unit to unit. We at the district level were not able to bring teachers together to talk about common issues or how we teach specific skills effectively, so we knew something needed to be done.

The summer before our common assessment implementation, a team was created to dive into interim assessments (NC Check-Ins), released EOCs, and the NCDPI unpacking document to create a breadth and depth of each individual sub-standard. Once created, items were cloned/written in order to accurately represent each sub-standard

and made into a unit assessment based on the district's pacing. Rubrics and answer keys were also created to support teachers' work.

Year 1. The first year of common assessment focused on honing the assessment. Each assessment served as a working document and was adjusted depending on multiple factors. The two biggest factors included whether: 1) the test assessed the skills that we want to assess, and 2) the way that the questions were being asked was ambiguous or confusing to students. If we found any issues, we revised the assessment for the following year. We also made notes of common student struggles and then talked about it as a group on how to provide additional learning opportunities on certain skills with students who have not demonstrated proficiency. Each of these meetings lasted 45-60 minutes, and we had them after each unit assessment was given, which was approximately every 4 to 6 weeks.

To be fully transparent, the first year for each grade was always the biggest struggle we faced at the district level, since we were establishing buy-in from teachers. As a district, we had to stress that these common assessments were for the betterment not only of our students, but also our teaching practice. The accountability piece was designed for the teacher to internally assess what and how they were teaching in hopes of fostering a growth mindset and becoming a reflective practitioner. Luckily, we had buy-in from a majority of our teachers. One in particular, who is a fantastic educator, happened to have students who did not score as high as her peers' students. This teacher had the pedagogical content knowledge to be successful, but needed to be shown the breadth and depth of the standards. After the first year of implementation, this teacher's scores rose over 30%, and they became one of the highest in our district.

Year 2. Year two, our focus shifted from being reactionary to being proactive. Based on the previous year's results, we knew how we taught a certain skill and the results that came from that teaching practice. We were able to bring teachers together to talk about what they found to be successful and also collaborate to find solutions to where they were not successful. The success of the previous year solved the buy-in problem, and now we were working as an entire district instead of individual school PLTs. Each meeting we would spend about 10-15 minutes talking about results, discussing if any items needed to be modified, and next steps on student deficiencies, but then shifted to focus for the next 45-50 minutes on what was coming up next and how we would modify our current teaching practice to anticipate student struggles.

Year 2 also brought about our first attempt in grades 6 through 8. We decided to build all three grade levels at once and ended up being at least a few units ahead of the teachers. The first year of middle school looked very similar to the first year of NC Math 1. Our focus was making the assessments as good as they could be, and grade-level teachers and district staff spent a lot of time looking back as we discussed next steps to get our students where they needed to be.

The second year of middle school, however, looked very different from what the second year of Math 1 did. Middle schools across the state, especially in smaller rural districts, have found that a lot of their teacher candidates do not go through the typical higher education/pre-service teacher path. A majority of middle school teachers are getting their teaching licenses through a more non-traditional path, such as lateral entry or a residency license, and we

have a very transient population of middle school teachers. Due to many teachers having a lack of training and experience, our meetings in year 2 for middle school had a different format. During the meetings, we were still looking ahead, but the meetings were more district-led. The group talked about student struggles from the previous year, then resources were given and discussed in depth to help support the teachers' pedagogical content knowledge. The purpose of the meetings was exactly the same; however, what the district provided was more hands-on to help support newer teachers.

Our Latest Work: Including Upper Elementary Grades

During the summer of 2024, we decided to take the work of middle and high school mathematics CFAs and scale the work to both fourth and fifth grades for the 2024-2025 school year. Although the format of our assessments is slightly different due to standards-based grading, our processes have been almost the same. Similar to our experience at the secondary level, we have been met with positive teacher feedback and success with teachers being more knowledgeable of their Standards and the level of rigor that is expected for students.

One new aspect of our project that has emerged through the work in fourth and fifth grade has been considering the most essential questions for teachers to consider before and during data meetings. This topic came up due to teachers' busy schedules, and our goal is for teachers to look at their data prior to our meetings, but in ways that provide momentum to use data to inform planning for their students' future learning opportunities.

The addition of upper elementary grades has also led us to embrace teachers' flexibility in the administration and use of this. This flexibility has led to an increase in teachers' buy-in and interest in using them, while also leading to variance and differences in their uses. Specifically, some teachers in upper elementary school grades use these CFAs formally, like the NC Check-ins and End-of-Grade tests; they provide read-aloud options for students who qualify for that accommodation, and they require independent work. Meanwhile, other teachers have used the CFAs in modified formats. Some teachers use the items for group work for students to work on during their math period, while some use them individually for exit tickets where they complete the CFA across a few days. These variations in how the CFAs are given are also the case with middle school teachers in order to provide flexibility. With these differences in CFA use, we have provided teachers with autonomy on how to use these quality items with the caveat that students' data on group work likely looks different than their data on independent work done in a more formal assessment context.

Recommendations and Looking Forward

Through this work with common assessments, teachers' instruction has improved, and students' data has benefited. As we look ahead, here are a few things that we feel are valuable considerations. These include: a) maintaining a collaborative culture, b) selecting rigorous additional learning opportunities, and c) reflecting on what specifics our students need.

The work involving the use of CFAs has included using the data from CFAs to better understand students' strengths and opportunities for growth and plan additional learning opportunities. The collaborative nature of the conversations between the teachers and us about the data and how to respond to it has been powerful for everyone involved. Teachers have become open and transparent about what has led to increasing

students' data and where students still demonstrate misconceptions. These conversations have carried over to the focus in Year 2 of using the CFAs to anticipate misconceptions and use the items to help inform planning of units. We have found that having teachers who are collaborative and embrace a culture of discussion and sharing to be essential in this work.

Another takeaway that we have learned is the need to select rigorous additional learning opportunities. Upon looking at the data the first time, teachers are often disappointed in their students' overall data on these assessments. Some of them report the need to tell their students that these are challenging problems and that they just want students to show their work and try their best. As a result of these rigorous items, a lot of teachers divert to plan and use easier activities and tasks in follow-up lessons. We have learned from our conversations with teachers in data and planning meetings that teachers just need a bit of guidance and opportunities to talk about what makes items rigorous and to ensure that those aspects of items are included in future activities.

Lastly, this CFA work has led teachers to be more student-focused than ever before, with a focus on reflection. Teachers in the data meetings often make comments about how they are pleased that they focused on specific concepts and types of mathematics problems while they were teaching. They also have referred to their students' process of learning to push through the struggle of rigorous tasks. One fifth grade teacher shared via email, "I think the more of these assessments they have taken, the better they are getting at taking them…they have learned to productively struggle through these questions."

Teachers have become more reflective through the CFA process about their planning, their teaching, and their students' data. As indicated previously, it is this process of teachers reflecting on the CFA data and their teaching that improves their teaching and, in turn, increases their students' data and signs of their learning.

References

Beesley, A. D., Clark, T. F., Dempsey, K., & Tweed, A. (2018). Enhancing formative assessment practice and encouraging middle school mathematics engagement and persistence. *School Science and Mathematics*, *118*(1-2), 4-16. https://doi.org/10.1111/ssm.12255

Blanc, S., Christman, J. B., Liu, R., Mitchell, C., Travers, E., & Bulkley, K. E. (2010). Learning to learn from data: Benchmarks and instructional communities. *Peabody Journal of Education*, *85*(2), 205-225. https://doi.org/10.1080/01619561003685379

Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record*, *117*(4), 1-26. https://doi.org/10.1177/016146811511700408

Datnow, A., & Park, V. (2014). Data-driven leadership (Vol. 12). John Wiley & Sons.

DuFour, R., & DuFour, R. (2013). *Learning by doing: A handbook for professional learning communities at work*. Solution Tree Press.

Martin, C. S., Polly, D., Wang, C., Lambert, R. G., & Pugalee, D. K. (2016). Perspectives and practices of elementary teachers using an internet-based formative assessment tool: The case of Assessing Mathematics Concepts. *International*

Journal for Technology in Mathematics Education, *23*(1), 3-13.

Maskos, K., Schulz, A., Oeksuez, S. S., & Rakoczy, K. (2025). Formative assessment in mathematics education: A systematic review. *ZDM–Mathematics Education*, *57*, 679-693. https://doi.org/10.1007/s11858-025-01696-x

Nortvedt, G. A., & Buchholtz, N. (2018). Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity. *ZDM*, *50*(4), 555-570. https://doi.org/10.1007/s11858-018-0963-z

Polly, D., Wang, C., Martin, C. S., Lambert, R. G., Pugalee, D. K., & Middleton, C. W. (2017). The influence of an internet-based formative assessment tool on primary grades students' number sense achievement. *School Science and Mathematics*, *117*(3-4), 127-136. https://doi.org/10.111/ssm.12214

Schoenfeld, A. H. (2015). Summative and formative assessments in mathematics supporting the goals of the Common Core standards. *Theory Into Practice*, *54*(3), 183-194. https://doi.org/10.1080/00405841.2015.1044346

Veldhuis, M., & Van den Heuvel-Panhuizen, M. (2020). Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Mathematics Education Research Journal*, *32*(3), 449-471. https://doi.org/10.1007/s13394-019-00270-5

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. ASCD.

Joshua Griffin is the Director of Testing and Accountability for Franklin County Schools in North Carolina. He has served as an educational testing and accountability consultant for the North Carolina Department of Public Instruction and as a high school mathematics teacher in Wake County Schools.

Brittany Hurley is the secondary mathematics specialist for Franklin County Schools in North Carolina. She has also served as a district-level middle school mathematics coach and a high school mathematics teacher.

Drew Polly is a professor in the elementary education program at UNC Charlotte. He spends time supporting mathematics teaching and learning by sharing free resources (tinyurl.com/nck5math) and offering support to local schools.