INFLUENCE OF CONTEXT ON TEACHERS' ASSESSMENT PRACTICES

Kamie K. Stack
University of Minnesota
stack180@umn.edu

Erin E. Baldinger
University of Minnesota
eebaldinger@umn.edu

The process of assessing students is a fundamental part of teaching and learning mathematics. The assessment practices a teacher chooses are shaped by their values while also being shaped by the context of the school, district, state, and country where the teaching takes place. This can result in gaps between teachers' values and practices. In this study, we use student work sample interviews with five secondary mathematics teachers to illustrate their values around assessment, the factors that influence their assessment practices, and how their agency influences their assessment decisions. We focus on the important role contextual factors can play in shaping teachers' agency and assessment choices. These findings have implications for teacher education and further research around how assessments are used.

Keywords: Assessment, teacher beliefs

The process of assessing students is a fundamental part of teaching and learning mathematics. The assessment practices a teacher chooses are shaped by their values while also being shaped by the context of the school, district, state, and country where the teaching takes place. Enacting high-quality assessment practices has been identified as a key factor in preparing future mathematics teachers (Association of Mathematics Teacher Educators, 2017). Prior research has highlighted teachers' assessment practices and their values related to assessment within and outside of mathematics education (e.g., Barnes et al., 2014; Beswick, 2011; Brown, 2004; Davis & Neitzel, 2011; Remesal, 2007). Just as there may be gaps between mathematics teachers' values and instructional practices more generally (Beswick, 2011), such gaps may extend to assessment practices in particular. Despite this, teachers' assessment practices are used as factors in measuring teaching effectiveness (e.g., Sato, 2014; Sato et al., 2008).

Given these conditions, researchers in the Network for Excellence in Teaching (NExT) sought to investigate multiple measures of teaching effectiveness (NExT Teacher Effectiveness Work Group, 2018). One such measure built on D'Souza's (2012) Teacher Assessment/Pupil Learning Protocol (TAPL) which was used to examine assessment practices and how they evolve with early career high school English teachers over a period of five years. Using D'Souza's TAPL, NExT created a student work sample interview protocol to measure teacher effectiveness across disciplines. As part of our participation in NExT, we implemented the student work sample interview with recent graduates of the secondary mathematics licensure program in a Midwest University.

In a student work sample interview, teachers select student work from a recent assessment they implemented and then discuss the pieces of student work with an interviewer. Through piloting this interview protocol, we gathered important data for the NExT consortium and the teacher preparation program. Beyond programmatic evaluation, we also noticed interesting patterns related to teachers' values and assessment practices. We therefore engaged in a process of open coding, drawing on the principles of grounded theory (Charmaz, 2006). We arrived at two research questions: (1) What factors influence how secondary mathematics teachers enact assessment practices? (2) How does their enactment align with their values about assessment?

Literature Review

In this section, we highlight relevant literature related to assessment and teacher agency. These two constructs are key to understanding the themes that emerged from our interviews.

Assessments Serve Multiple Purposes for Multiple Audiences

Assessments generate artifacts that serve multiple purposes for multiple audiences (Davis & Neitzel, 2011; Remesal, 2007). In addition to students and teachers, audiences for assessment artifacts include parents, caregivers and what Davis and Neitzel (2011) call the "higher-ups"—people such as school administrators, district leaders, and policy makers. With multiple and potentially conflicting audiences come multiple conceptions of the purposes of assessment. Teachers' conceptions exist along a continuum from a focus on the pedagogical purposes of assessment to a focus on the accountability purposes of assessment (Barnes et al., 2014).

At the pedagogy end of the continuum, students and teachers are the primary audiences for assessments. Pedagogical purposes of assessment include informing instruction, facilitating learning, and providing evidence of student understanding. For students, assessments align with the content in the curriculum (Davis & Neitzel, 2011), demonstrate their progress toward learning goals (Karp & Woods, 2008), and offer an opportunity to receive feedback. On the alternative extreme of the continuum, the "higher-ups" are the driving audience. Here, assessments are used as measures for teacher and school accountability (Brown, 2004). As an example, teachers may choose to use classroom assessments as intentional preparation for high stakes, standardized assessments to adhere to expectations from the "higher-ups" audience and their focus on accountability (Davis & Neitzel, 2011). In this example, the audience of the "higher-ups" and their beliefs about the purpose of assessment influence classroom assessment practices. The accounting role of assessment is punctuated in mathematics education because most students in the United States take a high stakes mathematics assessment at the end of every school year starting in third or fourth grade.

Enacting Assessment Values

Teacher values around assessment, including where such conceptions might fall along the continuum, influence classroom decisions (Biesta et al., 2015). However, a teacher may hold competing goals and values for their approach to assessment (Thomas & Yoon, 2014). For instance, a mathematics teacher may value teaching practices that support a deep conceptual understanding of essential learning goals. They may also value preparing their students for success on high stakes assessments. The first value may require teachers to allocate extended instructional time to a single learning goal. The second value may require teachers to follow a particular curriculum or set of standards within a specific time frame. These two values may come into conflict when students need more time to master a particular learning goal but also need to move on to the next learning goal to complete the curriculum before the high stakes assessment. How a teacher decides which value to prioritize depends on teacher agency. Teacher agency is the degree to which teachers can enact their values in classroom practices within the culture and context of the school (Biesta et al., 2015).

Teacher agency is situated by context and a complex relationship between educational theory, practice, and environment (Biesta et al., 2015). Thomas and Yoon (2014) found that teachers prioritized their pedagogical values when making instructional decisions until one of three conflicting values came into play: (1) requirements about time, curriculum, or assessment; (2) potential success for future learning; and (3) respect for students' cultures. For instance, a secondary mathematics teacher who was committed to student-centered learning might not enact that value when it came in conflict with preparing students for assessments by completing the

national curriculum in the available time. In this case, the teacher switched to teacher-led instruction and limited the time allocated for a specific learning goal, despite their personal feelings that the concept was important, because it was not part of the required curriculum.

When there is tension between conflicting assessment values, particularly between pedagogical and accountability purposes, teachers may feel less agency around the accountability purposes of assessment. As a result, teachers may prioritize accountability purposes even to the detriment of their pedagogical values. Teachers' choices related to navigating conflicting values are influenced by the amount of agency they perceive in decision making. In this study we connect the research by Barnes et al. (2014), Thomas and Yoon (2014) and Biesta et al (2015) to discuss agency and how it relates to assessment values and practices with early career teachers in mathematics.

Methods

This study emerged as a result of piloting the student work sample interview protocol from the NExT digital handbook (NExT Teacher Effectiveness Work Group, 2018). The interview protocol was based on D'Souza's (2012) case study of early career teachers' assessment practices. D'Souza found that a student work sample interview protocol encouraged reflective practices with early career teachers that supported growth and development around assessment practices. Those findings motivated our choice to pilot this interview protocol, in particular because we were interested in how a protocol designed to work across disciplines might be used in the context of early career mathematics teachers. Unlike D'Souza, who implemented the TAPL with participating teachers routinely over five years, we piloted the NExT student work sample interview protocol once with each of our participating mathematics teachers.

Data Collection

Participants for this study were drawn from the approximately 40 secondary mathematics teachers who were in their first three years of teaching and had graduated from the secondary mathematics teacher licensure program at a Midwest university. All 40 teachers were invited to participate in the study and five teachers elected to participate. Four teachers taught at middle or high schools in the same Midwest state as the licensure program; the fifth taught in a school in the southeast United States.

Each teacher selected a recent assessment they had used in their mathematics courses and submitted a set of student work samples to the researchers prior to an individual interview. The only parameters for the teachers in this study was to select an assessment that had enough student work to drive a conversation about the learning that was present. The selected assessments ranged from two questions on a single skill to a summative unit test or project. The student work samples were graded and represented a range of student mastery.

The semi-structured interview included five main sections: (1) context questions focused on understanding the school setting, (2) description of the assessment, (3) discussion of student work samples, (4) implications for advancing student learning, and (5) a final reflection. The context questions asked teachers to share anything they felt was important for the researchers to know about their students and school. In the second section, teachers described the assessment and the learning activities that led to or followed the assessment. The majority of the interview time was dedicated to the third section, reviewing and analyzing the student work samples. During this portion of the interview, teachers were asked to organize the student work into three categories: *demonstrated mastery, approaching mastery,* and *still needs support*. The teachers then described what they noticed in each sample and explained why they categorized the student work as they did. Teachers were then asked about the next steps they wanted to take to further

student learning. Finally, each interview concluded with an opportunity to reflect on the experience of participating in a student work sample interview, as well as any reflections about the licensure program. Each interview lasted approximately 60-90 minutes. All interviews were video recorded and transcribed for analysis.

Data Analysis

To analyze the interviews, we used two distinct but complementary approaches. The first approach followed the NExT digital handbook (NExT Teacher Effectiveness Work Group, 2018) which suggests using rubrics provided to measure the effectiveness of teachers' assessment practices. These rubrics were adapted from edTPA. They evaluated (1) assessment planning, (2) analysis of student learning, (3) feedback on student learning, and (4) analysis of student learning to inform teaching. Using these rubrics for the initial purpose of piloting the materials from the handbook, we began to notice additional patterns in the data.

These initial observations led us to engage in a more systematic, open analytic process. Our second approach applied grounded theory to code and analyze the interviews. Grounded theory starts with a general area of research interests and uses the coding of qualitative data to identify patterns (Charmaz, 2006). Those initial patterns inform subsequent research questions and coding to arrive at a theory from the data. We concluded by looking at the results from these two analytic approaches to identify any further patterns evident in the data.

Results

Our analysis resulted in three key findings. First, we found that the mathematics teachers demonstrated high-level indicators of effectiveness relative to the pre-existing rubrics. In particular, they demonstrated values related to assessment that align with high-quality instructional practices. Second, through our use of open coding, we identified a wide variety of factors that influence how teachers enact their assessment values. Finally, we describe how teachers revealed tension between their assessment values and their perceived agency to enact those values.

Teachers' Assessment Views

We found evidence that all five teachers in the study demonstrated assessment practices and values that aligned with high-level indicators described in the NExT handbook. Teachers demonstrated that they (1) believe all their students can do well on mathematics assessments, (2) can differentiate feedback and analysis based on knowledge of individual students, (3) view assessment as ongoing, and (4) can design and implement assessments that provide students with opportunities for deep mathematical learning.

Mastery and ongoing assessments. A high-level performance indicator states that teachers use an "assets-focused approach to describing student progress towards learning goals". We found that all five teachers demonstrated this indicator in how they responded to interview prompts to discuss the student work samples. The interview protocol asked teachers to categorize their work samples based on student mastery. Despite this prompt, all five teachers resisted grouping the samples and instead discussed each piece of student work individually. The teachers consistently tailored their discussion and ideas for further instruction based on their knowledge of each individual student. They described the student understanding they saw demonstrated in the work and discussed how that aligned with the individual's progress towards their overall mathematics goals. We found their collective decision to discuss students individually rather than as a group to be evidence that the teachers were focused on each students' mastery level on the assessment and their progress toward the learning goals. The

teachers discussed student work in a way that consistently indicated that they believed all students would master the learning goal eventually.

Tied to the concept that all students would eventually master the learning goal, we also found that teachers approached assessment as ongoing. Teacher A, for example, said, "I like giving at least two chances on each test." This teacher went on to describe how additional opportunities for students to take assessments combined with standards-based grading provides Teacher A with a more comprehensive understanding of student mastery. Like Teacher A, all participants discussed intentionally providing ongoing assessment opportunities that enable students to demonstrate mastery of the learning goals beyond the first assessment.

Assessments as opportunities for deep mathematical learning. Another indicator of productive assessment practices from the student impact rubrics was opportunities for deep learning. To demonstrate valuing deep mathematical learning through assessment practices, teachers might require students to communicate mathematical arguments through their work and/or develop mathematical ideas. Teachers in our study demonstrated this in multiple ways. Teachers C and D analyzed student learning based on the ability to communicate valid arguments to support their work. Teacher B designed assessments that encouraged students to construct new mathematical ideas alongside checks for understanding.

To Teacher D, having the correct steps in a geometric proof was only part of demonstrating learning mastery. They also felt students must be able to create and communicate a valid argument through their work in a way that others can understand. Teacher D explained that they were not only looking for an accurate answer. "I also want them to understand that this is communication and it should be written to be read. I'm looking to motivate the idea that we're doing math as both a deductive and social activity." To assess that, Teacher D designed an open-ended project with student choice that allowed students to prove geometric theorems using paragraphs, annotated diagrams, or a two-column format. Teacher B's assessment design also encouraged students to think critically and advance conceptual understanding. Their assessment included questions that asked how a data summary would change if a new value was added to the data table. Students were also asked to make predictions about outliers, which served as a pre-assessment to the next lesson. These examples demonstrate that these early career teachers were prepared to design assessments with opportunities for deep learning by developing mathematical ideas and communicating mathematical arguments rather than focusing assessments solely on procedural fluency.

Factors that Influence Enactment of Assessment Values

The teachers indicated that there are a variety of factors that influenced how they enact their assessment values in practice. These factors included influence from administration, influence from their teacher preparation program, and their perceived agency in classroom decisions.

Influence from the teacher preparation program. All five teachers discussed assessment practices and values that were influenced by their teacher preparation program. These influences include practical elements of teaching and learning of mathematics, such as implementing multiple forms of assessments, incorporating unit plans they designed during their university coursework, and considering specific examples of student misconceptions in assessment design. They also demonstrated theoretical understanding of assessment practices by citing specific readings they had studied and the socioemotional needs of students.

For example, Teacher C reflected on their assessment and discussed designing assessments that were accurate measures of student understanding, which requires questions that can tease out misconceptions. The assessment selected by Teacher C was part of a routine created by the

mathematics team at their school to improve student skills on selected standards. In this assessment, students solved two questions using the distance formula. Both questions involved finding the distance of a horizontal line segment from the origin to another point. Teacher C critiqued their assessment, referencing a specific task from their teacher preparation program which revealed that students can develop misconceptions about triangles when the base is always horizontal to the x-axis. They applied that experience to this assessment and explained that future versions of this assessment should include questions that solve for the distance of both horizontal and diagonal lines. This example highlights the high-level indicators of designing and implementing assessments that are accurate measures of student understanding and using evidence in student work to measure nuanced growth towards the learning goal.

Teacher C explained, however, that adjusting assessment questions created a new set of challenges for department collaboration and measuring student growth. Their assessment was designed as an intervention strategy with multiple opportunities for reassessment. As such, one of the goals was consistency both in assessments/reassessments and across classrooms. Teacher C described the complexity of wanting to design assessments collaboratively and implement assessments that were consistent enough to support student growth through reassessments, while also varying the questions to be a true measure of what the students know. Teacher C's commitment to each of these values were influenced by their program, but the reality of navigating those values while collaborating with a department sometimes created tension.

Influence from school administrators. The interview data revealed that the teachers' assessment practices were strongly influenced by their school administrators' view of assessment in mathematics teaching and learning. The teachers expressed feeling tension between their assessment values and the assessment practices encouraged or enforced by their school administrators. The teachers described a complex and contextualized reality that influenced their ability to implement high level assessment practices.

Multiple teachers described tension between the values developed during their program and their administrator's views on assessment. For example, Teacher E discussed how their district's "data driven" assessment plan felt at odds with the values learned during their program. "We were taught about all the awesome ways that you could do things. Which is great...But I don't know if that effectively prepared me for walking into a school where that's the exact opposite of what they do". Teacher E felt that pressure from the district to maximize student performance on standardized tests influenced instructional and assessment decisions that went against the practices learned as a pre-service teacher.

Similarly, Teacher C shared that their school leadership viewed mathematics learning through a procedural lens.

> The program prepared me wonderfully for actually teaching mathematics, but navigating [my] district's and administration's attitude towards mathematics is more difficult...the decision makers in our administration conceptualize mathematics as all procedural. [...] How to navigate that while pushing towards the ultimate goal of making mathematics education about understanding is a thing that I feel like I don't know.

Through their work in the teacher preparation program, Teacher C learned to value understanding in mathematics teaching and learning, yet their administration viewed mathematics as "all procedural". Both Teachers C and E described feeling frustrated about how the "higher ups" views of assessment purposes influenced their ability to enact practices that align with their personal assessment values. They offset those tensions by creating ongoing assessment opportunities that provided students with additional chances. The teachers explained

that in their respective departments, the mathematics teachers chose to integrate reassessments and test corrections to better align their values with the required assessment expectations.

Finite learning time. Limited learning time also appeared to be a significant factor for how teachers enacted their assessment values. We define *learning time* as available time for instruction and intervention based on scope and sequence constraints. All five teachers talked about learning time. The teachers described immense pressure to maintain scope-and-sequence pacing for the year. This is particularly notable since our interview questions focused strictly on assessment without any reference to instructional time. Teacher B reflected on this from their position as the sole mathematics teacher at their school, saying, "I feel kind of blessed almost to be the only teacher because I get to work on my own schedule...and really use data from assessments to help guide my teaching." This contrasts with teachers who felt bound by their pacing guides. These teachers described feeling pressure to "push through the content" faster than the students could handle and having to "hustle to convince" students to get extra help during lunch or after-school. These time barriers also affected how much time teachers felt they had to provide feedback and modify their instruction based on assessment data.

These influences from the teacher preparation program, school administrators, and learning time all contributed to situations where teachers found themselves with competing goals relating to assessment. In the next section, we turn to the construct of agency to help us understand how teachers navigated these competing values.

Complex and Contextualized Teacher Agency About Assessment

Despite our finding that these five teachers demonstrated high-level indicators with respect to their assessment values, how they were able to enact those values was contextualized through the complex realities of their schools. Our data highlights complexities around administrators' view of assessment purposes in mathematics and the relationship between learning time and instructional decisions. Teachers navigated these contexts based on their perception of agency over teaching practices. In some circumstances teachers felt agency to align their assessment values with practices, such as using assessment data to inform classroom decisions and designing assessments with opportunities for deep mathematical learning. At the same time, all of the teachers described situations where they felt pressure to enact assessment practices that went against their values. One area where teachers demonstrated agency was in designing year-long assessment plans with ongoing assessments. All of the teachers were required to meet assessment expectations set by the "higher ups". This included weekly department-wide skill assessments, a minimum of two summative assessments every four weeks, or administering assessments that were closely tied to high stakes standardized assessments. However, all five teachers also incorporated ongoing assessment practices or assessments designed to allow solutions in multiple representations within their school's larger assessment plan. Teachers were able to find space within those expectations to enact their assessment values based on the degree of agency.

Discussion

Our study had three findings. First, all five teachers demonstrated high-level performance indicators for assessment values. Second, factors influenced how teachers enacted their values - including tension created by the "higher ups" views and expectations about assessments in mathematics. Third, teachers navigated this tension by agency that was contextual. A key finding from this study showed that mathematics teachers recognized a misalignment between their assessment values and their agency to enact those values.

We argue that the degree of teacher agency to enact assessment values depends on where the conflicting values lie on the continuum of purposes of assessment (Barnes et al., 2014). When

the purpose of assessment was at the pedagogical end of the continuum, teachers were doing the assessing and had agency to align their values with classroom practices. In our study, we saw this when teachers used multiple forms of assessment, created ongoing assessment practices, and individualized feedback. However, when the conflicting values were between accountability purposes from the "higher ups" and pedagogical purposes, teachers yielded their values to conform to expectations. At this end of the continuum, teachers were *being assessed* while they were assessing student understanding. The "higher ups" use assessment as a form of control and accountability to evaluate teacher effectiveness (Barnes et al., 2014). When there is tension between the teachers' assessment values and higher ups views of mathematics assessments, teachers are put in a position where they must choose between their evaluation of teacher effectiveness or pedagogical practices that align with their values.

One of the most striking themes from our data was the influence learning time had on teachers' assessment practices. All of the teachers in our study cited the pressures of learning time as a driving factor for instructional decisions that misalign with their values. Thomas and Yoon (2014) found that time was one of three factors that influenced a teacher to abandon their pedagogical values in order to conform to accountability purposes of assessment. In that study, time was grouped with curriculum and assessment. Echoing Thomas and Yoon, we posit that required curriculum and high stakes assessments, as well as learning time, are factors that influence teachers' agency to enact their assessment values. Furthermore, we argue that teachers felt pressured to disregard their values because time, curriculum, and high stakes assessments are at the accountability end of the continuum of assessment purposes.

"Higher ups" make decisions and set policies that they believe will help students learn mathematics, such as scope-and-sequencing and departmental teaching strategies. However, their lens of understanding these practices is often driven by standardized assessments, which means that it is still largely procedural. Mathematics teachers, through their licensure programs, have a complex and nuanced understanding of teaching mathematics for conceptual understanding that align more strongly with pedagogical purposes of assessment. Our findings support research that continues to think about ways to use available learning time for deep learning rather than focusing solely on procedural skills. Our data shows that less pressure around maintaining a scope-and-sequence may enable teachers to better enact their assessment values.

Our study piloted a teacher interview protocol that was adapted from D'Souza's (2012) TAPL with early career mathematics teachers. Similar to D'Souza, we found that this protocol was an effective tool in understanding teachers' practices and values around assessment. We found that when implemented with secondary mathematics teachers, the student work protocol revealed both indictors of high-level assessment practices and contextual factors that might prevent teachers from enacting those practices. We posit that teacher agency is a key component to enacting assessment values and is tied to the continuum of purposes of assessment (Barnes et al., 2014). Since accountability purposes influence teachers' agency, evaluations of teacher effectiveness should include opportunities for teachers to identify when they felt pressured to act against their values. The NExT student work sample protocol is one alternative to evaluate teachers' assessment values and practices.

References

Association of Mathematics Teacher Educators. (2017). *Standards for preparing teachers of mathematics*. Author. http://amte.net/standards

Barnes, N., Fives, H., & Dacey, C. M. (2014). Teachers' beliefs about assessment. In *International Handbook of Research on Teachers' Beliefs*. Routledge.

Beswick, K. (2011). Teachers' beliefs about school mathematics and mathematicians' mathematics and their relationship to practice. *Educational Studies in Mathematics*, *79*(1), 127–147. https://doi.org/10.1007/s10649-011-9333-2

Biesta, G. J. J., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, *21*(6), 624–640. https://doi.org/10.1080/13540602.2015.1044325

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, *11*(3), 301–318. https://doi.org/10.1080/0969594042000304609

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications.

Davis, D. S., & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *The Journal of Educational Research*, *104*(3), 202–215. https://doi.org/10.1080/00220671003690148

D'Souza, L. A. (2012). Assessing student learning through guided inquiry: A case study of a beginning teacher. *Journal of Education*, *192*(2–3), 79–87. https://doi.org/10.1177/0022057412192002-310

Karp, G. G., & Woods, M. L. (2008). Preservice teachers' perceptions about assessment and its implementation. *Journal of Teaching in Physical Education*, *27*(3), 327–346. https://doi.org/10.1123/jtpe.27.3.327

NExT Teacher Effectiveness Work Group. (2018). *Preparing effective teachers: Multiple approaches to ensuring teaching quality and impact* (M. Sato & Lloyd, Eds.). Network for Excellence in Teaching. https://www.nexteachers.org/next-digital-handbook

Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam (2005). *The Curriculum Journal*, *18*(1), 27–38. https://doi.org/10.1080/09585170701292133

Sato, M. (2014). What Is the underlying conception of teaching of the edTPA? *Journal of Teacher Education*, *65*(5), 421–434. https://doi.org/10.1177/0022487114542518

Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board certification. *American Educational Research Journal*, *45*(3), 669–700. https://doi.org/10.3102/0002831208316955

Thomas, M., & Yoon, C. (2014). The impact of conflicting goals on mathematical teaching decisions. *Journal of Mathematics Teacher Education*, *17*(3), 227–243. https://doi.org/10.1007/s10857-013-9241-8