UNPACKING TEACHERS' PERSPECTIVES ON THE PURPOSE OF ASSESSMENT: BEYOND SUMMATIVE AND FORMATIVE

Rachael Kenney
Purdue University
rhkenney@purdue.edu

Lane Bloome
Purdue University
lbloome@purdue.edu

Yukiko Maeda
Purdue University
ymaeda@purdue.edu

The current age of standardization and accountability in education presents a need for educators to be proficient consumers and producers of data. To understand what mathematics teachers see as the purpose of data and assessment, as well as their perceived needs in these areas, we administered a survey to all mathematics teachers in two school districts in the Midwest. We analyzed perceptions of the purpose of assessment using a framework that breaks from the common formative/summative divide and considers the purpose(s) for assessment as being nuanced and potentially multi-faceted. The responding teachers demonstrated a robust working knowledge of assessment focusing on purposes that serve students, teachers, schools, and states. Responses also indicated that there are gaps in connecting the purposes of assessment and their knowledge on how to implement assessments in ways that meet these purposes.

Keywords: Assessment and Evaluation, High School Education, Teacher Knowledge

In the current climate of accountability, school administrators and teachers face unprecedented demands to be assessment-literate and use a wide range of data to inform educational decisions that document and promote student achievement. Accordingly, *data-driven decision-making* (DDDM) has become a central focus for educational policy and practice at all levels as an innovative strategy for school system and instructional reform (Gill, Borden, & Hallgren, 2014; Luo, 2008; Mandinach, 2012). DDDM is broadly defined as "the use of data analysis to inform choices involving policies and procedures" (Gill et al., 2014, p. 338). It is a complex undertaking, even for experienced educators and administrators equipped with quantitative skills (Mandinach, 2012). Developing the data literacy required for effective implementation of DDDM involves systematic experience with a variety of data (Mandinach & Gummer, 2013), including demographic, behavioral, achievement/ performance, attendance, financial, policy, programmatic, and compliance data. The data most under teacher and school control, however, are gained from classroom or school-level assessment. With this study we address a critical need to understand teachers' current assessment practices and current understanding of how to connect instruction, assessment, and related data in a way that can inform teaching practice (Datnow & Hubbard, 2015).

It is estimated that teachers spend up to one half of their classroom instruction time on some form of assessment (Stiggins, 1999). What then, do teachers see as the purpose(s) for devoting teaching time to this activity? In this paper, we share one piece of a larger study of DDDM by focusing on the critical component of assessment. We address the research questions: (a) *What do mathematics teachers see as the main purposes of assessment, and* (b) *what challenges impede teachers' use and understanding of assessment data?*

Background & Framework

Effective DDDM requires adequate *data literacy*, which can be broadly defined as the ability to understand and use data effectively to inform decisions (Mandinach & Gummer, 2013). Data-literate educators are a driving force of student learning, and ensure the continued success and funding of their schools in the era of accountability where data-based evidence plays a prominent role in both instructional and evaluative decision making (Orland, 2015). *Assessment literacy* is a critical component of data literacy, often defined broadly as being able to recognize sound assessment, evaluation, and communication practices to benefit student learning and achievement (Stiggins,

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

1999). As part of data literacy, educators integrate assessment results with other data reflecting content, context, perception, motivation, process, and behavior (Mandinach & Gummer, 2013).

Assessment, including both high-stakes and classroom assessments, can be one of the most influential activities in education. Research, however, suggests that teachers enter the profession with insufficient practice and knowledge for developing assessments for learning, evaluating student progress, and interpreting data (Bocala & Boudett, 2015; Popham, 2009; Volante & Fazio, 2007). When courses are offered in teacher education, topics tend to focus on the use of assessment for evaluating student outcomes rather than on the use of assessment as part of student learning (DeLuca & Klinger, 2010). Moreover, programs do not prepare preservice teachers to develop assessment skills that are adaptable to diverse student populations (DeLuca & Lam, 2014).

There is a large body of research on assessment, with significant attention given to clarifying the role of school assessment. For example, Griffin (2008) suggests that the most fundamental role of assessment is to help interpret observable behaviors in order to infer learning; "the more skills are observed, the more accurately generalized learning can be inferred" (p. 195). Assessment can be used to provide information to make decisions about student achievement (i.e., assessment *of* learning) and to support ongoing student learning (i.e., assessment *for* learning) (Stiggins, 2005). The distinction between summative and formative assessment, which are often distinguished by purpose, timing, or level of generalization (Black & Wiliam, 1998; 2003; Harlen, 2005; Sadler, 1989), is a prominent feature of the majority of this literature base. The fact that the purpose of assessment can be interpreted in a number of different ways is, according to Newton (2007) one of the most basic points for an educator to appreciate. Newton cautions, however, against reducing the purpose of assessment to two or three categories:

> We give the wrong message when we try to simplify assessment purposes by allocating them to a small number of categories (such as formative, summative and evaluative): we imply that the sub-purposes within those categories are importantly alike. This risks the impression that results which are fit for one sub-purpose within a category will be fit for the other sub-purposes as well…this is contrary to the impression that ought to be given to policy-makers, to ensure that wise decisions are made. (p. 161).

Newton suggests that the dichotomy between summative and formative assessment has been ineffective in understanding the nuanced nature of assessment and may hinder advancements in assessment theory. Bennett (2011) concurs that the use of assessments in support of learning is not limited to a certain kind of assessment such as formative or summative, because more than one type of assessment can contribute to judgments about students' achievement. Making distinctions of some sort, however, is useful because it helps us consider what assessment practices teachers are using and for what purposes.

In our work, we take the perspective of Newton (2007) who claims that, "to avoid getting ourselves confused, and to avoid confusing others, we need to use the language of assessment with greater precision" (p. 157). To this end, we choose to focus on the language that practicing mathematics teachers use to describe their understandings of the main purposes of assessment in their schools with the hopes of breaking the common borders between summative and formative distinctions. Newton shares at least three levels at which the purpose of assessment can be distinguished, including: (a) the *judgment* level—which concerns the technical aim of an assessment event (e.g., the purpose is to derive a standards-referenced judgment, expressed as a grade); (b) the *impact* level—which concerns the intended impacts of running an assessment system (e.g., the purposes are to ensure that students remain motivated, and that all students learn a common core); and (c) the *decision* level—which concerns the use of an assessment judgment, the decision, action or process which it enables. Newton further illustrates a range of purposes that may occupy discourse at

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

this last level. We use these (shown in Table 1 along with judgment and impact levels) as an initial framework for unpacking secondary mathematics teachers' descriptions of assessment purposes.

Table 1: Purposes of Assessment, Initial List (Newton, 2007; 2010)

Level	Categories	Description
Decision	Social Evaluation	Judge the social or personal value of students' achievements
	Formative	Identify proximal learning needs, guiding subsequent teaching
	Student monitoring	Decide whether students are making sufficient progress in attainment in relation to expectations or targets
	Transfer	Identify educational needs of students who transfer to new schools
	Guidance	Identify the most suitable courses, or vocations for students to pursue, given their aptitudes
	Institution monitoring	Decide whether institutional performance – relating to individual teachers, classes or schools – is rising or falling in relation to expectations or targets
	Resource allocation	Identify institutional needs and allocate resources
	Program Evaluation	Evaluate the success of educational programs or initiatives, nationally or locally
	Placement	Locate students with respect to their position in a learning sequence
Judgment		Derive a standards-referenced judgment, expressed as a grade
Impact		Ensure that students remain motivated, and that all students learn a common core

Methods

We addressed the research questions for this study by conducting a survey with teachers across two school districts in the Midwest.

Participants

School District A is a small district with 2,200 students. It is located in a rural area and it has implemented evidence-based decision making since the 1990s. There is one middle and one high school in District A. The district has a history of high academic performance, maintaining a higher passing rate of state mandated test than the state average for the last eight years. School District B is also rural with 6,700 students. It has two middle schools and one high school. District academic performance has declined somewhat over the last eight years. District B has not yet formulated a strong culture of DDDM. A total of 99 mathematics, science, and language arts teachers at the middle and high schools and 16 administrators participated in the online survey. Of these, 29 mathematics teacher survey results were analyzed for the current investigation.

Data collection and Analysis

The survey was designed by the research team by drawing on current research and assessment/data related reports from the U.S. Department of Education. The survey contained both open-ended and multiple-choice items and was designed to allow us to identify respondents'

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

perceptions of the purpose(s) of assessment and DDDM. The survey was distributed through Qualtrics®, an online survey tool. We focused our analysis on three survey questions:

1. What do you feel are the three main purposes of assessment in your school? (open-ended)
2. What do you see as your biggest areas of need?
3. What do you see as the biggest challenges related to using data from student assessment?

Survey responses were summarized descriptively for this paper to understand teacher and administrators' understandings of the purpose of assessment, challenges in meeting assessment goals, and teachers' current assessment practices. We used Newton's (2007, 2010) list of purposes (Table 1) to guide the axial coding analysis of open-ended items.

Findings

Purpose of assessments

Coding teachers' responses to the question, "What do you feel are the three main purposes of assessment in your school?" allowed us to expand our framework and further understand the variety of ways in which teachers view assessment use. From the 29 responding teachers, we received 75 distinct responses to this item. Some responses could be broken up into separate statements that fit different codes, resulting in a total of 87 total coded statements in our data. Table 2 provides the summary of the axial coding analysis these 87 statements from the mathematics teachers' responses. After initially coding responses to fit under the decision, impact, or judgment levels described above (Newton, 2007), we identified who seemed to the target of assessment decisions, impacts, or judgments – students, teachers, schools, or the state. We analyzed these codes further to identify where they fit in our framework and created additional codes that we posit to be beneficial in capturing the nuances of teachers' perceptions.

Table 2: Teacher Identified Major Purposes of Assessment

Levels	Focus	Codes	N
Decision Level	Students	Formative - Pre-assessment	2
		Formative - Instruction	9
		Formative - Student mastery	13
		Student monitoring (long term)	13
		Resource allocation - remediation	3
		Placement	2
	Teachers	Institution monitoring - Teacher accountability	12
	School	Institution monitoring - School accountability	9
Impact Level	Students	Encouragement	5
		State Exam Practice	4
Judgment Level	Students	Grading	5
	School	Collect data	6
	State	Meet a requirement	4

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

Focus on students. The majority of responses focused on the assessment purposes that targeted students (57 out of 87). These were found at all three levels. Coded responses at the decision level were broken down into three separate categories of what Newton (2010) referred to as formative (where the purpose of assessment is to "identify students' proximal learning needs, guiding subsequent teaching" (p. 392)): (a) to preassess students' prior knowledge, (b) to make decisions during class about whether students had mastered a concept, and (c) to make decisions about how to move forward with instruction. Some examples of responses coded into these categories include:

- "Determine if a student has mastered a concept." (*mastery*)
- "In my classroom I use them to find out students ability to recall and apply concepts." (*mastery*)
- Student scores are then used as a tool to adjust my lessons." (*instruction*)
- "To determine where students currently are with a certain standard before it is taught." (*pre-assessment*)

Student mastery formed the largest group under the formative code (n=13). Student monitoring was also coded for 13 responses, and was distinguished from the formative categories if the response implied that assessment served a purpose of determining student learning or growth over time (instead of within a lesson). Some examples include:

- "To measure growth of knowledge."
- "To make sure students are learning the state standards."
- "To see what students learn and retain throughout the year."
- "To determine how much students know after teaching standards."

At the impact level, which concerns the intended impacts of assessment, teachers reported assessment purposes that target students. This student-mindedness was evidenced by identifying assessment as a tool for keeping students motivated and allowing students to practice for high-stake state exams. For example:

- "Assessment gives students clearer goals to help put more emphasis and importance on learning the content; it works well for students as an incentive to grow." (*encouragement*)
- "Prepare students to succeed when taking state standardized tests." (*practice*)

One teacher added that assessments were "almost like a reward in the sense that students are given problems they've been training themselves for and are given a number or letter grade showing how much they know."

Only five teachers made mention of giving grades as a purpose of assessment. These responses were coded at the judgment level, as they described a technical aim of assessment.

Focus on Teacher/School/State. The survey results revealed that teacher, school, or state accountability is a common purpose that teachers see for assessment. At the decision level, these foci appeared as responses indicating that assessments are used to "compare teachers," "Determine if a teacher is an effective teacher", and "to make sure everyone is teaching the same content." These and similar responses were coded under institutional monitoring, but focused on teacher accountability. Additional responses claimed that assessments were to "compare to other schools," "calculate school performance," and "let the public how the school is doing." These were coded as institutional monitoring with a focus on school accountability. Additional comments coded as focusing on assessment purposes for the whole school included six responses that explained the

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

purpose of assessment in basic terms, "to get data," which is a technical aim of assessment (and thus coded at the judgment level) in their school.

Some negative connotations of assessment were also found at the judgment level, with four teachers indicating that assessment served the purpose of "making the state happy" or meeting state requirements. One teacher added that they served the state by "reporting numbers that don't tell the whole story", indicating frustration with the roles that s/he sees assessment playing at the school or state level.

Areas of Need and Challenges

Table 3 summarizes the areas of needs relative to assessment and data practices that teachers identified in their responses. As shown above, a large number of teachers identified long-term student monitoring as a significant purpose of assessment. However a majority also identified collecting useful growth data as a major area in which they need additional guidance or support. Similarly, teachers identified examining student data to identify which instructional practices work best for which students and analyzing with an equity lens as problematic areas. However, a majority also indicated little to no need with help on adapting instructional activities to meet students' individual needs in this same question. It is possible that teachers see this activity as disconnected from the use of data or assessment. It is also interesting to note that none of the teachers made mention of differentiation of any kind in their descriptions of the purposes of assessment above. It will be important to follow up on this issue in interviews with the teachers. A majority of the mathematics teachers did feel that they were well practiced with designing assessments aligned to state standards, which is not surprising given the current influence of state exams on teaching practices.

Table 3: Teacher Identified Area of Need

Need	Some to Great Need	No to Little Need
Collecting useful growth data.	18	7
Examining student data to identify which instructional practices work best for which students	18	7
Analyzing data through an equity lens.	16	9
Collaborating and sharing ideas with colleagues regarding data inquiry and analysis issues	16	9
Structuring the district organization and practices to support data-driven decision-making.	14	11
Communicating with families about student progress.	12	13
Developing curriculum-embedded formative assessments.	12	13
Interpreting assessment data to identify gaps in student achievement	12	13
Adapting instructional activities to meet students' individual needs	10	15
Designing assessments aligned to standards.	8	17

Additional challenges in using assessment data for instruction were identified in teachers' survey responses, including resource limitations, significantly, time. Teachers also perceive some major challenges in being able to access multiple sources of student data in useful ways to guide decision-

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

making, and many suggested that they were not prepared or did not have adequate guidance for making instructional decisions from assessment data.

Conclusions

The teachers who responded to our survey have a working knowledge of multiple purposes for assessment. This runs contrary to the sentiment in current literature that teachers enter the profession with deficient assessment literacy and that this skill deficit remains throughout their professional careers (Bocala & Boudett, 2015; Popham, 2009; Volante & Fazio, 2007). In discussions about what teacher education can do to provide assessment-related professional development, the focus is often on the presumption that teachers do not have a strong knowledge base for assessment literacy. The data presented here does not really support this. Through their professional experiences or training, teachers have well-developed ideas about the varied purpose of assessment and what they want to be able use them for. What there does seem to be is a gap in their ability or understanding of how to enact assessment and data use in a way that serves the purposes they know they can serve.

Our results suggest that the key to developing teachers' assessment literacy is not to inundate pre-service and in-service teachers with information about the technical uses of assessment. Instead, it would be beneficial to use teachers' existing assessment knowledge base to craft professional development opportunities targeted at developing their skills at linking this knowledge base to actionable instructional strategies. To this end, it seems critical to find ways to provide time and incentives for teachers to collaborate with their colleagues in sharing assessment data and share best practices for linking assessment data to instructional adjustments. The more advanced perceptions on the purpose of assessment came from statements that focused on the students as the benefactors of assessment (rather than on teacher or school accountability). Thus is also seems important to help teachers direct their discourse away from a primary focus on state exams as "the" assessment and reflect more directly on daily teaching practices that involve assessment, prompting teachers to continually reflect on the goodness of fit between their assessments and their purpose(s). These strategies are critical components to include when focusing on the development of a school and district-wide culture of DDDM and data sharing.

The teachers' descriptions of purposes of assessment support the need to broaden the distinctions that we make as a field beyond summative and formative. These words did not appear in teachers' descriptions and to not adequately capture the wide variety of roles they see assessment playing in their classrooms and schools. Our use of Newton's work (2007) as an initial framework proved useful in unpacking the teachers' perceptions in a meaningful way. We feel this is a positive step in helping teachers and teacher educators develop more productive discourse around the use of assessment for guiding instruction in classrooms.

There is a paramount need for teachers to be assessment-literate in this age of standardization and accountability. Our survey suggests that teachers have the skills and knowledge to answer this call. However, teachers report that they need support and guidance transforming their working knowledge of assessment into habits of using this data to inform dynamic instruction. As mathematics educators, we have the resources to provide this guidance, creating an opportunity for us to work together with teachers to ensure that their students are given every opportunity to receive quality experiences in the mathematics classroom.

References

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Black, P. J., & Wiliam, D. (2003). In praise of educational research: formative assessment. *British Educational Research Journal*, *29*(5), 623–637.

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

Bocala, C., & Boudett, K. P. (2015). Teaching educators habits of mind for using data wisely. *Teachers College Record*, *117*(4), 1-20.

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.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419-438.

DeLuca, C., & Lam, C. Y. (2014). Preparing teachers for assessment within diverse classrooms: An analysis of teacher candidates' conceptualizations. *Teacher Education Quarterly*, *41*(3), 3.

Gill, B., Borden, B. C., & Hallgren, K. (2014). *A conceptual framework for data-driven decision making* (No. 8142). Princeton, NJ: Mathematica Policy Research.

Griffin, P. (2008). Teachers' use of assessment data. In C. M. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 187-212). New York: Springer.

Harlen, W. (2005) Teachers' summative practices and assessment for learning—tensions and synergies. *The Curriculum Journal*, *16*(2), 207–223.

Luo, M. (2008). Structural equation modeling for high school principals' data-driven decision making: An analysis of information use environments. *Educational Administration Quarterly*, *44*(5), 603-634.

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71-85.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, *42*(1), 30-37.

Newton, P.E. (2007). Clarifying the purpose of educational assessment. *Assessment in Education*, *14*(2), 149-170.

Newton, P. E. (2010). The multiple purposes of assessment. In B. McGaw, P. L. Peterson, & E. Baker (Eds.), *International encyclopedia of education* (3rd ed.). Elsevier.

Orland, M. (2015). Research and policy perspectives on data-based decision making in education. *Teachers College Record*, *117*(4), 1-10.

Popham, W. J. (2009). Assessment literacy for teachers: faddish or fundamental? *Theory Into Practice*, *48*(1), 4-11.

Sadler, D. (1989) Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Stiggins, R. J. (1999). Are you assessment literate? *The High School Journal*, *6*(5), 20-23.

Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, *87*(4), 324-328.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749-770.