5-1-2017

Developing Measures of Teacher and Student Understanding in Relation to Learning Trajectories

Caroline Ebby
University of Pennsylvania

Philip Sirinides

Jonathan Supovitz

This research was funded by the National Science Foundation (grant no. DRK-12 1316527). Any opinions expressed in this paper are those of the authors' and do not necessarily reflect the views of the National Science Foundation.

Developing Measures of Teacher and Student Understanding in Relation to Learning Trajectories

Abstract
This paper describes the impact of the OGAP intervention on teachers' ability to use formative assessment data for instructional decision making. We measured this construct both before and after one and two years of the intervention with an instrument developed to measure teacher knowledge of student thinking in the activity of looking at and responding to student work. We begin with an overview of the design and development of the TASK instrument, and then present quantitative and qualitative findings on the impact on teacher responses.

Keywords
OGAP, NSF, Ongoing Assessment

Disciplines
Educational Methods | Elementary Education | Science and Mathematics Education

Comments
This research was funded by the National Science Foundation (grant no. DRK-12 1316527). Any opinions expressed in this paper are those of the authors' and do not necessarily reflect the views of the National Science Foundation.

The Impacts of OGAP on Teachers' Interpretation and Response to Student Thinking[1]

Caroline B. Ebby, Phil Sirinides, Jonathan Supovitz
Consortium for Policy Research in Education
Graduate School of Education, University of Pennsylvania

This paper describes the impact of the OGAP intervention on teachers' ability to use formative assessment data for instructional decision making. We measured this construct both before and after one and two years of the intervention with an instrument developed to measure teacher knowledge of student thinking in the activity of looking at and responding to student work. We begin with an overview of the design and development of the TASK instrument, and then present quantitative and qualitative findings on the impact on teacher responses.

Background and Rationale

Interpreting and responding to student thinking is central to recent characterizations of ambitious mathematics instruction that have emerged from mathematics education research (Lampert, Beasley, Ghoussenini, Kazemi & Franke, 2010; Stein, Engle, Smith & Hughes, 2008). Stein et al (2008) define five practices related to making student thinking central to mathematics instruction: anticipating, monitoring, selecting, sequencing, and making connections between student strategies and explanations, while at the same time ensuring that the learning of the whole class moves towards the mathematical goals. In addition, using information on student thinking formatively has proven to be a key factor for improving learning (Black & Wiliam, 1998; Wiliam, 2007). Effective formative assessment involves continually collecting and interpreting evidence of student thinking to formulate an instructional response targeted to help the learner move closer to the learning goal. Yet despite the importance of understanding student thinking to both current theories of mathematics instruction and formative assessment, research suggests that many teachers struggle to make effective use of evidence of student thinking (Goertz, Olah &Riggan, 2009; Supovitz, Ebby & Sirinidies, 2014).

A constructivist perspective on learning argues that teachers cannot directly know their students' mathematical understandings (von Glasersfeld, 1995); they can only draw from evidence of students' mathematical work to build a model of student knowledge. Research from the Cognitively Guided Instruction project highlights the importance of knowledge of children's solution strategies in this model-building process. In both experimental and case studies, teachers who were provided with research-based knowledge about children's thinking and problem solving in addition and subtraction were found to have higher levels of student achievement (Carpenter, Fennema, Peterson, & Carey, 1988; Carpenter & Fennema, 1992). More recent research begins to explore the process by which teachers develop and use the pedagogical content knowledge necessary for interpreting and responding to evidence of student learning. For example, Wilson, Lee, and Hollebrands (2011) investigated how pre-service teachers make sense of students' work on a data analysis task and found that in addition to

[1] This research was funded by the National Science Foundation (grant no. DRK-12 1316527). Any opinions expressed in this paper are those of the authors' and do not necessarily reflect the views of the National Science Foundation.

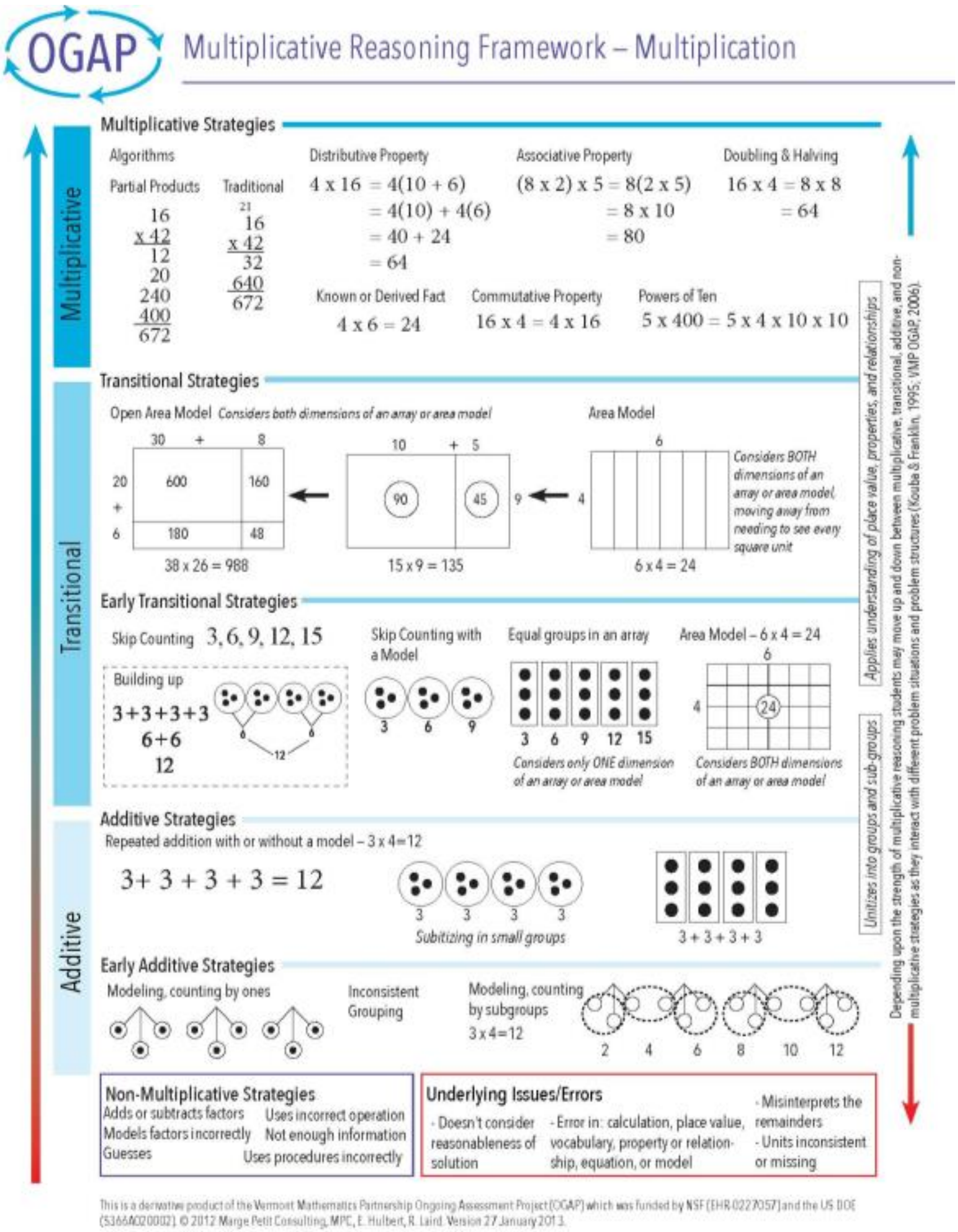
the actions of describing, comparing and inferring, teachers go through a process of restructuring their own mathematical understandings as they collect evidence of multiple approaches and develop models of student thinking around particular mathematical concepts. Kazemi and Franke (2004) studied facilitated conversations with elementary teachers around student work and found that over time teachers began to recognize sophistication of strategies, think about better ways to elicit student thinking, and develop "possible instructional trajectories" that built on student thinking. Key to this transformation was learning to focus on the details of student thinking as well as an increase in teachers' attempts to elicit student thinking in the classroom.

Learning trajectories, or developmental progressions of levels of student thinking in particular mathematical domains, are gaining increasing prominence in mathematics educational research (Battista 2011; Daro, Mosher, & Corcoran, 2011) and are at the core of current conceptualizations of both standards and instructional practice (Common Core State Standards Initiative, 2010; Szatjn, Confrey, Wilson, & Edgington 2012). Learning trajectories can provide an important link between research on learning and research on teaching by providing teachers with a clear articulation of learning goals, a framework for how students' thinking develops, and learning activities that are likely to move students along the path towards achieving those goals (Heritage 2008). Recent studies confirm that introducing teachers to these research-based frameworks of how students build mathematical understanding can enhance their ability to interpret evidence of student learning and respond productively in light of that evidence (Clements et al., 2011; Wilson 2009).

The OGAP formative assessment system is set of tools, resources, and routines to help teachers systemically and continuously monitor and respond to student understanding in relation to learning trajectories. In each year of the study, teachers experienced five days of intensive professional development focused on mathematical content, the research base on student thinking, and the use of formative assessment items, frameworks, and strategies. Teachers were also supported throughout the year in Professional Learning Communities (PLCs) focused on looking at student work. A central component of OGAP is a framework that synthesizes problem structures and the progression of student thinking in core content areas to help teachers analyze evidence in student work and make instructional decisions. (see Figure 1). This is both a *conceptual* framework that provides a general schema for how students learn the core content--i.e., developing procedural fluency through conceptual understanding, the use of visual models, and properties of operations-- and a *classification* structure that helps teachers categorize and label student strategies in order to make sense and act on the data to inform instruction.

As part of the larger RCT study on impacts of OGAP on teacher and student learning (See Supovitz et al., 2017), we used the TASK instrument, described below, to explore the impact of this intervention on grades 3-5 teachers' capacity to make sense of and respond to student work in multiplicative reasoning over two years.

Figure 1 OGAP Multiplication Progression[2]



[2] There is a corresponding progression for division and also for fractions, the focus of the intervention in Year 2.

Theoretical Framework

In conceptualizing the knowledge that teachers need to implement effective formative assessment in the classroom, we draw upon a conception of teaching as a complex activity that is dependent on distinct but interconnected bodies of knowledge. (Ball, Thames & Phelps, 2008; Putnam & Borko, 2000; Shulman, 1987). Arguing that teachers draw on knowledge that is distinct from either knowledge of subject matter, Shulman defines *pedagogical content knowledge* (PCK) as “the ways of representing and formulating the subject matter that make it comprehensible to others” (p. 9) and frames it as the intersection between content and pedagogy. Building on this work to study the work that teachers do when teaching mathematics in the classroom setting, Ball and colleagues have further defined and delineated mathematical knowledge for teaching (MKT) by breaking down the domain of content knowledge into *common content knowledge*, *specialized content knowledge* and *horizon content knowledge* and pedagogical content knowledge into *knowledge of content and students*, *knowledge of content and teaching,* and *knowledge of content and curriculum* (Ball et al, 2008).

More recently, Sztajn et al (2012) bring together research on learning trajectories with research on teaching to propose the construct of *learning trajectory based insruction* as "teaching that uses student learning trajectories as the basis for instruction (p. 147)." In addition to presenting a learning trajectory interpretation of the six MKT categories, they define a learning trajectory interpretation of formative assessment as the case where teachers are "guided by the logic of the learner" rather than only by disciplinary goals when eliciting student thinking and providing feedback to students. In developing the TASK instrument and analyzing the results of the field test, we draw on these frameworks to explore how teachers actually make sense of evidence of student thinking for their instruction.

We also draw on situated views of learning to prioritize sensemaking and conceptualize teacher learning in relation to the mediation of action by cultural signs and tools (Wertsch, 1998). Within the OGAP intervention, the *OGAP Progression* can be seen as a cultural tool; as a mediator of knowledge and action, it has both affordances and constraints (Wertsch, 1991, 1998). In using sociocultural theory to understand teacher’s use of pedagogical tools, Grossman and colleagues (1999; 2000) distinguish between *conceptual tools*, or principles, frameworks and ideas about teaching, learning and content that teachers use “as heuristics to guide their instructional decisions,” and *practical tools* that “have more local and immediate utility” (p. 634). Both conceptual and practical tools can be appropriated by teachers, a developmental process which involves internalization of specific culturally embedded ways of thinking through active participation in social practices (Leont'ev, 1981, Wertsch, 1991). Grossman et al. (1999) define five different levels of appropriation: lack of appropriation, appropriating a label, appropriating surface features, appropriating conceptual underpinnings, and achieving mastery. The level of appropriation can be influenced by the social context of learning and individual characteristics of the learner.

The TASK Instrument

The Teacher Assessment of Student Knowledge (TASK) instrument was developed, field tested, and validated to provide a contextualized measure of teachers’ ability to a) analyze students’ mathematical thinking within a grade-specific content area in relation to research-based learning trajectories, and b) formulate effective instructional responses (Ebby & Sirinides, 2015; Ebby, Sirinides, Supovitz & Oettinger, 2013; Supovitz et al., 2014). For this study we used three parallel forms of the Multiplicative Reasoning (MR) TASK for grades 3-5 that were designed to assess the following three domains of knowledge relevant for making sense of student work for instruction and assessment:

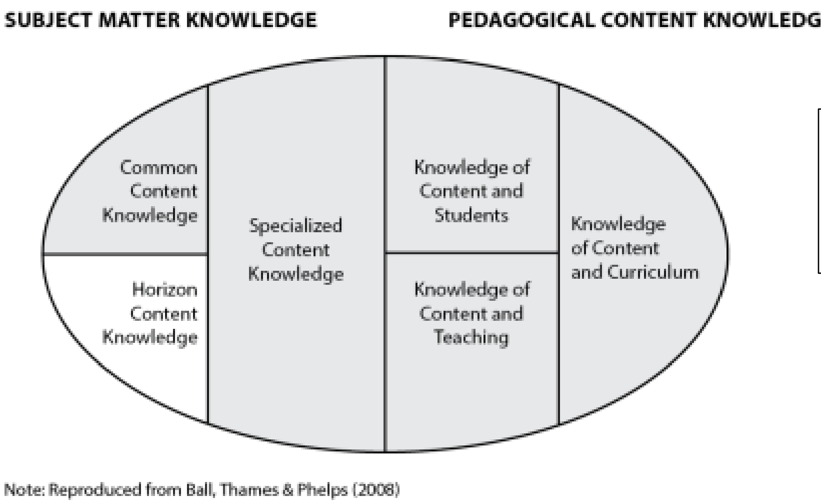
1. *Analysis of Student Thinking (AST)* – In order to build on student thinking, teachers need to be able to go beyond determining whether or not a response is correct or incorrect to identify the underlying conceptual understanding or misconceptions that are present in student work.
2. *Learning Trajectory Orientation (LTO)* – After analyzing the strategy a student uses to solve a math problem, teachers need to be able to position that strategy along a learning trajectory for the respective math content. Thus, teachers must have a sense of what the developmental progress looks like for the particular math concept and where to place students along that continuum and be able to use this as a framework to interpret and respond to student thinking.
3. *Instructional Decision Making (IDM)* **–** Finally, teachers must choose an appropriate instructional response and be able to describe why that instructional intervention is designed to move students from their current level of understanding along the developmental trajectory towards greater understanding.

TASK is situated in the activity of looking at and responding to a carefully designed set of typical student responses to a mathematics problem in the specified content area. The student responses characterize distinct levels of sophistication of student thinking as well as common misconceptions that are supported by mathematics education research. The set of student work for the Multiplicative Reasoning TASK contains three correct solutions and three incorrect solutions to a word problem involving equal grouping (1-digit x 2-digit) (See Appendix A). The correct solutions include various levels of sophistication: drawing out the groups and counting by ones, drawing an open area model and multiplying the tens and ones separately, and using a related fact and compensating. The incorrect solutions include a correct strategy (skip counting) with an error, a misuse of the traditional multiplication algorithm reflecting a lack of place value understanding, and an incorrect modeling of the problem. In this way, the student work represents some of the important landmarks that have been identified in current research on children's learning of multiplication, common conceptual and procedural errors, as well as an overall progression from additive to multiplicative reasoning. Thus, TASK is designed to provide a realistic context from which to elicit information about what teachers pay attention to when they examine student strategies that they are likely to come across in their own classrooms.

Through an online instrument, teachers are presented with the six samples of student work and then led through a series of ten questions designed to measure these three key domains of knowledge related to the mathematical concept. Respondents move through several screens where the student work is shown along with the respective prompts. Open-ended responses for analysis of student thinking, ranking rationale, and instructional decision making are entered into text boxes.

The three domains measured by the TASK can be located in domains of Knowledge of Content and Students (KCS) and Knowledge of Content and Teaching (KCT) in the framework of Mathematical Knowledge for Teaching (MKT), proposed by Ball, Thames, and Phelps (2008) (shown in Figure 2). While we do not aim to measure the MKT domains in their entirety, the TASK measures their application in the context of formative assessment that is informed by learning trajectories. Sztajn et al. (2012) also propose a Learning Trajectory-Based Instruction (LTBI) interpretation of the MKT categories that in many ways parallels our conceptualization of these domains. In order to show how the TASK aligns with MKT and LTBI interpretations of MKT, the specific prompts from the TASK are shown in Table 1 along with the corresponding domain of teachers' formative assessment capacity that are assessed by each set of prompts, as well as where these domains are located in MKT and the learning trajectory conceptualization of MKT.

Figure 2 Mathematical Knowledge for Teaching



Note: This will be reshaded so that only KCS and KCT are shaded

Table 1 TASK Domains, Prompts, and Correlation to MKT and LTBI

TASK Domain	Number of Prompts	TASK Prompt	MKT	LTBI interpretation of MKT
Analysis of Student Thinking	3	Comment on four students' solution processes in terms of what the work suggests about their understanding of number and operations.	Knowledge of Content and Students	Content knowledge intertwined with knowledge of how students think about, know, or learn particular content.
Learning Trajectory Orientation—Rationale	4	Explain the rationale for the rankings between selected pairings of student work.	Knowledge of Content and Students	Knowledge of the various levels of the trajectories through which learners progress; knowledge of the cognitive steps that support development and of the ways learners approach certain tasks.
Instructional Decision Making	2	Suggest instructional next steps and explain the rationale for those steps for a student who has a correct, but less-sophisticated response to the	Knowledge of Content and Teaching	Knowledge of ways to support learners' cognitive development along the trajectory to help students' voices develop into mathematical perspectives; knowledge of how to select and target tasks to promote

	problem, and a student who demonstrates conceptual weakness in the response.	individual movement along the trajectory and content-rich classroom discourse.

Scoring rubrics were developed for each domain based on a four- or five-point ordinal scale to capture the overall orientation toward teaching or student understanding. These rubrics were developed from pilot data through both an inductive and deductive process and then further refined after field testing. Each rubric scale captures a continuum of depth, from a focus on general or surface characteristics of student work (correctness, format, or unrelated to multiplicative reasoning), to a descriptive focus (what the student did to solve the problem), to a conceptual focus (what the strategy suggests about student understanding) and finally a developmental focus (situating strategy within the learning trajectory)

The shift from procedural to more conceptual views of mathematics has long been promoted in mathematics reform literature (e.g., Hiebert, 1986; National Council of Teachers of Mathematics, 1988; National Research Council, 2001), and so a conceptual orientation toward student work was rated as higher than one that was only procedural. More recently, research on learning trajectories has promoted a developmental view, where students' conceptual knowledge develops in relation to instruction along a predictable path toward more complex and sophisticated thinking (Battista, 2011). Therefore, in order for a response to be at the highest level of the rubric, we determined that a teacher's focus on conceptual understanding must have evidence of drawing upon a developmental framework. We then had four ordinal categories (general, procedural, conceptual, and learning trajectory) that applied to each question on the TASK.

The rubric shown in Table 2 describes each of the TASK rubric categories. These categories are not mutually exclusive and teacher responses were rated in relation to the highest level of analysis present. (Therefore a response that contained general, procedural, and conceptual elements would be rated as conceptual.) Specific rubrics were developed for each of the three domains. For Analysis of Student Thinking and Learning Trajectory Orientation, a fifth category of *Early Conceptual* was created to capture the distinction between responses that had some general reference to concepts but the teacher's reference to those concepts was vague or not sufficiently articulated (e.g., stating the student "understands multiplication"). Responses in this category indicate that teachers may be paying attention to conceptual understanding, but do not have the knowledge or vocabulary to articulate it clearly.

For this study, we created three parallel forms of the Multiplicative Reasoning TASK to mitigate testing threat. The forms contain samples of student work that have the same array of strategies and problem structure, but reflect different numbers, student names, and a different order of presentation.

Table 2 TASK Rubric Levels

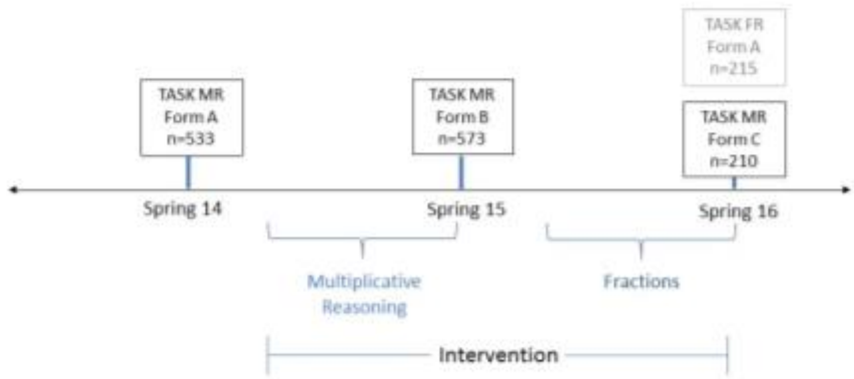
TASK Rubric Levels and Descriptions

Score	Category	Description
5	Learning Trajectory	Response draws on developmental learning trajectory to explain student understanding or develop an instructional response.
4	Conceptual	Response focuses on underlying concepts, strategy development, or construction of mathematical meaning.
3	Early Conceptual	Response contains some reference to conceptual understanding, but concepts are not sufficiently articulated to warrant conceptual rating.
2	Procedural	Response focuses on a particular strategy or procedure without reference to student conceptual understanding.
1	General	Response is general or superficially related to student work in terms of the mathematics content.

Methods

The Multiplicative Reasoning TASK was administered as part of an online baseline survey in the Spring of the year preceding the intervention and again in the Spring of each intervention year to all OGAP teacher leaders and grades 3-5 teachers, including Special Education and ESL, from 60 schools (30 treatment and 30 control). Response rates for all three surveys were over 80%. Figure 3 illustrates the timeline of the TASK administration and intervention. At the end of year 2, when the intervention focused on fractions, teachers were randomly assigned to take either the Multiplicative Reasoning TASK or the Fraction TASK.

Figure 3. TASK data collection timeline



As a school level intervention, we collected data from all grade 3-5 teachers, regardless of whether or not they attended all of the training or follow ups. Three raters who had masters degrees in elementary education and teaching experience were trained to code the 9 open-ended responses with the rubrics on training sets of actual teacher responses. TASK responses were scored by raters after they had established reliability of at least 80% direct agreement with an expert rater across all nine items.

At the end of Year 1, TASK was administered to 603 teachers from treatment and control schools. There were 79 teachers from the treatment group who completed the TASK in all three administrations. This sample reflects teachers from 27 of the 30 treatment schools with between 1 and 7 teachers from each school. The breakdown of this sample by grade level in comparison to the larger group is shown below in Table 3.

Table 3 Comparison of Year 1 Teacher Sample (*n*=603) and Treatment Sample (n=79)

	Treatment	**Control**	**Treatment BL, Y1, Y2**
Number of Teachers	285	248	79
ESL teacher	10.6%	10.0%	1.3%
SPED teacher	21.9%	15.8%	2.7%
3rd grade teacher	34.2%	36.9%	31.1%
4th grade teacher	34.8%	35.2%	36.5%
5th grade teacher	32.8%	29.3%	28.4%

Although this sample of 79 teachers was not intentionally constructed, it ended up being fairly representative of the treatment group overall. To explore the impact of the OGAP intervention on teacher knowledge over time, we analyzed the rubric scores and responses from these teachers across the three administrations.

[add Phil's analysis here]

We also conducted a qualitative analysis of the TASK responses to identify specific aspects of OGAP use. Responses from the three TASK administrations were grouped by teacher and then coded for patterns across the responses. The codes were developed inductively and deductively to hone in on aspects of OGAP use that were evident in responses to particular domains and questions. As shown in Table 4, for AST, all responses were coded for evidence of use of labels from the *OGAP Progression* (e.g., early additive, transitional), and the presence of conceptual analysis, but we also had codes that were specific to the nature of the student work. For example, in ASTC the student decomposes the factors and uses compensation (e.g., to solve 26 x 3, computes 25x 3 and adds 3). In addition to whether or not the teacher referred to this as a multiplicative strategy, we coded for whether it included analysis of the students conceptual understanding and/or mention of the distributive property, since this understanding was a focus in OGAP training.

This analysis offered slightly different information than the TASK coding in that we were not assessing the overall level of the response but rather evidence of use or take-up of the OGAP framework. For example, in the TASK coding, a response was given the learning trajectory code if the *OGAP Progression* was used correctly to determine the level of a student strategy. However, in our teacher level analysis, if a teacher misjudged the level it was still counted as use of the progression. For instructional decision making, we coded for whether or not the response suggested teaching a standard algorithm as the next step, which allowed us to identify cases where teachers' instructional recommendation remained stable over time or shifted from the standard algorithm to a strategy on the progression.

Table 4. Codes for Qualitative Analysis of TASK Responses

Question	Codes		Word search
ASTA	Use of *OGAP Progression* labels	Conceptual analysis	
ASTB	Use of *OGAP Progression* labels	Characterization of error as conceptual	Place value
ASTC	Use of *OGAP Progression* labels	Conceptual analysis	Distributive property
LTO (A-D)	Use of *OGAP Progression* labels	Change in ranking	Place value Distributive property
IDM (A,B)	Strategies from Progression	Standard algorithms	

Results: Impact on Teacher Knowledge

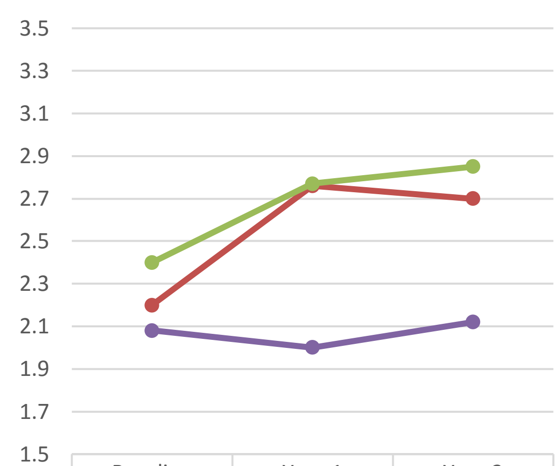
At the end of Year 1, on the TASK as a whole, as measured by the average score across the 9 open-ended items, the treatment group scored significantly higher at the end of the first year of treatment that control teachers when controlling for baseline (F(2,446)=64.54, p<.001, R2=.22; d=.77). Treatment teachers also scored significantly higher on each of the three domains. Data from the first year also indicates that implementation of OGAP in the treatment schools was highly variable. For example, on the survey administered at the end of Year 1 only 20% of teachers reported using OGAP assessment items with their students at least once a week, as the program model specifies. Half of the participating teachers (49%) reported using OGAP items about twice a month, and almost a third (31%) of teachers reported administering OGAP items to their students only rarely. Furthermore, only 22% of teachers in the treatment schools reported adhering to the twice-monthly PLC meetings specified in the program model (Supovitz, 2016). (Analysis on Year 2 TASK data is forthcoming.)

While these findings indicate that OGAP had an impact on teacher knowledge as measured by TASK, we conducted further qualitative analysis to understand more about the variation in these results and how teachers appropriated the learning trajectory framework as a conceptual and/or practical tool.

Teacher Growth Over Two Years

Within the group of 79 teachers who took the TASK at all three points in time, 51 (65%) demonstrated a positive change in the average TASK score from the baseline to the end of Year 2. 59 teachers (75%) showed growth from the baseline to the end of year 1 and 38 (48%) showed growth from the end of year 1 to the end of year 2. Figure 4 below shows the average TASK scores by domain for this group of teachers at each point in time. As the data show, the average score for Learning Trajectory Orientation (LTO) increased in both Year 1 and Year 2. While Analysis of Student Thinking increased after Year 1, there was a slight decrease in Year 2. Finally, Instructional Decision Making showed only slight changes in both directions. (While the IDM scores appear to be lower, IDM is out of a total of 4, while AST and LTO are out of 5.)

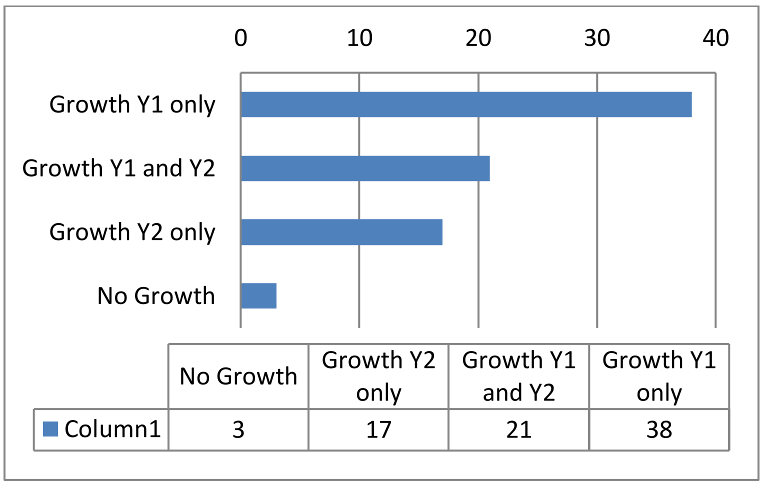
Figure 4. Average TASK scores for AST, LTO, and IDM over 3 years



	Baseline	Year 1	Year 2
AST Average	2.2	2.76	2.7
LTO Average	2.4	2.77	2.85
IDM Average	2.08	2	2.12

This sample of 79 teachers can be further divided into four groups to characterize the change in their individual TASK score over time: The largest group of 38 teachers only showed increase in TASK scores after the first year of the intervention. A smaller group of 21 teachers had a greater score after Y1 and then additional growth in Y2. There were also 17 teachers who only demonstrated growth in Y2. Two of the 3 teachers who showed no detectable growth had the same average score after Y1 and only a slight decline in Y2.

Figure 5. Relative Frequency of Growth in TASK Score Patterns over Two Years



Several themes emerged from the analysis of the 79 sets of teacher responses across all three TASK administrations that inform our understanding of what teachers may have learned from the intervention. These themes are summarized in Table 6 and then described below in relation to the three domains: analysis of student thinking, learning trajectory orientation, and instructional decision making. Since the student names and genders were changed in the different forms administered, those names and work samples are referred to with a letter (e.g., Student A).

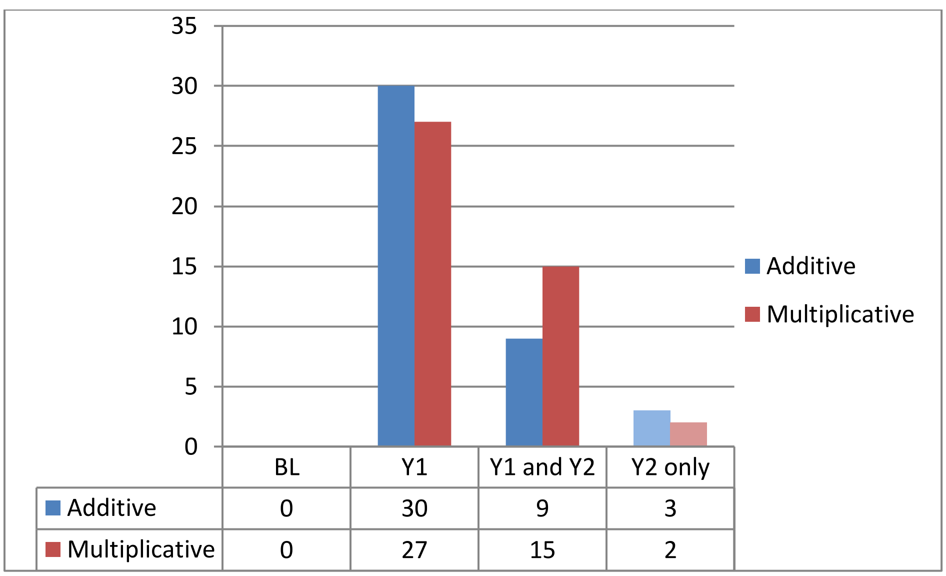
Table 6. Themes in Analysis across Teacher Responses

Domain	Theme
Analysis of Student Thinking	• Classification with labels from the progression • Paying close attention to evidence in student work • Recognizing limits in students procedural understanding • Instructional frame of mind (over evaluative)
Learning Trajectory Orientation	• Changing order of rankings • Using progression to justify rankings

	• Justifying rankings in relation to underlying conceptual understanding and reasoning • Balancing multiple cues from student work (e.g., strategy use and errors)
Instructional Response	• Using progression to identify next developmental step • Shift from teaching traditional algorithm to building on what student knows • Increased specificity in instructional response

Analysis of Student Thinking

The most prevalent change from the baseline to the Y1 and Y2 responses was evidence of teachers using labels from the *OGAP Progression* (e.g. early additive, transitional) to categorize and analyze the samples of student work. The words additive, transitional and multiplicative were not present in any of the baseline responses, but at the end of Y1 more than a third of the teachers correctly used the labels to categorize the two work examples that clearly showed an additive or multiplicative strategy.



The example below is illustrative of many of the cases where teachers showed this growth: they shifted from relatively vague descriptions of the student strategy (using pictures, making groups) to naming the strategy according to the progression: The teachers is analyzing Student A's strategy, which was to draw out each object in every group and count the total by ones, as evidenced by tick marks next to the objects.

BL		Y1
Student A drew a picture of the problem showing that there are 19 packs of erasers and 3 erasers in each pack. She understood the problem.	→	Student A is in the *early additive stage* modeling and counting by ones.

As this example, illustrates, the labels give teachers an efficient way to categorize student work in relation to learning trajectory. Interestingly, these labels were less prevalent at the end of Year 2, indicating that teachers may have been using the *OGAP Multiplicative Reasoning Progression* more in the first year than they were in the second year of the intervention. One explanation for this could be because the focus of the intervention during Year 2 was on fractions; alternatively, there may have been some fading out of the use of the progression. This does not necessarily indicate that these teachers were analyzing the student work at a lower level, only that they were not using the labels. A few teachers who did not use the labels at the end of Year 1 did use the labels at the end of Year 2, indicating that for some, additional take up of the *OGAP Progression* may have taken place in the second year.

A second theme that emerged was that many teachers were paying closer attention to the evidence in the student work, as shown by the following example. While in the baseline year, this teacher merely noted the presence of the picture and used that to conclude that the student "understands the concept" at the end of year 1, the teacher carefully analyzed that picture for evidence of understanding of multiplication and counting by ones. This teacher also recognized that the presence of the multiplication sentence (24 x 4 = 96) did not count as evidence that the student was actually using multiplication to solve the problem.

BL		Y1
Student A uses pictures to solve. Understands the concept	→	Student A makes 24 groups and draws 4 lines for each group to represent the legs on the elephants. I believe he is counting by ones because it looks like he went back and put a dot to count up all the legs. Student A gave a multiplication fact to find the answer, but did not get the answer by multiplying.

A third theme in analysis of the student work was recognizing the limits in student's procedural understanding. This was prevalent in the analysis of work where the student had attempted to use the traditional US algorithm but showed lack of understanding of place value in carrying out the steps and had an unreasonable solution. The student work also showed evidence of a less sophisticated level of understanding in the fact that repeated addition had been used to calculate a single digit multiplication fact. Many teachers initially analyzed the errors in the work as procedural and secondary to the fact that the student was using the algorithm, but then after the intervention recognized that the student was using a procedure without understanding. For the following teacher, this recognition only came at the end of the second year:

BL	Y1	Y2

Student C understand multiples. Difficulty carrying out the process to the traditional algorithm.	→	Student C has an understanding of the traditional algorithm for multiplication, but has made an error in the process of using it.	→	Student C doesn't understand that in 26 x 3 that he is multiplying 3 by both 6 and 20. He doesn't understand the value on the 2 due to his place value.

The response in Year 2 shows recognition that the error in the student work indicates a lack of place value understanding, something that was noted by only 9 teachers in the baseline, but then by 20 teachers in either Year 1 or Year 2. Several other teachers recognized that the student did not "have good number sense," might not be "ready" for the standard algorithm, and/or would benefit from more modeling or conceptual work.

Finally, we noted a tendency of teachers to talk about instructional implications of the student work even though this was not a focus of the prompt. For the example that had errors, several teachers suggested that the student might benefit from a model or more experience with equal groups. Likewise, for the student work that showed evidence of counting by ones, some teachers wrote about how this could be built upon to introduce addition or arrays.

Learning Trajectory Orientation-Ranking and Justifying Student Work

After the first year of the intervention teachers began to use the progression to help justify their rankings of the student work in terms of sophistication of reasoning. This resulted in changing the order of sophistication of the selected pairs of student work in relation to each other.

For example, when asked to compare two samples with correct solutions, Student B who used the open area model and Student E who used a more abstract strategy of using a related fact and adjusting through compensation, teachers who initially ranked B over E or ranked them equally, came to see E as more sophisticated after the intervention:

BL (B>E)		Y1 (E>B)		Y2 (E>B)
Student B is more sophisticated because he knew the lattice strategy of multiplication. Student E estimated to the nearest ten to make it easier for her to solve the problem. Then she found the related fact 2x3 to get 60 and then subtracted 3.	→	Student E uses multiplication and subtraction	→	Student E is multiplicative and Student B is transitional using open area model.

In addition to using the progression labels, this teacher seemed to develop a stronger familiarity with different strategies, learning to recognize the open area model (rather than thinking it was the lattice method) and seeing Student E's strategy as a multiplicative strategy rather than estimation.

In comparing Student E's multiplicative strategy to Student A, who drew out all the groups and counted by ones, most teachers could identify Student E as being more sophisticated even at the baseline. However, after the intervention, many were also able to articulate their reasoning in terms of the students' underlying level of reasoning and understanding. Note how the following respondent moves from stating that Student E understands the problem more abstractly to being able to articulate the student's underlying understanding of multiplication:

BL (E>A)		Y1 (E>A)		Y2 (E>A)
Student E understands the problem abstractly as 20 packages is 3 erasers less than 19 packages.	→	Student E understands 25 groups subtract one group is the same as 24 groups of four. He finds the fastest and easiest way to calculate and he may also be able to do this mentally without writing it down. A is counting by ones. Even though Student A drew equal groups of fours, there is no evident that he can count by fours. Student E can multiply 4 by 25 without any counting. Student E showed that he has a lot more experiences of family facts of four up to 25. Student E uses his prior knowledge to solve the problem creatively and quickly.	→	Student E is using multiplicative thinking. She understands that the conceptual algorithm of partial products of 25 x 3 and adding 1 group of three is equivalent to the product of 26 x 3. On the other hand, Student A still has to rely on drawing models of 26 equal groups of three to arrive at the answer. If Student A had just written the equation 26 x 3=78, he may be able to think as abstractly as Student E did. Even if Student A did that, Student E's ability to break 26 x3 into two equations proves to me that she has a better conceptual understanding of how multiplication and addition are related.

Similarly, the following respondent moves from stating that Student E didn't need to "draw it out" to using the progression to justify the rankings, and finally to identifying the use of the distributive property in this strategy. While there was no mention of the distributive property in any of the baseline responses for this question, it came up directly in 5 responses after the intervention and was described less directly by many others (e.g., "E understands she can break apart the number to create two easier problems").

BL (E>A)		Y1 (E>A)		Y2 (E>A)
Student E knew to use easier numbers to get the answer and didn't need to draw it all out.	→	Student E is in the multiplicative area and Student A is still transitional, or additive but using repeated addition.	→	Student E understood how to make the problem easier by using the distributive property. Student A had to draw it out.

Again, the use of labels from the progression was particularly prevalent at the end of Year 1, but by the end of Year 2, many teachers were also able to justify their rankings in terms of the underlying conceptual understanding in the student work without necessarily using the progression.

However, when teachers were asked to compare two students who had obtained the incorrect answer the results were slightly different: Student C used the traditional US algorithm incorrectly with place value errors while student D used skip counting but made an error in keeping track of the number to count. In other words, student C's error was conceptual while student D was computational and easier to address. On the other hand, student C was attempting to use a higher level strategy. In the following example, a teacher who initially ranked student C higher then used the progression after Year 1 and 2 to justify why student D was higher.

BL (C>D)		Y1 (D>C)		Y2 (D>C)
Student C attempts to multiply conventionally, even though she multiplies incorrectly. Student D does not attempt to multiply. Instead he relies on skip counting, which is not as sophisticated.	→	Student D uses a transitional strategy. Student C incorrectly uses the traditional multiplication strategy.	→	Student D uses an early transitional approach. He skip counts and is almost correct. Student C attempts to use a traditional algorithm but uses procedures incorrectly and is therefore non-multiplicative.

These responses from Y1 and Y2 reflect how teachers were taught to use the *OGAP Progression* in training. If a student attempted to use an abstract procedure but showed lack of understanding of that procedure, the response was classified a non-multiplicative. Making the comparison between these two pieces of student work requires paying attention to and balancing several different pieces of information: the level of the strategy attempted, the correctness of the answer, the nature of the error, the reasonableness of the answer, and additional information that Student C used repeated addition to find a simple fact. For example, another teacher explained how she balanced this information:

> I ranked Student D higher because even though they are both incorrect, Student D's answer is reasonable and she has just made an error in her skip counting. Student C is trying to use a traditional method without understanding and she is adding on the side.

However, many teachers continued to identify the use of the algorithm as being indicative of multiplicative thinking, either because the student was trying to use multiplication or because Student D was using a less sophisticated strategy : "Student D is LESS sophisticated in her thinking than Student C because she used an additive strategy, but skipped one group of 4 (from 40 to 48). Student C is multiplicative, but just regroups incorrectly." This example illustrates the complexity of analyzing and ranking student work in terms of sophistication, and the data suggests that many teachers may adopt the use of the progression somewhat procedurally, at least initially. Learning to classify student work by strategy is a new activity for most teachers, and as they make that shift they may do so to the extent that they ignore the other evidence on the page. Learning to balance and weigh strategy and errors may be a more complex activity that takes additional time to learn and master.

Instructional Implications

The most common theme in the instructional decision making responses was a shift from teaching students a standard algorithm (partial products, lattice, traditional US algorithm) to a developmental view. After the intervention, when teachers were asked what they would do next with Student A, who had drawn out all the groups and counted by ones to find the total, 57 teachers (72%) made an instructional recommendation that drew on the progression in some manner. 24 of those teachers had shifted from initially suggesting the student learn a standard algorithm or fact recall to recognizing that the student was not yet ready, and instead suggesting a more appropriate next step such as skip counting or introducing the open area model, as in the first example shown below. In other words, they drew on the progression to suggest transitional strategies rather than jumping all the way up to the multiplicative level. There were also several teachers who shifted from relatively vague instructional suggestions (moving away from using pictures, developing a quicker strategy) to recommending more specific strategies from the transitional level of the progression.

BL		Y1
I would instruct Student A on how to do lattice multiplication to solve this problem.	→	I would have Student A move from modeling to skip counting with modeling.

Even those teachers who seemed to take a developmental view before the intervention, over time were able to articulate the instructional next step more specifically in terms of the progression, such as in the example below.

BL		Y1
When I feel that she is ready I would start to move her from representational to abstract by showing her some methods to figure the problem out with just numbers and steps. It is clear that she understands the concept so helping her move to the next stage would be the next step.	→	I would try to move Student A from the additive stage to the early transitional stage by showing him how to use an area model to solve the problem. I would to do this to try to move him to the next stage of reasoning. I think the area model would be a good next step.

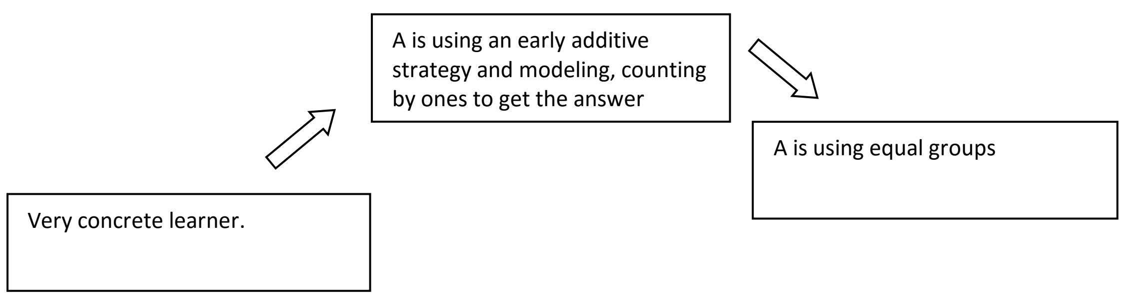
One teacher was able to go beyond merely suggesting the open area model, to explain how to transition the student from the drawing to the area model and the distributive property:

> I will have Student A visualize and communicate why he had made an array of 5 groups x 5 groups to get 25 groups. Treating the model as an array, I would ask him how many stickers are in 5 groups of 3, 10 groups of 3, 15 groups of 3, 20 groups of 3 and finally 25 groups of 3. Or how many stickers are in each row? each column? Example, if 26 x3 =78, what is 5 x 3, what is 10 x3 and what is 20 x3...I would also assess if Student A knows how to skip count by three and be able to physically match the number with the model as well as explain his thinking using his model.

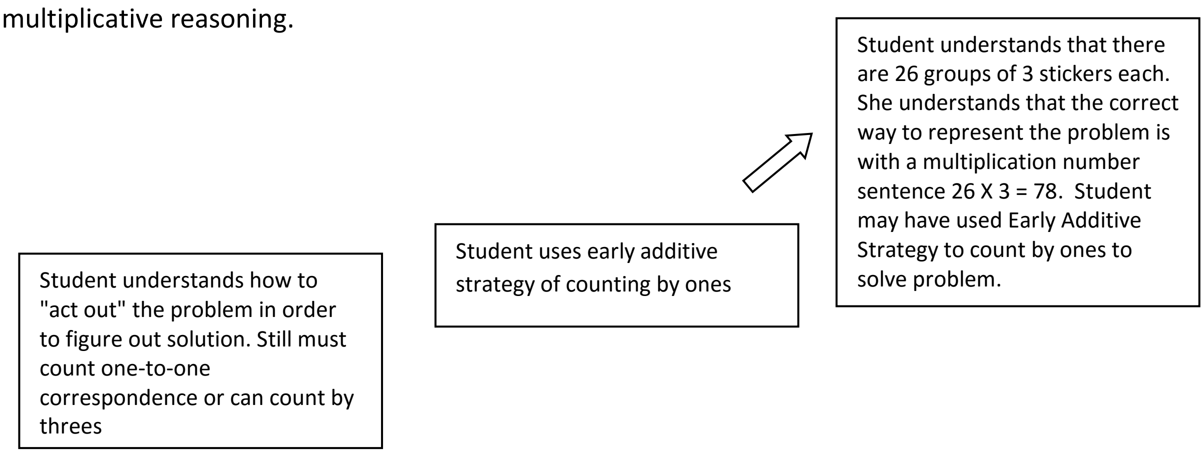
This shows a deeper level of instructional decision making that was relatively rare and goes beyond merely applying the progression, to demonstrate fuller understanding of the development of multiplicative reasoning. In other words, the teacher was appropriating conceptual underpinnings of the tool (Grossman, 1999).

Trajectories of Teacher Learning

Looking across TASK scores from all three time periods illustrates four distinct patterns of change over time. We illustrate these patterns with examples of teacher's analysis of Student A's work, the one that drew out all the groups and counted by ones to find the total. The first and most common path is characterized by teachers whose responses were rated higher at the end of Y1 but then slightly lower at the end of Y2. As the following example shows, many teachers in this group used the progression to analyze student work differently in Y1, but then not in Y2. (In this case, the teacher went from a general to a learning trajectory and then back to a descriptive or procedural analysis.) One possible explanation is that the Multiplicative Reasoning progression was forgotten with the focus being on fractions in Y2. Another explanation may be that the teacher or school put less focus on OGAP and formative assessment during Y2.



As stated earlier, a sizeable number (39) of teachers demonstrated increased scores from baseline to end of Y2 with more than half of those increases in both years. Looking more closely at responses from this group of teachers a common trajectory is an initial shift from general or procedural analysis to use of the progression to identify the student strategy and then in the second year, a shift from merely identifying the strategy to providing an analysis of what the student understands in relation to multiplicative reasoning.



In this case, and in several others, the teacher began by identifying that student is using drawings or pictures, but had nothing else to say about it, then moved to identifying strategy on progression in Y1, and then ultimately to understanding what that evidence (the picture) shows about both strategy and underlying understanding.

This pattern suggests a possible trajectory where introduction of tool supports categorization, focuses attention, and then over time leads to qualitatively different analysis in terms of student understanding. Could this be because they are paying more attention to student thinking? Recognizing student understanding in their classrooms?

A third trajectory that appears to be common is characterized by an initial shift to using progression in Y1 and that stays stable in Y2.

Student understands how to "act out" the problem in order to figure out solution. Still must count one-to-one correspondence or can count by threes

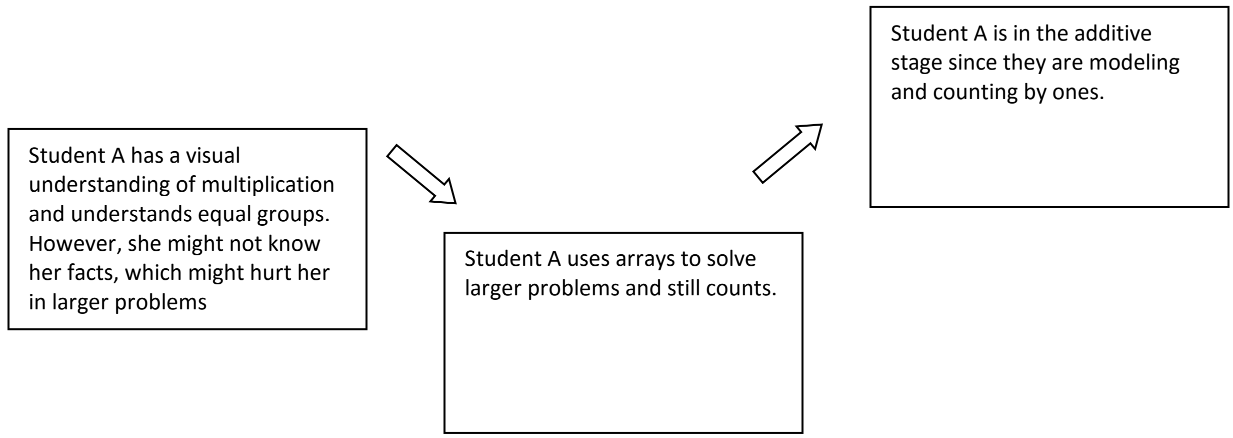
He knew the correct operation to use. He's appears to be in the early additive stages because he counted each leg by ones, but he knew how to write the equation.

Is in the additive stage of multiplicative reasoning

In this example, it may be that the categorization of student work in relation to the framework has become so automatic or routine that the teacher no longer needs to extract and explain all the evidence, or it may represent a slightly less sophisticated analysis. Either way, the framework is still in use at the end of Y2.

A fourth and final trajectory is characterized by no discernable or negative change in Y1 but then positive growth in Y2. In the example below, the teacher provided a descriptive and conceptual analysis of Student A's work at the baseline. At the end of year 1, the analysis seemed to draw on the progression but misidentified the strategy as an array and did not reference conceptual understanding.

It is only at the end of Y2 that the analysis correctly describes the strategy and places it in the additive stage, suggesting a stronger understanding of the progression.



This last example suggests that appropriation of the learning trajectory framework may take time, and may be procedural or focus only on surface features at first. Of course there were also those who demonstrated no significant change over time, indicating that the learning trajectory framework may have never appropriated--either because the teacher decided not take it up or did not have adequate time or opportunity to learn about it.

Discussion

Uncovering some of the patterns in how teacher responses in each domain changed after the intervention and the patterns of change in individual teachers' responses over two years, highlights both different levels of appropriation and some affordances and constraints of the learning trajectory as a tool for analyzing and responding to student thinking.

First, it is clear that the *OGAP Progression* gave teachers new categories to frame their interpretation and response to student work. The labels on the progression helped to both focus teachers' attention on student strategies and filter or organize that information in a progression. Several studies have found that classificatory talk pervades teacher workgroup conversations and that these categories tend to come from the way student performance is reported on high stakes standardized tests (e.g., below basic, basic, proficient, advanced) (Little, Bowker & Star, 1999; Horn, Kane & Wilson, 2013). This intervention introduced teachers to a very different way to categorize student work--not just on overall performance in relation to a standard but in relation to strategy, errors, and conceptual understanding. However, it is also clear that some teachers appropriated the OGAP Progression at a superficial level--what Grossman et al call "appropriating a label"--while others "appropriated conceptual underpinnings."

In addition, the *OGAP Progression* seems to have functioned as a tool that scaffolds teacher analysis and response to student work. Many teachers demonstrated shifts in their ways of thinking about student thinking. These shifts include the kind of evidence teachers pay attention to, what they value in student thinking, how they compare solution strategies, and the importance of building on student thinking rather than prescribing specific strategies. These shifts were particularly evident in relation to how teachers made sense of and responded to non-traditional strategies and in their views on procedural or algorithmic knowledge. Furthermore, many came to recognize both the importance of conceptual understanding and the limitations of procedural understanding.

The data also suggest that many teachers appropriated the tool in somewhat superficial ways, at least initially. Categories and labels can help to focus attention but also be constraining. As Weick describes, words are a powerful part of the sensemaking process: "Words constrain the saying that is produced, the categories imposed to see the saying, and the labels with which the conclusions of this process are retained" (p.106)." While teacher responses showed evidence of using the progression, it was not always used correctly or in the spirit of making sense of student thinking. Nevertheless, many teachers who initially used the progression to apply labels to student work also moved beyond strict categorization to demonstrate an ability to articulate student understanding from a conceptual lens. In these cases the progression functions as both a schema for categorization and a conceptual framework. This raises important questions--does using the progression for categorization of student work lead to conceptual appropriation over time? What other factors may have influenced the appropriation to focus on conceptual underpinnings rather than only surface features?

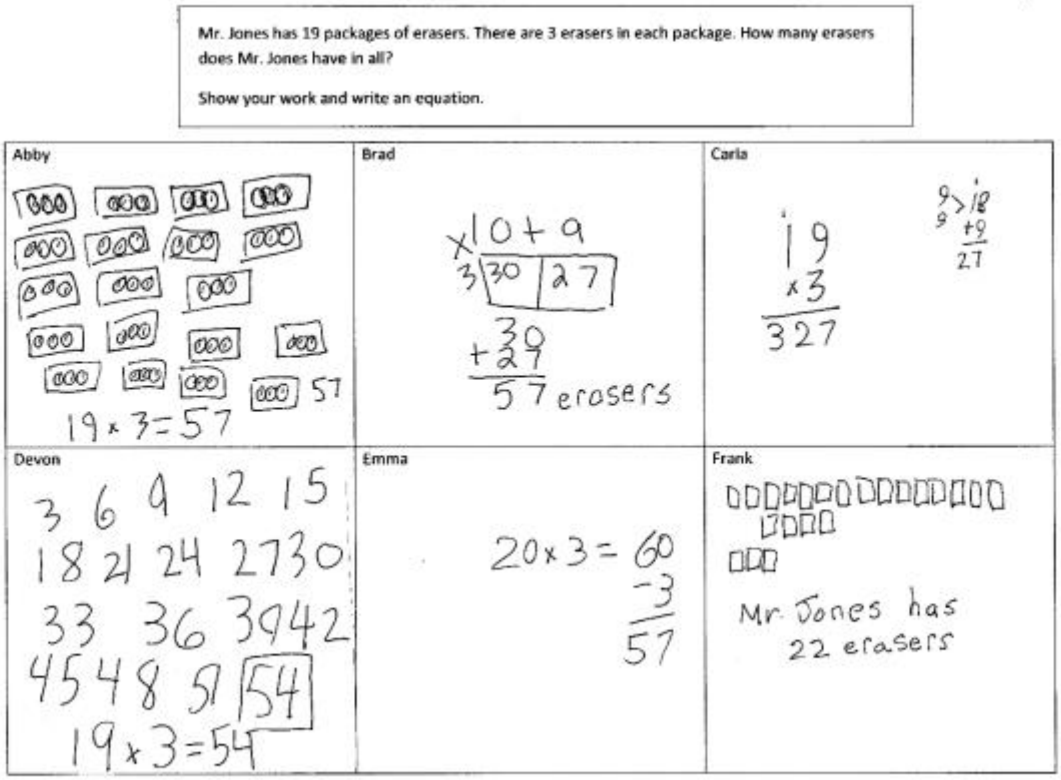
Finally, our analysis suggests that the TASK rubric, which was developed independent of this intervention and study, does not necessarily represent a developmental path--from general to procedural to conceptual to learning trajectory--as we might have predicted. Before the intervention, teachers were analyzing and responding to student work at general and procedural or descriptive levels. The intervention introduced teachers to a learning trajectory framework which led many of them to analyze student work in relation to this progression but some teachers appeared to do this without considering underlying conceptual understanding. Other teachers moved beyond this surface level usage to develop a more articulated analysis and response to student thinking, sometimes even dropping the labels by the second year. In this way the learning trajectory framework appears to act as a scaffold for deeper and more substantive interpretation rather than only serving as the endpoint of appropriation and mastery.

Implications

The findings of this study suggest that a learning trajectory framework can be a tool that supports growth in teacher knowledge. The shifts that are evident in teacher thinking are important not only for analyzing and responding to student work, but for developing ambitious instructional practices that center on student thinking. Teachers make sense of student thinking during whole class discussions, while working with individual students or small groups, and when planning for instruction. Current reform efforts, standards, and curriculum materials put value on having students use multiple strategies to solve problems. Teachers need the knowledge to be able to make sense of those strategies, know how to support student understanding of them, and move students to more sophisticated understanding over time. Teachers often remark that the newer curriculum materials introduce and have students practice all of the strategies represented on the *OGAP Progression*—e.g., repeated addition, skip counting, arrays, open area, partial products, standard US algorithm. A learning trajectory framework gives teachers a way to move beyond presenting students with a smorgasbord of strategies to organize and map these strategies in a way that guides and inform both formative assessment and instruction. The fact that so many teachers in this study shifted from seeing standard multiplication algorithms as a universal next step for students to seeing them as strategies that should be introduced only after students have developed the multiplicative reasoning to make sense of them is particularly important. The progression of student strategies based on understanding of place value and properties of operations culminating in standard algorithms is a central theme the Common Core Standards for Mathematics and important for more principled use of CCSS aligned curriculum materials.

The analysis reported in this paper was exploratory and raises several questions for further research. To explore change in teachers' analysis and interpretation of student thinking over time, we drew on averages of rubric scores where increases and decreases varied and were often small. Further analysis will draw on more sophisticated models to explore teacher growth and change over time. We also plan to further explore the trajectories that emerged from this study and the relationships between these patterns and implementation, student performance, and other measures of teacher knowledge (e.g., MKT).

Appendix A. Student Work Samples from MR TASK Form A. Form B and C had parallel but slightly different samples.



References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*, 398-407

Battista, M. T. (2011) Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiast, 8*, 507-570.

Black, P. B., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.

Carpenter, T.P., Fennema, E., Peterson, P. L., & Carey, D. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education*, *19*, 385-401.

Carpenter, T. P., & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Educational Research, 17*, 457-470.

Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. *Journal for Research in Mathematics Education*, *42*(2), 127-166.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and Council of Chief State School Officers.

Daro, P., Mosher, F. A., & Corcoran, T. (2011). Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction. (Research Report No.68). Madison, WI: Consortium for Policy Research in Education.

Ebby, C.B., & Sirinides, P.M. (2015). Conceptualizing teachers' capacity for learning trajectory-oriented formative assessment in mathematics. In J. A. Middleton, J. Cai, & H. Hwang (Eds.), *Large-scale studies in mathematics education*. New York: Springer.

Ebby, C.B., Sirinides, P., Supovitz, J. & Oettinger, A. (2013). Teacher Analysis of Student Knowledge (TASK): A measure of learning trajectory-oriented formative assessment. Technical Report. Consortium for Policy Research in Education

Goertz, M., Oláh, L., & Riggan, M. (2009). *From testing to teaching: The use of interim assessments in classroom instruction.* CPRE Research Report #RR-65. Philadelphia: Consortium for Policy Research in Education

Grossman, P.L, , Valencia, S. W., Evans, K., Thompson, C., Martin, S., & Place, N. (2000) Transitions into teaching: Learning to teach writing in teacher education and beyond. Journal of Literacy Research, 32, p. 631-62.

Grossman, P. L., Smagorinsky, P., Valencia, S. (1999). Appropriating tools for teaching English: A theoretical framework for research on learning to teach. *American Journal of Education, 108*, p. 1-29.

Heritage (2008). Learning progressions: Supporting instruction and formative assessment. Los Angeles, CA: National Center for the Research on Evaluation, Standards, and Student Testing (CRESST).

Hiebert, J. (Ed.) (1986). *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kazemi, E. & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education, 7*, 203–35.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, K., & Franke, M. L. (2010). Instructional explanations in the disciplines. In M. K. S. L. Kucan (Ed.), *Using designed Instructional activities to enable novices to manage ambitious mathematics teaching* (pp. 129-141). New York: Springer

Leont'ev, A. N (1981). *Development of the Mind.* Moscow: Progress Publishers,

National Council of Teachers of Mathematics (1988). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA; NCTM.

National Research Council. (Ed.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Putnam, R.T. & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*, 4-15.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22

Stein, M.K., Engle, R. A., Smith, M. & Hughes, E.K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, *10*, 313–340.

Supovitz, J., Ebby, C.B., Sirinides, P. (2014). Teacher Analysis of Student Knowledge: A measure of learning trajectory-oriented formative assessment. Retrieved from http://www.cpre.org/sites/default/files/researchreport/1446_taskreport.pdf.

Sztajn, P., Confrey, J., Wilson, P.H., & Edgington, C. (2012). Learning trajectory based instruction: Towards a theory of teaching. *Educational Researcher, 41* (5), 147-156.

Von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. New York: Falmer Press.

Weick, K. (1995). *Sensemaking in Organizations*. Sage Publications, Thousand Oaks: CA.

Wertsch, J. V. (1998). *Mind as action*. New York: Oxford University Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

Wilson, P.H., Lee, H.S., & Hollebrands, K. (2011). Understanding prospective mathematics teachers' processes for making sense of students' work with technology. *Journal for Research in Mathematics Education*, *42*, 39-64.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053-1089). Charlotte, NC: Information Age Publishing

Wilson, P.H. (2009). *Teacher's uses of a learning trajectory for equipartitioning* (Unpublished doctoral dissertation). North Carolina State University, Raleigh, N.C.