2020

Professional Noticing on a Statistical Task

Courtney Nagle
Penn State Erie

Stephanie Casey
Eastern Michigan University

Michele Carney
Boise State University

This is an Accepted Manuscript of an Article published in *Investigations in Mathematics and Learning* in 2020, available online at doi: 10.1080/19477503.2018.1534539

Professional Noticing on a Statistical Task

Courtney Nagle
Penn State Behrend

Stephanie Casey
Eastern Michigan University

Michele Carney
Boise State University

Post-Print Copy, please see and cite final manuscript:

Courtney Nagle; Stephanie Casey; & Michele Carney. (2020). Professional noticing on a statistical task. *Investigations in Mathematics Learning*, 12(1), 10-27. doi: 10.1080/19477503.2018.1534539

Introduction

Professional noticing of children's mathematical thinking is a core practice of high quality K-12 mathematics teaching (Jacobs & Spangler, 2017). Eliciting and using evidence of student thinking as a basis for instruction is one of eight effective teaching practices promoted by the National Council of Teachers of Mathematics (NCTM, 2014). A reason for this is that numerous research studies (e.g., Bobis et al., 2005; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Clements, Sarama, Spitler, Lange, & Wolfe, 2011) have demonstrated that teaching focused on students' mathematical thinking has positive outcomes for students and teachers.

Teachers' elicitation and use of student thinking is particularly important for topics where students are inventing their own approaches, such as when students develop their own processes for informally fitting the line of best fit. As stated in the Common Core State Standards for Mathematics (CCSS-M) (National Governors Association (NGA) and Council of Chief State School Officers (CCSSO), 2010), eighth grade students in the United States are expected to informally fit a straight line for scatter plots which display a linear association. When students are informally fitting a linear model, they are not using technological tools to automatically determine the line for them; instead, students devise their own processes for fitting the line. Thus, teachers must attend to student thinking regarding their processes, considering what understandings students have that are influencing how they place the line and interpret meaning from this process. Moreover, teachers must then formulate a response that builds on the student thinking and moves the student forward to a generalizable understanding of what the line of best fit is and how it is placed. The current study looked to gain a sense of this teaching practice, including studying teacher responses to students who use common processes for fitting a line of best fit. The purpose of this study is to investigate how secondary inservice teachers (hereafter, referred to as "teachers") interpret and respond to student thinking on a statistical task involving the informal placement of a line of best fit. We believe this is the first study to investigate teachers' professional noticing skills in the context of a statistical task.

Professional Noticing of Children's Mathematical Thinking

Professional noticing of children's mathematical thinking involves three skills that are connected both temporally and in practice: attending to children's strategies presented in their written or verbal work, interpreting this work to understand the thinking of the children, and responding to children based on their understandings (Jacobs, Lamb, & Philipp, 2010). Teachers' professional noticing skills are intimately tied to other teacher characteristics such as their orientations (including beliefs) and resources (including knowledge) (Schoenfeld, 2011). For our study, the teachers were given written student work in isolation (i.e., not embedded in documentation of a classroom session in general) that the research team selected; thus, the teachers did not have to attend to student thinking within the complexity of a classroom context. Instead, we narrowed our focus to how teachers interpret and respond when provided with students' written work on a statistical task.

While interpreting, teachers make sense of the student work to consider the student's understandings that informed the work. It is important that the interpretation be evidence-based and not extrapolate beyond the evidence at hand (Jacobs et al., 2010). The response is how the teacher replies to the student whose thinking was interpreted. Like Jacobs and colleagues (2010), we studied teachers' intended responses as they are not actually responding to the students who produced the work. Since interpreting and responding are consequential skills (Jacobs et al., 2010), it is important to investigate how a teacher's interpretation informs his or her response, as we did in this study.

Previous research found that teachers often interpret student work with a goal of responding in an evaluative way, such as telling the student whether their work is right or wrong (Franke, Kazemi, & Battey, 2007; Sherin and van Es, 2009). Teachers typically evaluate the correctness of the student's response, then in turn tell the student what to do. Rarely do they address conceptual aspects of student thinking or build from what the student understands (Franke et al., 2007; Jacobs et al., 2010). To address this, AUTHORS (2018) described four characteristics of a good response: a) works toward student learning objective; b) draws on and is consistent with the student thinking presented; c) draws on and is consistent with research on students' mathematical development; and d) proposed interaction with student leaves space for student's future thinking. These four characteristics synthesize work in the field regarding desirable teacher responses (e.g., Jacobs, Lamb, Philipp, & Schappelle, 2011) and include elements teachers commonly struggle with in crafting effective responses.

Teacher Knowledge Concerning Informal Line of Best Fit

Since teachers' professional noticing of student thinking is associated with teachers' knowledge, prior research by Casey and Wasserman (2015) concerning teachers' knowledge of the informal line of best fit is relevant for this study. Of the nineteen teachers in their study, most conceived of the line of best fit as the best representation of the sample data (7 teachers) or as a model to show the general relationship between the variables (5 teachers). When placing an informal line of best fit, the teachers primarily used three criteria that determined their process for placing the line: equal number of data points above and below the line, closest to all the points, or sum of deviations of points equal for points above and below the line.

Present Study

Our study builds on past literature to explore how teachers' interpretations of a student's work relates to their response to the student in the context of a statistical task requiring the placement of an informal line of best fit. In particular, our study addresses the following research questions: (1) How do teachers interpret student work on the line of best fit task?; (2) How do teachers respond to students on the line of best fit task and what characteristics of good responses do teachers use?; and (3) What relationships exist between how teachers interpret student work on the line of best fit task and how they respond to the student?

Methodology

Participants

The 42 participants were inservice mathematics teachers from the Northwest region of the United States enrolled in a professional development course on data analysis and statistics offered in the summers of 2016 and 2017. The teachers came from a mix of urban, suburban and rural school districts and most had at least three years of teaching experience. The teachers indicated the grade levels they taught during the previous school year: 24 indicated middle school (grades 6-8), nine indicated high school (grades 9-12), six indicated both (grades 6-12), and three indicated 'other.' Before the course began, they completed a knowledge inventory on their own time through a web-based survey; it included the instrument for the present study. The inventory directions requested they describe their response to each prompt and they had unlimited time to complete it.

Instrument

Teachers were presented with a task involving informally placing a line of best fit, the work of two students who had completed the task, and a series of prompts. The task and student work came from AUTHOR's (2015) previous study. In the earlier study, task-based interviews were conducted with thirty-three middle school students to investigate students' conceptualizations of the line of best fit. In the interviews, the students were presented with a series of five tasks that directed them to informally fit a straight line model to scatter plots and explain their reasoning as they did so. The two student work samples used in the present study were purposely selected because they included elements of student thinking that could be built upon for instruction, but also had elements that needed to be addressed or redirected. They also were chosen because they present common approaches students use when informally placing the line of best fit (AUTHOR, 2015). The first sample of student work comes from Student M, who explained that "I put the line through the first and last points." This strategy was also used by two other students to place the line on this task in the previous study. Student N's approach, given as the second sample of student work, entailed placing the line so that half the points were above the line and the other half below the line. This was one of the most common approaches taken by students in the previous study.

Teachers in the present study were asked to work through the task (see Figure 1) and respond to a prompt asking the important mathematical idea(s) the task targeted.

[Insert Figure 1].

Teachers were then presented with Student M's response to the task, including the transcription of the student's verbal response and an image of the placed line of best fit on the scatterplot (see

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning,* published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

Figure 2). Next, they were prompted to describe (a) what Student M's work indicated about her mathematical understanding (interpreting student work), and (b) how they would respond to her (responding to student work). These prompts are similar to those used in Jacobs and colleagues' (2010) study of elementary teachers' professional noticing skills. Next, Student N's work for the task (see Figure 2) with the corresponding prompts related to interpreting student work and responding to student work were presented. Our study's approach of having teachers engage with and share their professional noticing about researcher-selected artifacts of student work is one of three major approaches to investigate professional noticing (Sherin, Russ, & Colestock, 2011). An advantage of this approach is the facilitation of comparisons of professional noticing across teachers (Jacobs & Spangler, 2017). For more information on the validation work related to this style of assessment see Authors (2017).

[Insert Figure 2]

Analysis

Analysis of the teachers' replies provided an opportunity to study teachers' professional noticing skills for student thinking about a statistical task. After removing blank replies, there were 82 replies to the interpreting student work prompt (41 each for Student M and N) and 80 replies to the responding to student work prompt (40 each for Student M and N). The analysis was carried out in three phases.

Phase 1 – Open coding. Phase one of analysis involved inductive category formation, or open coding, to reveal potentially important aspects of teachers' replies to the prompts (Mayring, 2015). Open coding was initially conducted on a subset of replies with a goal of developing a coding schema that worked across both student work samples. Data were analyzed to identify

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning,* published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

categories of replies for the prompts related to (a) interpreting student work, and (b) responding to student work, with a focus on identifying categories likely to be relevant and/or impactful to classroom instruction. Enacting an open coding process, the authors read through the replies multiple times, made notes, and had iterative discussions regarding ways to categorize the data.

Through the open coding process, we noticed replies tended to be evaluative, noting perceived affordances and/or deficits in a student's understanding or process. In addition, while one prompt asked teachers to describe students' mathematical understanding, teachers' replies sometimes included a focus on process, either in isolation or in conjunction with understanding. Attention to process was indicated by a description of a student's technical approach to placing the line, while attention to understanding was indicated by a description of the conceptions students were perceived to have that influenced their process of placing the line.

Another element that rose to our attention through the open coding process was the wide variation in the quality of the teachers' responses to the students. We decided to use AUTHOR'S (2018) categorization schema to document characteristics of high quality responses to student work. We looked for three characteristics: (a) works toward student learning objective, (b) draws on and is consistent with the student thinking presented, and (c) proposed interaction with student leaves room for student thinking. We omitted AUTHOR's fourth characteristic-drawing on research on student development- because teachers participated in this study prior to completing the professional development course and we did not feel it was reasonable to assume teachers would be knowledgeable about this prior to the course.

Phase 2 - Application of coding scheme. Once the coding scheme was developed in Phase 1, it was applied to the entire data set. For the interpreting student work prompt, the coding scheme contained four categories: (a) *evaluative affordances*, (b) *evaluative deficits*, (c) *interprets student process*, and (d) *interprets student understanding*. For the response to student prompt, the coding scheme had three categories which noted whether the response met selected characteristics of a high quality response: (a) *works toward student learning objective*, (b) *draws on student thinking*, and (c) *open-endedness*. To facilitate examination of the relationships between categories, numeric scores were assigned to each category. Scores of 0 or 1 were used for categories where it was possible to say a characteristic was present or not (i.e., *evaluative affordances/deficits*, *draws on student thinking*, *open-endedness)* and scores of 0-2 were used for categories where it was necessary to distinguish quality of responses within the category (i.e., *interprets student process/understanding*, *works toward student learning objective*). See Tables 1 and 2 for the coding schemes for each prompt with category descriptions, levels, and examples.

[Insert Table 1 and Table 2]

All the data was coded individually by the three authors before comparing coding results. Intercoder reliability (indicating agreement by all three authors on initial coding) for the interpreting student work prompt was: (a) 98% for *evaluative affordances*, (b) 100% for *evaluative deficits*, (c) 93% for *interprets student process*, and (d) 88% for *interprets student understanding*. For the responding to student prompt, intercoder reliability was: (a) 79% for *works toward student learning objective*, (b) 86% for *draws on student thinking*, and (c) 96% for *open-endedness*. When discrepancies occurred in the level applied within a category, the authors discussed the discrepancy until consensus was reached. This occasionally resulted in modification to language in the rubric, allowing for iterative improvement of the coding scheme based on the data. All relevant data was recoded following changes to the coding scheme.

Through the application of the coding scheme, we also noticed potential areas for analysis of the interplay between teachers' interpretations of the student work and the associated response to the student.

Due to our interest in examining how teachers' interpretations of student work could be related to how they responded to the student, we created an overall response to student code with categories of low, medium, and high to more generally represent how effectively teachers responded to students. A teacher's overall response to student code was determined by summing scores across the characteristics of a high quality response – *works toward student learning objective* (0, 1, or 2), *draws on student thinking* (0 or 1), and *open-endedness* (0 or 1). These scores were summed separately for responses to student M and N. The resulting totals were then binned into three categories: low (total of 0-1), medium (total of 2), and high (total of 3-4). While binning the total loses some degree of specificity, it provides a more general picture to examine the trends in teachers' responses.

Phase 3. Analysis of the coded data. The codes were analyzed by looking at frequencies within the categories, and relationships between frequencies across categories, both within prompt and between the two prompts. Graphs were created to depict these relationships.

Results

This section presents the results of our analysis of teachers' professional noticing of student work on the line of best fit task as it pertains to our three research questions.

How do Teachers Interpret Student Work on the Line of Best Fit Task?

This section sheds light on our first research question by analyzing how teachers responded to the interpreting student work prompt. As a whole, teachers' interpretations of student work accurately depicted the process used by both Student M ("…they tried to have the line hit the first and last dot...") and Student N ("They understand that … there should be about an equal amount of data points on each side of the line."). However, interpretations of what that process reveals about student understanding varied a great deal and were often unsupported.

Some teachers broadly interpreted the work as indicating that Student M/N understands or misunderstands lines of best fit in general ("They have an idea of what best fit means but need some refinement."), without addressing what that understanding (or lack thereof) entails. Others more specifically described aspects of what Student M/N does or does not understand, but these aspects varied widely and sometimes focused on aspects that were irrelevant to the task at hand. Teachers had a particularly difficult time describing aspects of Student N's understanding. The teachers tended to either praise the approach as completely correct ("Student N is understanding the concept of best fitting line for the data given and is using the splitting of dots evenly to display this concept.") or only provided a descriptive account of the process the student used ("The student has placed the spaghetti with even amounts of data above the spaghetti as below the spaghetti."), never addressing what understandings the student had that resulted in the student using that method. Further analysis of teachers' interpretations of student work are outlined below, as they relate to the four coding categories for interpreting student work.

Evaluative Affordances and Deficits. Even though the prompt did not ask teachers to evaluate the student's work, the clear majority (78%) of teachers' replies made an evaluative statement. Recall that a teacher's reply was coded as evaluative if it acknowledged a perceived affordance/deficit in the student's work, regardless of whether the researchers agreed with the evaluation. Affordances were acknowledged (59% of teacher replies) more often than deficits

(45% of teacher replies) in interpretations of each student's approach to placing the line of best fit. Table 1 includes sample responses that do (Level 1) and do not (Level 0) include *evaluative affordances* and *evaluative deficits*, respectively. Despite the large percentage of replies that included an evaluative component, only 26% attended to both affordances and deficits, as seen in this example: "Student M's work indicated that he looks at the progression of the data points but does not understand the line of best fit. He is not seeing how all the points correlate." This response acknowledges that the student is doing something worthy ("looks at the progression of data points") but also names a deficit ("not seeing how all the points correlate"). In terms of evaluative responses, the largest category (33%) were those with an affordance-only evaluation of the student's approach. Although not as common, some replies (19%) included only descriptions of perceived deficits within a student's approach. The 22% of replies which did not provide any type of evaluative comment only went as far as describing the student's approach; for example: "The student is trying to use some type of average to make their trend line."

Student Process and Understanding. Since the prompt specifically asked teachers to describe what the work revealed about the student's mathematical understanding, it is not particularly surprising that 77% of the replies (17% at Level 2 and 60% at Level 1) did so (*interprets student understanding*). Somewhat less expected, 69% of the replies (45% at Level 2 and 24% at Level 1) attempted to describe the process used by the student to generate the solution (*interprets student process*). The Level 1 replies for *interprets student process* and *interprets student understanding* in Table 1 are illustrative of replies which attempted to describe the process or understanding used by the student but which are incomplete or not well supported by the student work. For instance, one teacher replied that Student M "...used the spaghetti and placed it close to the plotted points picking 2 points that he/she believed to be part of the actual line of best fit." While this teacher is attending to the process used by Student M, this description is incomplete because her work sample specified that she placed the line through the first and last data points. Another teacher claimed that Student M "...understands that there is a relationship between the drop height and bounce height." This reply is illustrative of interpreting student understanding in an unsupported way since Student M's approach of placing the line through the first and last points does not provide evidence of considering a relationship between variables.

The teachers were more successful in providing Level 2 interpretations of a student's process (nearly two-thirds of those who *interpret student process* were at Level 2) compared with the underlying understanding (less than one-fourth of those who *interpret student understanding* were at Level 2). Altogether, only 17% of the replies interpreted understanding in a supported way (Level 2), and the majority of all replies (60%) attempted to describe what the solution revealed about the student understanding in unsupported or incomplete ways (Level 1).

Further investigation into the 49 (60%) Level 1 *interprets student understanding* replies indicated that these replies could be classified as *incomplete*, *overreaching*, or *contradictory*. *Incomplete* replies were most common (32) and included descriptions of the student understanding that the researchers felt were consistent with the presented solution but not specific enough to help explain why the student gave the particular solution. For example, the reply that Student M "understands a straight line is determined by two points" is consistent but is incomplete in that it doesn't note which two points were chosen. A much smaller number of replies were *overreaching* (12) in that they attributed understanding (or a lack of understanding) that extended beyond what was reasonable to assume. "The student has an idea that a line will fit with an equal distribution of variability above and below the line, but lacking an understanding of how to do that." is an example of a reply that reaches beyond what can be reasonably drawn from Student N's solution to the problem. There is no indication from the work that Student N was attending to variability when placing the line of best fit. Finally, five replies were *contradictory* to the understanding demonstrated in the student solution. For instance, one reply stated that "[Student N]... was perhaps troubled that some dots didn't fit on the perfect line." Rather than just overreaching, this reply contradicts Student N's solution of placing the line to have an equal number of points above and below the line rather than trying to fit any points perfectly on the line. There was no evidence that the student was troubled by this.

Despite the large number of replies that attempted to interpret the student's process and understanding, less than 10% of all replies did so in a way that was well supported for both categories. An example of a reply that was well supported for both *interprets student process* and *interprets student understanding* is:

> "I think student M is only considering the first and last data points as important to create a line of best fit. The student does not appear to be taking the data set as a whole into consideration. They may know that a line is determined by two points, however if the student always chooses the first and last points to fit a line then they do not understand the line of best fit as it relates to the entire data set."

This reply acknowledges both a process (i.e., passing the line through the first and last data points) and interpretations of understanding (i.e., lines pass through two points but not recognizing that lines of best fit should relate to all points) that are well-supported by the student's solution.

To summarize, teachers' replies to the interpreting student work prompt indicated that they tended to be evaluative and attempted to describe both the process the student used and the understanding that supported it. However, teachers' descriptions of student understanding were often poorly supported by the student work. Teachers were somewhat more likely to point out the positive aspects of a student's approach rather than the negative aspects, and a relatively small percentage pointed to both aspects.

How Do Teachers Respond to Students on the Line of Best Fit Task?

This section reports the results for the second research question by analyzing teachers' replies to the responding to student work prompt. The results are reported for three desired features of teachers' responses to students: *works toward student learning objective*, *draws on student thinking*, and *open-endedness*.

Works Toward Student Learning Objective. A teacher's response needed to meet two conditions to be coded Level 2 for *works toward student learning objective*. First, the response must aim to convince the student that the strategy that he/she used would not work in all situations and second, the response must press the student toward a generalizable strategy. Most responses (51%) satisfied neither of these conditions for *works toward student learning objective* (Level 0). Many of the responses either attempted to convince the student that the strategy was not generalizable (15%) or to move toward a generalizable strategy (30%) but did not incorporate both components (Level 1). Only 3 out of 80 responses (4%) satisfied both conditions for *works toward student learning objective* (Level 2). Table 2 provides a sample Level 2 response that leverages a counter-example to address the non-generalizability of the strategy used by the student and provides a platform to find a more generalizable strategy.

Since most responses did not effectively *work toward the student learning objective*, additional qualitative analysis of the responses was conducted. In lieu of working toward the objective, most of the teachers' responses either provided evaluative statements that simply told the student they solved it correctly ("great job") or probed the student to describe their work further. This latter category was generally done in a generic way, with teachers claiming they would ask the student to "explain your thinking" but without providing any sense of how they would follow-up on the student's reply. In addition, some responses included irrelevant or unhelpful questions for the students, such as "Are there any other ways you could use the spaghetti?" Finally, some teachers prompted the students to consider ideas that were irrelevant to the task or unhelpful in meeting the learning objective. Some of these responses mentioned statistical topics irrelevant to the task at hand, such as suggesting the student collect more data, look for outliers, or calculate the standard deviation (e.g., "I would ask them to analyze the data and see if there appear to be any outliers. I might also ask them to find the standard deviation…"). Other irrelevant responses focused on mathematical aspects (e.g., slope or intercepts) that were unrelated to the objective of placing the line of best fit (e.g., "...I would need to ask him/her what points he/she will use to determine the slope of the line" and "I would ask them about the slope and to interpret its meaning, and what the y intercept of the line would be…").

Draws on Student Thinking and Open-Endedness. Most teachers (55%) did not *draw on student thinking*, particularly noticeable in teachers who provided identical responses to Student M and Student N despite the students' different approaches to solving the task. For instance, one teacher responded to both students by saying "I would ask why the student chose this line of fit and what this information would tell them. Also is there another line of fit that would work?" Despite not drawing on student thinking, this teacher's response was open-ended and a large percentage of teachers' responses had *open-endedness* (80%). It is noteworthy that many of the responses that were open-ended posed generic questions that only trivially left room for additional student thinking, such as "Could you find a better line?" Additional examples are provided in Table 2.

Interactions Within Characteristics of Teachers' Responses. There were some noteworthy relationships between the various characteristics of teachers' responses to the students, primarily related to how *draws on student thinking* related to the other two components of the responses. First, 89% of the teachers who drew on student thinking in their response also responded in an open-ended manner (up from 80% across all responses). Additionally, Figure 3 illustrates a relationship between *draws on student thinking* and *works toward student learning objective*. Of the teachers who did not *draw on student thinking*, 66% did not satisfy either condition for a response that *works toward the student learning objective* (Level 0), 34% satisfied one condition (Level 1), and none of them satisfied both conditions (Level 2). This distribution changed for teachers who did draw on student thinking: one-third of these teachers met neither condition while the remaining two-thirds satisfied either one (58%) or both (8%) conditions for *works toward student learning objective*.

The relationship between *open-endedness* and *works toward student learning objective* does not appear quite as strong; see Figure 4. Although all three teachers who satisfied both conditions for *works toward student learning objective* (Level 2) did have open-ended responses, teachers whose responses were open-ended were slightly less likely to meet at least one condition for *works toward student learning objective* (Level 1 or 2) compared with teachers whose responses were not open-ended (47% versus 56%).

[Insert Figure 3 and Figure 4]

To summarize, teachers' responses to student work tended to include broad, open-ended remarks, but were unlikely to draw on student thinking or work toward the learning objective. Drawing on student thinking seemed to be beneficial for teachers in that it was associated both with responding in an open-ended way and with satisfying at least one condition for working toward the learning objective.

What Relationships Exist Between How Teachers Interpret Student Work and How They Respond to the Student?

In this section, we address the third research question by describing associations between teachers' interpretations of and responses to student work.

Figure 5 illustrates the relationships between a teacher's evaluative interpretations and their overall response score. The trends in Figure 5, graph 1 suggest that teachers who noticed both evaluative affordances and deficits of a student's solution (labeled as 'Both' in this graph) tended to have higher overall response scores compared to teachers who did not make an evaluative comment or made only one type (labeled as 'Neither or Only 1' in this graph). Although the highest overall response score was equally distributed among teachers who did or did not attend to both affordances and deficits of the student's approach, the highest overall response score represented nearly 50% of all teachers who included both affordances and deficits but only 17% of all remaining teachers. Similarly, a much lower percentage of teachers who attended to both affordances and deficits had low response scores when compared with teachers who did not attend to both components.

Further analysis of the relationship between the evaluative interpretations of student work and the overall response score follows from analyzing graphs 2 and 3 in Figure 5. Teachers who noted *evaluative deficits* trended toward higher overall response scores. Figure 5, graph 2 shows that the largest percentage of teachers who did not attend to deficits had low overall response scores, while the largest percentage of teachers who did attend to deficits had high overall response scores. While a relationship between attending to deficits and response score is clearly present, Figure 5, graph 3 does not support the existence of a relationship between *evaluative affordances* and overall response score. This can be seen by the large percentage of teachers who attended to affordances but who had low overall response scores (versus medium and high scores), as well as by the relatively similar distribution of overall response scores regardless of whether teachers did or did not attend to affordances. This lack of relationship may be due to the generic nature of many of the evaluative affordance responses (e.g., “Nice job!”).

[Insert Figure 5]

Associations between how a teacher interpreted the student’s process or understanding and the teacher’s response to the student also emerged. Figure 6, graph 1 shows a relationship between *interprets student process* for placing the line of best fit and overall response score, with the distribution of overall response score shifting higher (greater score) as the *interprets student process* score increased. By contrast, the relatively consistent distributions of overall response score as *interprets student understanding* varies in Figure 6, graph 2 suggests that interpreting student understanding is not related to the overall response score. However, we read this finding

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning,* published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

cautiously and suspect this may be related to the relatively low number of teachers who attended to student understanding in a manner supported by the evidence (i.e., the small number of teachers with Level 2 *interprets student understanding* scores).

[Insert Figure 6]

Given the suggested relationship between interpreting the student process and overall response score, it is useful to consider whether that relationship is consistent across the various sub-components of the overall response score. Figure 7 provides an illustration of how teachers' scores for *interprets student process* related to whether their responses *draw on student thinking* (graph 1), *work toward the student learning objective* (graph 2), or have *open-endedness* (graph 3). The relationship between *interprets student process* and *draws on student thinking* appears strong, as can be seen by comparing the percentage of teachers who drew on student thinking as the process score increases from 0 to 2. Only 17% of those teachers who did not *interpret student process* (Level 0) *drew on student thinking* in their response. This increased to 47% and 62% of those teachers who partially (Level 1) and fully (Level 2) *interpret student process*. A relationship between *interpret student process* and a response that *works toward student learning objective* is also visible. As the *interprets student process* score increases, scores for *works toward student learning objective* tended to increase (Level 0 to Level 1). Recall that only three teachers met both conditions for *works toward student learning objective*, making it difficult to extend this relationship to level 2 scores.

Figure 7, graph 3 suggests there is not a strong relationship between *interprets student process* and *open-endedness* in the teachers' responses to the student. Although no interpretation of student process (Level 0) is slightly more suggestive of a response that is not open-ended, the

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning*, published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

percentage of teachers who responded in an open-ended manner was the same regardless of whether a teacher interpreted the student process in an unsupported (Level 1) or supported (Level 2) manner. Overall, the additional analysis into how *interprets student process* relates to the various components of a teacher's response suggests that the relationship between *interprets student process* and overall response score is largely due to the relationship between *interprets student process* and responding in a way that *draws on student thinking* and *works toward the student learning objective*.

[Insert Figure 7]

Discussion and Implications

Our study extended research on teachers' professional noticing to the domain of statistics. When presented with student work on the line of best fit task, teachers often responded to students in generic ways that did not reference the students' work and related understandings. This coincides with Jacobs, Lamb, and Philipp's (2010) finding that nearly all (97%) inservice teachers did not provide robust evidence of drawing upon students' work in their responses prior to professional development regarding children's mathematical thinking.

Most teachers' responses also did not address the lack of generalizability of each student's strategy for placing the line of best fit nor did they move the students toward a more generalizable strategy. Moreover, the teachers often responded by focusing the student's attention on tangential ideas, including mathematical concepts that were not directly related to the task at hand. Our analysis of how teachers interpreted students' work shed some light on possible reasons their responses fell short of moving students toward the learning objective. One reason could be that many teachers may not have recognized the deficits in students' work.

Although 88% of teachers' replies were evaluative (in line with past research findings from Franke et al., 2007; Sherin and van Es, 2009), teachers acknowledged deficits in students' work much less frequently than they acknowledged affordances. Common generic responses of teachers who only provided positive affirmation of the student's work (e.g., "Nice Job!") neglected important deficits in their work to address. Recognizing both affordances and deficits is important due to its relationship with higher quality responses to students.

The importance of recognizing deficits in the student approach was particularly apparent in the work by Student N, where the line of best fit was placed with half the points above and half the points below. While that strategy resulted in a reasonable line in this case, it is not a generalizable strategy. Very few teachers' replies acknowledged this limitation to the student's approach, perhaps because they use this criterion themselves (Casey & Wasserman, 2015). This highlights the importance of teachers considering whether an approach is generalizable and how to push students to think about whether a strategy is generalizable (e.g., present a counter-example). The teachers in our study tended to struggle with both components. Past research suggests that teachers' pedagogical decisions are strongly influenced by their own mathematical understanding (Kieboom, Magiera, & Moyer, 2014). The difficulty teachers had responding to students on this task may indicate a weakness in their knowledge related to line of best fit and perhaps to data analysis and statistics topics more generally.

Our finding that many teachers (a) focused on tangential mathematical features of the task (rather than on relevant statistical components) when interpreting and responding to students and (b) struggled to recognize shortcomings in students' solutions support the conjecture that limited understanding of statistics and data analysis may have played a role in teachers' responses to students on this statistical task. For example, some teachers' responses focusing on the slope and y-intercept seemed intended to get the student to write the equation of the line. While writing the equation of a drawn line is a common goal in mathematics, it is not relevant to the presented task where the students are using a piece of spaghetti to place the informal line of best fit and think about how well that line fits the data. We wish to stress that this finding is not meant as a criticism of the teachers in the study, but rather points to the need for professional development opportunities for teachers related to statistics. The increased emphasis on statistical topics in the secondary curriculum via the Common Core State Standards (NGA & CCSSO, 2010) and the lack of attention to statistical knowledge for teaching in many teacher preparation programs (Conference Board of the Mathematical Sciences, 2012) suggests this is an area that warrants further efforts by mathematics teacher educators.

Another important aspect of responding to student work highlighted in our results was drawing on student thinking. Teachers struggled with this aspect, yet when it was present teachers were more likely to both respond in an open-ended manner and work toward the student learning objective. An interesting relationship that emerged was that teachers who interpreted the process students used also tended to draw on student thinking in their response. Furthermore, interpreting the process students used tended to be accessible to teachers, with a large number of teachers doing so in supported ways even though the prompt didn't specifically ask for this information. Based on our findings that (a) teachers frequently and successfully interpreted the process used by the student to generate the solution and (b) interpreting the student process was strongly related to drawing on student thinking in their responses to student, we recommended that professional development with teachers focus on these aspects. This would involve leveraging a strength of teachers' noticing (interpreting student process) to build an important component of a response to students (drawing on student thinking) based on that information.

Our results suggest that teachers may benefit from professional development focused on interpreting student understanding that leads to a particular solution. Recall that 60% of teachers who addressed student understanding did so in a way that was not well-supported by evidence from the student's work. The additional analysis suggests that teachers' interpretations of student understanding were often on the right track but were either incomplete or overreaching. This highlights an area of need for future research. Future work may focus on gathering a large enough sample of respondents who are attentive to student understanding to further examine the relationship between *interprets student understanding* and the quality of the teacher response. Although we did not find that teachers who were highly attentive to student understanding had high quality responses, we suspect this is because we had a limited number of teachers who were highly attentive to student understanding. Conclusions from our study are also limited since teachers' responses to students were hypothetical in nature. Future work in this area may involve studying teachers' responses to students working on line of best fit tasks in real time. Finally, our results showed that teachers' responses to students often failed to meet the criteria for *works toward student learning objective* and *builds on student thinking*. The prompt for this question asked teachers to describe "how they would respond" to the student, but did not specify the goal of advancing student understanding. It is possible that some of the limitations in teachers' responses may have resulted from the wording of the prompt.

Interpreting and responding to student work are critical components of eliciting and using evidence of student thinking, one of the key teaching practices for school mathematics (NCTM, 2014). Our study highlights the importance of considering secondary mathematics teachers' preparation concerning interpreting and responding to student work on statistical tasks and highlights key interactions between interpretation and consequential response to students.

References

AUTHOR (2015).

AUTHOR (2016).

AUTHOR (2017).

AUTHORS (2018).

Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, R., Young-Loveridge, J., & Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: Three large-scale cases. *Mathematics Education Research Journal, 16*(3), 27–57.

Burrill, G., and Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, and C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education, A Joint ICMI/IASE Study: The 18th ICMI Study* (pp. 57-69). New York, NY: Springer.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499–531.

Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learning 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.

Conference Board of the Mathematical Sciences (2011). *The Mathematical education of teachers II.* Providence, RI and Washington, D.C.: American Mathematical Society and Mathematical Association of America.

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225-256). Charlotte, NC: Information Age Publishing.

Franklin, C., Kader, G., Bargagliotti, A., Schaeffer, R., Case, C., & Spangler, D. (2015). *SET: Statistical Education of Teachers.* Alexandria, VA: American Statistical Assocation.

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, *41*(2), 169-202.

Jacobs, V. R., Lamb, L. L., Philipp, R. A., & Schappelle, B. P. (2011). Deciding how to respond on the basis of children's understanding. In M. Sherin, V. Jacobs & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 97-116). New York, NY: Routledge.

Jacobs, V.R., & Spangler, D.A. (2017). Research on core practices in K-12 mathematics teaching. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 766-792). Reston, VA: National Council of Teachers of Mathematics.

Kieboom, L., Magiera, M., & Moyer, J. (2014). Exploring the relationship between K-8 prospective teachers' algebraic thinking proficiency and the questions they pose during diagnostic algebraic thinking interviews. *Journal of Mathematics Teacher Education, 17*(5), 429-461.

Mayring P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping C. & N. Presmeg (Eds.), *Approaches to Qualitative Research in*

Mathematics Education. Advances in Mathematics Education (pp. 365-380). Dordrecht, Netherlands: Springer.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematics success for all.* Reston, VA: Author.

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

Qualifications and Curriculum Authority (2007). *The National Curriculum 2007*. Earlsdon Park, Coventry: Author.

Schoenfeld, A. (2011). Noticing matters. A lot. Now what? In M. Sherin, V. Jacobs, & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 223-238). New York: Routledge.

Sherin, M.G., Russ, R.S., & Colestock, A.A. (2011). Accessing mathematics teachers' in-the-moment noticing. In M.G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79-94). New York, NY: Routledge.

Sherin, M.G., & van Es, E.A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20-37.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning*, published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

Table 1

Coding scheme for interpreting student work prompt (n=82)

Categories		Levels		
		0	1	
Evaluative Affordances	Description	Response does not provide an evaluative judgment in regards to the merits or affordances of the student process or understanding	Response acknowledges the merits or affordances of the student process or approach	
	Example	*The student did not use mathematical thinking when choosing the first and last points. They were randomly chosen and aren't a good representation of the overall data.*	*They understand that the range end points is a good place to identify. Then, they can look to see how closely all of the other data points are related to the range. Student M has a good understanding about data and how it is related to the other values in the set.*	
	Freq (%)	34 (41%)	48 (59%)	
Evaluative Deficits	Description	Response does not provide an evaluative judgment in regards to the deficits of the student process or understanding	Response acknowledges the deficits of the student process or approach	
	Example	*They were able to connect along a diagonal plane and determine outliers.*	*The student wrongly believes that the best fit line perfectly separates the points into two equal groups.*	
	Freq (%)	45 (55%)	37 (45%)	
Interprets Student Process	Description	Response does not attend to student process for placing the line of best fit	Response attempts to acknowledge student process for placing the line of best fit but is not well supported by the student work and/or statement or is somewhat incomplete in describing the student's process	Response ac placing the student wor
	Example	*Student M understands that the information shows a positive trend.*	*[Student M] It appears that the student used the spaghetti and placed it close to the plotted points picking 2 points that he/she believed to be part of the actual line of best fit. This would tell me they have an understanding of line of best fit.*	*Student M i points (first*
	Freq (%)	25 (30%)	20 (24%)	
Interprets Student Understanding	Description	Response does not attend to student understanding that informed the process for placing the line of best fit	Response attempts to acknowledge student understanding that informed the process for placing the line of best fit but is not well supported by the student work and/or response	Response ac that informe fit and is su response
	Example	*They looked for patterns and ran the spaghetti through the designated pattern.*	*The student understands that there is a relationship between the drop height and bounce height.*	*[Student M] should go th should appr the first and understand contain any*
	Freq (%)	19 (23%)	49 (60%)	

Table 2
Coding scheme for response to student work prompt (n=80)

		0	1	2
Works Toward Student Learning Objective	Description	Response is not focused on: (A) convincing student strategy is not generalizable NOR B) pressing student towards a correct, generalizable strategy	Response is either focused on: (A) convincing student strategy is not generalizable OR (B) pressing student towards a correct, generalizable strategy	Response focused on: (A) convincing student strategy is not generalizable AND (B) pressing student towards a correct, generalizable strategy
	Example	*I would let the student know that they did a great job and ask them if there were any other ways to use the spaghetti.*	Only Meets 1A: *... connecting first and last dot is not true representation of best way to proceed, and can give very wrong answers, especially since values at the ends often turn out to be outliers (numbers that don't quite fit with everything else)...* Only Meets 1B: *I would tell them that they don't have to have the noodle go through any of the points. instead try to have an even amount of points on each side of the noodle and the line will be perfect.*	*I might construct an example scatter plot where a line connecting the first and last point clearly do[es] not [show the] overall trend of the data points. Then I would ask the student, "Does this line accurately represent the trend we see in the data?"*
	Freq (%)	41 (51%)	36 (45%) Met 1A: 12 (15%) Met 1B: 24 (30%)	3 (4%)
Draws on Student Thinking	Description	Response may acknowledge the line but not the student thinking in terms of the placement of the line	Response acknowledges aspects of the student thinking in terms of placement of the line (response must go beyond describing the process or placement of the line)	
	Example	*I would tell them that they are right on track with their thinking.*	*I would respond by saying that the best fit line does not have to pass through points in a scatter plot, but instead should best incorporate all points.*	
	Freq (%)	44 (55%)	36 (45%)	
Open-Ended ness	Description	Response does not leave room for varied student responses or a singular student response is likely	Response provides an opportunity for the student to engage in reasoning about the problem	
	Example	*Consider what the height would be when*	*I would ask the student if there were any other factors that could have influenced the bounce height. I would*	

		the golf ball is dropped 0 cm.	*ask them to explain their thinking about their line placement.*	
	Freq (%)	16 (20%)	64 (80%)	

An eighth grade student in a physics class asked the teacher whether the same object dropped from different heights would bounce different heights. Specifically, the student was interested in how the drop height and the bounce height are related. She hypothesized that the higher the drop height, the higher the bounce would be. The class decided to investigate this question by using a golf ball. Students were given eight set heights to drop the ball from. Then they dropped a golf ball from each of those heights and measured how high the ball bounced back up.
Below is a scatterplot that shows the data about how high a golf ball bounces when it is dropped from different heights, as collected by the class.

Using the piece of spaghetti, determine the line of best fit for the data. Be cognizant of your thoughts as you decide where to place the line on the graph.

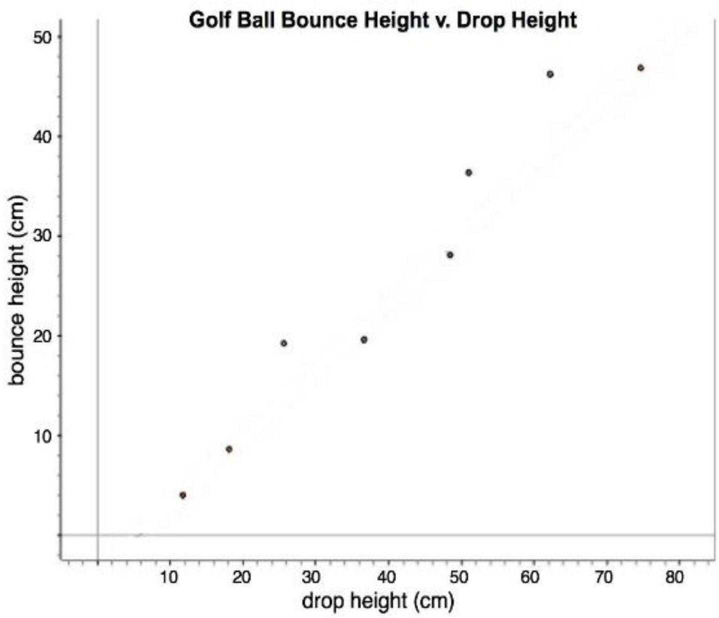

Figure 1. Line of best fit task.

Student M's response is shown below:

Student N’s response is shown below:

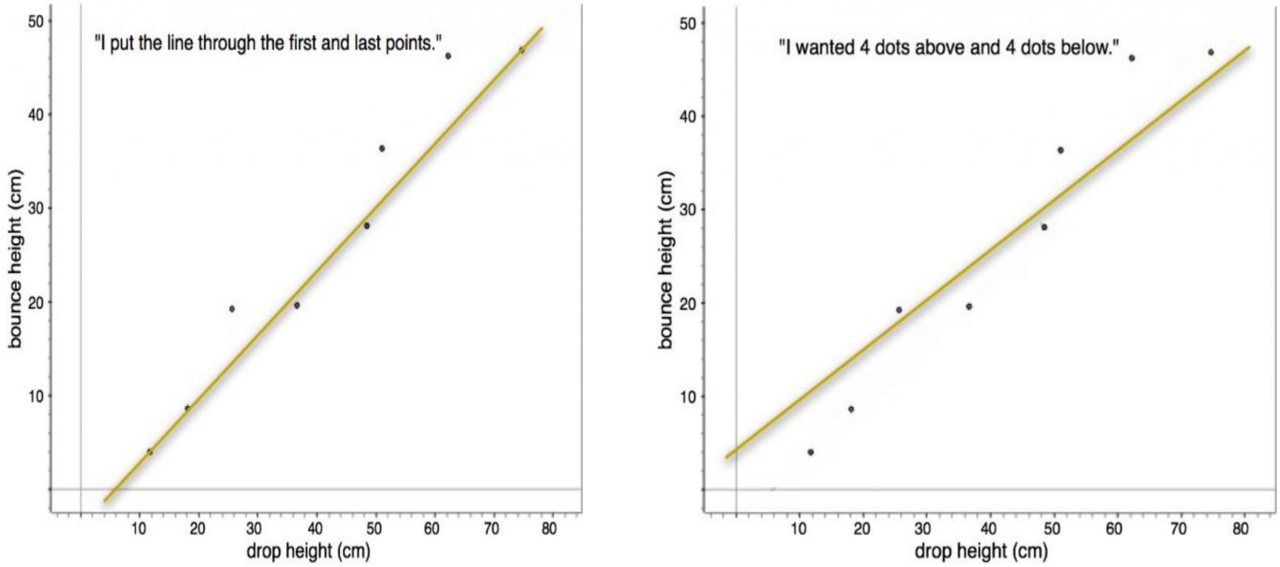

Figure 2. Student M and N’s work on the line of best fit task.

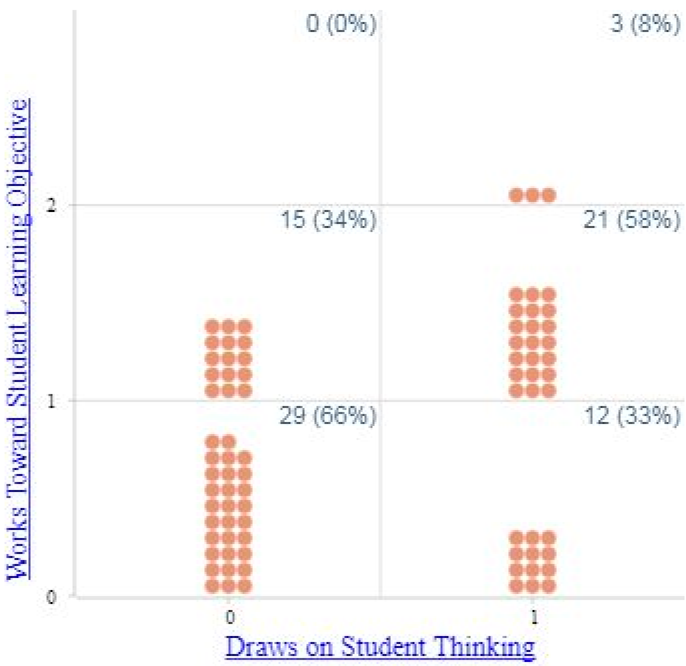

Figure 3. Relationship between the *Works Toward Student Learning Objective* and *Draws on Student Thinking* features of teachers’ responses to students.

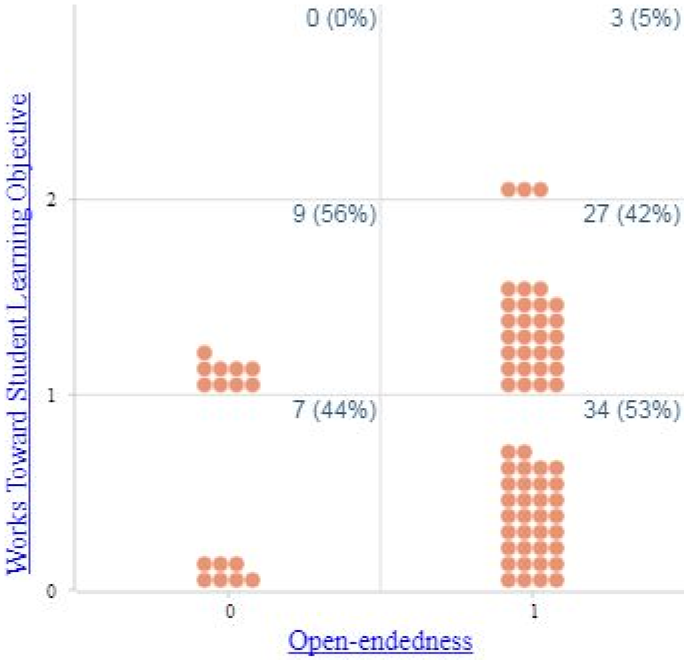

Figure 4. Relationship between the *Works Toward Student Learning Objective* and *Open-endedness* features of teachers' responses to students.

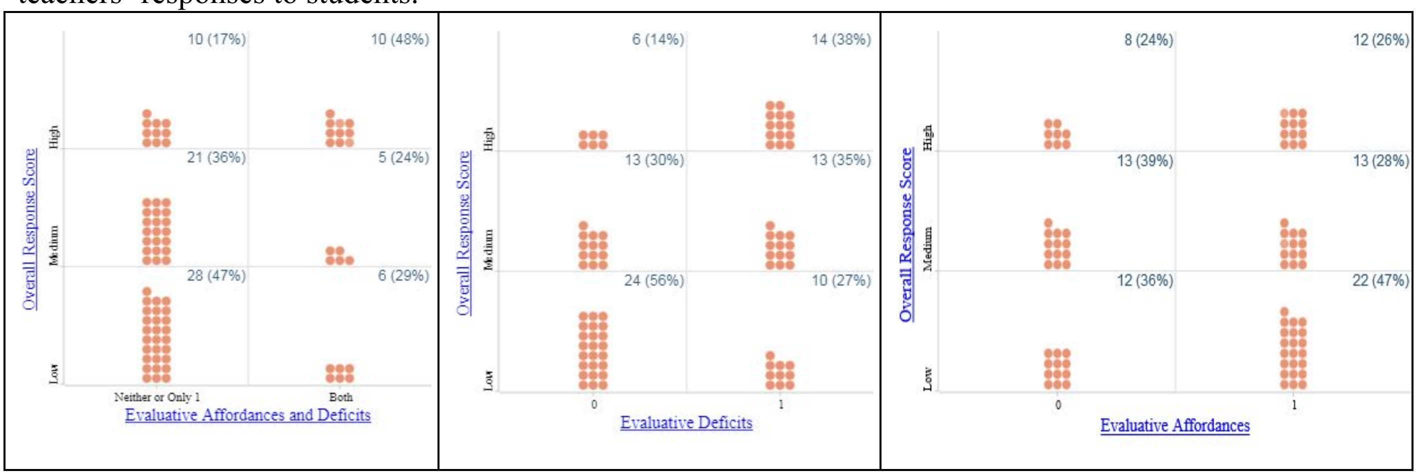

Figure 5. Relationship between Evaluative interpretations of student work and the Overall Response Score.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Investigations in Mathematics Learning*, published by Taylor & Francis. Copyright restrictions may apply. doi: 10.1080/19477503.2018.1534539

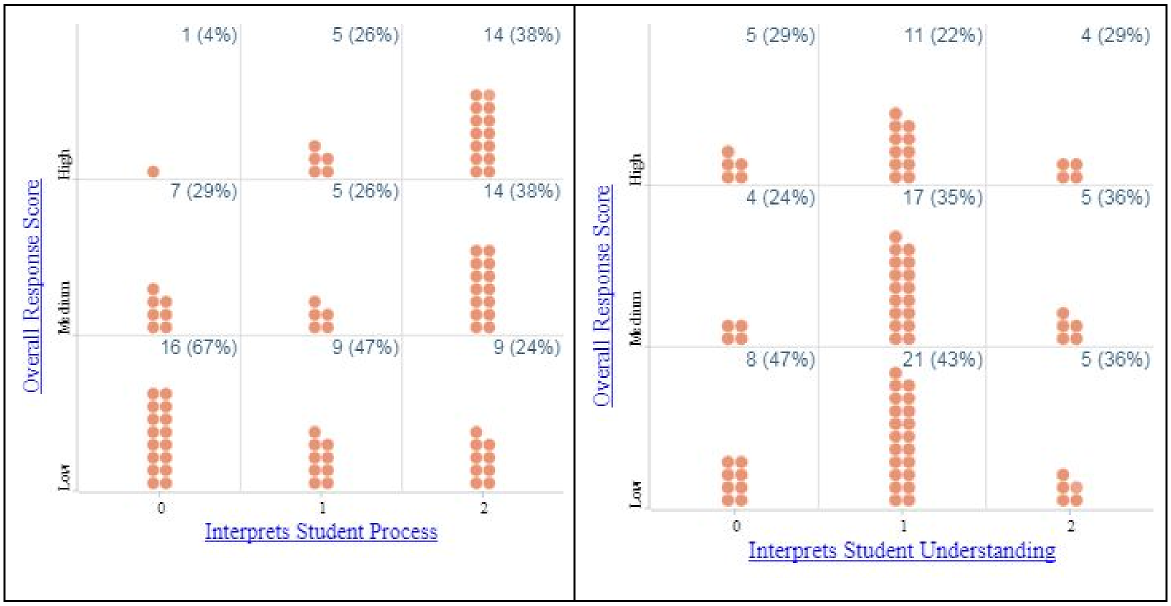

Figure 6. Relationship between *Interprets Student Process* and *Interprets Student Understanding* components of teachers' interpretations of student work and the Overall Response Score.

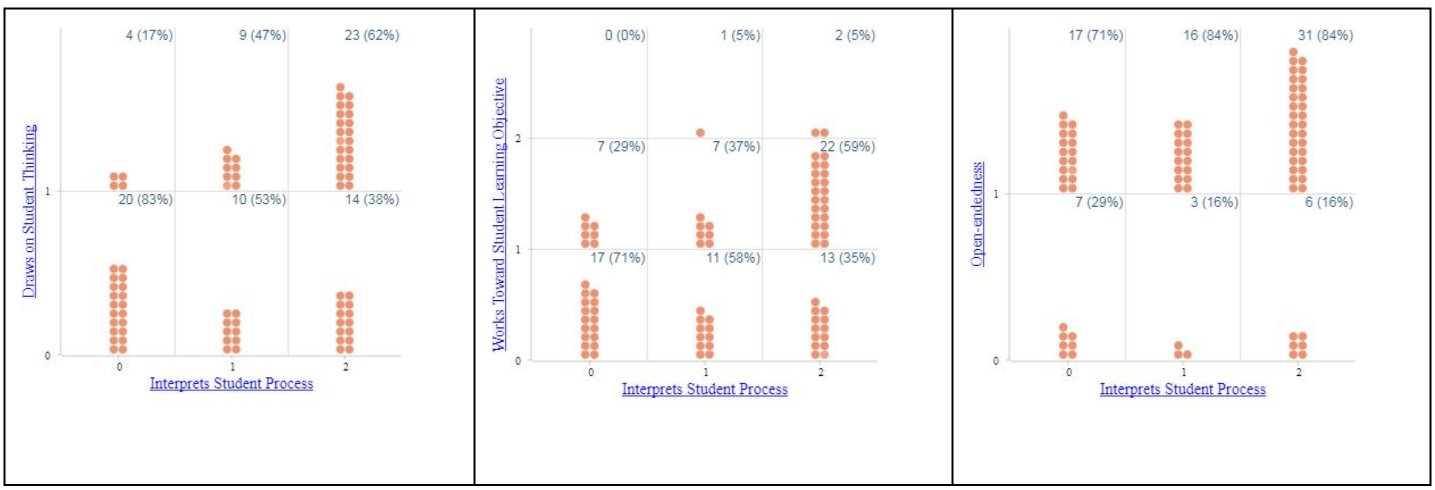

Figure 7. Relationship between *Interprets Student Process* and the various sub-components of the Overall Response Score.