LOOK ACADEMIC PUBLISHERS
OPEN ACCESS

IEJME — MATHEMATICS EDUCATION
2016, VOL. 11, NO. 4, 991-1012
Article number: mathedu.2016.090

Primary Teachers' Choice of Probing Questions: Effects of MKT and Supporting Student Autonomy

Karl W. Kosko[a]

[a]Kent State University, Ohio, USA

ABSTRACT

The present study explored whether primary grades teachers chose probing questions, given two hypothetical mathematics lesson scenarios. After responding to the mathematics lesson scenarios, participating teachers completed the Problems in Schools survey assessing dispositions to support student autonomy, and the Mathematical Knowledge for Teaching (MKT) assessment for primary grades patterns, functions and algebra. Logistic multiple regression was used to examine the influence of teachers' MKT and dispositions for supporting student autonomy. Results differed by format of scenario. In the scenario where the choice of a probing question would act as an initial prompt for description, results showed this choice was influenced more strongly by MKT score. In the scenario where a choice of probing question followed an already embedded student description, choosing a probing prompt as a follow-up question was more strongly influenced by support for student autonomy. Additionally, a negative, statistically significant interaction effect was found across both scenarios. Implications for these findings are discussed.

KEYWORDS
Teacher knowledge, mathematical discussion, teacher questioning

ARTICLE HISTORY
Received 04 January 2016
Revised 20 April 2016
Accepted 25 April 2016

Introduction

Mathematical discussion is a hallmark of reform oriented pedagogy. There are several pedagogical moves a teacher must make to facilitate effective mathematical discussions, such as revoicing (e.g., Forman, Larreamendy-Joerns, Stein, & Browns, 1998; O'Connor & Michaels, 1993), selection of meaningful and appropriate mathematical tasks (e.g., Stein & Lane, 1996; Zahner, 2012), and facilitating a positive social environment (e.g., Jansen, 2006; NCTM, 2000), among others. Yet, of such actions, teacher questioning is perhaps the most single identifiable pedagogical approach to facilitating mathematical discussions. Teachers' questioning strategies have been observed to strongly influence the

CORRESPONDENCE Karl Wesley Kosko kkosko1@kent.edu

manner in which students engage in mathematical discussions, as well as their success in developing mathematical understandings (e.g., Franke et al., 2009; Hufferd-Ackles, Fuson, & Sherin, 2004; Kazemi & Stipek, 2001; Martino & Maher, 1999), and there is some quantitative research to support these qualitative findings (e.g., Hiebert & Wearne, 1993). Specifically, such research confirms the general effectiveness of soliciting explanation and justification from students; a form of questioning that Boaler and Brodie (2004) refer to as probing questions.

Several qualitative studies have examined classroom data to identify teacher and classroom related factors that influence use of probing questions. Although numerous factors have been identified, I focus attention on two factors in particular: supporting student autonomy (also referred to as ceding teacher control) and mathematical knowledge for teaching (MKT). Teachers who are generally supportive of students' autonomy have been observed to ask more probing questions during mathematical discussions (Hufferd-Ackles et al., 2004; Kazemi & Stipek, 2001; Turner, Meyer, Midgley, & Patrick, 2003). The rationale for this observed relationship is that by asking probing questions, teachers are ceding a portion of the mathematical meaning-making process to students such that students, and not the teacher, are providing explanations and justifications for the mathematics at hand. The influence of MKT on teacher questioning has been observed in some studies (Cengiz, Kline, & Grant, 2011; Kim, 2011). These observations suggest that higher levels of MKT supported teachers "in deciding which mathematical ideas to pursue and how to pursue them during whole-group discussions" (Cengiz et al., 2011, p. 372). While both supporting student autonomy and higher MKT have been observed to contribute to teachers' questioning in mathematical discussions, it is unclear how consistently these factors influence such pedagogy. Acknowledging the benefits of supporting student autonomy, Chazan and Ball (1999) suggest that there are situations in which teachers should not cede control of descriptions or justifications to students. Rather, they suggest that the mathematics at hand may warrant the use of teacher descriptions or clarifications instead of allowing students such options. It stands to reason that teachers with higher levels of MKT would recognize such contexts and act accordingly. However, what if the teacher is also highly supportive of students' autonomy? What decision might they make in such a scenario, where the mathematics at hand warrants a more teacher centered approach? Such questions drive the purpose of the current study: to determine the degree to which primary grade teachers' MKT and disposition to facilitate student autonomy predict their choice of probing question in mathematics classroom scenarios, and whether these two factors interact.

Posing probing questions

In their analysis of primary teachers' questioning, Boaler and Brodie (2004) distinguish probing questions from others in that such questions ask students to "articulate, elaborate or clarify ideas" (p. 776). In essence, such questions encourage students to explain their mathematical thinking. Hiebert and Wearne (1993) include a similar question type in their study of primary grades classrooms. Such a question encourages students to "explain why a procedure is chosen or why it works" (p. 402). Thus, a probing question can be defined as a question soliciting explanation or justification of mathematics from a student or students. Use of such questions have been found to be more prevalent in reform oriented teaching (Boaler & Brodie, 2004) and corresponds to higher mathematics achievement

(Cross, 2009; Hiebert & Wearne, 1993; Kosko, 2012b). However, some research suggests that for a probing question to be the most effective, it must be included in a sequence of questions that include two or more specific questions that, taken as a sequence, serve to press students for mathematical meaning (Franke et al., 2009; Kazemi & Stipek, 2001).

Examining the questioning sequences of primary teachers participating in a professional development, Franke et al. (2009) found that students whose teachers used probing sequences of specific questions, following an initial student explanation, more often had fully complete and correct descriptions of mathematics elicited from students. By specific question, Franke et al. refer to a question that addressed something specific in a student's original mathematical description. Kazemi and Stipek (2001) referred to such sequences as a press for meaning by the teacher. A press for meaning involves questions that seek explanations and/or justifications that go beyond simple procedural descriptions. Likewise, Martino and Maher's (1999) observations of mathematical argumentation in primary grades suggest a sequence of questions that solicit explanations and justifications helps to elicit mathematical argumentation.

Franke and colleagues (Franke et al., 2009; Webb et al., 2008) have observed that although teachers may receive the same professional development on encouraging probing of student description, there is a large degree of variance regarding their use of probing questions versus probing sequences. Franke et al. (2009) suggest that such differences may be related to teachers' pedagogical content knowledge stating at one point that singular usage of probing questions may be related to "incorrect assumptions" (p. 390) made by the teacher regarding students' mathematical thinking. Such a potential relationship is also identified by Webb et al. (2008). However, Webb et al. (2008) also suggest that teachers' dispositions towards controlling of the mathematical discourse, referred to in this paper as their degree of support for student autonomy, also may interact with their use of probing sequences and probing questions. Thus, the observed and hypothesized interactions described by Franke and colleagues suggest that teachers' posing of isolated probing questions is likely due to variations in teachers' knowledge and their dispositions towards controlling mathematical discourse (identified in this paper as MKT and autonomy support, respectively).

Influence of teachers' MKT and support of student autonomy

Teachers' MKT and autonomy support are two important individual resources that influence how teachers pose questions in mathematical discussions. Students with more mathematical autonomy are described as having more freedom to make decisions about the strategies they use and mathematical decisions they make (Kosko, 2012a; Kosko, 2015; Kosko & Wilkins, 2015; Yackel & Cobb, 1996). By their nature, probing questions and sequences support student autonomy. This feature of such prompts is apparent in descriptions provided by studies where an increased use of probing questions is aligned with moving from a teacher centered to student centered classroom (Kazemi & Stipek, 2001; Martino & Maher, 1999). Hufferd-Ackles et al. (2004) provide, perhaps, the most informative description of this relationship. Hufferd-Ackles et al. observed a grade 3 U.S. teacher and found that as the teacher developed a math-talk community, the teacher slowly shifted control of questioning and explaining to students. As she did so, she asked increasingly more probing questions. This increase in probing questions correlated with a decrease in the frequency the teacher would add on her own explanations or justifications to what students provided, thus signifying an increase in supporting students' mathematical autonomy. Although an increase in student autonomy, and teacher support of it, appear to correlate with increased use of probing questions, such autonomy support is associated with particular features of the classroom environment. Yackel and Cobb's (1996) seminal study of how students' autonomy related to their mathematical descriptions focused on two contextual factors as inclusive of autonomy support: sociomathematical norms for how to engage in the classroom's mathematics, and social norms for how students were to interact with the teacher and one another. Examining how Geometry student's mathematical autonomy influenced their hedging in class discussions, Kosko (2012a) suggests that such classroom norms (social and sociomathematical) support students' mathematical autonomy either directly or indirectly. Thus, it is likely that teachers who are more supportive of student autonomy are likewise scaffolding how students should talk about mathematics (via establishment of certain sociomathematical norms) and providing a supportive social context (via establishment of certain social norms). Indeed, several observational studies suggest this to be the case (i.e., Hufferd-Ackles et al., 2004; Kazemi & Stipek, 2001; Kosko, 2012a; Kosko, 2015; Yackel & Cobb, 1996).

Although facilitating student autonomy appears to be related to increased use of probing questions and sequences, various researchers suggest there are times where it is more appropriate for the teacher to be the one providing explanations and justifications (Chazan & Ball, 1999; Lobato, Clarke, & Ellis, 2005) Lobato et al. found in their teaching interviews with young children, certain telling acts can serve to increase the amount of description from students. However, opportunities to open or close opportunities for mathematical discourse are dependent on the context. Ball, Thames, and Phelps (2008) suggest that this task of 'asking mathematically productive questions' is facilitated by teachers' MKT.

Developing the Mathematical Quality of Instruction assessment, and pairing it with their MKT assessment, Hill et al. (2008) observed 10 primary grades teachers' classroom instruction. Although findings generally indicated that higher levels of MKT correlated with more appropriate facilitation of justification and explanation, this was not always the case. One participant with higher MKT offered "students the opportunity to engage with mathematical justification and explanation, cognitively rich tasks, and the careful use of language *in the moment*, these moments rarely build toward a greater purpose, synthesis, or closure" (p. 469). More simply, teachers with higher MKT may pose more probing questions, but they may not pose them purposefully to consistently press students for mathematical meaning. Cengiz et al.'s (2011) observation of six primary grade teachers, and Kim's (2011) observation of three middle grades teachers confirm one observation of Hill et al. (2008). Specifically, teachers in the aforementioned studies who demonstrated higher levels of MKT more frequently extended opportunities for student thinking, press for generalization, and adapt questioning strategies in response to students' mathematical statements. Contrasting these findings are Grassetti's (2010) observations of three novice teachers. Grassetti found that teachers' levels of MKT did not necessarily correspond with their use of effective questioning strategies. Rather, Grassetti suggests that teachers' confidence levels may have played a stronger role in predicting their questioning.

The studies examining the relationship between MKT and facilitation of discussion suggest a mixed view of whether MKT influences teacher questioning, and in what manner. This may simply be a characteristic of the limited sample sizes present in each of the studies, thus indicating the need for a study of larger sample size to provide a certain degree of confirmation for these qualitative findings. However, the varying findings may indicate the influence of other, confounding factors that may regulate the effect MKT has on teachers' effective use of questioning. One possible factor may be the influence of teachers' support, or not, of student autonomy. If, as Ball et al. (2008) suggest, MKT allows teachers to determine questioning approaches given particular scenarios, then it is logical to conclude that a teachers' support for student autonomy will influence their decision whether to give students the opportunity to explain and/or justify mathematics. Yet, if such individual resources (MKT and support of student autonomy) influence whether certain questioning approaches are used in association with specific contexts, I suggest that the specific context brings with it specific expectations that also inform teachers' decision making in regards to questioning. It is this latter possibility, along with the influence of teachers' MKT and autonomy support that the present study focuses.

Another possible reason for the mixed findings regarding MKT's relationship with probing questions and sequences lay in the differences existing in teachers' levels of MKT for one mathematical domain over another. Examining Geometry teachers' decision to open or close opportunities for discourse in two hypothetical classroom scenarios, Kosko and Herbst (2013) found that "in the face of providing opportunities for discussion, more experienced and knowledgeable teachers will select the option with more instructional scaffolding" (p. 8). Kosko and Herbst's (2013) study focused particularly on experience and MKT for Geometry (see Herbst & Kosko, 2014 for a description of this construct and instrument); the course in U.S. secondary mathematics with the predominant focus on mathematical proof. The present study, however, focuses on the elementary setting, of which early algebra has been most often identified as a context for facilitating richer mathematical argumentation (Tall et al., 2011). Although probing questions may be used across multiple contexts in elementary classrooms, the present study focuses on MKT for patterns, functions, and algebra (i.e., key early algebra topics) given the propensity for probing questions to be used in such contexts for scaffolding argumentation.

Assessing Teacher Questioning

In order to determine the influence of autonomy supportiveness and MKT on primary teachers' probing questions, a pragmatic and meaningful way of assessing teacher questioning is necessary. Several studies have assessed such questioning through observation of teachers' lessons (e.g., Franke et al., 2009; Hiebert & Wearne, 1993; Hufferd-Ackles et al., 2004; Kazemi & Stipek, 2001). Yet, analysis of teacher questioning with observational data has certain limitations. The most obvious of which is the demands of sample size. Specifically, to ascertain teachers' general habits of using probing questions, several lessons would need to be observed over a period of time. Further, if one desires to assess teachers' questioning beyond a small sample, the amount of observational data, and the time spent to collect it, is multiplied several times over. Beyond considerations of sample is, perhaps, the more pertinent issue of context. Teachers do not teach in a vacuum. They are influenced by a myriad of factors that press them to facilitate mathematical instruction in various ways. Herbst and Chazan (2012) describe a portion of these factors as professional obligations and include, among other obligation sets, an obligation to the institution (of education). This obligation manifests through influences of various curricula, as well as policy mandates from various government levels, among other institutional stakeholders. For example, a primary teacher may know that the use of probing questions and sequences is beneficial to student learning, and may know how to do so effectively, but if they teach in a school where they are expected to adhere to a rigid curriculum guide, then what they know may not manifest in their classroom practices (Kosko & Gao, 2014).

Other means for assessing teachers' instruction have utilized various representations of practice. Such representations include written vignettes (e.g., Jochums & Pershey, 1993), video of actual classrooms (e.g., Kersting, 2008; Kersting, Givvin, Sotelo, & Stigler, 2010; Moyer & Milewicz, 2002), and animated or comic-based representations (Herbst, Chazan, Chen, Chieu, & Weiss, 2011; Kosko, Rougee, & Herbst, 2014; Moreno & Ortegano-Layne, 2008). As with observational data, there are various benefits and limitations to using representations of practice to assess teachers' perceptions and knowledge of pedagogy, and these limitations and benefits can vary depending on the format of representation. Herbst et al. (2011) provide a useful description of such features, suggesting that written and animated / cartoon-based vignettes allow for more specification of scenarios than video-based representations. Specifically, for an event to be represented in video, it must have happened. Written and animated vignettes allow for scenarios to be created in specific ways by the researcher. Animated and cartoon-based vignettes allow for a visual representation of a classroom context that written vignettes cannot convey in the same meaningful way. For example, a gesture or facial expression included in a cartoon-based vignette must be described in written vignettes, which draws attention to such an occurrence (see Herbst et al., 2011 for a more complete description of comparisons between various representations of practice).

For the present study, cartoon-based vignettes were used to assess primary teachers' choice of probing question given specific mathematical classroom scenarios. Use of cartoon-based vignettes allowed for creation of scenarios where use of probing questions was more appropriate for instruction, while allowing for comparison across teachers, given the same scenario. Yet, the use of cartoon-based vignettes to assess teachers' decision-making in pedagogical contexts is a novel approach (Herbst et al., 2016). Herbst et al. (2016) note that such items may assess individual teachers' decision making, but move "beyond individual factors in the study of teachers' decision making by considering other resources that pull teachers toward both reproducing and changing their practice" (p. 15). In other words, the context which cartoon-based vignettes matters, and the degree to which it matters may be due to smaller or larger grain size factors. Therefore, the present study included only two cartoon-based vignettes in which the significant difference in structure was that one item included a descriptive response from a student following an included probing question and the other did not. Researchers examining probing questions via observation have found that teachers are less likely to ask probing questions of the same student back-to-back (e.g., Franke et al., 2009). Therefore, the differentiation in context allowed for a deeper analysis of how MKT and disposition to support student autonomy related to choice of probing question in either context.[1] Responses to both scenarios were used to answer the following research question:

Do primary grade teachers' assessed MKT and disposition to facilitate student autonomy predict their choice of probing question in two hypothetical mathematics teaching scenarios?

Method

Sample

The Data were collected from 45 elementary teachers in a Midwestern U.S. state in fall 2012. The sample was predominately White and female (95.6%), with 3 teachers self-classified as outside of this group (2 White males and 1 Black female). Teachers had a wide range of teaching experience ($M = 14.50$, $SD = 8.20$) and taught across grades K-4 (Kindergarten = 22.2%, Grade 1 = 20.0%, Grade 2 = 13.3%, Grade 3 = 20.0%, Grade 4 = 15.6%, Multiple Grades = 8.9%). Regarding educational background, all teachers reported having taken a minimum of 3 mathematics teaching methods courses ($M = 3.67$, $SD = .58$), and 71.1% reported holding a graduate degree[2]. Teachers were sampled from 17 different schools in both rural and suburban settings. Participating teachers were asked to complete a survey packet that focused on teachers' perceptions, knowledge, and habits regarding facilitation of mathematical discussion.

Measures for dependent variable

To assess participating primary teachers' choice of probing questions, two cartoon-based scenarios were created using the Depict tool on the LessonSketch.org platform (Herbst & Chieu, 2011). Although teachers' decision making within the context of teaching involves a multitude of choices within relatively short time frames, and the nature of facilitating effective discussions involves sequences of questioning and other pedagogical moves, the scenarios used in the present study examined participating teachers' choice of probing questions in particular instances. As particular instances, the scenarios shared some common contextual features, but were not designed to assess the same context.

The choice to focus on the smaller grain size of singular prompt choices, as opposed to choices within sequences of such prompts, was based on several factors. For example, one objective of this study is to examine teachers' questioning choices in such a way as to verify qualitative findings in the literature, while providing some degree of authenticity to practice. Yet, teachers' questioning choices have rarely been examined quantitatively with representation-based items. This suggested a more cautious approach and, therefore, I elected to use items that included singular decision points with the assumption that lessons learned from the present study would inform future work in developing vignettes that would assess multiple decision points, or decision chains.

---

[1] Statistical analyses of single sets of items is common in other scientific fields such as agriculture (Clason & Dormody, 1994). Although such approaches are rarely generalizable to a population, they can be informative for various other reasons. The present study's focus allows for examination of teachers' MKT and autonomy support in two slightly different contexts for posing questions to students.

[2] Until recently, teachers in the Midwestern state were required to obtain a graduate degree after a certain length of time teaching.

The novelty of quantitatively examining teachers' questioning choices with context-rich vignettes warranted a smaller grain size of analysis in terms of the number of items. Rather than composing numerous vignette-based items, I elected to focus on responses to two such items to allow for a comparison across two specific contexts. Although it is more conventional to consider construction of an assessment with numerous items representing similar, but distinct contexts, prior work on context-rich items such as those used here suggests a more cautious examination is warranted (Herbst, Kosko, & Dimmel, 2013).

The two cartoon-based vignettes depicted scenarios of elementary classrooms at two different grade levels (grade 1 and grade 2), and each vignette was situated in early algebra content (i.e., patterns). Within each vignette, a student provides a mathematical statement with each statement containing a mathematical error of some form. Variance between the scenarios was introduced particularly in regards to whether the student provided a description of their problem solving strategy. Such differentiation in the scenarios was purposeful considering Franke et al.'s (2009) findings that many teachers did not ask follow-up probing questions once students provided a description of their strategy (regardless of whether it was correct or not).

Each scenario concluded in such a way as to invite the depicted teacher to prompt students with a question. Participants were presented with four potential options in which they were asked to select the best question for the depicted teacher to ask in the scenario. All four options were questions structured following Boaler and Brodie's (2004) scheme (see Table 1). Only four question types were included to limit the complexity of each item: *Gathering Information*, *Probing*, *Generating Discussion*, and *Orienting & Focusing*. Probing questions were included primarily due to the focus of the present study. The other categories were selected due to their prevalence various studies using the same scheme (Boaler & Brodie, 2004; Gokbel & Boston, 2015; Kosko et al.'s, 2014). Specifically, I elected to include question types which appeared likely to garner more responses than not. Ideally, all nine of Boaler and Brodie's question types would be assessed, but this would increase the complexity of the item itself. Following are brief descriptions of both scenarios, and the question options for each.

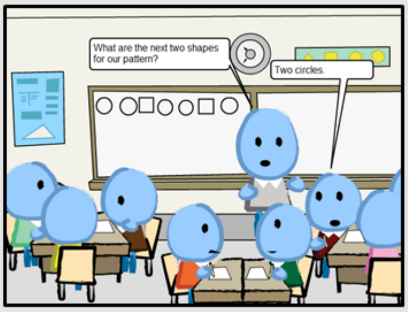
Grade 1 scenario

The grade 1 scenario is shown in Figure 1. In the vignette, a repeating pattern of two circles and one square is depicted on the board when the teacher asks for the next two shapes. The depicted student's response of "two circles" does not fit the depicted pattern. Each of the questions provides the opportunity for soliciting more description from the student, but allow for different kinds of description simply by the nature of the question itself (1 = *Gathering Information*; 2 = *Probing*; 3 = *Generating Discussion*; 4 = *Orienting & Focusing*).

Table 1. Teacher Questions Coding Scheme from Boaler and Brodie (2004).

Gathering information, leading students through a method	Requires immediate answer rehearses known facts/ procedures Enables students to state facts/procedures
Inserting terminology	Once ideas are under discussion, enables correct mathematical language to be used to talk about them
Exploring mathematical meanings and/or relationships	Points to underlying mathematical relationships and meaning. Makes links between mathematical ideas and representations
Probing, getting students to explain there thinking	Asks student to articulate, elaborate or clarify ideas
Generating Discussion	Solicit contributions from other member of the class
Linking and applying	Points to relationships among mathematical ideas and mathematics and other areas of study/life
Extending thinking	Extends the situation under discussion to other situations where similar ideas may be used
Orienting and focusing	Helps students to focus on key elements or aspects of the situation in order to enable problem solving
Establishing Context	Talks about issues outside of math in order to enable links to be made with mathematics

In this scenario, Mrs. Jenson is working on repeating patterns with her first-grade students when Xavier suggests a pattern:



Mrs. Jenson wants her students to have a solid understanding of the math involved.
Of the options below, which would be the best question for Mrs. Jenson to ask?

① "And what comes after the two circles?"

② "Why are the next two shapes circles?"

③ "Did anyone get different shapes than Xavier?"

④ "What about what's happening at the end here, where we have a square then circle?"

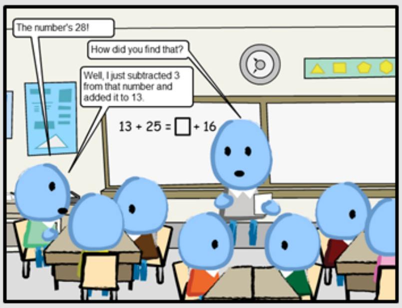
Figure 1. Grade 1 survey item assessing teacher questioning preference.

The *Gathering Information* question (Option 1), by its nature, solicits facts and procedures of an immediate nature. However, the prompt could potentially be used to help Xavier realize his error. Option 3, *Generating Discussion*, solicits information from other students and could potentially be used to garner the correct response for Xavier to compare to his incorrect response. Option 4, *Orienting & Focusing*, redirects Xavier's attention to where the sequence ends and could be used to help Xavier see his error. Option 2, *Probing*, solicits a rationale from Xavier and could be used to help Xavier see his misapplication of the pattern. Each question holds the potential for helping Xavier understand his mistake, but solicits different information to do this. However, the literature base suggests that probing questions hold the most potential for facilitating students' conceptual understanding were included as the focus of the present study. Therefore, the probing question (Option 2) was designated as the prompt of focus for the Grade 1 Scenario.

Grade 2 scenario

The grade 2 scenario depicts an equation with two addends on both sides of the equals sign (see Figure 2). However, one of the addends is missing. A student, Katie, shouts out the incorrect solution of 28. When the depicted teacher, Mrs. Patel, prompts Katie for an explanation of her procedures, Katie describes her solution strategy. The grade 2 scenario is different from the grade 1 scenario because it already includes a probing question from the depicted teacher. Thus, it allows for assessment of whether participants would elect choosing a second probing question following the one the teacher initially used. Each of the four potential teacher prompts provides a different opportunity for soliciting more description from the student (1 = *Generating Discussion*; 2 = *Orienting & Focusing*; 3 = *Probing*; 4 = *Generating Discussion*).

In this scenario, Mrs. Patel is playing "guess my number" with her second-grade class when Katie tells her guess:



Mrs. Patel wants her students to have a solid understanding of the math involved.
Of the options below, which would be the best question for Mrs. Patel to ask?

① "Does anyone have an opinion about what Katie did here?"

② "So Katie subtracted 3 from 25 and then added it to 13. Why do you think she subtracts and then adds the same number over here?"

③ "Katie, why did you subtract 3 from that number?"

④ "Katie, what number did you subtract 3 from?"

Figure 2. Grade 2 survey item assessing teacher questioning preference.

As with the grade 1 scenario, each of the types of questions provides an opportunity for the student to understand their error, but solicits different information to do this. Since the options are similar enough in structure to those described for the grade 1 vignette, I limit such details here. Yet, the *Orienting & Focusing* prompt (Option 2) has slightly different features than the previous example described. Option 2 includes a "why" question which would, at its surface, seem to classify the prompt as a *Probing* question. However, the question's primary features solicit attention from the class to the procedures Katie used. By revoicing Katie's description, the prompt allows Mrs. Patel to focus the discussion on the procedures used and not the incorrect response provided by Katie. Thus, while the prompt does contain features of a *Probing* question by soliciting a rationale, it is, at its core, an *Orienting & Focusing* prompt. The prompt categorized as the *Probing* question (Option 3), solicits a rationale, but from Katie and not the whole class.

Nature of variable measuring prompts

The distribution of responses per vignette is shown in Table 2. As can be seen from this distribution, there is no consistent preference for prompt across the two scenarios. Additionally, a preliminary Chi-Square analysis found that the distribution of responses for each vignette is not independent from chance (Grade 1, χ_2 (df=3) = 53.58, p < .001; Grade 2 χ_2 (df=3) = 18.20, p < .001). A post hoc analysis of the standardized residuals in each cell of the 1 x 4 contingency tables suggests specific trends relevant to how the data is analyzed. *Probing* prompts were consistently found to be the most prevalent response in the grade 1 scenario and the least prevalent response in the grade 2 scenario ($Z_{Grade\ 1}$ = 6.19, p < .001; $Z_{Grade\ 2}$ = -1.86, p < .10). *Gathering Information* prompts followed a similar trend, but where the most prevalent response when *probing* was least prevalent and vice versa ($Z_{Grade\ 1}$ = -3.35, p < .001; $Z_{Grade\ 2}$ = 3.50, p < .10). The other two types of prompts could effectively be considered as useful distractors given the context of the items and the nature of this study. Therefore, the frequency distribution of responses lends support to dichotomizing the data such that selecting a *Probing* option is compared to all other options. This dichotomized response was used as the outcome measure for the present analysis.

Table 2. Distribution of Question Types across Scenarios

	Probing	Orienting & Focusing	Generating Discussion	Gathering Information
Grade 1	71.1%	15.6%	13.3%	10.0%
	n = 32	n = 7	n = 6	n = 0
Grade 2	11.1%	24.4%	11.1%	51.1%
	n = 5	n = 11	n = 5	n = 23

Independent variables

Three independent variables were included for analysis. The first independent variable included was teachers' relative autonomy support for students (*Autonomy*). As outlined in the review of literature, a key feature of effective mathematical discussion is allowing for students to have a measure of control in their mathematics. Thus, a measure of how supportive teachers are of students' autonomy was particularly useful for this study. *Autonomy* was assessed with the Problems in School survey (Deci, et al., 1981). This survey presents a series of vignettes to teachers and asks them to rate four potential actions in response to the vignette. The calculated outcome measure provides an indicator as to the teachers' disposition to how controlling versus autonomy-supportive they are in their instruction (M = 4.80, SD = 1.74, $Range$ = -.25 to 9.13).

For the present sample, participants were more supportive of student autonomy than not.

For the second independent variable, a 2006 revised version of Hill, Schilling, and Ball's (2004) MKT assessment for 'patterns, functions, and algebra' was used (α=.78, *M*= .00, *SD* = .83, *Range* = -1.49 to 1.72). This particular version of the MKT assessment allowed for a more topic-specific alignment with the dependent variables, as all three vignettes were embedded in the same mathematical topic. The assessment used allows for scores calculated for 1-parameter Item Response Theory (IRT), and these scores were used in the present study. IRT assumes that responses to particular items in a test assess particular latent traits, in this case teachers' mathematical content knowledge. This latent trait is represented by the statistic theta (θ) and is calculated both for specific items and as a general test score. The benefit to using IRT scores over raw test scores (number correct ÷ total number items) is that it takes into account that some items are more difficult than others (Crocker & Algina, 2006; Wilson, 2005). MKT scores reported here are the theta scores for individuals' tests. A theta score of 0.00 represents an individual with 'average' ability as defined by the population targeted (in this case, elementary teachers). The current sample is fairly representative of the population in regards to both the mean MKT score and standard deviation (*M*= .00, *SD* = .83, *Range* = -1.49 to 1.72).

The final independent variable included was an interaction variable for MKT scores and *Autonomy* (*MKT·Autonomy*). As discussed in the literature review, teachers who have a higher degree of MKT may facilitate discussions more effectively. Also, teachers who support student autonomy should may also facilitated more effective mathematical discussions. Yet, as Chazan and Ball (1999) note, there are instances where a teacher should take control and not the students. There are also instances where it is more mathematically productive for students to discuss incorrect approaches; such as when that discussion will lead to the validation of a correct solution strategy for a task. Determining which scenario is which may require an interaction between teachers' support for student autonomy and their degree of MKT. Thus, the variable *MKT·Autonomy* allows for a measure of general indicator of this interaction.

Analysis & Results

Logistic multiple regression was used to examine participants' choice of question for each mathematics teaching scenario. As described in a previous section, responses were treated as dichotomous (*1 = Probing*; *0 = Not Probing*), and is thus suitable for logistic multiple regression. Logistic multiple regression is, essentially, the logistic form of multiple regression. However, beta values for the intercept and effects of predictors are estimated as logits, as is the outcome measure (see Hosmer & Lemeshow, 2000 for a detailed description). Equation 1 was used as the preliminary final model for each cartoon-depicted scenario.

Equation 1:

$$g_m(Probing) = \beta_0 + \beta_1(Autonomy)_1 + \beta_2(MKT)_2 + \beta_3(MKT \cdot Autonomy\ )_3$$

In modeling each scenario, it is assumed that the conditional probability that a participant selects a *Probing* prompt is denoted by $P(Y = 1\mid Probing) =$

$\pi(Probing)$. Within the equation modeling this probability, $g_m(Probing)$ represents the logit outcome, and is therefore a continuous measure. This logit outcome can be converted to an indicator of probability using Equation 2.

Equation 2:

$$\pi(Probing) = \frac{e^{g(Probing)}}{1 + e^{g(Probing)}}$$

In Equation 1, the intercept β_0 represents the average logit for a participant selecting a *Probing* prompt, where the participant has average MKT (score = 0.00), is neither more nor less supportive of students' autonomy (score = 0.00), and has a score of 0.00 for the interaction variable (*MKT·Autonomy*). Essentially, one can convert this logit with Equation 2 to denote the probability of a participant with average MKT and score 0.00 for *Autonomy* to select a *Probing* prompt. β_1, β_2, and β_3 represent the average effects, in logits, of each variable on the outcome measure.

Constructing the logistic regression model

While Equation 1 represents the posited final model, the selection process of variables to include when building a logistic regression model follows certain steps. As suggested by Hosmer and Lemeshow (2000), each independent variable considered for inclusion should first be examined prior to such inclusion. For continuous variables, such as those used in the present study, t-tests are suggested for this evaluation with the criteria for inclusion being a p-value of .25 or less (see Hosmer & Lemeshow, 2000, p. 95 for further details). Following these steps, each of the variables selected for inclusion in the logistic regression model should be examined using the Wald statistic. Next, the slope of graph comparing the logit and each continuous variable are examined. Completion of these steps yields the *main effects model*. The main effects model does not include interactions between the main variables. For the present study, an interaction between MKT scores and *Autonomy* is logical, and should therefore be examined. Including an interaction effect creates what is called the *preliminary final model*. The final step in constructing the logistic regression model is to assess the fit of the preliminary final model. For the present study, I used Hosmer and Lemeshow's Goodness of Fit statistic to indicate model fit. These general steps for constructing a logistic regression model were applied to each individual vignette for analysis. The description of results that follows provides a summary of the outcomes for each vignette, per guidelines outlined here.

Results

Grade 1 Scenario

Prior to evaluating the main effects model, both continuous variables were evaluated for inclusion using independent samples t-test. MKT was found to have a t-statistic of 1.30 and p-value of .20, thus meeting the .25 p-value criteria for inclusion. For *Autonomy*, a t-statistic of .24 was found at a p-value of .81. This typically would signal possible exclusion of the variable from the model. However, it was retained for further evaluation with potential inclusion of the interaction variable later. Examination of the main effects model showed Wald statistics that were not statistically significant at the .05 level for the intercept (*Wald* = 1.51, $p = .22$), the effect of Autonomy (*Wald* = .04, $p = .84$) or MKT scores (*Wald* = 1.62, $p = .20$). This information suggests that modeling responses for the grade 1 vignette is not useful. However, when we compare logits of the outcome with the continuous variables (see Figure 3), an interaction between the variables seems likely. Therefore, both variables were retained to examine the inclusion of the interaction variable *MKT·Autonomy*.

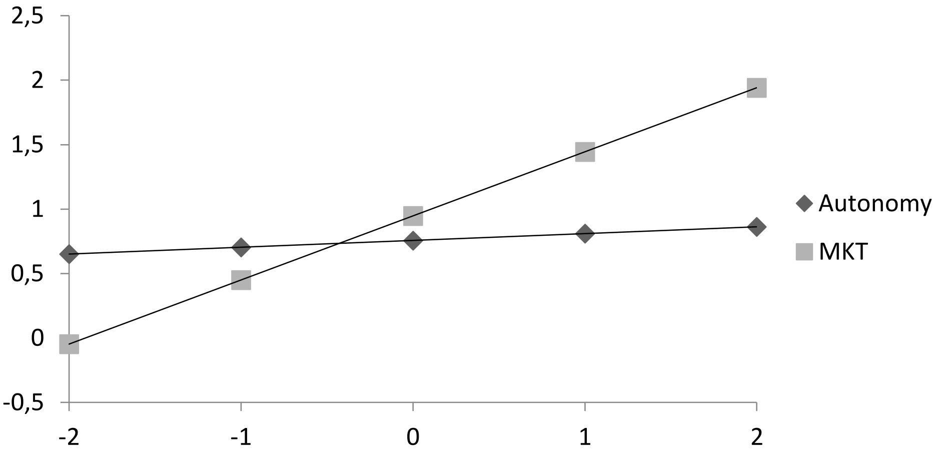


Figure 3. Relationship between logit outcome and independent variable scores for grade 1 vignette.

After including it into the preliminary final model, examination of the Wald statistic for *MKT·Autonomy* shows the variable is a statistically significant predictor of the outcome logit measure (*Wald* = 4.26, $p < .05$). Further, MKT also was found to be statistically significant (*Wald* = 4.86, $p < .05$). *Autonomy* was not found to be statistically significant, but was retained due to the statistical significance of the interaction variable. Before finalizing the model, I examined the Hosmer and Lemeshow test for Goodness of Fit. Results indicated that the model predicts results not statistically significant than what was observed (χ^2 (df=7) = 10.63, $p = .16$). This indicates good model fit. Results from the final model are presented in Table 3.

Table 3. Final Model for Grade 1 Scenario.

	β (logits)	S.E.	Wald	p-value
Intercept	1.50	.96	2.43	.12
Autonomy	-.15	.29	.26	.61
MKT Score	2.36	1.07	4.86	.03
MKT·Autonomy	-.72	.35	4.26	.04

Results indicate that participants' *autonomy* score has little direct influence on their decision to select a *probing* question for the grade 1 vignette. Following inclusion of the interaction effect, MKT scores did have a statistically significant effect such that an increase in level of MKT relates to an increase in likelihood that a probing question would be selected. Interestingly, the interaction effect between MKT and *autonomy* was found to be statistically significant but was negative in direction. In other words, while higher levels of MKT, alone, increased the likelihood participants would select *probing* for the grade 1 vignette, if this higher level of MKT corresponded with a higher sense of autonomy support for students, the likelihood actually decreased. Figure 4 provides a three-dimensional representation of the effects of the interaction variable on the probability of teachers selecting a probing question. Those individuals with higher MKT scores, such as 2.0, but had relatively lower dispositions towards autonomy support were more likely than those with similar MKT scores but much more strongly supportive of students' autonomy. By contrast, those with lower MKT scores, such as -2.0, but had relatively lower dispositions towards autonomy support were less likely than those with similar MKT scores but higher *autonomy* scores. Those most likely to select probing prompts either had very high MKT and very low *autonomy* scores, or vice versa. However, the majority of participants were more likely than not (50% chance or higher) to select a probing question.

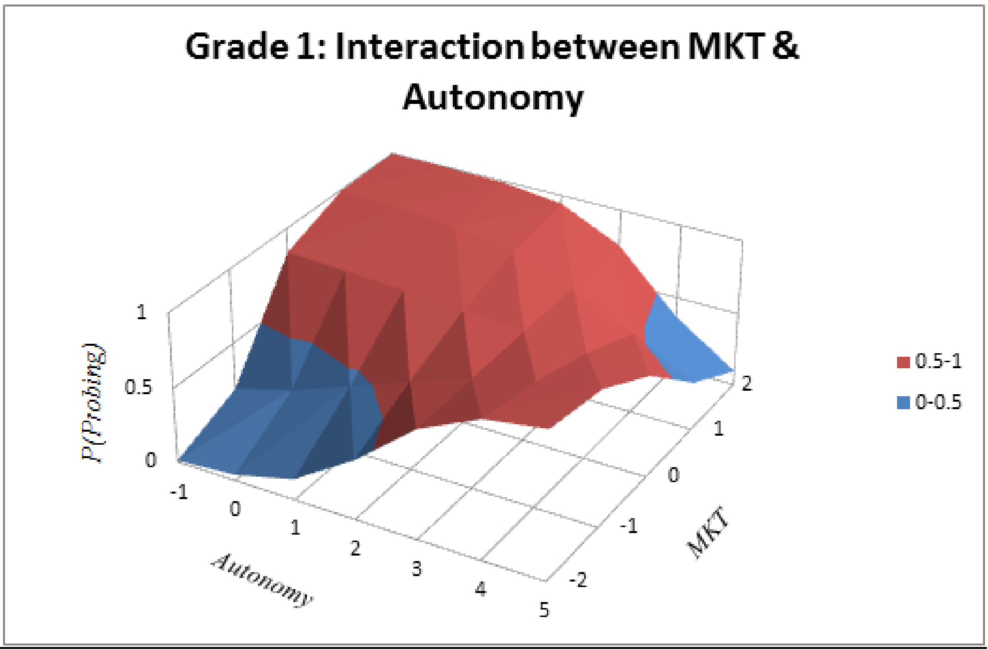


Figure 4. Three-dimensional graph representing the change in probabilities of selecting probing questions in the grade 1 scenario. The graph accounts for autonomy, MKT, and the interaction of these two variables.

Grade 2 vignette

As was done with the grade 1 model, both continuous variables were evaluated for inclusion into the main effects model using independent samples t-test. *Autonomy* met criteria for inclusion ($t = 2.60$, $p = .01$), but MKT did not ($t = -.78$, $p = .44$). As was similarly done with *autonomy* in the grade 1 model, MKT was retained to evaluate the inclusion of the interaction variable (*MKT·Autonomy*). Wald statistics for the main effects model showed the intercept

(*Wald* = -5.56, *p* = .00) and *autonomy* (*Wald* = 4.87, *p* = .03) were statistically significant, but MKT was not found to be statistically significant (*Wald* = .62, *p* = .43). Again following the steps used for the grade 1 vignette, a graphical comparison of outcome logits with continuous variables scores (see Figure 5) suggests an interaction effect, and both variables were retained to examine the inclusion of the interaction variable *MKT·Autonomy*.

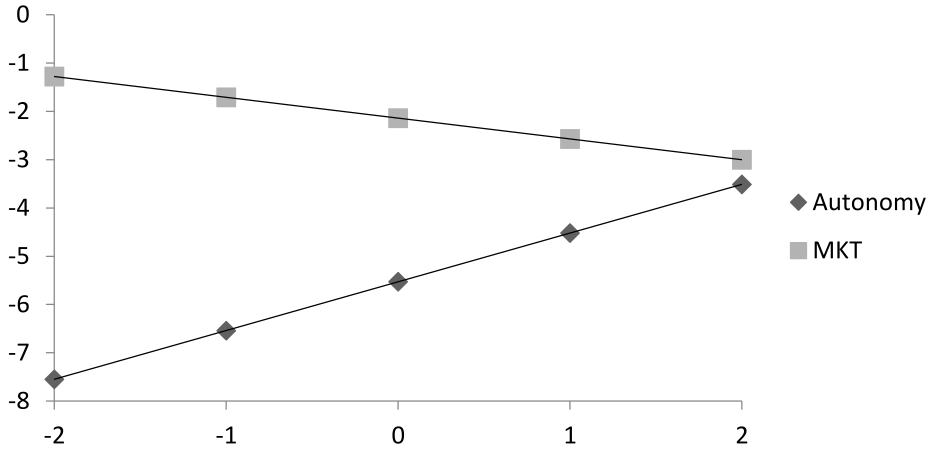


Figure 5. Relationship between logit outcome and independent variable scores for grade 2 vignette.

Results for the preliminary final model are displayed in Table 4. *MKT·Autonomy* was found to be a statistically significant predictor of the outcome logit measure (*Wald* = 2.72, $p < .10$). *Autonomy* was also found to be statistically significant (*Wald* = 3.17, $p < .10$), but MKT was found not to be statistically significant. Examination of the Hosmer and Lemeshow test for Goodness of Fit indicated that the model has sufficient fit (χ^2 (*df*=7) = 8.45, *p* = .29).

Table 4. Final Model for Grade 2 Scenario.

	β (logits)	S.E.	Wald	p-value
Intercept	-5.88	2.34	6.32	.012
Autonomy	1.01	.57	3.17	.075
MKT Score	2.91	2.20	1.75	.186
MKT·Autonomy	-.95	.57	2.72	.099

Results indicate that similar to results for the grade 1 vignette, the interaction effect *MKT·Autonomy* was found to be statistically significant and negative. However, where in the grade 1 vignette, MKT and not *autonomy* was found to be the statistically significant and meaningful indicator of whether a probing question was selected, findings for the grade 2 vignette reflect the opposite. Specifically, participants' support for student autonomy was a critically influential factor as to whether they selected a probing question, but their level of MKT had little meaningful or statistical effect.

So, higher *autonomy* increased the likelihood participants would select a probing question for the grade 2 scenario, but if this higher level of autonomy corresponded with a higher MKT score, the likelihood decreased to a degree. The three-dimensional graph shown in Figure 6 helps illustrate the effect of the interaction. As with the grade 1 scenario, individuals with higher MKT scores but lower *autonomy* scores, and vice versa, were more likely to select probing questions. Yet, the interaction in the grade 2 model suggests a much lower likelihood for selecting a probing question, in general. This suggests that while the negative interaction between MKT and *autonomy* was consistent across both scenarios, the difference associated with the context of each scenario may have interacted with the general probability of selecting a probing question.

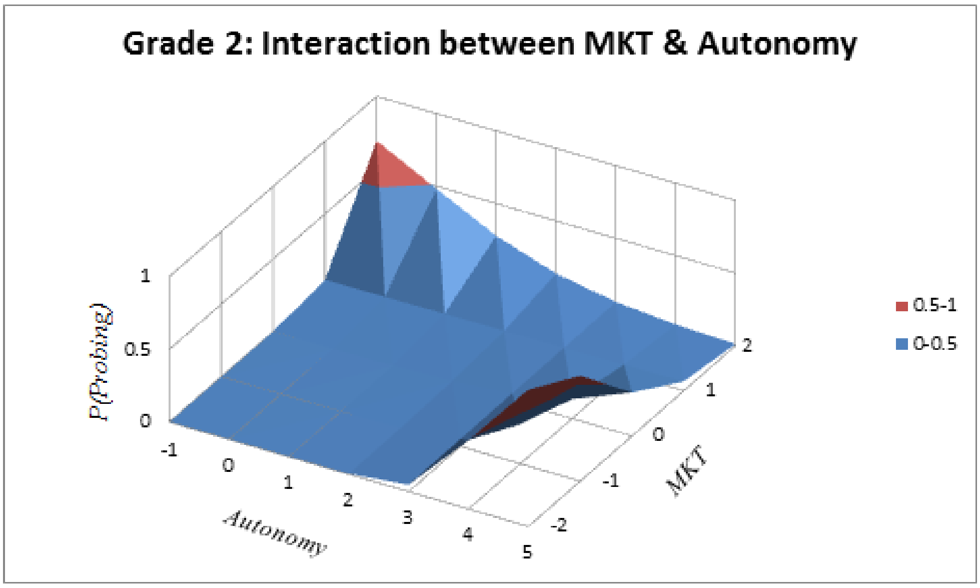


Figure 6. Three-dimensional graph representing the change in probabilities of selecting probing questions in the grade 2 scenario. The graph accounts for autonomy, MKT, and the interaction of these two variables.

Discussion

In their assessment of the MQI instrument, Hill et al. (2008) noted that although teachers with higher MKT scores generally used questioning effectively in their instruction, one teacher with higher MKT did not use questioning in a manner that would facilitate what Kazemi and Stipek (2001) would refer to as a press for meaning. Another means of viewing this is that for the particular teacher Hill et al. described, higher MKT translated to having useful questions that did not necessarily fit together in a sequence of questions that pressed for meaning. The present study found evidence to explain a part of the phenomenon observed by Hill et al. Specifically, in in the grade 1 scenario, higher levels of MKT increased the likelihood participants would select a probing question following an answer-only response from a student and support for autonomy had no statistical effect. However, MKT did not have a statistically significant effect on the likelihood participants would choose a probing question in a context where the student described their solution strategy (grade 2). In the latter case, a probing question was already included in the vignette and selection of another probing question would provide more confirmatory evidence of developing a probing sequence (or press for meaning) than selection of probing questions in the grade 1 scenario. Although MKT did not statistically influence participants' choice of probing question in the grade 2 scenario, *autonomy* did. Therefore, results of the present study provide preliminary evidence to suggest that MKT and support of student autonomy may both affect teachers' choice of questioning, but potentially in distinctly different ways potentially related to features specific to the context of instruction. Although informative, these results are preliminary and much needs to be done both to confirm these findings and examine other facets not examined here.

The results concerning MKT and *autonomy* are intriguing and informative, yet the negative interaction effect found for grade 1 and 2 scenarios illustrates an important tension between the two factors. Teachers who were likely to select probing questions in the scenarios generally have either higher demonstrated MKT or stronger dispositions towards supporting student autonomy, but not both. There are several potential reasons why this interaction exists in this data, and may exist more generally among primary grades teachers. Rather than speculate, however, the findings presented here suggest a strong need for further study of this phenomenon. These findings suggest a potential complex interaction not observed in prior research, and both confirmation and further explanation are needed for the field to better understand their implications. Additionally, although the negative interaction effect was consistent across scenarios, the overall probability of selecting probing prompts was substantially different (see Figures 1 and 2). This suggests that in addition to the influence of individual resources teachers bring to the task of deciding what prompts to use, further study of the contextual factors and the social resources that may be associated with them should also be further studied.

Implications and future study

Primary teachers' responses to the scenarios used in this study provided interesting statistical results that confirmed and expanded upon findings from various qualitative studies. The particular features of the scenarios themselves allowed for examination of specific questions participating teachers considered most pedagogically appropriate at particular moments of instruction. Although context to the depicted scenarios was provided, the breadth of the scenario was somewhat brief and available responses were single decisions, which may not best represent some teachers' use of sequences of questions. Therefore, a necessary next step is to examine not only specific decision points, but decision chains related to teachers' choice of question prompts in mathematical discussion scenarios. Further, creation of composite scores for items of similar contexts to the grade 1 and grade 2 vignettes is needed. Findings from the present study help to provide useful information for what such scenarios might look like. Available decision chains for particular scenarios could include more than one path that would represent a probing sequence of questions, as well as other questioning sequences that represent lower press for meaning. Such scenarios might best be considered as expanded versions of those used in the present study.

Findings from the present study suggest certain relationships between participating teachers' responses and their MKT. Since, participants' responses represent their considerations of appropriate pedagogical practice in mathematics teaching context, the relationship between Hill et al.'s (2004) MKT measure may suggest that these vignettes tap a particular domain of teacher knowledge. To better understand the implications of such a domain of knowledge, future research is necessary to better understand the role of MKT in teachers' knowledge for facilitating mathematical discussions. Within such a line of inquiry, future study can examine the connection between such teacher knowledge and their actual practice in whole class mathematical discussions.

Although this study has focused specifically on the teacher practice of questioning in mathematical discussions, use of the scenarios joins a promising line of research using hypothetical (e.g., Herbst & Miyakawa, 2008; Kosko & Herbst, 2012; Moreno, & Ortegano-Layne, 2008) and actual (e.g., Kersting, 2008; Kersting et al., 2010) vignettes to examine teachers' decision making. Although uses of representation in this manner do not examine actual practice, such approaches allow for examination of teachers' conceptions and knowledge of practice given a shared context among participants. Such shared contexts allow for statistical examinations not available with certain observational study, with items more directly relatable to the practice of teaching.

Conclusion

Statistical results from this study confirm findings from qualitative research which has observed certain relationships with why teachers use probing questions in mathematical discussions: namely that both supporting student autonomy and having higher MKT facilitates use of probing questions and sequences. However, results from this study suggest that these two factors may affect teachers' use of questioning in distinct yet interacting ways. For the scenarios examined here, higher MKT scores were more strongly associated with selecting a probing question when no student description was provided. However, higher *autonomy* scores were more strongly associated with selecting such prompts when a student description was provided. Further, MKT and *autonomy* scores had a negative interaction across both scenarios. Although further study is needed to both confirm and explore these results, these findings provide important clues as to why teachers vary in their use of probing questions. To some degree, the observed patterns may provide more questions than answers, but they are important questions that can help guide further study in teachers' facilitation of mathematical discussion. Kent State University, Ohio, USA

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Karl W. Kosko is an assistant professor in mathematics education at Kent State University, Ohio, USA.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Boaler, J., & Brodie, K. (2004). The importance of depth and breadth in the analyses of teaching: A framework for analyzing teacher questions. In D. E. McDougall & J. A. Ross (Eds.), *Proceedings of the 26th meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 773-782). Toronto, Canada.

Chazan, D., & Ball, D. (1999). Beyond being told not to tell. *For the Learning of mathematics, 19*(2), 2-10.

Cengiz, N., Kline, K., & Grant, T. J. (2011). Extending students' mathematical thinking during whole-group discussions. *Journal of Mathematics Teacher Education, 14*, 355-374.

Clason, D. L., & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education, 35*(4), 31-35.

Cross, D. I. (2009). Creating optimal mathematics learning environments: Combining argumentation and writing to enhance achievement. *International Journal of Science and Mathematics Education, 7*, 905-930.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Mason, OH: Wadsworth.

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*, 642-650.

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Browns, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction, 8*(6), 527-548.

Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education, 60*(4), 380-392.

Gokbel, E. N., & Boston, M. D. (2015). Considering students' responses in determining the quality of teachers' questions during mathematical discussions. In T. G., Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez (Eds.), *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (p. 1172). East Lansing, MI: Michigan State University.

Grassetti, M. T. (2010). *Engaging students in mathematics conversations: Discourse practices and the development of social and socialmathematical norms in three novice teachers' classrooms* (doctoral dissertation). Amherst, MA: University of Massachusetts – Amherst.

Herbst, P. & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM, 44*(5), 601-612.

Herbst, P., Chazan, D., Chen, C. L., Chieu, V. M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM, 43*(1), 91-103.

Herbst, P., Chazan, D., Kosko, K. , Dimmel, J., & Erikson, A. (2016). Studying decision making in instructional situations: How multimedia questionnaires can help. *ZDM-Mathematics Education, 48*(1), 167-183.

Herbst, P., & Chieu, V. M. (2011). Depict: A tool to represent classroom scenarios technical report. University of Michigan. Retrieved from http://hdl.handle.net/2027.42/87949.

Herbst, P., & Kosko, K. W. (2014). Mathematical knowledge for teaching and its specificity to high school geometry instruction. In J. Lo, K. R., Leatham, & L. R. Van Zoest (Eds.), *Research trends in mathematics teacher education* (pp. 23-46). New York: Springer.

Herbst, P., Kosko, K. W., & Dimmel, J. (2013). How are geometric proof problems preented? Conceptualzing and measuring teachers' recognition of the diagrammatic register. In A. C. Superfine, M. Martinez, G. Larnell, T. Stoelinga, & D. Martin (Eds.), *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 179-186), Chicago, IL: University of Illinois at Chicago.

Herbst, P., & Miyakawa, T. (2008). When, how, and why prove theorems? A methodology for studying the perspective of geometry teachers. *ZDM, 40*(3), 469-486.

Hiebert, J. & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*(2), 393 – 425

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430-511.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105*, 11-30.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd Ed.). New York: John Wiley & Sons, Inc.

Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*(2), 81-116.

Jansen, A. (2006). Seventh graders' motivations for participating in two discussion-oriented mathematics classrooms. *The Elementary School Journal, 106*(5), 409-428.

Jochums, B. L., & Pershey, E. J. (1993). *Using the vignette method in formative evaluation. Evaluation Practice, 14*(2), 155-161.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal, 102*(1), 59-80.

Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. Educational and Psychological Measurement, 68(5), 845-861.

Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education, 61*(1-2), 172-821.

Kim, H. J. (2011). *An exploratory study of teachers' use of mathematical knowledge for teaching to support mathematical argumentation in middle-grades* (doctoral dissertation). Austin, TX: The University of Texas at Austin.

Kosko, K. W. (2012a). Geometry students' hedged statements and their self-regulation of mathematics. *The Journal of Mathematical Behavior, 31*(4), 489-499.

Kosko, K. W. (2012b). Student enrollment in classes with frequent mathematical discussion and its longitudinal effect on mathematics achievement. *The Mathematics Enthusiast, 9*(1&2), 111-148.

Kosko, K. W. (2012). Geometry students' self-determination and their engagement in mathematical whole class discussion. *Investigations in Mathematics Learning, 8*(2), 17-36.

Kosko, K. W., & Gao, Y. (2014). Perceptions and reality: One teacher's use of prompts in mathematical discussions. In P. Liljedahl, S. Oesterle, C. Nicol, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 4, pp. 41-48), Vancouver, Canada: PME.

Kosko, K. W., & Herbst, P. G. (2013). *Opening opportunities for discourse in the face of norms for posing proof problems*. Round table paper presented at the 2013 Annual Meeting of the American Educational Research Association, Division C, San Francisco, CA.

Kosko, K. W., & Herbst, P. (2012). Evaluating teachers' decisions in posing a proof problem. In L. R. Van Zoest, J. J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 813-820), Kalamazoo, MI: Western Michigan University.

Kosko, K. W., Rougee, A., & Herbst, P. (2014).). What actions do teachers envision when asked to facilitate mathematical argumentation in the classroom? *Mathematics Education Research Journal, 26*(3), 459-476.

Kosko, K. W., & Wilkins, J. L. M. (2015). Does time matter in the growth of student-reported engagement in mathematical discussions? The influence of mathematical autonomy. *Journal of Experimental Education., 83*(3), 368-385.

Lobato, J., Clarke, D., Ellis, A. B. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education, 36*(2), 101-136.

Martino, A. M., & Maher, C. A. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *Journal of Mathematical Behavior, 18*(1), 53-78.

Moreno, R., & Ortegano-Layne, L. (2008). Do classroom exemplars promote the application of principles in teacher education? A comparison of videos, animations, and narratives. *Educational Technology Research and Development, 56*(4), 449-465.

Moyer, P. S., & Milewicz (2002). Learning to question: Categories of questioning used by preservice teachers during diagnostic mathematics interviews. Journal of Mathematics Teacher Education, 5(4), 293-315.

National Council of Teachers of Mathematics [NCTM]. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly, 24*(4), 318-335.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation: An International Journal on Theory and Practice, 2*(1), 50-90.

Turner, J. C., Meyer, D. K., Midgley, C., & Patrick, H. (2003). Teacher discourse and sixth graders' reported affect and achievement behaviors in two high-mastery / high-performance mathematics classrooms. *The Elementary School Journal, 103*(4), 357-382.

Webb, N. M., Franke, M. L., Ing, M., Chan, A., De, T., Freund, D., & Battey, D. (2008). The role of teacher instructional practices in student collaboration. *Contemporary Educational Psychology, 33*, 360-381.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum.

Zahner, W. C. (2012). "Nobody can sit there": Two perspectives on how mathematics problems in context mediate group problem solving discussions. *Journal of Research in Mathematics Education, 1*(2), 105-135.