



HAL
open science

Variation in mathematics teacher questioning across instructional settings

Reidar Mosvold

► **To cite this version:**

Reidar Mosvold. Variation in mathematics teacher questioning across instructional settings. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04398644

HAL Id: hal-04398644

<https://hal.science/hal-04398644>

Submitted on 16 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Variation in mathematics teacher questioning across instructional settings

Reidar Mosvold¹

¹University of Stavanger, Faculty of Arts and Education, Stavanger, Norway;
reidar.mosvold@uis.no

Research on mathematics teacher questioning often distinguishes between the types of questions teachers ask, but rarely considers variation in the type of questions across instructional settings. To investigate this, an exploratory case study analyzed the questioning of a Norwegian mathematics teacher who taught the same lesson in two different grade 5 classes. Analysis indicates significant difference in teacher questioning between the two classes, and there were also significant differences between questioning across different ways of organizing the discourse inside a lesson. Questions that call for participation, elicit students' thinking or invite them to take a stance are more often asked in discussion settings, whereas check-in questions are more often asked in seatwork settings.

Keywords: Teaching, questioning, settings, enactment.

Introduction

Everyone agrees that questioning is important in (mathematics) teaching, and that it matters what type of questions teachers ask. Open-ended questions are often valued over closed questions, higher-order questions are valued over low-order questions, and why-questions over what-questions. Yet, as indicated in a recent review of literature on mathematics teacher questioning (DeJarnette et al., 2020), there is little consensus about typologies of questioning in research.

Several categorizations of types of teacher questions are referenced in the literature (DeJarnette et al., 2020). One common distinction is between questions of higher and lower cognitive order, and Redfield and Rousseau (1981) concluded that higher cognitive questioning positively influenced student learning. Perry et al. (1993) listed six types of questions, and they found that teachers in Asia asked significantly more questions that tapped into conceptual knowledge or problem-solving strategies than teachers in the United States. In another study, Di Teodoro et al. (2011) distinguished between “deeper” questions and more “surface” questions, and they suggested that the former type is suitable for supporting students' problem solving. In yet another study, Aziza (2018) differentiated between open-ended and closed questions and suggested that open-ended questions better stimulate students' mathematical creativity. The list goes on.

DeJarnette et al. (2020) conclude their review of research on mathematics teacher questioning by suggesting that continued research on the interactions between teacher questions and student responses is needed. Since certain types of questions seem to be suitable for different purposes, “there is opportunity to explore when such questions are necessary and productive” (p. 10). The word “when” points to different circumstances around a particular situation or action. These circumstances can relate to variations in site, historical context and more. In mathematics teaching, settings can also refer to circumstances regarding the teacher, the students, the content, as well as influences from the environment (Cohen et al., 2003). This study views teaching as managing instructional interactions,

and instructional interactions are inextricably linked to settings. This study considers two types of variations in settings: variations across groups of students, and variations between ways of organizing the instructional discourse (Cohen, 2011).

Surprisingly, research on mathematics teacher questioning has not highlighted differences across settings, and this study aims at contributing to this. From analysis of the case of a Norwegian 5th grade teacher, who is teaching the same lesson in two parallel classes, the study approaches the following research questions:

1. Does teacher questioning vary between separate groups of students?
2. Does teacher questioning vary between discussions and seatwork within a lesson?

Theoretical framework

To explore the research questions, the present study applies a framework of questions and questioning practices in mathematics teaching by Enright et al. (2016). This framework was developed from and for analysis of classroom data. Where other frameworks focus on types of questions and their *purpose*, which is relevant in studies of the planning of questioning practices (e.g., Mosvold & Wæge, 2022), this framework focuses instead on types of questions and their instructional *function*.

A focus on function of questions corresponds with a view of teaching as management of instructional interactions, where the aim is to understand the intricacies of such interactions more than measuring their effects. The instructional triangle, as presented by Cohen et al. (2003), is a conceptual foundation for the framework, and for this study. This study also draws on Cohen’s (2011) conceptualization of teaching. Enright et al. (2016) consider teacher questions to be instructional moves that “are situated in broader environments that enter the classroom through the teachers, content and students and influence, among other factors, beliefs, actions, and perceptions in the classroom” (p. 2).

Table 1: Framework of question types (Enright et al., 2016, pp. 4–5)

Question type	Example	Code
Call for participation	Who will give it a try?	<i>ic_call</i>
Check-in	Are you done?	<i>ic_checkin</i>
Orienting	On which problem?	<i>ic_orient</i>
Survey	People who did it with this same method?	<i>ic_survey</i>
Clarification	Not a straight line?	<i>amc_clarify</i>
Confirmation	Did you extend it like this?	<i>amc_confirm</i>
Direct answer	What is the length in this case here?	<i>amc_direct</i>
Eliciting explanation	Why not?	<i>amc_explain</i>
Eliciting process	What did you do?	<i>amc_process</i>
Eliciting thinking	What do you think?	<i>amc_think</i>
Eliciting stance	Do we agree with that?	<i>amc_stance</i>

Enright et al. (2016) distinguish between two broad types of questions. The first type has the function of information collection (IC), and the codes for this type are indicated in this study with the prefix “ic.” The other type has the function of accessing mathematical content (AMC), and the codes for these questions have been given the prefix of “amc” in the present study (see Table 1).

Teachers may ask questions in different instructional settings. This study draws on Cohen's (2011) distinction between different ways of organizing instructional discourse. Cohen first differentiates between direct and indirect discourse, and the focus on teacher questioning in this study is primarily related to direct discourse. Cohen then distinguishes between four types of organization of direct instructional discourse: seatwork, lecture, recitation, and discussion. The present study investigates and compares questioning in seatwork and discussion.

Methods

The study is an exploratory case study, drawing on data from a larger project where a Norwegian 5th grade mathematics teacher and her students were observed for two weeks. The teacher taught two parallel classes in grade 5 (referred to here as 5A and 5B). Both classes had 16 students. The teacher taught the same lessons in the two classes, and she normally taught 5B before 5A. There are several reasons why this was an interesting and relevant site to study. First, comparison of a teacher's questioning practice between these two different classes provides a site for exploring potential differences in teacher questioning between classes of varying ability levels. The teacher considered 5A to have larger differences between high-performing and low-performing students than 5B. Second, this school followed a curriculum of "Developmental Education in Mathematics," which emphasizes systematic development of all students, high level of difficulty, promoting students' awareness of their own learning process, and rapid pace of progression (Gjære & Blank, 2019). It was therefore expected that the lessons would include many tasks with high cognitive demand. DeJarnette et al. (2020) note that there appears to be a correlation between higher-order questions and cognitively demanding tasks; this was a promising site for exploring higher-order questioning. Third, this curriculum also emphasized classroom dialogue, and lessons often involved much whole-class discussion. This corresponded well with my own interest in questioning in discussions as opposed to more traditional teaching that follows an IRE or IRF pattern (cf. Cazden, 2001).

Ten lessons were observed over a period of two weeks—five lessons in each class. Each lesson was recorded with audio and video, and the recordings were transcribed verbatim. To select a case for analysis in this study, the frequencies of teacher questions across all ten lessons were first considered. The arithmetic mean for the number of teacher questions across the lessons was $M=143.9$ ($SD=37.2$). I then selected the lesson that came closest to the average number of teacher questions as an initial case lesson (it had 146 teacher questions and a similar ratio of teacher-student questions as the overall average). The assumption was that this would be a 'typical' lesson with respect to teacher questioning in this sample—and thus suitable for an exploratory case study. For comparison, I selected the corresponding lesson in the parallel class. This lesson was taught by the same teacher, directly following the lesson in the first class, and with the same lesson plan.

The lesson was taught on a Tuesday morning, six weeks into the school year, first in 5B and then in 5A in the subsequent lesson. Each lesson lasted about 70 minutes. The lesson started with a warm-up problem presenting two addition pyramids, followed by tasks that involved turning expressions with repeated addition into multiplication. These tasks were discussed in whole-class discussions. This first half of the lesson was concluded by the students being asked to write down the commutative law of multiplication in their books. After a short break that involved some physical activity, the lesson

continued with a weekly 5-minute multiplication test, where the students scored and recorded their own progression. The last part of the lesson was devoted to individual seatwork before the lesson concluded with a check-out task.

The author coded the data in three phases. First, each teacher question was identified from the lesson transcripts. Second, the questions were assigned a value of 0 or 1 to indicate if the question was asked in a whole class setting or not. Third, the questions were coded according to the framework by Enright et al. (2016). Each question was assigned a value of 0 or 1 for each question type. The question types were considered exclusive. Coding was mostly straightforward, but some categories were challenging. For instance, the teacher asked several questions that were formulated “What do we/you think?” To decide the category, questions were always considered in relation to their function, as displayed in the following student responses. Sometimes such “think-questions” prompted students’ elaboration of their thinking (*amc_think*), but, in other instances, such questions prompted students to show if they agreed or disagreed (*amc_stance*). For instance, if the teacher asked, “What do we all think?”, the students would often show their stance by giving a sign of agreement or disagreement. In other situations, a question that might appear to prompt thinking instead prompted students to elaborate on the process they used (*amc_process*) or the answer they got (*amc_direct*). This testified to the importance of considering a question in connection with its *function* and not only by its formulation or plausible purpose. Whenever a question did not fit any of the given categories, it was coded as *other*. Examples are questions that were asked about the observers present, or about practical organization of the classroom or students. The transcripts were coded in Emacs Org-mode (e.g., Schulte & Davison, 2011). Frequency tables were automatically generated and imported to R, where each column was assigned to a variable, and these variables were further analyzed in R (R Core Team, 2022). Since the data were nominal, hypothesis testing was mainly done with chi-square tests (Reynolds, 1977), or with Fisher exact tests when the frequencies were too low.

Results

The first research question asks if there is variation in teacher questioning across groups of students. The null hypothesis was that there is no difference in the teacher’s questioning across student groups, whereas the alternative hypothesis was that the teacher questioning systematically differs across groups. As a first step in the analysis, I identified frequencies of distinct types of questions used, and the framework by Enright et al. (2016) served as analytic framework. Questions that did not fit into any of the predefined categories of this framework were coded “other.” As shown in table 2, the teacher used many direct questions (*amc_direct*) in both lessons, and there were also multiple questions that were coded as “other” in both lessons. An initial comparison of the two classes indicates that the teacher asked more check-in questions (*ic_checkin*) and call for participation questions (*ic_call*) in 5A, and she asked more *amc_stance* questions in 5B. Although the numbers were lower, the teacher asked more questions about explanation (*amc_explain*) and process (*amc_process*) in 5B as compared to 5A. On the other hand, she asked more orienting questions (*ic_orient*) in 5A.

Table 2: Frequencies and percentages of questions by type in each class

Type	5A		5B	
	<i>f</i>	%	<i>f</i>	%
IC-questions				
<i>ic_call</i>	22	12.1	12	8.2
<i>ic_checkin</i>	38	20.9	16	11.0
<i>ic_orient</i>	13	7.1	3	2.1
<i>ic_survey</i>	2	1.1	0	0
AMC-questions				
<i>amc_clarify</i>	1	0.5	1	0.7
<i>amc_confirm</i>	9	4.9	4	2.7
<i>amc_direct</i>	30	16.5	36	24.7
<i>amc_explain</i>	2	1.1	6	4.1
<i>amc_process</i>	0	0	6	4.1
<i>amc_stance</i>	16	8.8	24	16.4
<i>amc_think</i>	15	8.2	9	6.2
Other questions	34	18.7	29	19.9
<i>Sum</i>	<i>182</i>	<i>100%</i>	<i>146</i>	<i>100</i>

Based on the frequencies in Table 2, there appears to be variation between the two classes. Since the conditions for a chi-square test were not met, I ran a Fisher exact test to check for independence in questioning between the two classes. This gave a p-value of .001, which indicates that the null hypothesis can be rejected, and there is a significant difference in teacher questioning between the two classes. As a next step, I summarized the frequencies of questions in the AMC and IC categories, to check for independence of question categories between the two classes. In 5A, there were 73 AMC-questions and 75 IC-questions, whereas there were 86 AMC-questions and 31 IC-questions in 5B. I ran both chi-square tests and Fisher exact tests and got $p < .001$ for both, indicating that there is a significant difference in the types of questions asked in the two classes. An odds ratio of .35 indicated a negative connection between the two classes, and that the teacher asked fewer AMC-questions and more IC-questions in 5A than in 5B. However, when I ran Fisher exact test with Bonferroni-adjusted significance levels to check for independence of individual question types across classes, there were no significant differences. This indicates that although there is a tendency of difference with respect to AMC-questions and IC-questions between classes, there are no significant differences for any individual question types between classes.

The second research question asks if there are differences in teacher questioning across instructional settings. I had already observed that the first part of the lessons in the two classes was normally organized as whole-class discussions, whereas the second part of the lessons was typically organized as individual seatwork. I therefore decided to focus on these two types of organization of instructional discourse (Cohen, 2011). Based on the observations, I wondered if there were any connections between the types of questions asked in whole-class discussion settings and when approaching students during individual seatwork. The null hypothesis is that there is no such difference, whereas the alternative hypothesis is that the teacher asks different types of questions in different instructional settings. Since the setting (discussion or seatwork) and the different question types are categorical variables in this study, and since the conditions for chi-square tests were not met, I ran Fisher exact tests to check for differences in questioning between settings within the same lesson. Since the number of observations were low, and since some cells contained values of less than 5, I ran these tests with Monte Carlo simulation. In both classes, I got $p < .001$, which indicates that there is a

significant difference in questioning between settings. As a next step, I checked for differences in questioning with respect to the overall question categories across settings. In 5A, there were 58 AMC-questions and 32 IC-questions in discussions, and 15 AMC-questions and 43 IC-questions in seatwork. Chi-square test gave $p < .001$, which indicates that there is a significant difference between settings in 5A. Odds ratio was 5.13, which indicates that there is a clear tendency of more AMC-questions in discussions as compared with seatwork in 5A. When making the same comparison in 5B, however, there were no significant differences between settings. In 5B, the teacher asked 53 AMC-questions and 17 IC-questions in discussion settings, and 33 AMC-questions and 14 IC-questions in settings where students engaged in individual seatwork, and the p-values were .53 and .66 for the Fisher exact test and the chi-square test respectively, which indicates no significant differences between settings in 5B.

As a follow up, I ran Fisher tests for each of the question types across settings. With Bonferroni-adjustment, the significance level was .0022. Four question types stood out, but there were differences across classes here as well (see Table 3). In both classes, there were clear indications that questions to call on students to participate (*ic_call*) were most frequently used in discussion settings (p-values of .002 in both classes). An example of such questions is when the teacher asks, “Do you want to try, Adam?” For the other three question types, there was variation across classes. Check-in questions (*ic_checkin*) had $p < .001$ in 5A, but only $p = .025$ in 5B. Questions that prompted students to think (*amc_think*) were clearly more common in discussions with low p-values in both 5A and 5B. Questions that prompted students to take a stance (*amc_stance*)—often posed as, “What do we think?”—were also more common in discussions than in seatwork in both classes.

Table 3: Relationships between key variables in discussion settings in each class

	<i>ic_call</i>	<i>ic_checkin</i>	<i>amc_think</i>	<i>amc_stance</i>
5A:				
p-value	.002	.000	.000	.005
Odds Ratio	2.03	-2.44	3.13	2.39
5B:				
p-value	.002	.025	.011	.000
Odds ratio	3.10	-1.19	2.78	3.17

To further measure the strength of the associations between each of these pairs of variables, I used odds ratio (Pearce, 1993). Table 3 shows that the odds ratio for the association between whole-class setting and *ic_checkin* questions is negative (-2.44 in 5A and -1.19 in 5B). This means that check-in questions are less likely to be posed in discussion settings than when students are engaged in individual seatwork. Associations between discussion settings and the other three question types displayed, however, were positive. Ratios with values higher than 1 indicate that there is a positive connection between two variables, and we notice that odds ratio values were well above 1 for *ic_call*, *amc_think*, and *amc_stance*, and in both classes. In other words, the teacher asked more questions to call on students, to prompt students to take a stance, or to elaborate on their thinking in the discussion setting as compared to settings where the students were engaged in individual seatwork. In one of the lessons, there was a significant negative association between discussions and *amc_direct* questions, but in the other lesson there was not.

Discussion

Some studies in the review by DeJarnette et al. (2020) indicate that teacher questioning varies between groups of students, for instance according to their race. Other studies indicate that teachers' beliefs or expectations might influence their teaching. This was confirmed in the present study, where teacher questioning was significantly different when the same teacher taught two similar lessons in two different classes. This indicates that teacher questioning might indeed vary based on the teacher's expectations of the students.

When comparing teacher questioning across instructional settings, some significant differences were found. For instance, questions that called on students to participate, invited them to share their thinking, or to take a stance with respect to the mathematical content, were more common in discussion settings. On the other hand, check-in questions were more common when the teacher walked around and helped students who were engaged in seatwork. Such differences might be expected, and they can be explained by the nature of the questions in these categories. Questions like "Do you want to share your solution?" or "Do you agree with X?" are by nature more geared toward discussion settings. Still, I wonder why few, if any, other studies have considered question types in connection with different instructional settings. Enright et al.'s (2016) definition of teacher questions as instructional moves that are situated in particular environments seems to imply that research on teacher questioning must be sensitive to variations in settings. Yet, except for a few studies that indicate differences in teacher questioning between countries (e.g., Enright et al., 2016; Perry et al., 1993), the tendency in the research on teacher questioning reviewed by DeJarnette et al. (2020) is that question types are considered without regard for instructional settings. This study suggests that such considerations should be made.

The findings in this study are only based on data from an exploratory case study of one teacher and her teaching of the same lesson in two classes, and they cannot be generalized to a larger population. Still, the study provides existence-proof that variation in question types across settings is possible. Further investigations, both qualitative and quantitative, would be interesting here. Quantitative studies could explore differences in teacher questioning between different mathematical topics, or between different age groups or grade levels. Longitudinal studies could explore if teacher questioning changes naturally over time, for instance over the course of a school year. Qualitative studies could explore what is involved in questioning practices with different mathematical topics, and how the demands of teacher questioning might differ between age groups and grade levels.

Acknowledgment

I would like to acknowledge the assistance from a group of master students who participated in the process of collecting data and transcribing the recordings.

References

- Aziza, M. (2018). An analysis of a teacher's questioning related to students' responses and mathematical creativity in an elementary school in the UK. *International Electronic Journal of Elementary Education*, 10(4), 475–487. <https://doi.org/10.26822/iejee.2018438138>
- Cazden, C. B. (2001). *Classroom discourse: the language of teaching and learning*. Heinemann.

- Cohen, D. K. (2011). *Teaching and its predicaments*. Harvard University Press.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142. <https://doi.org/10.3102/01623737025002119>
- DeJarnette, A. F., Wilke, E., & Hord, C. (2020). Categorizing mathematics teachers' questioning: The demands and contributions of teachers' questions. *International Journal of Educational Research*, 104, 101690. <https://doi.org/10.1016/j.ijer.2020.101690>
- Di Teodoro, S., Donders, S., Kemp-Davidson, J., Robertson, P., & Schuyler, L. (2011). Asking good questions: Promoting greater understanding of mathematics through purposeful teacher and student questioning. *The Canadian Journal of Action Research*, 12(2), 18–29. <https://doi.org/10.33524/cjar.v12i2.16>
- Enright, E. A., Hickman, L., & Ball, D. L. (2016, July). *A typology of questions by instructional function*. Paper presented at the 13th International Congress on Mathematical Education, Hamburg, Germany.
- Gjære, Å. L., & Blank, N. (2019). Teaching mathematics developmentally. *For the Learning of Mathematics*, 39(3), 28–33.
- Mosvold, R. & Wæge, K. (2022). Entailments of questions and questioning practices in ambitious mathematics teaching. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti (Eds.), *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)* (pp. 3434–3441). Free University of Bozen-Bolzano and ERME.
- Pearce, N. (1993). What does the odds ratio estimate in a case-control study? *International Journal of Epidemiology*, 22(6), 1189–1192. <https://doi.org/10.1093/ije/22.6.1189>
- Perry, M., VanderStoep, S. W., & Yu, S. L. (1993). Asking questions in first-grade mathematics classes: potential influences on mathematical thought. *Journal of Educational Psychology*, 85(1), 31–40. <https://doi.org/10.1037/0022-0663.85.1.31>
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of educational research*, 51(2), 237–245. <https://doi.org/10.2307/1170197>
- Reynolds, H. T. (1977). *Analysis of nominal data*. Sage Publications.
- Schulte, E., & Davison, D. (2011). Active documents with org-mode. *Computing in Science & Engineering*, 13(3), 66–73.