Journal of Mathematical Behavior 72 (2023) 101087

Contents lists available at ScienceDirect

Journal of Mathematical Behavior

journal homepage: www.elsevier.com/locate/jmathb

Impact of prompts on students' mathematical problem posing

Jinfa Cai *, Hua Ran **, Stephen Hwang, Yue Ma, Jaepil Han, Faith Muirhead

University of Delaware, USA

ARTICLE INFO

Keywords:
Problem posing
Problem-posing prompt
Problem-posing processes
Task variables
Task characteristics
Teaching mathematics through problem posing
P-PBL

ABSTRACT

This study used three pairs of problem-posing tasks to examine the impact of different prompts on students' problem posing. Two kinds of prompts were involved. The first asked students to pose 2–3 different mathematical problems without specifying other requirements for the problems, whereas the second kind of prompt did specify additional requirements. A total of 2124 students' responses were analyzed to examine the impact of the prompts along multiple dimensions. In response to problem-posing prompts with more specific requirements, students tended to engage in more in-depth mathematical thinking and posed much more linguistically and semantically complex problems with more relationships or steps required to solve them. The findings from this study not only contribute to our understanding of problem-posing processes but also have direct implications for teaching mathematics through problem posing.

1. Introduction

Mathematical problem posing (MPP) has been at the forefront of discussion for the past few decades (Brown & Walter, 1983; Cai, 1998, 2022; Ellerton, 1986; English, 1998; Kilpatrick, 1987; Silver, 1994; Silver & Cai, 1996; Singer et al., 2015). This increased research activity on problem posing is reflected in not only journal special issues (Cai & Hwang, 2020; Cai & Leikin, 2020; Liljedahl & Cai, 2021; Singer et al., 2013) but also books (e.g., Felmer et al., 2016; Singer et al., 2015) and conferences (e.g., ICME-14: TSG 17, PME45: Problem-Posing Research Forum). Such increased attention toward problem posing is at least partly due to problem posing's reputation as an activity with a low floor and high ceiling and its potential to offer opportunities for mathematical sensemaking for all students.

One important angle of problem-posing research has been attempts to understand problem-posing processes. Some researchers (e.g., Silber & Cai, 2017) have focused on the cognitive and affective processes of posing mathematical problems whereas others (e.g., Zhang & Cai, 2021) have focused on the processes of integrating problem posing into classrooms. In this study, our focus is on understanding the impact of prompts on students' problem-posing processes. The research questions for this study are: How do problem-posing prompts influence middle school students' problem-posing responses? More specifically, how do the prompts influence the solvability and complexity of problems that students pose? How do prompts influence the similarity among students' posed problems when they are provided a sample problem? Based on these research questions, we make four hypotheses about how different prompts influence students' responses on problem-posing tasks, which we discuss in detail in the next section.

* Correspondence to: University of Delaware, 437 Ewing Hall, Newark, DE 19716, USA.
** Corresponding author.
E-mail addresses: jcai@udel.edu (J. Cai), hxr240@miami.edu (H. Ran).

https://doi.org/10.1016/j.jmathb.2023.101087
Received 17 May 2023; Received in revised form 15 August 2023; Accepted 17 August 2023
Available online 8 September 2023

2. Theoretical bases and hypotheses

2.1. Problem-posing tasks

By *problem posing* in mathematics education, we refer to several related types of activity that entail or support teachers and students formulating (or reformulating) and expressing a problem based on a particular context (which we refer to as the *problem context* or *problem situation*; Cai & Hwang, 2020; Silver, 1994). Problem-posing tasks, then, are those tasks that position students as generators or shapers of new problems based on real-life and mathematical situations. As we have discussed elsewhere, a problem-posing task usually includes two distinguishable parts: a situation and a prompt (Cai, 2022; Cai & Hwang, 2023). The problem situation is what provides problem posers the contextual information and data that they may draw from to craft problems. The contextual information and data may be presented in words, pictures, graphs, patterns, tables, and mathematical expressions. Problem situations can be based on real-world referents or be purely mathematical or abstract. Accompanying the problem situation is a prompt that lets posers know what they are expected to do in response to the problem-posing task. The prompt may be quite general, such as "Pose a problem" or "Write questions using this information." Alternatively, the prompt may include more specific guidance for the problem poser (e.g., "Pose a mathematical problem that can be solved using this information.") or specify particular characteristics that the posed problem (s) should exhibit (e.g., specifying the number of problems to pose or the kinds of problems to pose). Moreover, a prompt may be quite concise, or it may comprise multiple guiding statements and conditions for the posed problems. Ultimately, there are many choices with respect to prompts, and many different prompts may be paired with the same problem situation to create problem-posing tasks with different characteristics. For example, some choices of prompt may result in more open tasks and others in more closed tasks. The focus of this study is to investigate the impact of prompts on students' problem posing.

2.2. Task characteristics and problem posing

Prior research has examined the cognitive processes involved in problem posing as well as the impact of a number of problem-posing task characteristics on problem-posing cognition and the kinds of problems that are posed. Moreover, some researchers (e.g., Baumanns & Rott, 2022) have proposed frameworks for categorizing different kinds of problem-posing situations, positing that different situations induce different kinds of posing activity. In this section, we review some key research that has investigated problem-posing task characteristics and their interactions with problem-posing cognition. Ultimately, however, we will show that the extant research has not yet given sustained attention to the detailed characteristics of problem-posing prompts and how they interact with and affect how posers think.

Focusing on problem-posing cognition, Christou et al. (2005) proposed and tested a theoretical model of young students' problem-posing thinking. The model comprises several different cognitive processes that act on quantitative information, namely editing, selecting, comprehending and organizing, and translating quantitative information. Note that not all these processes are required in every instance of problem posing, and depending on its characteristics, a particular problem-posing task may more prominently invoke one of the processes than the others. Indeed, the researchers empirically tested the model using sets of tasks they had designed to correspond to each of the processes. Notably, the tasks varied substantially in both problem situations and prompts. Thus, although the tasks targeted specific problem-posing cognitive processes, it would not have been feasible to identify the specific characteristics of the prompts that influenced the posers' cognition and posed problems. Rather, each individual combination of problem situation and prompt would have had to be evaluated holistically as a unit to evaluate which problem-posing processes the tasks were likely to invoke.

Other researchers have focused their attention on the impact of more specific characteristics of problem-posing tasks on problem-posing cognition, although much of this work has attended to characteristics of the problem situations rather than characteristics of the prompts. For example, Leung and Silver (1997) developed and analyzed a Test of Arithmetic Problem Posing (TAPP) which they used to examine how the presence of specific numerical information in a problem situation affected preservice teachers' (PSTs') problem-posing abilities. The TAPP problem-posing prompt that guided the PSTs' problem posing included the following six statements:

[1] Consider possible combinations of the pieces of information given and pose mathematical problems involving the purchase and operation of the house. [2] Do not ask questions like "Where is the house?". because this is not a mathematical problem. [3] Set up as many problems as you can think of. Think of problems with a variety of difficulty levels. Do not solve them. [4] Set up a variety of problems rather than many problems of the same kind. [5] Include also unusual problems that your peers might not be able to create. [6] You can change the given information and/or supply more information. When you do so, note the changes in the box with the problem to which they apply (Leung & Silver, 1997, p. 8).

Although the term "prompt" was not explicitly used by Leung and Silver (1997), this set of statements collectively serves the function of the problem-posing prompt in this task: It provides guidance and conditions to satisfy during participants' problem posing. The posed problems were analyzed along two dimensions: quality (mathematical or nonmathematical, plausible or implausible, and contains sufficient or insufficient information) and complexity (the number of steps required to solve the problem). The results from the TAPP indicated that the teachers performed better with respect to both the quality and complexity measures on tasks that included specific numerical information than on tasks without specific numerical information. The researchers postulated that this might be because having specific numbers provided in the problem situation made it unnecessary for the poser either to generate their own numbers or to have to create purely qualitative reasoning problems without numerical information.

Zhang et al. (2022) replicated and extended the study by Leung and Silver (1997), focusing on elementary school students' problem posing. They divided the cognitive process of MPP into three stages: (a) input—understanding the task, (b) processing—constructing the problem, and (c) output—expressing the problem. They also found that the provision of specific numerical information in the problem situation was associated with better problem-posing performance. However, rather than a general effect throughout the problem-posing process, they localized the effect to only the stages of understanding the task and constructing the problem. Together with the work of Leung and Silver (1997), these findings identify a particular impact of one specific characteristic of problem situations on problem-posing cognition.

English (1998) also studied the impact of a characteristic of problem situations, specifically comparing the effect of including formal and informal contexts in problem situations. She designed a problem-posing program for 8-year-old students that incorporated activities with formal contexts (i.e., addition and subtraction number sentences such as 12–8 = 4) and activities with novel, informal, non-operation-based problem situations that included non-goal-specific statements, stimulus pictures, and literature-based tasks. The non-goal-specific statements were sentences such as, "Sarah has five dolls on one shelf in her room and four toy cars on another shelf." After reading the sentence, students were asked to make the statement into a problem they could solve. The stimulus picture showed a large photograph of children playing with sets of colored items; students were asked to make up story problems about something that could be seen in the photograph. The third informal context was a piece of literature supported by a list of numbers of native animals. English (1998) employed pre- and post-tests of students' problem-posing performance that included both formal and informal contexts in the problem situations. The results showed that the students displayed a significant increase in the proportion of part-part-whole problems created in the informal context between the pretest and posttest as well as increases in multistep problems in the formal context.

The nature of the context in the problem situation was also the subject of a study by Guo et al. (2021) of Chinese junior high school students' problem posing. Guo et al. sought to evaluate the problem-posing performance (and trajectory of that performance across grade levels) of students using three problem-posing tasks that differed in contextual knowledge drawn on in the problem situation. One task drew on a common everyday life experience of clothes shopping; another drew on a medical context; the third drew on a purely mathematical context. Although the context differed, all three situations involved underlying algebraic relationships. Moreover, the prompt used was consistent across all three problem-posing tasks, requiring the students "to pose as many mathematical problems as possible, to make them as different as possible, and as difficult as possible, based on the given situation" (Guo et al., 2021, p. 909). The researchers compared the performance across grades, finding variation in problem posing (fluency, flexibility, and profundity) by context: Performance decreased with grade level for the first two contexts and varied irregularly for the third context.

One strand of research on characteristics of problem-posing situations has drawn on the framework of Stoyanova and Ellerton (1996), who classified problem situations as free, semi-structured, and structured. This framework has proven somewhat challenging to operationalize consistently (e.g., Baumanns & Rott, 2022, found it necessary to combine the free and semi-structured categories into a spectrum of less restrictive problem situations), but a relatively reliable distinction has been that a structured problem-posing situation includes or is predicated on the existence of a specific problem. Research using this framework has, for example, compared problem-posing performance across the three categories. In a study of Chinese eighth graders' problem-posing abilities, Xu and Li (2021) analyzed the students' performance on six problem-posing tasks—two structured, two semi-structured, and two free—finding that student performance decreased as the level of structure decreased. However, the problem situations and the prompts used in this study varied greatly, making it impossible to pinpoint the effect of specific differences aside from the categorization of the task as free, semi-structured, and structured.

An issue with Stoyanova and Ellerton's (1996) framework is that, although it is explicitly defined with respect to problem situations, deciding whether a problem situation is free, semi-structured, or structured may depend on the nature of the prompt as well. For example, a problem situation can provide information in the form of a story with numerical data. The prompt might then ask the poser to generate problems based on the information in the story. With that prompt, this problem situation would likely be classified as free (or perhaps semi-structured). However, if the prompt were changed to include the condition that the posed problem must have a particular numerical answer (that also happens to be the answer to a predetermined problem based on the story), the problem situation would now be considered structured. Thus, it is possible to draw on Stoyanova and Ellerton's (1996) framework to analyze the effects of both prompt and problem situation characteristics.

Indeed, in one of the very few studies examining the impact of prompt characteristics on problem-posing cognition, Silber and Cai (2017) varied a problem situation from free to structured to compare PSTs' problem posing under the two conditions. In addition, the prompt in that study asked participants to pose multiple problems. Previous research has suggested that, when asked to pose multiple problems in response to a problem-posing task, individuals will first pose simpler mathematical problems and then pose more complex mathematical problems (Christou et al., 2005; Silver & Cai, 1996; Stickles, 2011). Thus, Silber and Cai (2017) proposed three hypotheses related to prompts: (1) PSTs would pose more complex mathematical problems under structured posing conditions than under free posing conditions; (2) the problems that PSTs posed would increase in complexity level in their later responses; and (3) structured posing conditions would encourage PSTs to attend to the underlying mathematical concepts of the tasks. The researchers tested these hypotheses using a mixed-method design conducted with a written assessment that collected 61 PST's problem-posing responses and a task-based interview that probed four PSTs' thinking processes while they engaged with the problem-posing tasks. For the written assessment, 30 PSTs were given tasks with structured and free posing situations, and 31 PSTs were given tasks with only free posing situations. To match the structure of the written assessment, two PSTs participated in interviews with both structured and free posing situations, and the remaining two participated in interviews with only free posing situations. Four problem-posing tasks were used: the Road Trip task, the Doorbell task, the Food Drive task, and the Pattern of Dots task. The two former tasks were adapted into both structured and free versions. The influence of the structured posing situation on PSTs' problem posing diverged based on the task. No significant relationships between the posing situations and the complexity of the PSTs' posed problems were found for one task (Doorbell). For another task (Road Trip), PSTs in the structured posing condition posed significantly more complex problems than those in the free posing condition but only for the first response. This effect was not found for the second or third responses for this task. This means that, after posing their first problem, PSTs in the structured condition subsequently wrote problems almost equally as complex as those in the free condition. For the third and fourth tasks used in Silber and Cai (2017; Food Drive and Dots), PSTs in both the free and structured groups increased the complexity of their posed problems with each response. Regarding the relationship between the posing situation and attention to mathematical concepts, PSTs in the structured condition showed visual evidence of attending to the mathematical concepts underlying the task situations across all four tasks.

In another study that included a component related to problem-posing prompt characteristics, Yao et al. (2021) investigated PSTs' understanding of fraction division. Part of the study involved asking the participants to complete a problem-posing task wherein they generated problems related to a given fraction division (1¾ divided by ½). One of the conditions that Yao et al. tested was the inclusion of a conceptual cue in the prompt that reminded participants of a typical meaningful interpretation of the division statement. They found that the inclusion of this conceptual cue greatly increased the likelihood that the participants would pose problems that exhibited conceptual understanding. Thus, this characteristic of the prompt appears to have had a strong impact on the participants' problem-posing cognition.

In summary, thus far, researchers have worked to identify the effects of problem-posing task characteristics on problem-posing cognition and posed problems. However, most of these studies have either treated problem-posing tasks holistically, making it difficult to attribute effects to particular components of the task, or they have investigated the effects of characteristics of the problem situations. In fact, very few studies have targeted the impact of specific characteristics of problem-posing prompts. Looking across the literature above, we highlight two general findings. The first is with respect to problem situations: Informal problem situations and problem situations with specific information (rather than only abstract information) tend to encourage more complex problem-posing responses. The second is with respect to problem-posing prompts: Problem-posing prompts that include additional specific conditions or guidance such as specifying an answer to the problem to be posed or including conceptual information tend to support the posing of more complex mathematical problems compared to problem-posing prompts that are entirely general and do not specify guidelines.

However, there remain next to no direct comparisons in the literature of students' problem posing under different prompts. This study aims to investigate the influence of different prompts on middle school students' problem posing. In the following section we detail our research purpose and the hypotheses we propose related to the effects of prompt characteristics.

2.3. Problem-posing prompts and hypotheses

In prior problem-posing studies, researchers have used different types of prompts, such as "pose three different mathematical problems" (Pose-3) and "pose an easy, a moderately difficult, and a difficult mathematical problem" (Pose-EMD; see Cai, 2022, and Cai & Hwang, 2023, for a list of prompts). The original motivation of those studies for using the Pose-EMD prompt was to better engage students in problem posing. However, neither a separate analysis nor a comparison of the Pose-EMD prompt with other prompts have been conducted. Several prior studies have shown that when students are asked to pose multiple problems, later problems seem to be more complex than earlier problems, even when there is no particular demand placed on students to do so (e.g., Silver & Cai, 1996). The Pose-EMD prompt specifically demands different difficulty levels, so it is quite possible that students would consciously think about the difficulty levels when posing problems when presented with this prompt compared to the Pose-3 prompt. Thus, it is possible that the Pose-EMD prompt might foster more complex problems. Because no studies have directly compared the impact of the Pose-3 prompt with that of the Pose-EMD prompt, this study will directly investigate the impact of the two prompts. To that end, we hypothesize that the Pose-EMD prompt will better facilitate problem posing than the Pose-3 prompt. Given that problem posing can be measured in a variety of ways, we propose the following specific hypotheses with respect to the Pose-3 and Pose-EMD prompts:

(1) Middle school students will be able to pose more solvable mathematical problems under the Pose-EMD prompt than under the Pose-3 prompt. The proportion of solvable mathematical problems will increase in students' later responses for both prompts, but the Pose-EMD prompt will encourage a sharper increase in the proportion than will the Pose-3 prompt.
(2) Among the solvable mathematical problems, middle school students will pose more complex mathematical problems under the Pose-EMD prompt than under the Pose-3 prompt. The problems that middle school students pose will increase in complexity level in later responses for both prompts, but the Pose-EMD prompt will encourage a sharper increase in the complexity level than will the Pose-3 prompt.

Prior research has also shown that students tend to pose related problems when they are asked to pose multiple problems. For example, Silver and Cai (1996) found that almost one half of students' posed problems were either parallel or chained responses. This is not surprising given that "what if not" is a commonly used problem-posing strategy (Brown & Walter, 1983; Cai & Cifarelli, 2005). What if a sample problem is given but with prompts only asking for students to pose additional problems or to pose similar or different problems? In this study, we have also examined the impact of another pair of prompts that arise when a sample posed problem is given: (1) Pose two additional problems given a sample problem (Pose-2) and (2) pose one problem similar to and one problem different from a given sample problem (Pose-SD). The Pose-SD prompt provides additional conditions that require students to think about the relationship between the problems they will generate and the sample problem (similar and different); therefore, it is natural that the Pose-SD prompt could help students pose more different and possibly more complex problems. Like the Pose-3 and Pose-EMD comparison, we propose the following two additional hypotheses:

(1) Middle school students will pose more solvable mathematical problems under the Pose-SD prompt than under the Pose-2 prompt. The proportion of solvable mathematical problems will increase in students' later responses for both prompts, but the Pose-SD prompt will encourage a sharper increase in the proportion than will the Pose-2 prompt.
(2) Among the solvable mathematical problems, the first posed problem will be more similar to the given sample problem under the Pose-SD prompt than under the Pose-2 prompt, and the second posed problem will be more different from the given sample problem under the Pose-SD prompt than under the Pose-2 prompt.

3. Method

3.1. Participants and tasks

This study is based on data gathered as part of a large research project aiming to support middle school mathematics teachers to teach mathematics through problem posing. The participants in the first year of the larger project (when these data were collected) were 15 middle school mathematics teachers in two schools and the students in 14 classes taught by these teachers (one class was taught by a team of two teachers). The problem-posing task responses analyzed in this study came from a problem-posing assessment administered to the participating students as part of the project. Responses to three pairs of problem-posing tasks were analyzed in this study. The tasks were based on three problem situations: the Road Trip situation, the Doorbell situation, and the Animal situation.

The Road Trip situation is as follows: *Jerome, Elliot, and Arturo took turns driving home from a trip. Arturo drove 80 miles more than Elliot. Elliot drove twice as many miles as Jerome. Jerome drove 50 miles.* Two prompts were used with this problem situation, thereby generating the first pair of problem-posing tasks. The first prompt asked students to pose "three different mathematical questions that can be answered from this information" (Pose-3). The second prompt asked students to pose "three different mathematical questions that can be answered from this information: an easy question, a somewhat difficult question, and a difficult question" (Pose-EMD).

The Doorbell situation is as follows: *Sally is having a party. The first time the doorbell rings, 1 guest enters. The second time the doorbell rings, 3 guests enter. The third time the doorbell rings, 5 guests enter. The fourth time the doorbell rings, 7 guests enter. This pattern continues in the same way. On the next ring a group enters that has 2 more guests than the group that entered on the previous ring.* The same two prompts that were used with the Road Trip situation were also used with the Doorbell situation to produce the second pair of problem-posing tasks.

The Animal situation includes a short table (Table 1) and is as follows: *One problem that can be answered based on the information in the table below is, "How many times as fast is a spider compared to a tortoise?"* In this situation, a sample problem was given. Again, two prompts were used to generate the third pair of problem-posing tasks. The first prompt was, "Pose two more problems that can be solved using the information in the table" (Pose-2). The second was, "Pose a similar problem that can be solved using the information in the table" followed by "Now pose a very different problem that can be solved using the information in the table" (Pose-SD).

The problem-posing assessment had two forms, Form A and Form B, which differed only in the type of problem-posing prompt used. In Form A, Pose-EMD prompts were used for the Road Trip and Doorbell tasks, and the Pose-SD prompt was used for the Animal task. In Form B, Pose-3 prompts were used for the Road Trip and Doorbell tasks, and the Pose-2 prompt was used for the Animal task. Within each class, we randomly assigned students to take Form A or Form B. In total, 151 and 153 students were assigned Forms A and B, respectively. Due to absences and some students arriving late or leaving early during the assessment, the final numbers of students who completed each task varied from 127 to 129 for Form B and from 137 to 139 for Form A. For this paper, we aggregated data from the three grades because there was no differential effect of the prompts across the three grade levels. Table 2 shows the number of students who responded to each task and the numbers of responses produced for each of the three tasks. Ultimately, there were 129 students who completed the Road Trip task using the Pose-3 prompt and 137 students who did so using the Pose-EMD prompt. Meanwhile, 128 students completed the Doorbell task using the Pose-3 prompt and 139 students did so using the Pose-EMD prompt. For the Animal task, 127 students used the Pose-2 prompt and 137 used the Pose-SD prompt.

3.2. Data coding and analysis

We followed the procedures used in Silver and Cai (1996) and Cai and Hwang (2002) to code the data: (a) examining response types, (b) examining if posed problems are solvable, and (c) examining the complexity of solvable mathematical problems. Table 3 summarizes the data coding across the three tasks. The details of the coding process are explained in the next section.

Students' responses were first coded according to the following three general categories: statements, non-mathematical questions, and mathematical questions/problems. A statement is a sentence or numerical expression that does not include any mathematical demands or questions. A non-mathematical question is a question that does not involve any mathematical relationship, computation, or content. Mathematical questions were further coded as solvable mathematical problems if there was sufficient information to

Table 1
Table presented in the Animal situation.

Animal	Crawling speed
Snail	50 m per hour
Tortoise	1000 m per hour
Spider	6000 m per hour

Table 2
Number of student assessments by tasks and prompts.

Tasks	Form B: Pose-3/Pose-2		Form A: Pose-EMD/Pose-SD	
	NS	NR	NS	NR
The Road Trip task	129	387	137	411
The Doorbell task	127	381	139	417
The Animal task	127	254	137	274

Note. NS is the number of students who took the assessment; NR is the number of responses.

Table 3
Summary of data coding for the three tasks.

	Road Trip task	Doorbell task	Animal task
Problem-posing types (i.e., solvability)	X	X	X
Linguistic structure	X		X
Semantic structure	X		X * (number of operations)
Extension		X1 * (figure extension) X2 * (number pattern)	X * (topical extension)
Similarities/differences			X * (numerical information)

generate solutions and as non-solvable mathematical problems if there was not enough information available to find a solution or the responses were self-contradictory with the given information. The solvable mathematical problems were further coded as mathematical expressions that only contained numbers (e.g., $50 + 1000 + 6000 = ?$) and mathematical word problems that included story contexts.

At least 20% of students' responses for the three tasks were randomly selected and independently coded by two raters. The inter-rater reliability was calculated as Krippendorff's alpha (Krippendorff, 2004), which was 91% for the Road Trip task, 92% for the Doorbell task, and 89% for the Animal task. Any disagreement during the coding process was discussed with the research team, and agreements were reached.

Subsequently, each solvable mathematical problem was further analyzed for its complexity level. For the Road Trip and Animal tasks, the complexity level was first evaluated based on Marshall's (1995) five semantical structural relations of word problems: change, group, compare, restate, and vary. Change refers to a difference in a measurable quantity of a particular thing over a period of time. However, in the current data set, no such example was found. *Group* refers to the combination of a number of small groups into a large group. *Compare* refers to determining the difference in quantity between two things or items. *Restate* refers to a specific relationship between two different things at a fixed point in time. Finally, *Vary* refers to a situation that has altered the conditions of the given situation. Frequently, a vary situation entails "if-then" statements. In addition, we considered a mathematical problem to have a "zero" relation with respect to complexity if the problem could be answered directly from the given information. Table 4 provides descriptions and examples of the structural relations for the Road Trip tasks.

Silver and Cai (1996) and Silber and Cai (2017) considered a problem more mathematically complex if it contained more relations. Thus, after coding the structural relations, we also counted the number of relations for each response to evaluate its complexity level. Table 5 presents examples of how the number of relations for each response was coded.

Another type of complexity we examined for the Road Trip task was the linguistic structure of the word problems. We followed the categories used in Silver and Cai (1996) based on Mayer's et al. (1992) three propositions: assignment, relational, and conditional. An assignment proposition is a question asking for the assignment of a specific value to an unknown quantity in the problem, such as "How many miles did Elliot drive?" A relational proposition compares two quantities, such as "How many miles less did Elliot drive than Arturo?" A conditional proposition is a question that changes a given condition, such as "If Stacy drives 3 times as much as Jerome, how many miles does she drive?" Silver and Cai (1996) viewed mathematical questions with relational or conditional proportions as more linguistically complex.

For the Doorbell task, which was fundamentally based on a numerical pattern, the complexity level was considered from two perspectives. The first, based on Cai and Hwang (2003), was to consider whether the posed problem extended beyond the given numbers in the task (i.e., beyond the fourth number in the pattern). These were categorized as extension problems. In contrast, non-extension problems were restricted solely to the first four given numbers in the pattern (e.g., "How many guests came when the

Table 4
Descriptions and examples of structural relations.

Relations	Description	Examples (Road Trip task)
Zero	Answer can be directly read from the task statement	Then how many more miles did Arturo drive than Elliot?
Group	Combination of two values together	How many miles in all
Compare	Determining the difference in quantity between two things	How many more did Jerome drive than Elliot?
Restate	Find a quantity that is not directly given in the task	How many miles did Elliot drive
Vary	Change the original task conditions	If they had to travel there and back how many miles would they each drive

Table 5
Number of relations for the Road Trip task.

Relations	Examples (Road Trip task)	Coding and explanations
Zero	How many miles did Jerome drive?	Directly read the answer from the task
One	How many miles did Elliot drive?	Restate: Find miles for Elliot
Two	How many miles did Elliot and Jerome drive together	Restate/group: Find the miles for Elliot and group Elliot and Jerome
Three	How many more miles did A drive than E	Restate/restate/compare: Find the miles for Elliot, find the miles for Arturo, and compare the miles between Arturo and Elliot
Four	If they went 10 more miles, how many miles would they all drive?	Restate/restate/group/vary: Find the miles for Elliot, find the miles for Arturo, group all three, and add 10 more miles
Five	If they had to travel there and back how many miles would they each drive	Restate/restate/vary/vary/vary: Find the miles for Elliot, find the miles for Arturo, times two for Jerome, times two for Elliot, and times two for Arturo
Six	If they each had to drive there and back how many miles would they drive and how much total miles would there be	Restate/restate/vary/vary/vary/group: Find the miles for Elliot, find the miles for Arturo, times two for Jerome, times two for Elliot, times two for Arturo, and group all three together

doorbell rang the fourth time?"). The extension problems included both extensions of the number of guests, such as "How many people entered on the 12th ring?" and of the number of rings, such as "If there are 19 guests, how many times did the doorbell ring?" The second perspective was to consider whether the posed problems included a number pattern or not. Pattern-related problems contained some kind of number-pattern relationship, which could be the same as the given one (i.e., increasing by 2 each time) or different (e.g., increasing by 55 each time). In contrast, non-pattern problems did not explicitly include a number pattern, such as, "There are four pancakes at the table and 8 people, so how will everyone get the pancake?"

For the Animal task, which included a sample problem in the task statement, we coded along four dimensions to explore whether the posed problems were similar to or different from the original situation and the sample problem. These four dimensions included numerical information, the extension of math topics (adapted from Cai & Hwang, 2002), linguistic structure (adapted from Silver & Cai, 1996), and the number of operations. Numerical information refers to whether the problem used numbers from the given information. There were three categories for numerical information: (a) no numerical information; (b) related numerical information, where the numbers in the posed problem were taken from the given information; and (c) not-related numerical information, where the numbers in the posed problem were not taken from the given information. The extension of math topic coding involved two categories: (a) within, where the posed problem involved the same topic as the given one (which involved speed) and (b) beyond, where the posed problem asked about a different math topic, such as distance, number of hours, or depth. The linguistic structure coding followed the same codes as the Road Trip task. The number of operations involved coding the operations required for solving the problem and then counting the number of operations. For instance, "how fast would all then be combined?" is solved by two additions.

Regarding the data analysis, the chi-square test for association determines whether there is a significant association between two nominal variables (Arnholt, 2007). The chi-square test is reported in the following form: χ^2 (df, N) = chi-square statistic value, $p = p$ value. The degrees of freedom (df) for the chi-square test are calculated using the formula: $df = (r - 1)(c - 1)$ where r is the number of rows and c is the number of columns. The effect size is calculated through eta square $\left(\widehat{\eta}^2\right)$ based on the chi-square coefficient, the p value, and the number of observations (Ben-Shachar et al., 2021).

4. Results

4.1. *Influence of prompts on problem-posing responses: Pose-3 versus Pose-EMD*

The first hypothesis was to test whether the Pose-EMD prompt helped students pose more solvable mathematical problems than the Pose-3 prompt and whether the Pose-EMD prompt encouraged more solvable mathematical problems in students' later responses. This section mainly presents and discusses the results of this hypothesis from the Road Trip and Doorbell tasks.

Table 6 shows the different response types for the Road Trip and Doorbell tasks for the two different prompts. Overall, there was no difference between the two prompt conditions with respect to the distribution of the response types. A chi-square test showed that the different types of posed problems were not significantly associated with the two types of prompts for the Road Trip task (χ^2 (5, 798) = 9.68, $p = .08$, $\widehat{\eta}^2 = 0.02$). The results for the Doorbell task aligned with those for the Road Trip task. A chi-square test showed that the different types of responses were not significantly associated with the two types of prompts for the Doorbell task (χ^2 (6, 798) = 5.22, $p = .51$, $\widehat{\eta}^2 = 0.01$).

It is encouraging that students with very limited problem-posing experience provided more than half of the responses as solvable mathematical word problems for the Road Trip and Doorbell tasks. Moreover, there was no difference regarding mathematical word problems between the two prompts for the Road Trip task (57.36% vs. 56.93%) and the Doorbell task (51.97% vs. 52.04%). The only noticeable difference was in the no response rates for the Road Trip task (9.30% vs. 5.35%, χ^2 (1, 58) = 4.04, $p = .02$).

For both tasks, we also investigated whether there was an association between students' response types and the positions of the responses—in other words, whether a response was given as the first, second, or third posed problem (or as the easy, moderately difficult, or difficult problem, respectively). Table 7 shows the different response types for the Road Trip and Doorbell tasks broken down by position in the sequence of posed problems (First/Easy, Second/Moderately Difficult, or Third/Difficult) for the two prompts. We ran the chi-square tests separately for each prompt—that is, for the Pose-3 and Pose-EMD prompts respectively. A chi-square test

Table 6
Percentage of response types by prompt for the Road Trip and Doorbell tasks.

Response types	Road Trip ($N = 798$)		Doorbell ($N = 798$)	
	Pose-3	Pose-EMD	Pose-3	Pose-EMD
	($N = 387$)	($N = 411$)	($N = 381$)	($N = 417$)
No response	9.30	5.35	14.96	17.51
Incomprehensible	–	–	1.57	2.16
Statement	9.82	12.41	11.81	11.99
Non-math questions	1.55	2.68	0.79	1.44
Mathematical expressions	18.60	16.79	15.75	13.43
Word problems	57.36	56.93	51.97	52.04
Non-solvable math questions	3.36	5.84	3.15	1.44

showed that the different types of problem-posing responses were not significantly associated with position in the sequence under the Pose-3 prompt condition (χ^2 (10, 387) = 6.98, p = .72, $\widehat{\eta}^2$= 0.01) but were significantly associated with position in the sequence under the Pose-EMD prompts (χ^2 (10, 411) = 18.67, p = .04, $\widehat{\eta}^2$= 0.04) for the Road Trip task. Specifically, the proportion of solvable mathematical word problems for the Road Trip task significantly decreased in students' later responses[1] for the Pose-EMD prompt (from 63.50% to 53.28%, χ^2 (1, 160) = 2.53, p = .05). However, the proportion of non-responses significantly increased in students' later responses for the Pose-3 prompt (from 6.20%. to 13.18%, χ^2 (1, 25) = 2.83, p = .04) and for the Pose-EMD prompt (from 1.46% to 8.76%, χ^2 (1, 14) = 6.09, p = .006). Thus, our hypothesis that the Pose-EMD prompt would help students pose more solvable mathematical problems and encourage a sharper increase in the proportion of solvable mathematical problems is rejected.

Based on chi-square tests, no significant differences in students' response types were found to be related to the posed problem position for the Pose-3 prompts (χ^2 (12, 381) = 15.97, p = .19, $\widehat{\eta}^2$= 0.03) or the Pose-EMD prompts (χ^2 (12, 417) = 12.78, p = .38, $\widehat{\eta}^2$= 0.03) for the Doorbell task. However, the proportion of no responses significantly increased from the Pose-First (9.45%) to the Pose-Third (21.26%, χ^2 (1, 39) = 5.93, p = .007) and from the Pose-Easy (12.59%) to the Pose-Difficult (21.58%, χ^2 (1, 48) = 3.04, p = .04) prompts. This seems to be a downside of sequential problem posing—that is, students chose not to generate a mathematical question for the later position. Aside from this difference, no other obvious changes or differences were observed for the Doorbell task.

4.2. Influence of prompts on complexity levels of posed solvable mathematical problems: Pose-3 versus Pose-EMD

The second hypothesis was to test whether the Pose-EMD prompt helped students pose more complex mathematical problems than the Pose-3 prompt and whether the Pose-EMD prompt encouraged a shaper increase in students' later responses. For the Road Trip task, the complexity level was examined through linguistic and semantic structure relations; for the Doorbell task, the complexity level was examined through the extension of the given figures and number-pattern relationships.

4.2.1. Linguistic and semantic structure relations for the Road Trip task

Table 8 presents the results of the complexity level analyses in terms of linguistic structure and semantic structure relations for solvable mathematical problems for the Road Trip task based on prompt. A chi-square test showed that the linguistic structures of solvable mathematical problems were statistically significantly associated with the two prompts for the Road Trip task (χ^2(2, 448) = 6.68, p = .03, $\widehat{\eta}^2$= 0.02). Noticeably, the proportion of solvable mathematical problems with relational and conditional propositions under the Pose-EMD prompt was much larger than under the Pose-3 prompt, and mathematical problems with conditional propositions showed a sharper increase under the Pose-EMD prompt (13.18% vs. 21.05%, χ^2 (1, 77) = 4.33, p = .01) than under the Pose-3 prompt. As mentioned earlier, the relational and conditional propositions are considered indicators of problems with more mathematical complexity. In this case, it was concluded that the Pose-EMD prompt encouraged more complex mathematical problems for middle school students' problem posing than the Pose-3 prompt.

In terms of the semantic structure relations, a chi-square test showed that the semantic structure relations of solvable mathematical problems were not statistically significantly associated with the two prompts for the Road Trip task (χ^2(3, 448) = 4.44, p = .21, $\widehat{\eta}^2$= 0.01). Despite the non-significant association between prompts and semantic structure relations, it is encouraging that the proportion of word problems with three or more relations was much larger for the Pose-EMD prompt (43.42%) than for the Pose-3 prompt (38.18%, χ^2 (1, 183) = 1.06, p = .15).

Table 9 presents the results of the complexity level analyses in terms of linguistic structure and semantic structure relations for solvable mathematical problems for the Road Trip task based on the position of the posed problem in the sequence of posed problems. Two chi-square tests based on the prompt conditions were run separately to examine whether the position of the posed problem would

[1] For students' later responses, we examined the influence on response type of the position of the posed problem in the sequence, achieved through comparing the first-posed problems with the third-posed problems for the Pose-3 prompt and the problems posed as easy problems with the problems posed as difficult problems for the Pose-EMD prompt.

Table 7
Percentage of problem-posing types by position for the Road Trip and Doorbell tasks.

Response types	Road Trip (*N* = 798)						Doorbell (*N* = 798)					
	Pose-3 (N = 387)			Pose-EMD (N = 411)			Pose-3 (*N* = 381)			Pose-EMD (*N* = 417)		
	First (*N* = 129)	Second (*N* = 129)	Third (*N* = 129)	Easy (*N* = 137)	Moderate (*N* = 137)	Difficult (*N* = 137)	First (*N* = 127)	Second (*N* = 127)	Third (*N* = 127)	Easy (*N* = 139)	Moderate (*N* = 139)	Difficult (*N* = 139)
No response	6.20	8.53	13.18	1.46	5.84	8.76	9.45	14.17	21.26	12.95	17.99	21.58
Incomprehensible							3.15	0.00	1.57	2.16	1.44	2.88
Statement	11.63	10.08	7.75	13.14	13.87	10.22	14.17	11.81	9.45	13.67	12.23	10.07
Non-math questions	0.78	1.55	2.33	3.65	1.46	2.92	0.79	0.79	0.79	3.60	0.00	0.72
Math expressions	18.60	19.38	17.83	16.79	18.25	15.33	17.32	16.54	13.39	13.67	14.39	12.23
Word problems	60.47	55.81	55.81	63.50	54.01	53.28	53.54	54.33	48.03	53.24	52.52	50.36
Non-solvable math questions	2.33	4.65	3.10	1.46	6.57	9.49	1.57	2.36	5.51	0.72	1.44	2.16

Note. The solvable mathematical problems included math expressions and mathematical word problems.

Table 8
Percentage of linguistic structures and semantic structure relations by prompt for the Road Trip task.

Dimensions	Road Trip ($N = 448$)	
	Pose-3 ($N = 220$)	Pose-EMD ($N = 228$)
Linguistic structures		
Assignment	73.18	62.28
Relational	13.64	16.67
Conditional	13.18	21.05
Semantic structure relations		
0	3.64	3.51
1	34.09	36.84
2	24.09	16.23
≥3	38.18	43.42

Table 9
Percentage of linguistic structures and semantic structure relations by position for the Road Trip task.

Dimensions	Road Trip ($N = 448$)					
	Pose-3 ($N = 220$)			Pose-EMD ($N = 228$)		
	First ($N = 78$)	Second ($N = 72$)	Third ($N = 70$)	Easy ($N = 85$)	Moderate ($N = 72$)	Difficult ($N = 71$)
Linguistic structures						
Assignment	80.77	73.61	64.29	69.41	62.50	53.52
Relational	10.26	15.28	15.71	16.47	19.44	14.08
Conditional	8.97	11.11	20.00	14.12	18.06	32.39
Semantic structure relations						
0	6.41	4.17	0.00	5.88	1.39	2.82
1	50.00	25.00	25.71	56.47	30.56	19.72
2	20.51	26.39	25.71	9.41	20.83	19.72
≥3	23.08	44.44	48.57	28.24	47.22	57.75

influence the complexity level of the solvable mathematical problems for the Road Trip task. The chi-square tests showed that the linguistic structures of the solvable mathematical problems were not statistically significantly associated with the position of the posed problem in the sequence for the Pose-3 prompt ($\chi^2(4, 220) = 6.14, p = .18, \widehat{\eta}^2 = 0.05$) or for the Pose-EMD prompt ($\chi^2(4, 228) = 8.77, p = .06, \widehat{\eta}^2 = 0.06$) for the Road Trip task. Although there was no overall difference, there were significant differences for conditional propositions. In particular, students posed solvable mathematical problems with increasing complexity levels, reflected in the increase of conditional propositions in the later responses compared to the earlier ones, which was true both for the Pose-3 prompt (from 8.97% to 20.00%, χ^2 (1, 21) = 2.83, $p = .04$) and the Pose-EMD prompt (from 14.12% to 32.39%, χ^2 (1, 35) = 6.41, $p = .005$). Furthermore, there was a relatively sharper contrast regarding the conditional proportions between the Pose-3 prompt and the Pose-Difficult prompt (Pose-3, 20.00% vs. Pose-Difficult, 32.39%, χ^2 (1, 37) = 2.19, $p = .06$). Thus, it seems that the Pose-Difficult prompt stimulated more linguistically complex mathematical problems.

Lastly, the chi-square tests showed that the semantic structure relations of solvable mathematical problems were statistically significantly associated with the position of the problem in the sequence for the Pose-3 prompt (χ^2 (6, 220) = 21.25, $p = .001$, $\widehat{\eta}^2 = 0.14$) and for the Pose-EMD prompt (χ^2 (6, 228) = 29.71, $p < .001$, $\widehat{\eta}^2 = 0.18$) for the Road Trip task. Specifically, there were significant increases in the proportions of three or more relations in students' later responses for both prompts. Again, the Pose-EMD prompt appeared to exhibit a sharper increase in students' later responses (from 28.24% to 57.75%, χ^2 (1, 65) = 12.67, $p = .0001$). Therefore, the later responses tended to be somewhat more linguistically and semantically complex than the earlier responses, and the

Table 10
Percentage of extensions of given figure and pattern by prompt for the Doorbell task.

Dimensions	Pose-3	Pose-EMD
Extension of the given figures		
Total	$N = 159$	$N = 175$
No extension	34.59	45.14
Extension	65.41	54.86
Pattern relation		
Total	$N = 198$	$N = 217$
Pattern-related	95.45	97.24
Not-pattern related	4.55	2.76

Pose-EMD prompt helped students pose more complex solvable mathematical problems.

4.2.2. Extension complexity and number pattern for the Doorbell task

Table 10 presents the complexity level analysis results by prompt for the Doorbell task. A chi-square test showed that the two prompts were not statistically significantly associated with the extension of given figures for the Doorbell task (χ^2 (1, 334) = 3.43, $p = .06$, $\widehat{\eta}^2 = 0.01$). However, the Pose-EMD prompt produced fewer extension problems than the Pose-3 prompt (54.86% vs. 65.41%, χ^2 (1, 200) = 3.43, $p = .03$). In addition, the number-pattern relationship of solvable mathematical problems was not significantly associated with the two prompts for the Doorbell task.

Table 11 presents the complexity level analysis results by position of the posed problem (First/Easy, Second/Moderately Difficult, or Third/Difficult) for the Doorbell task. The chi-square tests showed that the position of problems for the Pose-3 and Pose-EMD prompts were not significantly associated with the extension of given figures for the Doorbell task (χ^2 (2, 159) = 0.23, $p = .89$, $\widehat{\eta}^2 = 0.002$; χ^2 (2, 175) = 3.98, $p = .13$, $\widehat{\eta}^2 = 0.04$, respectively). Looking closer at the position data, students posed more problems that extended beyond the given numbers in the Pose-Difficult position than in the Pose-Easy position (64.29% vs. 45.76%, χ^2 (1, 118) = 3.26, $p = .03$).

However, the number-pattern relationship of solvable mathematical problems was not significantly associated with the position of the posed problem in the sequence for the two prompts for the Doorbell task. Most of the solvable mathematical problems for the two Doorbell task prompts contained a number-pattern relationship. This may have been influenced by the Doorbell task also containing a number pattern.

To recap: The Pose-EMD prompt did not help students pose more solvable mathematical problems than did the Pose-3 prompt for the Road Trip and Doorbell tasks. However, the Pose-EMD prompt helped students pose more linguistically and semantically complex mathematical problems for the Road Trip task. Additionally, both prompts helped increase the complexity levels of students' later responses, but the Pose-EMD prompt showed a sharper increase in the complexity levels than did the Pose-3 prompt for the Road Trip task. In the Doorbell task, the Pose-EMD prompt did not help students pose more solvable mathematical problems that extended the given figures than did the Pose-3 prompt, but students posed more mathematical problems with extensions of the given figures for their difficult problem than they did for their easy problem.

4.3. Influence of prompts on problem-posing responses: Pose-2 versus Pose-SD

Table 12 presents the distribution of students' response types by the two prompts for the Animal task. Recall that the Animal task included one of two possible prompts. The first prompt was, "Pose two more problems that can be solved using the information in the table" (Pose-2). The second was, "Pose a similar problem that can be solved using the information in the table" followed by "Now pose a very different problem that can be solved using the information in the table" (Pose-SD). As needed below, we will refer to the two parts of the Pose-SD prompt as Pose-Similar and Pose-Different, respectively. A chi-square test showed that the different types of responses were not significantly associated with the two prompts (Pose-2 vs. Pose-SD) for the Animal task (χ^2 (6, 528) = 11.83, $p = .06$, $\widehat{\eta}^2 = 0.04$). Although there is no overall difference, there are significant differences regarding particular response types between the two prompts. For instance, more mathematical statements were found in the responses to the Pose-SD prompt (8.03%) than in responses to the Pose-2 prompt (1.97%, χ^2 (1, 27) = 8.76, $p = .001$), whereas fewer mathematical word problems were found under the Pose-SD prompt (44.16%) than under the Pose-2 prompt (51.97%, χ^2 (1, 253) = 2.91, $p = .04$).

Table 13 presents the distribution of students' response types by position in the sequence of posed problems for the Animal task. A chi-square test showed that the different types of problem posing were not significantly associated with position for either the Pose-2 prompt (χ^2 (6, 254) = 1.42, $p = .96$, $\widehat{\eta}^2 = 0.01$) or the Pose-SD prompt (χ^2 (6, 274) = 9.72, $p = .13$, $\widehat{\eta}^2 = 0.06$) for the Animal task. There was a significant decrease in the proportion of mathematical word problems in the later responses than in the earlier ones for the Pose-SD prompt (from 49.64% to 38.69%, χ^2 (1, 111) = 2.90, $p = .04$). However, even though the Pose-SD prompt did not help students pose more solvable mathematical problems, it may have helped students pose problems different from the sample problem provided by the original task. Thus, we next examined whether the Pose-SD prompt helped students pose problems that were different from the given problem example according to the four dimensions of linguistic structure, number of operations, numerical information used, and extension of the given math topics.

Table 11
Percentage of extensions of given figure and pattern relation by position for the Doorbell task.

Dimensions	Pose-3			Pose-EMD		
	First	Second	Third	Easy	Moderate	Difficult
Extension of the given figures						
Total	$N = 55$	$N = 57$	$N = 47$	$N = 59$	$N = 60$	$N = 56$
No extension	36.36	35.09	31.91	54.24	45.00	35.71
Extension	63.64	64.91	68.09	45.76	55.00	64.29
Pattern relation						
Total	$N = 68$	$N = 69$	$N = 61$	$N = 74$	$N = 73$	$N = 70$
Pattern-related	98.53	95.65	91.80	98.65	97.26	95.71
Not-pattern related	1.47	4.35	8.20	1.35	2.74	4.29

Table 12
Percentage of problem-posing types by prompt for the Animal task.

Problem-posing types	Animal task ($N = 528$)	
	Pose-2 ($N = 254$)	Pose-SD ($N = 274$)
No response	22.83	22.99
Incomprehensible	0.79	1.46
Statement	1.97	8.03
Non-math questions	1.18	1.09
Math expression	16.54	16.79
Word problem	51.97	44.16
Non-solvable math questions	4.72	5.47

Note. The solvable mathematical problems included mathematical expressions and word problems.

Table 13
Percentage of problem-posing types by for the Animal task.

Problem-posing types	Animal task ($N = 528$)			
	Pose-2 ($N = 254$)		Pose-SD ($N = 274$)	
	First ($N = 127$)	Second ($N = 127$)	Similar ($N = 137$)	Different ($N = 137$)
No response	20.47	25.20	20.44	25.55
Incomprehensible	0.79	0.79	1.46	1.46
Statement	1.57	2.36	10.22	5.84
Non-math questions	1.57	0.79	0.00	2.19
Math expression	16.54	16.54	14.60	18.98
Word problem	54.33	49.61	49.64	38.69
Non-solvable math questions	4.72	4.72	3.65	7.30

Note. The solvable mathematical problems included mathematical expressions and mathematical word problems.

4.4. *Influence of prompts on similarity of solvable mathematical problems: Pose-2 versus Pose-SD*

Table 14 presents the distribution of students' response types along the four dimensions of similarity by prompt and position of posed problems for the Animal task. Regarding the differences between the Pose-2 and Pose-SD prompts, chi-square tests showed no significance for the linguistic structure, number of operations, numerical information, or extension of the given topics. However, the linguistic structure for the Animal task was not statistically significantly associated with position of posed problems for the Pose-2 prompt (χ^2 (2, 131) $= 4.78$, $p < .09$, $\widehat{\eta}^2 = 0.06$) but was statistically significantly associated with the position of posed problems for the Pose-SD prompt (χ^2 (2, 119) $= 29.06$, $p < .001$, $\widehat{\eta}^2 = 0.28$). The given problem example involved relational propositions. Thus, the solvable mathematical problems students posed in response to the Pose-Similar part of the prompt mostly had relational propositions (83.58%). In contrast, the solvable mathematical problems students posed in response to the Pose-Different part of the prompt were much more likely to use conditional propositions compared to the solvable mathematical problems they posed for the Pose-Similar part of the prompt (25.00% vs. 2.99%, χ^2 (1, 15) $= 10.96$, $p = .0004$) or for the first-posed problem for the Pose-2 prompt (25.00% vs. 11.59%, χ^2 (1, 21) $= 2.83$, $p = .04$). Considering conditional propositions as an indicator of more complex mathematical problems, the Pose-Different prompt helped students pose mathematical problems with more complex linguistic structures for the Animal task.

Second, the number of operations for the Animal task was not statistically significantly associated with the position of posed problems for the Pose-2 prompt (χ^2 (1, 114) $= 1.69$, $p = .42$, $\widehat{\eta}^2 = 0.02$) but was significantly associated with position of posed problems for the Pose-SD prompt (χ^2 (2, 106) $= 20.02$, $p < .001$, $\widehat{\eta}^2 = 0.24$). The Pose-Different part of the prompt helped students pose more solvable mathematical problems that required multiple operations than did the Pose-Similar part of the prompt (53.33% vs. 8.20%, χ^2 (1, 29) $= 24.32$, $p < .0001$) or the Pose-2 prompt, whether looking at the second-posed problem (53.33% vs. 19.61%, χ^2 (1, 34) $= 10.45$, $p = .0006$) or the first-posed problem (53.33% vs. 19.05%, χ^2 (1, 36) $= 12.38$, $p = .0002$); thus, the Pose-SD prompt helped students pose solvable mathematical problems that required multiple operations to solve.

Third, the numerical information for the Animal task was not statistically significantly associated with the position of the posed problem for the Pose-2 prompt (χ^2 (1, 132) $= 0.005$, $p = .99$, $\widehat{\eta}^2 = 0.00$) but was significantly associated with the position of the posed problem for the Pose-SD prompt (χ^2 (1, 121) $= 24.32$, $p < .001$, $\widehat{\eta}^2 = 0.25$). Regarding numerical information, the Pose-Different part of the prompt was especially helpful in encouraging students to pose more solvable mathematical problems in which the numerical information was different from the given information compared to the Pose-Similar part of the prompt (35.85% vs. 10.29%, χ^2 (1, 26) $= 10.06$, $p = .0007$) and the first-posed problem for the Pose-2 prompt (35.85% vs. 15.94%, χ^2 (1, 30) $= 5.37$, $p = .01$). This suggests that the Pose-Different prompt was beneficial for students to generate problems with numbers that were not given in the original task.

Last, the extension of given topics was not statistically significantly associated with the position of the Pose-2 prompts (χ^2 (1, 132)

Table 14
Percentage of response types by prompt and position based on the four dimensions for the Animal task.

Dimensions	First	Second	Pose-2	Similar	Different	Pose-SD
Linguistic structure						
Total	$N = 69$	$N = 62$	$N = 131$	$N = 67$	$N = 52$	$N = 119$
Assignment	18.84	33.87	25.95	13.43	38.46	24.37
Relational	69.57	51.61	61.07	83.58	36.54	63.03
Conditional	11.59	14.52	12.98	2.99	25.00	12.61
Number of operations						
Total	$N = 63$	$N = 51$	$N = 114$	$N = 61$	$N = 45$	$N = 106$
One operation	80.95	80.39	80.70	91.80	46.67	72.64
Multiple operations	19.05	19.61	19.30	8.20	53.33	27.36
Numerical information used						
Total	$N = 69$	$N = 63$	$N = 132$	$N = 68$	$N = 53$	$N = 121$
No number	76.81	66.67	71.97	82.35	43.40	65.29
Not related	15.94	22.22	18.94	10.29	35.85	21.49
Related	7.25	11.11	9.09	7.35	20.75	13.22
Extension of the given math topics						
Total	$N = 69$	$N = 63$	$N = 132$	$N = 68$	$N = 53$	$N = 121$
Within	81.16	76.19	78.79	91.18	62.26	78.51
Beyond	18.84	23.81	21.21	8.82	37.74	21.49

$= 0.23, p = .62, \widehat{\eta}^2 = 0.003$) but was significantly associated with the position of the Pose-SD prompts (χ^2 (1, 121) $= 12.65, p < .001, \widehat{\eta}^2 = 0.15$). The Pose-Different prompt helped students pose more solvable mathematical problems that contained topics different from the given ones compared to the Pose-Similar prompt (37.74% vs. 8.82%, χ^2 (1, 26) $= 13.09$, $p = .0001$) and the Pose-First prompt (37.74% vs. 18.84%, χ^2 (1, 33) $= 4.50$, $p = .01$). This suggests that the Pose-Different prompt did help students think about mathematical problems with diverse math topics different than the given ones.

In sum, the Pose-SD prompt did not necessarily help students pose more solvable mathematical problems, and it also did not increase the proportion of solvable mathematical problems in students' later responses. However, when examining the influence of the Pose-SD prompt on the similarity of posed problems compared to the given sample problem, the Pose-SD prompt appears to have helped students generate mathematical problems that were different from the given ones along multiple dimensions, including different linguistic structures, multiple operations, different numerical information, and diverse math topics.

5. Discussion

This study used three pairs of problem-posing tasks to examine the impact of problem-posing prompts on students' problem posing. Two kinds of prompts were involved. The first asked students to pose two or three different mathematical problems without specifying other requirements for the problems. The second kind of prompt asked students to also pose two or three mathematical problems but with additional specifications. For the first two tasks (Road Trip and Doorbell), students were asked to pose easy, somewhat difficult, and difficult problems (Pose-EMD). For the third task (Animal), students were asked to pose one problem similar to and another problem different from the sample problem (Pose-SD). The findings of the study show that there was no significant difference between the two prompts across the three tasks with respect to the number of responses and the distribution of the different types of responses students provided. This finding is not surprising because we only asked students to pose three mathematical problems, and it would seem natural, based on prior research (Cai et al., 2015), that students would be able to pose the three required problems despite these students having little prior experience with problem posing.

Looking more deeply, students tended to pose more difficult problems in their later responses than in their earlier responses related to the two kinds of prompts; however, the difference was greater in the case of Pose-EMD than in the case of Pose-3. Most importantly, in reviewing the direct comparisons of the third responses with respect to the two prompts for the Road Trip and Doorbell tasks, when the prompt asked students to pose difficult problems (Pose-Difficult), the posed problems tended to be much more linguistically and semantically complex and involved more relationships or steps to solve them. In addition, for the Animal task, students were asked to first pose a problem similar to and then pose a problem different from the given sample problem using the Pose-SD prompt. Compared to the prompts without this specification, students tended to pose much more similar problems in the first response and more different problems in the second response. Thus, in summary, even though the prompts had subtle differences in terms of specifying the kinds of problems students were asked to pose, such as problems with different difficulty levels or problems that were similar, these prompts did have an impact on students' problem posing. This finding has confirmed what we initially hypothesized, which is that additional specification in the prompts engages students more in the problem-posing process. That is, when students are given a problem-posing prompt with more specific requirements, they tend to think more about what to pose, such as the solvability and complexity of the problems. This finding has also confirmed the finding from Silber and Cai (2017) as well as extended it with more pairs of tasks and more in-depth analyses.

This finding is significant in at least three ways with respect to teaching mathematics through problem posing (Cai, 2022). The first is that it suggests that even though the two kinds of prompts used in this study were appropriate to elicit problem-posing activity, the more specific requirements invited students to think more with respect to the problem situation and therefore engaged them in more in-depth mathematical thinking. Second, given the fact that textbooks do not currently include many problem-posing tasks, the findings of this study suggest that teachers or curriculum developers designing problem-posing tasks may want to include prompts with additional requirements to further engage students. The third implication echoes a finding from earlier studies (Cai et al., 2013; Yao et al., 2021): Problem posing can serve as an assessment tool to understand students' thinking. Even though earlier studies have already shown that requiring different difficulty levels of students' posed problems can better reveal students' thinking (e.g., Cai & Hwang, 2002), this study provides empirical evidence that, in requiring different difficulty levels, problem-posing tasks can provide better insight into students' mathematical thinking.

One aspect of research on MPP is to better understand problem-posing processes. As indicated in other studies, we do not yet have a robust understanding of problem-posing processes. However, the findings from this study (focusing on prompts) and earlier studies, (e.g., Leung & Silver, 1997; Zhang et al., 2022, focusing on situations) advance our understanding of problem posing from the perspective of task variables. That is, to pose mathematical problems, students need to clearly understand the problem-posing task. The task, in turn, includes all the information in the problem situation and the prompt, which includes the requirements for students' posed problems. In problem-solving processes, understanding the problem to be solved is a key first step; similarly, understanding the requirements and parameters of the problem-posing task is a key first step of the problem-posing process (Baumanns & Rott, 2022).

One possible limitation of this study is related to the fact that we only collected paper-and-pencil student data. Thus, although we have the products of the students' problem-posing activity, we do not have access to their cognitive processes and interpretations of the tasks. Ultimately, even though we assume that being able to pose a difficult or complex problem is a superior outcome, we do not have a true understanding of how the students themselves interpreted the two kinds of prompts. What is their interpretation of an easy problem, a somewhat difficult problem, and a difficult problem? What makes a problem difficult to them, and how does that map onto the complexity measures that we used in this study? A parallel set of questions pertains to the Pose-SD prompt. How do the students make sense of what counts as a similar problem or a different problem compared to a given problem? How do they think about varying the sample problem, and how does that influence the mathematics they think about and include in their own posed problems? Future problem-posing studies are needed that will more directly examine problem posers' thought processes. A natural avenue to pursue would be task-based interview studies in which problem posers narrate their thinking as they read and interpret problem-posing prompts. In doing so, we could not only understand how students interpret the two different prompts (e.g., Pose-3 vs Pose-EMD or Pose-2 vs Pose-SD) but also understand problem-posing processes under the two kinds of prompts.

Declaration of Competing Interest

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this manuscript.

Data availability

Data will be made available on request.

Acknowledgments

The preparation of this manuscript was supported by a grant from the National Science Foundation (DRL-2101552). Any opinions expressed herein are those of the authors and do not necessarily represent the views of the National Science Foundation.

References

Arnholt, A. T. (2007). Statistics: The art and science of learning from data. *The American Statistician, 61*(2), 181–182.

Baumanns, L., & Rott, B. (2022). The process of problem posing: Development of a descriptive phase model of problem posing. *Educational Studies in Mathematics, 110*, 251–269. https://doi.org/10.1007/s10649-021-10136-y

Ben-Shachar, M.S., Makowski, D., Lüdecke, D., Kelley, K., & Stanley, D. (2021). *Package 'effectsize'*.

Brown, S. I., & Walter, M. I. (1983). *The art of problem posing.* Franklin Institute Press.

Cai, J. (1998). An investigation of US and Chinese students' mathematical problem posing and problem solving. *Mathematics Education Research Journal, 10*(1), 37–50. https://doi.org/10.1007/BF03217121

Cai, J. (2022). What research says about teaching mathematics through problem posing. *Éducation & Didactique, 16*, 31–50.

Cai, J., & Cifarelli, V. (2005). Exploring mathematical exploration: How two college students formulated and solved their own mathematical problems. *Focus on Learning Problems in Mathematics, 27*(3), 43–72.

Cai, J., & Hwang, S. (2002). Generalized and generative thinking in US and Chinese students' mathematical problem solving and problem posing. *The Journal of Mathematical Behavior, 21*(4), 401–421. https://doi.org/10.1016/S0732-3123(02)00142-6

Cai, J., & Hwang, S. (2003). A perspective for examining the link between problem posing and problem solving. *International Group for the Psychology of Mathematics Education, 3*, 103–110.

Cai, J., & Hwang, S. (2020). Learning to teach through mathematical problem posing: Theoretical considerations, methodology, and directions for future research. *International Journal of Educational Research, 102*, 101391. https://doi.org/10.1016/j.ijer.2019.01.001

Cai, J., & Hwang, S. (2023). Making mathematics challenging through problem posing in the classroom. In R. Leikin (Ed.), *Mathematical challenges for all* (pp. 115–145). Springer.

Cai, J., Hwang, S., Jiang, C., & Silber, S. (2015). Problem-posing research in mathematics education: Some answered and unanswered questions. In F. M. Singer, N. F. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 3–34). Springer.

Cai, J., & Leikin, R. (2020). Affect in mathematical problem posing: Conceptualization, advances, and future directions for research. *Educational Studies in Mathematics, 105*(3), 287–301. https://doi.org/10.1007/s10649-020-10008-x

Cai, J., Moyer, J. C., Wang, N., Hwang, S., Nie, B., & Garber, T. (2013). Mathematical problem posing as a measure of curricular effect on students' learning. *Educational Studies in Mathematics, 83*(1), 57–69. https://doi.org/10.1007/s10649-012-9429-3

Christou, C., Mousoulides, N., Pittalis, M., Pitta-Pantazi, D., & Sriraman, B. (2005). An empirical taxonomy of problem posing processes. *ZDM–Mathematics Education, 37*(3), 149–158. https://doi.org/10.1007/s11858-005-0004-6

Ellerton, N. F. (1986). Children's made-up mathematics problems—A new perspective on talented mathematicians. *Educational Studies in Mathematics, 17*, 261–271. https://doi.org/10.1007/BF00305073

English, L. D. (1998). Children's problem posing within formal and informal contexts. *Journal for Research in Mathematics Education, 29*(1), 83–106. https://doi.org/10.2307/749719

Felmer, P. L., Pehkonen, E., & Kilpatrick, J. (2016). *Posing and solving mathematical problems*. Springer.

Guo, Y., Yan, J., & Men, T. (2021). Chinese junior high school students' mathematical problem-posing performance. *ZDM–Mathematics Education, 53*, 905–917. https://doi.org/10.1007/s11858-021-01240-7

Kilpatrick, J. (1987). Problem formulating: Where do good problems come from? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 123–147). Lawrence Erlbaum Associates.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Leung, S. S., & Silver, E. A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers. *Mathematics Education Research Journal, 9*(1), 5–24. https://doi.org/10.1007/BF03217299

Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: A look at the state of the art. *ZDM–Mathematics Education, 53*(4), 723–735. https://doi.org/10.1007/s11858-021-01291-w

Marshall, S. P. (1995). *Schemas in problem solving*. Cambridge University Press.

Mayer, R. E., Lewis, A. B., & Hegarty, M. (1992). Mathematical misunderstandings: Qualitative reasoning about quantitative problems. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 137–154). Elsevier.

Silber, S., & Cai, J. (2017). Pre-service teachers' free and structured mathematical problem posing. *International Journal of Mathematical Education in Science and Technology, 48*(2), 163–184. https://doi.org/10.1080/0020739X.2016.1232843

Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics, 14*(1), 19–28.

Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for Research in Mathematics Education, 27*(5), 521–539. https://doi.org/10.2307/749846

Singer, F. M., Ellerton, N., & Cai, J. (2013). Problem-posing research in mathematics education: New questions and directions. *Educational Studies in Mathematics, 83*(1), 1–7. https://doi.org/10.1007/s10649-013-9478-2

Singer, F. M., Ellerton, N. F., & Cai, J. (Eds.). (2015). *Mathematical problem posing: From research to effective practice*. Springer.

Stickles, P. R. (2011). An analysis of secondary and middle school teachers' mathematical problem posing. *Investigations in Mathematics Learning, 3*(2), 1–34. https://doi.org/10.1080/24727466.2011.11790301

Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. *Technology in Mathematics Education, 4*(7), 518–525.

Xu, B., & Li, M. (2021). Chinese eighth graders' competencies in mathematical problem-posing. In B. Xu, Y. Zhu, & X. Lu (Eds.), *Beyond Shanghai and PISA* (pp. 67–86). Springer.

Yao, Y., Hwang, S., & Cai, J. (2021). Preservice teachers' mathematical understanding exhibited in problem posing and problem solving. *ZDM–Mathematics Education, 53*(4), 937–949. https://doi.org/10.1007/s11858-021-01277-8

Zhang, H., & Cai, J. (2021). Teaching mathematics through problem posing: Insights from an analysis of teaching cases. *ZDM–Mathematics Education, 53*, 961–973. https://doi.org/10.1007/s11858-021-01260-3

Zhang, L., Cai, J., Song, N., Zhang, H., Chen, T., Zhang, Z., & Guo, F. (2022). Mathematical problem posing of elementary school students: The impact of task format and its relationship to problem solving. *ZDM–Mathematics Education, 54*, 497–512. https://doi.org/10.1007/s11858-021-01324-4