Instr Sci (2018) 46:61–76
https://doi.org/10.1007/s11251-017-9435-z

Examining the preparatory effects of problem generation and solution generation on learning from instruction

Manu Kapur[1]

Received: 22 March 2017 / Accepted: 13 November 2017 / Published online: 24 November 2017

Abstract The goal of this paper is to isolate the preparatory effects of problem-generation from solution generation in problem-posing contexts, and their underlying mechanisms on learning from instruction. Using a randomized-controlled design, students were assigned to one of two conditions: (a) problem-posing with solution generation, where they generated problems and solutions to a novel situation, or (b) problem-posing without solution generation, where they generated only problems. All students then received instruction on a novel math concept. Findings revealed that problem-posing with solution generation prior to instruction resulted in significantly better conceptual knowledge, without any significant difference in procedural knowledge and transfer. Although solution generation prior to instruction plays a critical role in the development of conceptual understanding, which is necessary for transfer, generating problems plays an equally critical role in transfer. Implications for learning and instruction are discussed.

Keywords Problem posing · Preparatory activities · Math learning · Transfer

Introduction

Eminent mathematicians and scientists have long underscored the importance of problem-posing in advancing and deepening knowledge and understanding. For example, Einstein and Infeld (1938) argued: "formulation of a problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old problems from a new angle, require creative imagination and marks real advance in science." (p. 95).

✉ Manu Kapur
manukapur@ethz.ch

[1] Institute of Learning Sciences and Higher Education, ETH Zurich, Clausiusstrasse 59, RZ J3 Zurich, Switzerland

An increasing number of mathematics educators and researchers have also argued for the importance in getting students to pose problems, so that their learning experiences afford them opportunities not just for problem-solving, but also problem-posing (e.g., Brown and Walter 2005; Ellerton 1986; Moses et al. 1990; Silver and Cai 1996; Silver 1994, 2013). Although both problem-solving and problem-posing are important part of mathematical education, interest and active research in the latter only started some 30 years ago. As with problem-solving, there remain challenges in defining the what constitutes problem posing (Singer et al. 2013). Nonetheless, it is possible to glean its key features from Silver (1994)'s widely-accepted definition, which defines problem posing as either the reformulation of a given problem, or the generation of new problems from a given situation.

The former type of problem posing happens during the problem solving process, in which the solver reformulates the problem, and transforms a given statement of a problem into a new version that becomes the focus of problem solving. For example, given the situation in Fig. 1, suppose the problem, which is currently unspecified, was specified as finding the better of the two football strikers. Although the situation is already specified in Fig. 1, the problem of finding the better of the two strikers specifies the problem as well. The goal for the problem solver is to solve the specified problem, and in the process of doing so, may reformulate the problem as one of finding the total goals scored by each striker, or perhaps how many times one striker outperformed the other, and so on.

In the latter type of problem posing, the goal is not to solve a given problem per se. It cannot be because the problem is not specified to begin with. It is left open, and the goal of problem posing is to create a new problems from the given situation or context. This can occur prior to problem solving, such as the coming up of a problem from a given contrived or naturalistic situation, or after a problem is solved, where one might examine the conditions of the problem and generate alternative related problems. In this paper, problem posing refers to this latter type, that is, the generation of new problems from a given situation, where students may generate either only the problems, or the associated solutions as well (see Fig. 1).

Among the benefits, problem-posing is seen as important for nurturing students' mathematical thinking (Silver 1994). Not only does it afford students opportunities to be actively engaged in both creating solution strategies and the problems that demand them (Moses et al. 1990), problem-posing is invariably a reflective process that may stimulate students' overall abilities to mathematize and develop understanding within new situations (Cifarelli and Sheets 2009). In addition, research has suggested that students-generated problems provides one with an insight to students' mathematical understanding (e.g., English 1998), problem-solving abilities (e.g., Ellerton 1986; Silver and Cai 1996), and creativity and originality (Silver 1997; Perrin 2007).

Given problem-posing's generative nature and its potential to provide a window into students' mathematical knowledge and conceptions, it seems well placed as a preparatory activity that prepares students to learn from subsequent instruction (Schwartz and Bransford 1998; Schwartz and Martin 2004; Schwartz et al. 2011).

From a discovery learning perspective, opportunities to generate problems and solutions prior to learning a new concept, especially in the context of reasoning with data, may potentially help students discover features of the concept that they may otherwise not be able to. In turn, although these features in and of themselves may not be sufficient to learn the concept, they constitute the knowledge that is necessary to "learn with" during subsequent instruction (Bransford and Schwartz 1999).

Fig. 1 The problems given to problem-posing students during the problem-posing phase

Problem Posing with Solution Generation
Football Strikers

Mike and Dave are the top two strikers in a Football league. The table shows the number of goals scored by Mike and Dave in 11 games in the league.

Generate as many different mathematics questions or problems that can be answered from the information provided in the table.

Where possible, answer or solve the problems/questions you have generated.

Game	Mike	Dave
1	14	13
2	11	11
3	15	14
4	12	16
5	16	14
6	12	12
7	16	14
8	13	15
9	17	14
10	14	17
11	14	14

Problem Posing without Solution Generation
Football Strikers

Mike and Dave are the top two strikers in a Football league. The table shows the number of goals scored by Mike and Dave in 11 games in the league.

Generate as many different mathematics questions or problems that can be answered from the information provided in the table.

You are NOT to answer or solve the problems/questions you generate.

Game	Mike	Dave
1	14	13
2	11	11
3	15	14
4	12	16
5	16	14
6	12	12
7	16	14
8	13	15
9	17	14
10	14	17
11	14	14

Yet, although there is much work documenting the preparatory effects of problem-solving prior to instruction, problem-posing's efficacy as a preparatory activity to help prepare students to learn new mathematical concepts remains severely under-researched by comparison. As the following section will reveal, although Kapur's (2015) study compared the preparatory effects of problem-posing with that of problem-solving, students in the problem posing condition generated both problems and solutions. Therefore, the study was not able to isolate the preparatory effects of problem generation from solution generation in problem-posing contexts.

Therein lies the aim of this paper: to build on Kapur's (2015) study to isolate the preparatory effects of problem-generation from solution generation in problem-posing contexts, and their underlying mechanisms on learning from instruction.

I start by briefly reviewing research on the preparatory effects of problem-solving, followed by a review of problem-posing in math learning. I then describe the attendant preparatory mechanisms of problem-posing to derive hypotheses for effects on learning. These hypotheses are then tested in a randomized-controlled experiment. I end by discussing the findings and drawing implications for math learning and instruction.

Preparatory effects of problem-solving

The goal of a preparatory activity is to prepare students to learn from instruction (Schwartz and Bransford 1998; Schwartz and Martin 2004; Schwartz et al. 2011). Preparatory activities are designed to engage students in generating and inventing solutions to problems that target concepts they have yet to learn formally. This is then followed by direct instruction on those very concepts. To the extent that students are able to generate a range of solutions to the problem, even if incorrect, they are better prepared to learn from subsequent instruction on the targeted concepts, compared to the case where students start with instruction (Kapur 2016).

There is now a growing body of evidence that preparatory activities such as generating solutions to novel problems *prior* to instruction can help students learn better from the instruction. Evidence comes not only from quasi-experimental studies conducted in the real ecologies of classrooms (e.g., Kapur 2012, 2013; Schwartz and Bransford 1998; Schwartz and Martin 2004), but also from controlled experimental studies (e.g., DeCaro and Rittle-Johnson 2012; Kapur 2014; Loibl and Rummel 2013; Roll et al. 2011; Schwartz et al. 2011).

For example, in a study with eight-grade students, Schwartz et al. (2011) compared students who invented solutions with contrasting cases before receiving instruction on the concept of density with those who were instructed first and then practiced with the same cases. They found that the guided invention activities prepared students to learn the deep structure of density better than those who received instruction first. Likewise, DeCaro and Rittle-Johnson (2012) had second- to fourth-grade students solve unfamiliar math problems on number sentences before or after receiving instruction on number sentences. Once again, students who solved problems first developed better conceptual understanding than those who first received instruction. More recently, in a randomized-controlled experiment with ninth-graders learning the concept of standard deviation (SD), Kapur (2014) had students individually generate solutions to a novel problem before or after receiving instruction. He found that students who engaged in problem-solving prior to instruction demonstrated significantly better performance on conceptual understanding and transfer than those who engaged in problem-solving after instruction.

These studies collectively point to the efficacy of preparatory activities that engage students in solving novel problems prior to instruction. However, these studies present students with situations where the problem is given to the students, and they have to generate only the solutions. In other words, these studies speak only to the preparatory effects of problem-solving on learning from instruction. They remain silent on if and how problem-posing might prepare students to learning from instruction.

Given that math education research emphasizes problem-posing as an instructional and learning goal, I now turn to review literature on problem-posing in math learning, especially as a preparatory activity for learning new math concepts. However, as we will see in the following section, the efficacy of engaging in problem-posing to learn new math concepts has not been systematically tested because empirical evidence to date remains largely descriptive in nature.

Problem-posing in math learning

A review of research on problem-posing in math learning revealed that past research is largely theoretical or descriptive in nature (e.g., English 1998; Silver and Cai 1996). Purported benefits of problem-posing include the development of greater learner agency and reflection (Kilpatrick 1987), responsibility and insight that helps understanding and reduces anxiety (Brown and Walter 2005), and ownership and engagement that can potentially help math learning (Perrin, 2007; Silver, 1994). Because evidence for these benefits comes mainly from descriptive studies, the absence of any comparison or controls limits what one can infer from these studies about the preparatory effects of problem-posing.

The closest experimental comparison in math learning, though not from a preparatory lens, comes from the work of Sweller and colleagues (Mawer and Sweller 1982; Sweller et al. 1982; Sweller et al. 1983). They showed that a reduction in or elimination of goal specificity of a given problem situation, a move that essentially requires students to pose and answer as many questions for solving a particular problem, helped schema acquisition for problem-solving.

For example, Sweller et al. (1983) first taught students targeted geometry concepts before assigning them to solve either no-goal or goal-specific problems. The no-goal problem required students to ask and answer as many unknown angles as possible in a geometry diagram. The goal-specific problem required students to solve for a particular unknown in the same diagram. As hypothesized, findings suggested that the no-goal students performed better than goal-specific students on subsequent problem-solving on similar geometry problems.

Note that students in Sweller's studies were first taught the targeted concepts before they solved problems, with or without goals. Hence, whereas these studies may speak to the cognitive load mechanisms of using problem-posing as a problem-solving strategy *after* learning a concept, they do not directly speak to the preparatory mechanisms of problem-posing *before* learning and for learning a new concept.

From the lens of preparatory activities (Schwartz and Martin 2004; Schwartz et al. 2011), one would want to examine how problem-posing can prepare students to learn new concepts in the first place, where participants engage in problem-posing *before* learning the targeted concepts.

To my knowledge, only one study (Kapur 2015) has experimentally examined the preparatory effects of problem-posing. However, Kapur's study compared the preparatory effects of problem-posing in comparison with that of problem-solving. Findings revealed students engaged in problem-posing prior to instruction achieved better transfer outcomes than those in problem-solving prior to instruction, even though the latter developed better conceptual knowledge. In other words, although problem-solving helped in the development of conceptual knowledge, problem-posing played a more critical role in transferring that knowledge to novel problems.

This trade-off between the development of conceptual knowledge and transfer as a function of the preparatory effects of problem-solving and problem-posing is most intriguing. Yet, given that Kapur's study compared the preparatory effects of problem-posing with that of problem-solving, it was not able to isolate the preparatory effects of problem generation from solution generation.

Therein lies the need for a study that experimentally isolates the preparatory effects of problem generation with versus without solution generation on the development of conceptual understanding and transfer. The following section unpacks the underlying preparatory mechanisms of problem-posing and their hypothesized effects on the development of conceptual understanding and transfer.

Preparatory mechanisms of problem posing

There are several preparatory mechanisms at work when students engage in problem-posing with or without generating solutions.

First, one could expect problem-posing with solution generation to afford greater prior knowledge activation and differentiation than problem-posing without solution generation. By knowledge activation, I refer to the problems and/or solutions students generate during problem-posing. Of course, this activated knowledge is a function of math ability, but it is also a function of the design of the problem-posing context and how students persist in problem-posing. Therefore, a distinction needs to be made between pre-existing differences between math ability and the knowledge activation that occurs during problem-posing prior to instruction.

In problem-posing with solution generation, both the problem and solution spaces are not only open, but also the problem space is potentially generative with the solution space. Whereas for the latter, only the problem space is open. Following previous work (e.g., Kapur 2012, 2014; Wiedmann et al. 2012), greater activation simply means the numbers of problems and solutions students are able to generate.

However, there is likely to be a trade-off between greater knowledge activation and how relevant the activated knowledge is, that is, conceptually related to and beneficial in learning the targeted concept. Therefore, the extent to which students benefit from greater prior knowledge activation may be contingent upon whether such activation is relevant to the learning of the targeted concept.

Further support for the trade-off hypothesis comes from research on the role of goal specificity and learning. This research suggests that the benefits of problem-posing (akin to having no goal or low goal specificity) are derived mainly if the lack of a goal actually affords students the opportunities to attend to the deep structure of the problem and solution spaces (Burns and Vollmeyer 2002; Miller et al. 1999; Vollmeyer et al. 1996).

Prior knowledge activation also affords opportunities to: (a) notice inconsistencies in and the limits of prior knowledge, and (b) compare and contrast (that helps student notice and encode the critical features of the concept better) between a learner's prior knowledge and the correct knowledge (Alfieri et al. 2013; Rittle-Johnson and Star 2009; Schwartz et al. 2011). However, consistent with the above-mentioned trade-off, these benefits are contingent upon the extent to which prior knowledge activation in problem-posing is relevant to the targeted concept.

Therefore, from a knowledge activation perspective, problem-posing with solution generation may lead to better learning than without solution generation, provided the activated knowledge is relevant to the targeted concept.

However, activation is not the only mechanism at play. One may hypothesize the reverse effect from a cognitive load perspective, where research suggests problem-solving may impose a greater load than problem-posing (Sweller et al. 1983; Wirth et al. 2009). Problem-posing reduces or eliminates the goal by turning a goal-specific problem into a goal-free problem. This reduces the burden on limited working memory resources to reduce differences between problem and goal states. Therefore, problem posing without solution generation may impose a lower cognitive load, and consequently, enable more cognitive processing capacity to be available for schema acquisition. However, goal-free effect of problem posing is contingent upon the degree to which students generate and explore relevant problem states that help them notice the deep structure of the problem space. If not, a reduction in cognitive load may not translate into effective schema acquisition.

Finally, on the development of transfer, problem-posing may afford greater contextual flexibility in encoding and assembling new knowledge by virtue of having students generate multiple ways of contextualizing the data that was given to them. Bransford and Schwartz (1999) argue that such generation may help learners discover relevant and irrelevant features of the domain—knowledge that they can use to learn the targeted concept during subsequent instruction. Generating different problems for the same data may also be a useful mechanism for reducing the likelihood of functional fixedness (Duncker 1945; Frank and Ramscar 2003) and set effects (Lurchins and Lurchins 1959), which in turn may allow for more pathways for such knowledge to be cued, retrieved, and used in a novel context, that is, aid transfer performance. At the same time, one cannot transfer without conceptual knowledge. If solution generation is critical to the development of conceptual knowledge, then it is reasonable to hypothesize that problem-posing with solution generation would result in better transfer than problem-posing without solution generation.

Purpose

The purpose of this study is to isolate the preparatory effects of problem generation with versus without solution generation on the development of conceptual understanding and transfer.

Participants

Participants were 72 ninth-grade mathematics students (14–15 year olds; 37 boys, 35 girls) from a co-ed private school in the national capital region of India. All students were of Indian ethnicity. English is the medium of instruction in the school.

Research design and procedures

A posttest-only randomized controlled design was used. Students were randomly assigned to engage in either problem-posing with solution generation ($n = 36$) or problem-posing without solution generation ($n = 36$) prior to receiving instruction on the concept of SD. For three reasons, there was no pretest:

(a) students had not had any instruction on SD because this topic is not taught until the eleventh grade;

(b) past research with a similar cohort of students suggested that ninth grade students did not know the concept of SD and were not able to solve problems requiring this concept when given a pretest on the same concept (e.g., Kapur 2014);
(c) having students solve problems on SD on the pretest would have meant both problem-posing conditions engaging in problem-solving on SD prior to instruction, thereby confounding experimental attribution of effects.

Figure 1 presents the problem-posing situation—Football Strikers—students were given. Students in the problem-posing with solution generation condition were asked to answer or solve the problems they generated. Students in the problem-posing without solution generation condition were asked not to answer or solve the problems they generated.

On the day of the experiment, students experienced two 1-h phases: a problem-posing phase followed by an instruction phase. The instruction phase was the same for all students, where the same teacher taught both the conditions together in the same lecture hall. Neither the teacher nor the students were made aware of the experimental hypotheses being tested.

For the problem-posing phase, students were seated in a classroom, provided with blank A4 sheets of paper together with the problem in Fig. 1 as appropriate to their assigned condition.

Students worked individually. They were asked to clearly number and demarcate their solutions, or problems and solutions as appropriate to their assigned condition. Because students could only rely on their prior knowledge to generate problems and/or solutions, the number of problems and/or solutions generated was taken as a proxy measure of their prior knowledge activation (e.g., Kapur 2012, 2014; Wiedmann et al. 2012).

To determine the number of problems and solutions generated, student work artefacts were analyzed using an analytical scheme developed by Kapur and Bielaczyc (2012). Accordingly, student work artefacts were segmented into different solutions and problems as appropriate to the condition. Because students were instructed to clearly number and demarcate their problems and solutions by using a separator (e.g., a line, numbering etc.), segmenting was made easier by the presence of clear transitions in the artefacts. For example, when a student moved from one problem (e.g., what is the average of each of the players?) to another (e.g., what is the pattern in the scores over time?), or from one solution (e.g., calculating the average) to another (e.g., drawing graphs). Repeated problems or solutions were not double counted. For example, if a student generated the problems "what is the average of Mike?" and "what is the average of Dave?" these were counted as one problem, not two. Likewise, if the student generated the average of each of the two players, it was counted as one solution, not two different solutions. The problems and solutions generated were segmented independently of each other. This process was repeated for all students. Two research assistants independently segmented the solutions and the problems with inter-rater reliabilities (*Krippendorff's* α) of .93 and .94 respectively. All conflicts were resolved through discussion with the author.

In the instruction phase, students were seated in a classroom, and their teacher—an experienced mathematics teacher at the high-school level—taught the concept and procedures of SD. The teaching of SD was organized around four problems that included cycles of teacher modeling through worked-out examples demonstrating the concept and procedures, student practice, and feedback. The design of the four problems in the form of simultaneously presented, contrasting cases was done in line with the well-established finding that contrasting cases help students attend to critical features of the problem, and therefore aid learning (Rittle-Johnson and Star 2009; Schwartz et al. 2011). A detailed plan can be found in Appendix.

From an educational standpoint, instruction ideally ought to build upon the student production (Loibl and Rummel 2013). However, for the purposes of a clean experimental comparison, the teacher did not make any reference to or build upon any of the problems or solutions generated by the students from the two conditions. Throughout this phase, the teacher directed attention to the critical features of SD, and highlighted common errors and misconceptions.

Immediately after each phase, all students estimated their amount of mental effort using a 9-point rating scale that is commonly-used in the cognitive load literature as a measure of cognitive load (Paas 1992). Thus, each student reported two mental effort scores.

Immediately after the second phase, all students took a 1 h posttest comprising 20 items targeting:

(a) *Procedural knowledge* Five multiple-choice items ($\alpha = .81$) testing the basic procedure for computing and interpreting SD, e.g.,
Calculate the standard deviation (SD) of the following set of marks on a statistics test: 30, 60, 50, 40, 70.

A. 10
B. $10\sqrt{2}$
C. 20
D. $20\sqrt{2}$

(b) *Conceptual knowledge* Ten multiple-choice items ($\alpha = .89$) testing understanding of critical features of SD and deducing its mathematical properties, e.g.,
A data set consisting of *five* numbers has mean, M = 7, and standard deviation, SD = 4. If each of the five numbers is increased by 2, what are the new mean and SD?

A. M = 7, SD = 4
B. M = 9, SD = 4
C. M = 7, SD = 6
D. M = 9, SD = 6

(c) *Transfer* Five multiple-choice items ($\alpha = .84$) testing whether students can adapt knowledge of SD to solve problems on the concept of normalization not taught during instruction), e.g.,
An equal number of students competed in the 100 m sprint and 100 m swim finals. The timings (in s) of the champions of the 100 m sprint and 100 m swim are shown below, as are the average timings and the SDs of the finalists in the two competitions.

	100 m sprint (s)	100 m swim (s)
Champion	11	40
Average of the finalists, M	12	45
SD of the finalists	1	10

Assuming all else being equal, between the two champions, who is the better performer?

A. The sprint champion
B. The swim champion
C. Both are equally good
D. Not enough information to decide

Each correct answer was awarded one mark. For ease of comparison, composite scores for each type of item were scaled linearly to 10. This score upon 10 for the three types of items, namely procedural knowledge, conceptual knowledge, and transfer, formed the three dependent variables.

Results

Table 1 presents the descriptive statistics for math ability, mental effort, process measures, and posttest performance.

Pre-existing math ability differences

An ANOVA with experimental condition as the between-subjects factor revealed no significant difference between the two conditions on math ability, $F(1, 70) = .008$, $p = .930$.

Student production: problem and solution generation

Table 1 shows that problem-posing without solution generation students produced on average about just under eleven problems. By comparison, problem-posing with solution generation students posed on average under six problems and between 3 and 4 solutions.

Mental effort

A mixed ANOVA was carried out with mental effort as the within-subjects dependent variable, and experimental condition as the between-subjects factor. For mental effort, there were no significant differences between the two phases, $F(1, 70) = .212$, $p = .647$,

Table 1 Summary of math ability, mental effort, process measures, and posttest performance

	Problem-posing w/o solutions		Problem-posing w/solutions	
	M	SD	M	SD
Math ability	4.56	1.21	4.58	1.44
Mental effort 1	6.28	1.23	6.75	.87
Mental effort 2	6.38	1.61	6.78	1.20
# Solutions	–	–	3.58	.94
# Problems	10.75	1.30	5.78	1.61
Posttest				
Procedural	8.23	1.37	8.11	1.35
Conceptual	5.78	1.05	6.53	1.13
Transfer	3.61	1.15	3.83	1.30

or of condition, $F(1, 70) = .109$, $p = .742$, or interaction between phase and condition, $F(1, 70) = 2.773$, $p = .100$.

Posttest results

A MANCOVA with scores on procedural knowledge, conceptual knowledge, and transfer as the three dependent variables, experimental condition as the between-subjects factor, and math ability as the covariate revealed significant multivariate main effects of experimental condition, $F(3, 67) = 4.083$, $p = .010$, and math ability, $F(3, 67) = 9.445$, $p < .001$. Covariate-by-condition interaction effect was not significant.

Univariate ANCOVAs suggested that the multivariate effect of math ability was largely due to its effect on procedural knowledge, $F(1, 69) = 26.108$, $p < .001$. The effect of math ability on conceptual knowledge and transfer did not reach significance.

Furthermore, there was no significant difference between the conditions on procedural knowledge, $F(1, 69) = .438$, $p = .510$, $d = .12$, or transfer, $F(1, 69) = .578$, $p = .450$, $d = .18$. However, problem-posing students with solution generation significantly outperformed their counterparts on conceptual knowledge, $F(1, 69) = 8.445$, $p = .005$, $d = .65$.

To better understand the effect on transfer, another ANCOVA was carried out with transfer as the dependent variable, condition as the between subjects factor, and math ability, procedural knowledge, and conceptual knowledge as the three co-variates. There was a marginally significant effect of conceptual knowledge on transfer, $F(1, 67) = 3.230$, $p = .075$. None of the other effects were significant.

Finally, Table 2 presents the correlations between the numbers of problems and solutions generated and posttest scores.

General discussion

The study reported in this paper examined the preparatory effects of problem-posing with versus without solution generation on learning from subsequent instruction on three measures: procedural knowledge, conceptual knowledge, and transfer.

There were no significant differences between the two conditions on procedural knowledge or transfer. However, students in the problem-posing condition with solution generation significantly outperformed their counterpart on conceptual knowledge.

That was no significant difference on procedural knowledge can be explained in part due to the relatively straightforward nature of computing and interpreting SD. Indeed, this pattern is consistent with earlier work on preparatory effects.

Table 2 Correlations between numbers of problems and solutions generated and posttest scores

	Problem-posing w/o solutions	Problem-posing w/solutions	
	#Problems	#Problems	#Solutions
Procedural	$r(33) = .23$, $p = .191$	$r(33) = .26$, $p = .193$	$r(33) = .34$, $p = .045$
Conceptual	$r(33) = .33$, $p = .051$	$r(33) = .39$, $p = .020$	$r(33) = .67$, $p < .001$
Transfer	$r(33) = .42$, $p = .012$	$r(33) = .52$, $p = .001$	$r(33) = .36$, $p = .032$

Also consistent with the hypothesis was the finding on conceptual knowledge. As expected, a lack of opportunity for solution generation in the problem-posing without solution generation condition potentially constrained the activation of prior knowledge, which may have adversely affected the development of conceptual knowledge. Said another way, only generating problems is neither sufficient nor optimal as a preparatory activity for the development of conceptual knowledge.

Evidence for the activation mechanism is further supported by the significant correlation between the number of solutions generated and conceptual understanding. Two patterns are noteworthy in Table 2: (a) the number of problems generated by students in the problem-posing without solution generation condition was only marginally significant, $r = .33$, $p = .051$. However, the same correlation in the problem-posing with solution generation reached significance, $r = .39$, $p = .020$, and (b) these two correlations ($r = .33$ and .39) were still not as high as that between the number of solutions generated and conceptual knowledge ($r = .67$). This preliminarily suggests an interplay between problem generation and solution generation on the development of conceptual knowledge. Problem generation alone is not sufficient, but when combined with solution generation, it does seem to play a role in the development of conceptual knowledge.

That there was no significant difference between the two conditions in either of the phases on cognitive load does not support the cognitive load mechanism. Even though solution generation was hypothesized to put greater cognitive load on the students, adversely affecting learning, data do not support this mechanism. Therefore, any explanatory basis for the findings on conceptual knowledge cannot rely on cognitive load differences. One must acknowledge here the obvious limitation of a single-item self-report measure of mental effort that was used in this study. One must also note that the applicability of the goal-free effect of problem posing is contingent upon the degree to which students generate and explore relevant problem states that help them notice the deep structure of the problem space. If not, a reduction in cognitive load may not translate into effective schema acquisition.

Finally, contrary to the transfer hypothesis, there was no significant difference between the two conditions. Although the transfer performance of problem-posing with solution generation students was descriptively better by a small effect size, $d = .18$, it was not significant. Controlling for math ability, condition, and procedural knowledge, the effect of conceptual knowledge was marginally significant on transfer. In other words, conceptual knowledge is necessary but not sufficient for transfer.

Once again, Table 2 shows that even though the number of solutions generated correlated significantly with transfer, the numbers of problems generated correlated with transfer more strongly in both the conditions. Interestingly, when students generated both problems and solutions, the numbers of problems generated correlated more strongly with transfer than the case when students generated only the problems. These findings suggest that whereas solution generation and its attendant preparatory mechanisms play an important role for transfer, problem generation also plays a critical role. Much like the effect of problem generation on conceptual knowledge was stronger when coupled with solution generation, the same seems to be true for transfer as well.

As hypothesized, one mechanism is that problem generation allows for greater contextual flexibility for encoding and assembling of schemas. One can infer this by noting that the transfer items targeted normalization. This required students to reason by flexibly assembling the data point (x), average (x-bar), and SD into a single metric (the normalized score), either qualitatively or quantitatively. Thus, successful performance on these items required not only conceptual understanding of the various components but also aflexible assembling and reasoning with the components. The noticing and encoding of these components may have been further enhanced when problem generation was coupled solution generation. On the one hand, students need conceptual understanding of these components. On the other, if such understanding is not flexibly encoded, then it may not transfer well. By backward inference, the opportunity to generate different problems from the same data/situation must have been a useful mechanism for the development of such flexibility in assembly, while at the same time, reducing the likelihood of functional fixedness and set effects. Taken together, this may have allowed for more pathways for such knowledge to be cued, retrieved, and used in a novel context such as the ones targeted by the transfer items.

Finally, it is noteworthy that an alternative interpretation of these findings stems from a comparison with earlier studies on Productive Failure (PF), which have shown how and why generating solutions prior to instruction works (for a review, see Kapur 2016; also see Loibl et al. 2017). In this study too, students generated solutions. The difference however was is in terms of the problems to which solutions were generated; the problems being generated by students themselves as opposed to being specified in previous PF studies. Taking Kapur's (2015) study that compared problem solving with problem posing together with this one, findings suggest that boundary conditions for the applicability of PF, which was previously limited only to solution generation, seem broader than previous research has identified.

Conclusion

The work reported in this paper demonstrates that problem-posing with solution generation is a more beneficial preparatory activity than problem-posing without solution generation. It builds on Kapur's (2015) study that compared the preparatory effects of problem-posing with problem-solving to more clearly isolate the relative effects of problem and solution generation on the development of procedural knowledge, conceptual knowledge, and transfer. More broadly, the work builds upon the largely theoretical and descriptive studies on problem-posing in the math education literature to experimentally demonstrate the efficacy of problem-posing for learning new math concepts.

One must note however that this is only one study, which presents the start of an experimental program of research on problem posing. Therefore, findings cannot be generalized beyond the context, domain, topic, and the sampled profile of students. Clearly, these findings need to be replicated before deriving more robust conclusions. Furthermore, the study sheds no light on the nature of student production, and how its quality and relevance influence learning from subsequent instruction. Such finer grained process analysis is currently underway.

Given that the bulk of student experience in schools is largely centered on problem-solving, and an equally strong focus can be found on problem-solving in the cognitive and learning sciences, work such as the study reported in this paper advocates for the need for more research on problem-posing, with the hope that better understanding of problem-posing may translate into the design of learning in practice.

Acknowledgements A shorter version of this manuscript has been submitted to the 2017 Annual Meeting of the Cognitive Science Society. The author would like to thank the principal, teachers and students for their support and participation, as well as the research assistants who helped with data collection and coding.

Appendix: additional description of instruction phase

The approximate timeline for the activities in the instruction phase were as follows:

1. *First 10 min* The teacher engaged students in a qualitative examination of Problem 1 to disambiguate the concept of mean from SD, and to ensure that students understood the concept of SD qualitatively first before learning its quantitative formulation.
   Problem 1: Math grades on six tests for students, S1–S3

S1	A, B, A, B, A, B
S2	C, C, C, C, C, C
S3	A, E, A, E, A, E

   (Problem 1 qualitatively contrasted the mean from the spread in a data. It was designed to emphasize that S1 is better than S2 but S2 is more consistent than S1. Also that S1 is not only better than S3 but also more consistent. Finally, S2 and S3 may have the same average grade but S2 is more consistent.)

2. *Next 20 minutes* The teacher modeled and explained the solution to Problem 2 by using a step-by-step procedure for calculating SD. Each student was also provided with this step-by-step procedure typed on an A4 sheet of paper.
   Problem 2: Marks on five tests out of 20 for students S1 and S2

S1	12, 13, 14, 15, 16
S2	12, 14, 14, 14, 16

   (Problem 2 was designed to contrast how two data sets with the same mean and range can have different SDs.)

   Printed step-by-step procedure for calculating SD given to students:

xx

1. Calculate the **mean**	$\bar{x}$
2. Calculate the **deviation** between each point and the mean, and **square** this difference	$(x-\bar{x})^2$
3. **Sum** the squared deviation	$\sum (x-\bar{x})^2$
4. Take the **average**, i.e., divide the sum of the squared differences by the total number of values	$\frac{\sum (x-\bar{x})^2}{N}$
5. Take the **square root** of the average of the deviations. The square root of the average deviations is the **standard deviation**	$\sqrt{\frac{\sum (x-\bar{x})^2}{N}}$

3. *Next 15 min* the teacher gave students 5 min to solve Problem 3 on their own. Students were allowed to use the procedure sheet. After the 5 min were up, the teacher invited students to share their solutions, discussed and explained the solution, and provided corrective feedback where needed.
   Problem 3: Runs scored in five innings by two batsmen B1 and B2

B1	20, 40, 60, 80, 100
B2	0, 40, 60, 80, 120

(Problem 3 was designed to contrast how changing the range affects the spread. Said another way, even though B1 and B2 have the same mean, their SD is different because the end points of B2 are further away from the mean, and the range is greater.)

4. *The last 10 min* The teacher gave students 5 min to solve Problem 4 on their own but without the procedure sheet this time. Student solutions on Problem 4 were collected for analysis. The teacher then discussed and explained the solution.
   Problem 4: Daily temperature over 5 days in two cities C1 and C2

C1	36, 38, 40, 42, 44
C2	36, 38, 40, 42, 54

(Problem 4 was designed to contrast data sets with and without an outlier. C1 and C2 are exactly the same except 44 is replaced with 54.)

The remaining 5 min were left as buffer to be used as appropriate throughout the instruction phase.

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48*(2), 87–113.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61–100.

Brown, S. I., & Walter, M. I. (2005). *The art of problem posing* (3rd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *The Quarterly Journal of Experimental Psychology, 55A,* 241–261.

Cifarelli, V., & Sheets, C. (2009). Problem posing and problem solving: A dynamic connection. *School Science and Mathematics*, *109*, 245–246.

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology, 113*(4), 552–568.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*(5), i.

Einstein, A., & Infeld, L. (1938). *The Evolution of Physics*. New York: Simon and Schuster.

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

English, L. D. (1998). Children's problem posing within formal and informal contexts. *Journal for Research in mathematics Education, 29,* 83–106.

Frank, M. C., & Ramscar, M. (2003). How do presentation and context influence representation for functional fixedness tasks? In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science, 40*(4), 651–672.

Kapur, M. (2013). Comparing learning from productive failure and vicarious failure. *The Journal of the Learning Sciences, 23*(4), 651–677.

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008–1022.

Kapur, M. (2015). The preparatory effects of problem solving versus problem posing on learning from instruction. *Learning and Instruction, 39,* 23–31.

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist, 51*(2), 289–299.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences, 21*(1), 45–83.

Kilpatrick, J. (1987). Problem formulating: Where do good problems come from. *Cognitive Science and Mathematics Education* (pp. 123–147).

Loibl, K., & Rummel, N. (2013). Delaying instruction alone doesn't work: Comparing and contrasting student solutions is necessary for learning from problem-solving prior to instruction. In N. Rummel, M.

Kapur, M. Nathan, & S. Puntambekar (Eds.), *Proceedings of the 10th international conference on computer-supported collaborative learning* (*CSCL 2013*) (Vol. 1, pp. 296–303). International Society of the Learning Sciences, Inc.

Loibl, K., Roll, I., & Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction support learning. *Educational Psychology Review*, *29*, 693–715.

Lurchins, A. S., & Lurchins, E. H. (1959). *Rigidity of behaviour: A variational approach to the effects of einstellung*. Eugene: University of Oregon Books.

Mawer, R. F., & Sweller, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 8,* 252–259.

Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science, 23*(3), 305–336.

Moses, B., Bjork, E., & Goldenberg, E. P. (1990). Beyond problem solving: Problem posing. In T. J. Cooney & C. R. Hirsch (Eds.), *Teaching and learning mathematics in the 1990s* (pp. 82–91). Reston, VA: NCTM.

Paas, F. (1992). Training strategies for achieving transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*(4), 429–434.

Perrin, J. R. (2007). Problem posing at all levels in the calculus classroom. *School Science and Mathematics, 107*(5), 182–192.

Rittle-Johnson, B., & Star, J. R. (2009). Compared to what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*(3), 529–544.

Roll, I., Aleven, V., & Koedinger, K. (2011). Outcomes and mechanisms of transfer in invention activities. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2824–2829). Austin, TX: Cognitive Science Society.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–522.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759–775.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129–184.

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

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM*, *29*(3), 75–80.

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

Singer, F. M., Ellerton, N., & Cai, J. (2013). Problem-posing research in mathematics education: New questions and directions. *Educational Studies in Mathematics*, *83*, 1–7.

Sweller, J., Mawer, R. F., & Howe, W. (1982). Consequences of history-cued and means-end strategies in problem solving. *American Journal of Psychology, 95,* 455–483.

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 112,* 463–474.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. *Cognitive Science, 20,* 75–100.

Wiedmann, M., Leach, R. C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science., 40,* 711–730.

Wirth, J., Kunsting, J., & Leutner, D. (2009). The impact of goal specificity and goal type on learning outcome and cognitive load. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2008.12.004.