[Journal of Humanistic Mathematics](https://scholarship.claremont.edu/jhm)

[Volume 12](https://scholarship.claremont.edu/jhm/vol12) | [Issue 1](https://scholarship.claremont.edu/jhm/vol12/iss1) January 2022

Middle School Students Generating Mathematical Problems from a Real-life Situation

David Coffland
Idaho State University

Ying Xie
Northern Illinois University

Recommended Citation

David Coffland & Ying Xie, "Middle School Students Generating Mathematical Problems from a Real-life Situation," *Journal of Humanistic Mathematics*, Volume 12 Issue 1 (January 2022), pages 76-97. DOI: 10.5642/jhummath.202201.06. Available at: https://scholarship.claremont.edu/jhm/vol12/iss1/6

Middle School Students Generating Mathematical Problems from a Real-life Situation

David Coffland

Department of Organizational Learning & Performance
Idaho State University, Pocatello, Idaho, USA
coffdavi@isu.edu

Ying Xie

Department of Department of Educational Technology, Research and Assessment
Northern Illinois University, DeKalb, Illinois, USA
yxie@niu.edu

Abstract

In this study, we examined the effect of different presentation formats of a realistic situation on students' mathematical problem-posing behavior. We divided thirty-six middle school students into two groups, gave them a pretest, and then showed them a realistic, problem-posing situation in Artifact or Video format. We used Silver's core dimensions of creativity, namely fluency, flexibility, and originality, to measure participants' problem-posing activity. The results for the fluency measures showed that the Artifact group wrote more questions than the Video group but the same number of mathematics problems. The Video group posed problems in more mathematical domains than the Artifact group. Overall, our results indicate that mathematics instructors should align the presentation format with goals of the problem-posing exercise.

Keywords: mathematics education, problem posing, middle school students, realistic education.

1. Introduction

The ultimate purpose of mathematics is to help humans understand and describe the physical world around us. Some branches of mathematics and the related mathematical skills, such as counting, measurement, shape, calculation, problem solving, etc. have been abstracted from the whole to facilitate students' systematic math learning. However, mathematical skills are emphasized to such an extent in the K-12 curriculum that students often end up associating school mathematics with skill development without context. The drill and practice teaching method could hamper students from seeing mathematics as an organized study of the appearances and motions of physical objects in our environment [11]. Since many problems in mathematics texts are artificial and unrelated to real lives of students [2], middle school students have viewed mathematics as boring and difficult because school mathematics usually lacks the personal and social aspects children enjoy [41]. Although advocates (e.g., [19, 40, 38]) of Realistic Mathematics Education stressed that mathematics learning in school should involve problem situations experientially real to students, it could be stated that little success was achieved [18].

Esmonde *et al.* [17] argued that to reform mathematics classrooms, teachers should encourage students to mathematize problems they encountered in their daily lives through model-building projects in an effort to foster mathematical inquiry and discourse. An aspect of mathematical inquiry, problem posing, requires students to view and analyze their world through a mathematical lens. Yet, middle school mathematics seems to emphasize procedural efficiency [9].

The goal of this study was to have students examine a real-life scenario within the context of their mathematics class, thereby aiming to create a mathematical moment overlapping with their out-of-school lives. Presenting a realistic scenario in two different formats (Artifacts or a Video), we compared students' posed mathematical problems on three measures: fluency, flexibility, and originality. Our major focus of the study was to discover if different formats, representing different modalities of represented reality, could affect students' ability to describe situations mathematically by creating mathematical problems to solve.

2. Theoretical Framework

2.1. Mathematical Problem Posing

Mathematical problem posing refers to the generation and/or re-formulation of problems that may lead to a mathematical answer [31]. According to Silver, problem-posing tasks can provide researchers with both a window into students' mathematical thinking and a mirror reflecting students' mathematical experiences. Cai *et al.* [7] found that participants often lacked the ability to create quality mathematical problems. Ellerton's study [15] also attested that students felt creating problems was more difficult than solving them.

Foci of studies have included both the nature of posed problems [33, 34, 7, 25] and problem-posing strategies and processes of both students and teachers [34, 8, 35]. Other studies addressed the problem-posing activities by examining the pedagogical considerations of students posing problems [21, 20, 37, 27]. Other research has indicated that the context of the task situation could affect problem-posing activities [6, 36, 30, 23, 14].

2.2. Real life and Problem Posing

Students' lives outside of school were found by research to be rich in mathematics [22, 12, 24, 29, 28]. Yet school mathematics has been rarely seen as applying to real life [18]. In an attempt to bridge this divide, Bonotto [3, 4] used cultural artifacts (e.g, menus, brochures, etc.) as tools in classroom problem posing activities. In particular, she had fifth-grade students pose problems based on a number of cultural artifacts [4], finding that the students could use artifacts to identify the meaning from the real-world situations and create problems.

According to Bonotto [3], cultural artifacts can serve as a means for students' interpretation and understanding of reality. Our study extends Bonotto's research by presenting a real-life situation with two different modalities of a same situation to middle students: some context (in artifacts) and richer context (in video clips). According to Watson and Mason [39], when constructing a problem or solution, people depend on the wording of the prompt and the circumstances under which the prompt was presented. Based on this reasoning, we chose these two formats because presenting the prompt in different modalities could possibly have an effect on one's ability to generate mathematical problems from the presented situation.

Thus, the purpose of this study was to examine if different representations of the problem-posing situation could have an effect on the fluency, flexibility, and originality of middle school students' posed mathematical questions. Our research questions were:

- Would different representations of the same situation affect students' posed mathematical problems on the measures of fluency, flexibility, and originality?
- What was the nature of the problems students generated from the realistic situations?

3. Method

We employed the mixed-methods sequential explanatory design (see for example [13]). We first analyzed the problems generated by the participants quantitatively, and then qualitatively, to substantiate and explain the quantitative findings.

3.1. Participants

The participants were drawn from two middle school classes in a small city located in the Intermountain West of the United States. The sample consisted of 36 seventh- and eighth-grade students. Because we worked with the whole class in each case, we made use of a quasi-experimental design involving a pretest and a posttest.

The demographic breakdown of the school from which the participants were drawn was approximately 52% female and 48% male. The school was predominantly Caucasian (92%). Hispanics were the next most commonly represented group, at 3%. Black and Asian Americans each represented 2% of the sample, with Native Americans making up 1% of the school population. The sample of 36 students was representative of these percentages while not meeting them exactly. Due to the small number of non-Caucasian students, we did not collect racial or ethnic data from the students.

3.2. Procedure

Pretest. The participants were given a pretest to determine if there were differences between the two classes. The pretest consisted of a geometric diagram for which participants were asked to pose mathematical problems. We selected a geometric diagram so as not to sensitize participant students to either the video or artifact presentation format.

Pre-treatment. One of the researchers gave the classes instruction on posing problems from real or realistic situations. This was done to orient participants to posing mathematical problems. Since each group received the same verbal (discussion-based) instruction, it did not constitute a treatment, and again, avoided sensitization to either presentation format. The instruction was based on the "What-if?" portion of Brown and Walter's process for problem posing [5] and took the majority of two class periods for each group.

Treatment and Posttest. Following the training, the participants were given the treatment and posttest. More specifically, we provided the students with a realistic situation that described two people driving to a pair of restaurants to buy lunch in one of two presentation formats: Artifacts or a Video. We provided relevant information, such as the time and distance traveled, the route taken, gas prices, the menu options, and the price of menu items, to both groups. The artifacts were actual objects that provided abstracted evidence of the reality of the situation. The video provided visual evidence, somewhat less abstracted, because the situation was based on real places and possibilities.[1]

The Artifact group received copies of three artifacts: menus from two local restaurants and a map of the route taken to get from a house to the restaurants. One restaurant had only three menu options — a single item, two items, or three items from a list of selections. With these meals, the drink was included in the price. The other restaurant had approximately 20 menu items, representing different styles or sizes of hamburgers, hot dogs, French fries or drinks. The map showed all the streets in an area approximately 5 miles by 5 miles. Some locations on the map were labeled. A scale was included to allow for the computation of actual distances.

[1] Access to the artifacts and the video can be requested by contacting the authors.

Additionally, gas price information, identical to what was visible in the movie, was added to the map artifact.

The Video group watched a short (approximately 1 minute and 15 seconds) montage of still images of the menu boards at both restaurants and nine other pictures that showed a trip from a house to the restaurants. Each image was on screen for 6–7 seconds. The menu boards at the restaurants displayed the same price information as in the menus seen by the Artifact group. Four images showed the vehicle dashboard information — specifically, mileage and time. The remaining five images showed the trip — pictures of houses, streets and businesses along the route. These showed roads, traffic, a lottery billboard, and a gas station price sign. The video was shown to the participants and then the group was given time to pose problems based upon it. The video was repeated continuously during the problem-posing period.

Participants worked individually to pose as many problems about the situation as they could on the post-test. We allowed them to continue until 75%–80% of the class had stopped writing questions because we wanted to allow most participants to exhaust their questions. Each class required about 10 minutes to pose their problems.

3.3. Instruments

We then read and transcribed the participant-created mathematical responses for both formats. We used Silver's core dimensions of creativity [32] — fluency, flexibility, and originality — to measure participants' problem-posing activity and examined their responses for each dimension.

Fluency was operationalized as both the total number of questions written applicable to the situational context and the number of mathematical problems posed by the participants in the given time period. At this point in the study, the student questions were identified as being in one of three categories: irrelevant questions, information seeking questions, and mathematical problems. Questions in the first category (irrelevant questions) were contextually inappropriate and did not include a mathematical operation or concept. An example was "Was the food worth the drive?" Information-seeking questions were situation-relevant questions that sought new data, perhaps for use in creating a mathematical problem. Questions such as "What did you order?"

fell into this category. Mathematical problems were contextually relevant and solvable questions from the information given in the problem and the artifacts or video. Both raters examined the written problems and made a determination of which problems were rated as irrelevant, information seeking, or mathematical problems.

Flexibility was defined as the number of different and pertinent ideas created in a given time period. In order to evaluate the flexibility of the participants, we categorized the mathematical problems based on the domains of the mathematical problems participants generated (e.g., speed, distance, time, cost, etc.) [4].

Lastly, originality was operationalized as the number of original questions participants created. Questions were considered original relative to other questions in the same treatment group. If a question was written by ten percent or less of the students in each treatment group (rounded to the nearest whole number), we labeled it original as per van Harpen and Presmeg [38]. Examples of the coding can be found in Appendix A.

Both raters coded 8% of 381 responses and compared their results, resulting in a Cohen's kappa = .871. One rater coded the rest of the questions participants generated.

3.4. Data Analysis

Since this study used intact classes as groups, a pretest was administered to all the participants to examine their problem-posing starting point. The pretest consisted of writing mathematics problems when given a geometric diagram. A MANOVA procedure was used to compare the two groups on the measures of fluency, flexibility, and originality. All of the differences were found to be statistically equivalent (see Table 1). Therefore, these two groups were regarded as equal in terms of their prior ability to generate mathematical problems.

Then, we used a MANOVA to analyze the posttest results for the Artifact and Video groups. Following the statistical analysis of the posttest, we once again analyzed the problems participants in these two groups generated by using an interpretive qualitative data analysis approach to provide further and possible explanations to the findings of the quantitative analysis [13].

Table 1: MANOVA Analysis of Pretest.

Problem Posing Skills	Operationalized Dependent Variable	Group	Mean	St. Dev.	N	F df=(1,34)	*p*
Fluency	Number of Math Problems	Artifact	7.37	3.56	19		
		Video	6.00	2.96	17	1.55	.22
	Number of Total Questions	Artifact	8.32	3.58	19	3.09	.09
		Video	6.47	2.58	17		
Flexibility	Number of Domains	Artifact	3.26	1.41	19		
		Video	3.00	1.17	17	0.37	.55
Originality	Number of Original Problems	Artifact	1.11	1.52	19		
		Video	0.71	0.92	17	0.88	.36

According to Elliot and Timulak [16], an interpretive analysis method is appropriate when the research focus is to explain why the phenomenon could possibly come about and how it evolves over time.

4. Results

The data analysis of the posttests for the artifact and video groups resulted in several findings. The MANCOVA data analyses of the four measurements for the two groups are shown in Table 2.

We used two measures of participant fluency: total number of questions and the number of mathematical problems. The results for the total number of questions showed a statistically significant difference $F(1, 33) = 4.87$, $p = .036$ with the Artifact group asking more total questions than the Video group. However, the analysis for the number of mathematical problems showed no statistically significant differences $F(1, 33) = 1.63$, $p = .210$.

Table 2: MANCOVA Analyses of Post Test Summary Table.

Problem Posing Skills	Operationalized Dependent Variable	Group	Mean	St. Dev.	N	F df= (1,33)	p	$\eta 2$
Fluency	Number of Math Problems	Artifact	3.37	2.65	19	1.63	.210	.047
		Video	4.29	2.86	17			
	Number of Total Questions	Artifact	13.84	9.94	19	4.87	.036*	.127
		Video	6.47	3.45	17			
Flexibility	Number of Domains	Artifact	1.89	1.41	19	5.90	.021*	.152
		Video	3.23	1.85	17			
Originality	Number of Original Problems	Artifact	0.73	1.05	19	0.81	.375	.024
		Video	1.06	2.11	17			

Note: * $p < .05$

Taken together, these results indicate that the Artifact group asked more total questions than the Video group, but the additional questions were not usually mathematical problems.

The measure of flexibility was the number of domains represented by the participants' questions. The analysis resulted in a statistically significant difference $F(1, 33) = 5.90$, $p = .021$, showing that the Video group posed problems in more domains than the Artifact group.

The number of questions that were evaluated as original showed no statistically significant differences between the two groups $F(1, 33) = 0.81$, $p = .375$. Thus, we found no evidence that the presentation formats had any effect on the number of original problems posed by participants.

The qualitative analysis results about the nature of the problems participants posed are presented in the discussion section below. These results provide explanations for and substantiate the interpretations of the quantitative data.

5. Discussion

The quantitative analysis showed differences between the Artifact and Video groups. We discuss both the quantitative and the qualitative results below.

5.1. Difference in the Total Number of Questions (Fluency)

The data in Table 2 show that the Artifact group asked a statistically significant average of 7.3 more total questions than the Video group; however, there was no statistical difference between the numbers of mathematical problems written by the two groups. Since there was no difference on the pretest, these findings indicate that the presentation formats led to different behaviors.

The Artifact group averaged 13.8 total questions while an average of 3.4 were considered mathematical problems. The Video group asked a mean of 6.5 total questions with 4.3 judged as mathematical problems. The Artifact group posed more than twice as many total questions as the Video group but nearly five times as many were classified as information seeking. An examination of the questions that were not judged to be mathematical problems showed that they were relevant, i.e. the students asked questions related to the problem-posing situation.

One possible explanation for the differences we saw might be that the paper artifacts were constantly and immediately available to the participants. To find a piece of information, students needed only to read it from the artifacts. A participant could pose a problem about the cost of three hamburgers at $5.49 each and with a glance back at the menu pose another problem about the cost of three cheeseburgers at $6.29 each. This finding is also consistent with systematic variation as explained by Silver *et al.* [34]. This type of process occurs when “a critical aspect of a problem is held constant while other critical aspects are varied systematically” (pages 303–304).

In contrast, the Video group saw imagery in the video that also contained text (restaurant menu boards, signage on the travel route, clock displays, and odometer images), but these pieces of information were not constantly available, so students couldn’t easily generate a larger number of similar questions through systematic variation. As a result, participants in the Artifact group were able to generate significantly larger number of questions than the video group.

Although the difference in total questions participants in both groups asked was statistically significant, the number of mathematical problems did not significantly differ. More than two-thirds of the questions submitted by the Video group were judged to be mathematical problems, compared to approximately one-fourth for the Artifact group. The Artifact group participants spent more of their time on writing information-seeking questions than the Video group. Figure 1 shows examples of questions coded as information seeking.

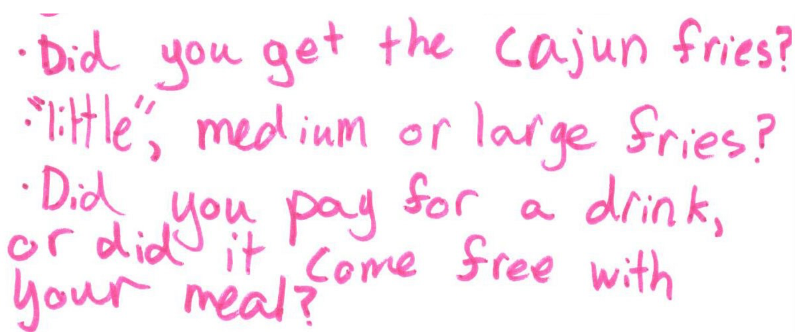

Figure 1: Examples of questions coded as information seeking.

The Video group did not show a pattern of creating a large percentage of information-seeking questions. It is possible to interpret these information-seeking questions asked by the Artifact group as the participants seeking to clarify their understanding of the situation. Since an understanding of the situation would be needed to develop a mathematical problem, the seeking of more information can be considered a precursor step in the problem-posing process. Christou *et al.* [14] proposed a model that described participants' problem-posing thinking as one of four processes: editing, selecting, comprehending, or translating information from the situation. Building on the work of Mamona-Downs [26], they suggested that in order for participants to edit tasks that are defined as "pose a problem without any restriction from provided information, stories or prompts", they need to first engage in "extracting information from the story context" (page 156).

In our study, the artifacts (two menus and a map) presented a relatively large amount of information at one time (menu offerings, prices, streets, distances, etc.); however, they did not tell a story of exactly what happened on the trip to get lunch. Instead, any piece of the data presented by the artifacts was potentially useful in posing a mathematical problem if the participants created (edited) a specific context in which to use it. The questions asked by some in the Artifact group sought specific data that would be of use in describing a situation about which participants could develop solvable mathematical problems. Figure 2 shows an example of a meaningful mathematical problem since the question posed a solvable problem requiring multi-step mathematical operations.

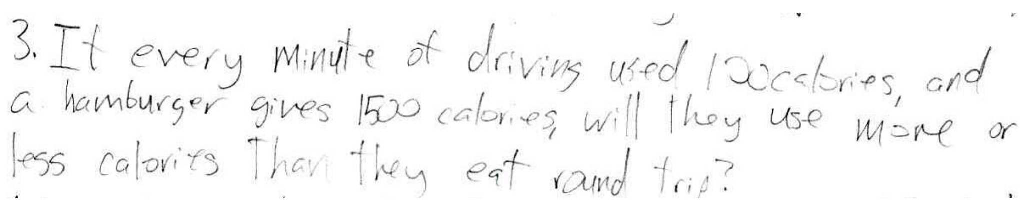
3. If every minute of driving used 100 calories, and a hamburger gives 1500 calories, will they use more or less calories than they eat round trip?

Figure 2: An example of a "good" mathematical problem: "If every minute of driving used 100 calories, and a hamburger gives 1500 calories, will they use more or less calories than they eat round trip?"

5.2. Difference in the Number of Domains (Flexibility)

The Video group posed mathematical problems in an average of 3.2 domains. The Artifact group had a mean of 1.9 mathematical problem domains. This difference was statistically significant, indicating that the presentation format had an effect on the number of domains for which participants posed problems.

The most common domains of the problems generated by participants in both groups were *distance* and *cost.* An example of a distance domain question is "In kilometers, how far did they travel round trip?" And here is a cost domain question: "Would it be a better deal to just get 5 drinks and pay for them or buy 1 meal and get free refills?" Participants in the Artifact group rarely generated problems in the *time* domain while many participants in the Video group posed problems related to the passage of time, such as "How long did it take you to get to the destination?" A variety of other miscellaneous domains were also covered by problems generated by the Video group participants. Examples of the *number* domain were, "How many cars were at Cole Chevrolet?" and "How many customers were at the Great Wall?" An example of *speed* domain was "How many avg. mph [average miles per hour] were they traveling?"

Several participants in the Artifact group posed a series of related mathematical problems, often modifying only a single element of the previous problem to create a new problem. This is representative of systematic variation as described by Silver *et al.* [34]. Figure 3 shows an example of this.

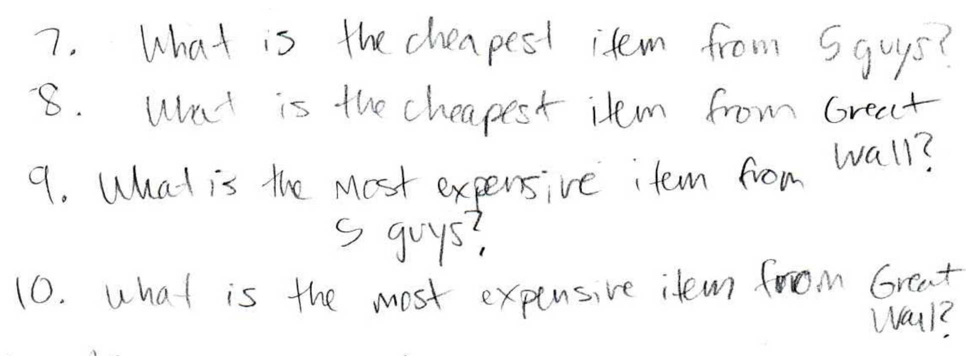

Figure 3: Examples of related questions identified as systematic variation.

Since systematically varied problems were closely related, we could expect that they would address the same domain. And indeed this was the case. In the example shown in Figure 3, the four problems showed all address the cost of food. Each of the problems was classified as belonging to the *cost* domain. Overall, when participants exhibited systematic variation in this way, we noted that the resulting problems were always within a single domain. We also noted that participants in the Artifact group exhibited more systematic variation than did the Video group participants, both in frequency of systematic variation and in the lengths of the problem sets.

5.3. Level of Mathematical Problems

It is worth noting that many of the mathematical problems written by the students were at a lower level than the grade level of the participants. One example was “If you ate for 30 min. and drove for eight minutes to and from and a 9 min. wait for food, how long did it take total?” This question was solvable with the information provided and on topic, so it was judged to be a mathematical problem. However, the problem required merely adding a series of whole numbers to find the solution.

The participants in this study were novice problem posers and the problem-posing unit was not an integrated part of their normal curriculum. As a result of their inexperience as problem posers, as well as the perceived separation of problem-posing tasks from their regular assignments, some of the middle school participants may have posed problems requiring only mathematics with which they were comfortable and confident in finding solutions.

6. Conclusions, Limitations, and Future Research

We found that presenting different modalities of a same situation made a difference in the number of questions asked and domains of the problems participants posed. The artifacts that one group received, the menus and the map, are abstracted reality, in the sense that mathematical information was already distilled from a typical ill-structured, real-life situation into familiar textual menu lists and a graphical map. In contrast, when viewing the video, participants were faced with the ill-structured and realistic situation, packed with multiple modes of less organized information. The familiar and predictable nature of information presented in the artifacts largely encouraged systematic variation [34].

According to dual-coding theory [10] and schema theory [1], the artifacts, the menus, were likely to activate only the verbal channel. As a result, the spread of activation of nodes in the brain was limited to primarily that channel using associative connections. These connections process data sequentially. The ready availability of large amounts of similar data in the menus — prices — and the relatively limited spread of activation may have influenced some Artifact group participants to write groups of related problems. In contrast, the Video group was presented information utilizing both the verbal and the imagery channels. This resulted in a larger spread of activation that consisted of both associative and referential connections. Unlike the Artifact participants who systematically varied some of their problems, the wider nodal activation of the Video group may have allowed then to evoke images of more elements of the situation and thus write problems in more domains. From our results, it seems plausible that the same real-life situation represented in different number of modalities (with both imagery and textual information vs. predominantly textual information) affected participants' ability to pose mathematical problems differently. The video format could have sparked participants to activate their schema in more channels, leading to problem generation in more domains.

Our study has several limitations. First, we used a quasi-experimental design, and in an attempt to address this issue, we used a pretest-posttest design with analyses of covariance. The pretest data showed no differences at the $a = .05$ level between the Artifact and Video groups. The pretest scores on each measure were then used as covariates when comparing posttest scores between groups.

A second limitation of our study is the different level of details available in the artifact and video formats. The video information had full pictures of parts of the trip while the map showed additional roads that weren't seen in the video. Even though this limitation was inherent to the design and research questions in this study, we took special care to ensure that the extraneous information did not affect the main findings of the study. For example, we found that very few participants in the Video group used their extra information to create math problems and none of the Artifact participants posed mathematical problems using their extra information.

A third limitation is inherent in our data collection. Participants were not interviewed about their experiences of using different types of prompts when generating mathematical problems. Future research should consider using a purposeful sampling strategy to inquire into experiences of participants of interest and therefore provide richer descriptions and even explanations about their thinking processes when generating problems from two different formats.

The results of this study suggest several avenues for future research. First, as seen in the literature review (Section 2), there is no complete theoretical framework to explain students' problem-posing processes from real life situations. Qualitative research that provides students with situations where they can create problems, reflect on their thinking, and discuss their process would be a step toward developing a theoretical framework. Second, the Artifact group posed more information seeking questions than mathematical problems in this study. Future research could re-examine this result to determine if the structure provided by the artifacts contributed to this result or if the fact that the participants were novice problem posers had an impact. Lastly, the Video group accessed their information from a continuously looping video playback. They were forced to either write down notes for later use, remember the information they needed, or wait for it to reappear.

Another study could examine if a more personal and controllable technology (such as a video available on individual tablet computers) would allow students to create more mathematical problems.

The study is significant in several different ways. First, many researchers have called for a close marriage between mathematics classrooms and participants' real lives outside school. Yet, addressing this call and examining the effect of reality-based representations in different representations on students' mathematical behaviors is difficult and a void remains in our understanding. With this study we believe we filled in a portion of this void.

Second, we have identified through this study how the presentation format in a problem-posing task affects the number, type, and domain of the questions participants ask. The results of this study indicate that the formats of the problem situation affected participants' problem posing about that situation. Thus, instructors and researchers of mathematics interested in having students pose problems from reality should be more attentive to the ill-structured messiness of the problem situation presented to the participants. Presentations in more modalities, although a good source for generating mathematical problems, could require extra cognitive power to edit and abstract the information. As a result, extra support, such as a handy device to capture important albeit transient information, could be used to reduce students' cognitive load and therefore facilitate the mathematical problem-posing process.

In summary, our work here is one attempt to bring realistic situations to problem-posing activities. We hope that our findings will contribute to future research in the topic and further support and encourage the introduction of real-life situations in mathematics classrooms.

References

[1] Bigenho, F. W., *Conceptual Developments in Schema Theory*, ERIC Report, 1992. https://eric.ed.gov/?id=ED351392.

[2] Boaler, J., *What's math got to do with it? Helping children learn to love their least favorite subject – and why it's important for America*, Penguin, New York, 2008.

[3] Bonotto, C., "Engaging students in mathematical modeling and problem posing activities," *Journal of Mathematical Modeling and Applications*, Volume **1** Issue 3 (2010), pages 18–32.

[4] Bonotto, C., "Artifacts as sources for problem-posing activities," *Educational Studies in Mathematics*, Volume **83** (2013), pages 37–55.

[5] Brown, S. & Walter, M., *The Art of Problem Posing*, 3rd edition, Routledge, New York, 2005.

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

[7] Cai, J., Moyer, J., Wang, N., Hwang, S., Nie, B., & Garber, T., "Mathematical problem posing as a measure of curricular effect on students' learning," *Educational Studies in Mathematics*, Volume **83** (2013), pages 57–69.

[8] Cifarelli, V. V. & Cai, J. "The evolution of mathematical explorations in open-ended problem-solving situations," *The Journal of Mathematical Behavior*, Volume **24** (2005), pages 302–324.

[9] Civil, M., "Everyday mathematics, mathematicians' mathematics, and school mathematics: Can we bring them together?" *Journal for Research in Mathematics Education Monograph* Volume **11** (2002), pages 40–62.

[10] Clark, J. M. & Paivio, A., "Dual coding theory and education," *Educational Psychology Review*, Volume **3** Issue 3 (1991), pages 149–210.

[11] Coffland, D. & Xie, Y., "The 21st century mathematics curriculum: A technology enhanced experience," pages 311–329 in *Full Steam Ahead: Emerging Technologies for STEAM*, edited by X. Ge, D. Ifenthaler, & J.M. Spector (Springer, New York, 2015).

[12] Cole, M., *Cultural Psychology: A Once and Future Discipline*, Harvard University Press, Cambridge, MA, 1996.

[13] Creswell, J., *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4$^{\text{th}}$ edition, Sage, Los Angeles, 2014.

[14] Christou, C., Mousoulides, N., Pittalis, M., Pitta-Pantazi, D., & Shiraman, B., "An empirical taxonomy of problem posing processes," *ZDM*, Volume **37** Issue 3 (2005), pages 149–158.

[15] Ellerton, N., "Engaging pre-service middle-school teacher-education students in mathematical problem posing: Development of an active learning framework," *Educational Studies in Mathematics*, Volume **83** (2013), pages 87–101.

[16] Elliot, R. & Timulak, L., "Descriptive and interpretive approaches to qualitative research," pages 147–159 in *A Handbook of Research Methods for Clinical and Health Psychology* edited by J. Miles & P. Gilbert (Oxford University Press, 2005).

[17] Esmonde, I., Blair, K. P., Goldman, S., Martin, L., Jimenez, O., & Pea, R., "Math I Am: What we learn from stories that people tell about math in their lives," pages 7–27 in *LOST Opportunities: Learning in Out-of-School Time* edited by B. Bevan, P. Bell, R. Stevens & A. Razfar (Springer, Netherlands, 2013).

[18] Goldman, S. & Booker, A., "Making math a definition of the situation: Families as sites for mathematical practices," *Anthropology & Education Quarterly*, Volume **40** Number 4 (2009), pages 369–387.

[19] Gravemeijer, K. P. E., *Developing Realistic Mathematics Education*, CD-Beta-Press, 1994.

[20] Koichu, B. & Kontorovich, I., "Dissecting success stories on mathematical problem posing: A case of the Billiard Task," *Educational Studies in Mathematics*, Volume **83** (2013), pages 71–86.

[21] Kul, Ü. & Çelik, S., "A Meta-Analysis of the Impact of Problem Posing Strategies on Students' Learning of Mathematics," *Romanian Journal for Multidimensional Education*, Volume **12** Issue 3 (2020), pages 341–368.

[22] Lave, J., *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*, Cambridge University Press, Cambridge, UK, 1988.

[23] Leung, S. S. & Silver, E. A., "The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers," *Mathematics Education Research Journal*, Volume **9** Issue 1 (1997), pages 5–24.

[24] Li, X., Song, N., Hwang, S., & Cai, J., "Learning to teach mathematics through problem posing: Teachers' beliefs and performance on problem posing," *Educational Studies in Mathematics*, Volume **105** Issue 3 (2020), pages 325–347.

[25] Liu, Q., Liu, J., Cai, J., & Zhang, Z., "The relationship between domain-and task-specific self-efficacy and mathematical problem posing: A large-scale study of eighth-grade students in China," *Educational Studies in Mathematics*, Volume **105** Issue 3 (2020), pages 407–431.

[26] Mamona-Downs, J., "On analyzing problem posing," pages 41–47 in Volume 3 of *Proceedings of the 17th Conference of the International Group for the Psychology of Mathematics Education*, edited by I. Hirabayashi, N. Nohada, K. Shigematsu, & F. L. Lin (Tsukuba, Japan, 1993).

[27] Olsen, J. & Knott, L., "When a problem is more than a teacher's question," *Educational Studies in Mathematics*, Volume **83** (2013), pages 27–36.

[28] Pea, R. D. & Martin, L., "Values that occasion and guide mathematics in the family," *The Teachers College Record*, Volume **109** Number 1 (2010), pages 34–52. doi:10.1177/016146811011201303

[29] Saxe, G. B., "Candy selling and mathematics learning," *Educational Researcher*, Volume **17** Issue 6 (1988), pages 14–21.

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

[31] Silver, E., "On mathematical problem posing," *For the Learning of Mathematics*, Volume **14** Issue 1 (1994), pages 19–28.

[32] Silver, E., "Fostering creativity through instruction rich in mathematical problem solving and problem posing," *ZDM*, Volume **97** Issue 3 (1997), pages 75–80.

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

[34] Silver, E., Mamona-Downs, J., Leung, S., & Kenny, P., "Posing mathematical problems: An exploratory study," *Journal for Research in Mathematics Education*, Volume **27** Issue 3 (1996), pages 293–309.

[35] Singer, F. M., Ellerton, N. & Cai, J., editors, *Mathematical Problem Posing: From Research to Effective Practice*, Springer, New York, NY, 2015.

[36] Stoyanova, E. & Ellerton, N. F., "A framework for research into students' problem posing in school mathematics," pages 518–525 in *Technology in Mathematics Education*, edited by P. Clarkson (Mathematics Education Research Group of Australasia, Melbourne, 1996).

[37] Tichà, M. & Hospesovà, A., "Developing teachers' subject didactic competence through problem posing," *Educational Studies in Mathematics*, Volume **83** (2013), pages 133–143.

[38] Van Harpen, X. & Presmeg, N., "An investigation of relationships between students' mathematical problem-posing abilities and their mathematical content knowledge," *Educational Studies in Mathematics*, Volume **83** (2013), pages 117–132.

[39] Watson, A. & Mason, J., *Mathematics as a Constructive Activity: Learners Generating Examples*, Lawrence Erlbaum Associates, Mahwah, NJ, 2005.

[40] Webb, D. C., van der Kooij, H., & Geist, M., "Design Research in the Netherlands: Introducing Logarithms using Realistic Mathematics Education," *Journal for Mathematics Education at Teachers College*, Volume **2** Issue 1 (2011), pages 47–52.

[41] Wijers, M., Jonker, V., & Kerstens, K., "MobileMath: the Phone, the game and the math," pages 507–516 in *Proceedings of the European Conference on Game Based Learning*, (Barcelona, 2008).

A. Data Analysis Coding Examples

In this section four examples of how the participant's problems were coded are given.

1. How long was the wait?
   - **Mathematical problem or not?** This was an information-skeeing question but *not* a mathematical problem. Rationale: The question sought information from the reader rather than asking the reader to solve a problem. Also, no mathematical operation was needed to answer the question.
2. In kilometers, how far did they travel round trip?
   - **Mathematical problem or not?** This was a mathematical problem. Rationale: the question used information in the artifacts or video to pose a solvable problem requiring at least one mathematical operation.
   - **Domain** Distance.
   - **Originality** Not original since more than 10% of the participants posed problems related to the trip's total distance.
3. How many cheeseburgers could you buy with $20?
   - **Mathematical problem or not?** This was a mathematical problem. Rationale: the question used information in the artifacts or video to pose a solvable problem requiring at least one mathematical operation.
   - **Domain** Cost.
   - **Originality** Not original since more than 10% of the participants posed problems realted to the cost of the food items.

4. If every minute of driving used 100 calories, and a hamburger gives 1500 calories, will they use more or less calories than they eat round trip?

    - **Mathematical problem or not?** This was a mathematical problem. Rationale: the question used information in the artifacts or video to pose a solvable problem requiring at least one mathematical operation.
    - **Domain** Calories.
    - **Originality** Original since fewer than 10% of the participants in each group posed problems related to the food's calorie content and the multi-step nature of the problem.