Creative and algorithmic reasoning – the role of strategy choices in practice and test

TOMAS BERGQVIST AND MATHIAS NORQVIST

This study is based on a framework of algorithmic and creative mathematical reasoning and focuses on students' strategy choices in both practice and test. Previous research indicates that students that practice mathematics with tasks with given solution methods are outperformed in later test by students that have to construct solution methods during practice. Video recordings, students' written solutions, and student interviews from ten university students provides data on strategy choices. The analysis was carried out to capture students' strategy choices and reasons for these choices. The results showed that there was no real difference in how the students solved the tasks in the test. Regardless of practice condition, more or less the same solution strategies were used in the test situation.

Many students practice mathematical methods without connecting them to conceptual aspects. Memorizing a set of solution methods is one of the most efficient learning strategies in a short-term perspective, while it could be harmful in the long run (Hiebert, 2003; Hiebert & Grouws, 2007). It is possible to solve the majority of all tasks in upper secondary school textbooks by imitating solution templates (e.g. Jäder et al., 2020). It is even possible to pass most exams in the first term of university studies by memorizing a set of solution strategies (Bergqvist, 2007). However, when you look at mathematics learning in a longer perspective, it is doubtful if learning a set of algorithms results in any deeper conceptual understanding (Hiebert, 2003).

Of course, without a given template, solving a task will be considerably more difficult. Brousseau (1997) clarifies that when students fail to complete a solution process it is a normal step in the students' knowledge building. It is in fact not a failure, but a necessary step towards concept acquisition (Kapur, 2010). Bjork and Bjork (2011) also argue that these

Tomas Bergqvist, *Umeå university*
Mathias Norqvist, *Umeå university*

Bergqvist, T. & Norqvist, M. (2022). Creative and algorithmic reasoning – the role of strategy choices in practice and test. *Nordic Studies in Mathematics Education*, 27 (1), 5–25.

difficulties are desirable when learning, but the desirable difficulties must be surmountable and concern the subject that is supposed to be learned. Solving tasks by imitating solution templates removes most difficulties and will therefore reduce the students' opportunities to understand the concepts involved.

In multiple studies the research program *Learning by imitative and creative reasoning*, LICR, (Lithner, 2017) have shown that students who practice mathematics by constructing solutions, outperform students who are given a solution template during practice (e.g. Jonsson et al., 2014; Karlsson Wirebring et al., 2015; Norqvist, 2018; Olsson & Granberg, 2019). In the study by Jonsson et al. (2014), the students who used a solution template scored about 95 % on the practice tasks, whereas students who had to construct their own solutions scored just above 60 %. On the test, the group who used a solution template during practice scored significantly lower than the other group. These results have been verified in other studies (e.g. Karlsson Wirebring et al., 2015; Norqvist, 2018; Olsson & Granberg, 2019). The reasons for the difference in performance between the two student groups is however not clear. One possible explanation is connected to differences in the students' need to struggle with the tasks during the practice session (Jonsson et al., 2016), but several other possible explanations exist.

There is a lack of information about what the students actually are doing in the practicing and test sessions. In this study we therefore focus on understanding more about the students' activities in two ways. Firstly, by characterizing students' strategy choices in both practice and test sessions, and analyzing the relation between these choices. Secondly, by interviewing students to understand more about their experiences of the practice and test session.

Background

In this section we will explore the foundation of the present study. The background will start with the importance of mathematical reasoning, which leads to the notions of strategy choices and the use of algorithms when learning mathematics, and finally, a summary of the earlier studies in the LICR-program will lead us to the aim and research questions of this study.

Mathematical reasoning

One of the competencies in mathematics is the ability to reason, that is, to give mathematically founded arguments for a solution method or strategy (e.g. Niss, 2003; Kilpatrick et al., 2001). This ability has, in one form or another, been a part of the mathematics syllabi in many countries (e.g. USA, Denmark, Singapore and Sweden) for at least two decades. In the late 1980's the American teacher association *National council of teachers of mathematics* first included reasoning as part of the mathematics curriculum (NCTM, 1989), and this is elaborated in the 2011 *Common core state standards* for mathematics (NCTM, 2011). The Danish KOM-project defined eight competencies needed to get a full grasp of mathematics, where reasoning is one of them (Niss & Jensen, 2002; Niss, 2003). The KOM-project has also influenced the Swedish syllabus. In the current Swedish syllabi, learning to reason mathematically is part of the aim of mathematics teaching together with procedural knowledge, communication, conceptual understanding, and problem solving (Skolverket, 2011a, 2011b).

Kilpatrick et al. (2001) defines five strands of mathematical proficiency, where "adaptive reasoning" is one of them. Adaptive reasoning is seen as "the glue that holds everything together" (p. 129) and includes a broad view of what reasoning entails. Here, both intuitive and inductive reasoning, such as finding patterns, and more formal mathematical reasoning, such as proofs, are included. This indicates that mathematical reasoning can be done by all, even young children, but that the argumentation may be of a different type. As long as a solution or theorem can be justified by logical arguments (based on the knowledge the child has) reasoning has occurred. This broader view of reasoning matches up with how reasoning is viewed by Lithner (2008, 2017). Lithner (2008, 2017) argues that mathematical reasoning always occurs when a task is solved, even though it might be rudimentary or is provided by a textbook, peer, or teacher. To distinguish between different sub-types of mathematical reasoning, Lithner (2008, 2017) suggests definitions for several types, and two of these, *Algorithmic reasoning* (AR) and *Creative mathematically founded reasoning* (CMR), are central for the present study. The definitions of these two types of reasoning are as follows:

> *Algorithmic reasoning* is an attempt to "solve a task by applying a given or recalled algorithm" (Lithner, 2017, p. 939), where an algorithm is defined as a "fixed set of step-by-step procedures for solving (mathematics) problems" (Fan & Bokhove, 2014, p. 486). This also includes pre-defined steps without calculations (e.g. measuring an angle in a figure). When students work with tasks designed for algorithmic reasoning (AR), they use a given or recalled solution strategy.
>
> *Creative mathematically founded reasoning* is when the solver creates or recreates a solution sequence that is new (or forgotten) for the

solver (Lithner, 2008, 2017). The solution should be supported by predictive and verificative arguments, and the arguments should be anchored in the intrinsic mathematical properties that are involved. When the students work with tasks designed for creative mathematically founded reasoning (CMR), they create a new solution strategy or recreate a forgotten one.

Strategies and strategy choice

In research literature, the word ”strategy” is most often used in its everyday meaning, a way to address a task. For example, Andrews et al. (2021) presents a thorough literature review on computational estimation and uses the word strategy to mean the ways students perform computational estimations. Polya writes in his famous book (1945) about ”heuristic strategies” as a tool for problem solving. Schoenfeld (1985) expands this concept to mean ”rules of thumb for successful problem solving, general suggestions that help an individual to understand a problem better or to make progress towards its solution” (p. 23). This indicates that a strategy can be of different types and should be given a rather broad definition.

In this study we will base our definition on Bergqvist and Lithner (2012), so that a ”strategy” can vary from very local procedures to general approaches. The ”choice” is then used in a broad sense, for example, recalling a strategy, constructing a strategy, or choosing between two or more known strategies. In many situations there might in fact not be a choice involved since the students only have one strategy available. However, a strategy will be used when trying to solve mathematics tasks.

When students work with mathematical tasks, they can often imitate a template or use a given algorithm, that is, use a solution strategy presented in the book or by the teacher (e.g. Newton & Newton, 2007; Shield & Dole, 2013). Jäder et al. (2020) found that 79 % of all tasks in the common textbooks in twelve countries (including, for example, USA, Singapore and Finland) could be solved by imitating a given procedure. Tasks used in tests look for the most part the same as the tasks in the textbook. However, there is one large difference: there is nothing to imitate during the test. This means that students must rely on memory from the training or create a new solution during the test (Lithner, 2017). Neither strategy is something that students normally practice.

Brousseau (1997) makes it very clear why it is inefficient to use readymade algorithms in the learning process. He explains that an algorithm is a task-solving method that guarantees success if you follow all the steps, and that the algorithm is created to avoid having to regard the mathematical properties of the task. The idea with the algorithm is actually that you don’t need to understand, only blindly follow the procedure. There is clear evidence that such activities will not lead to deeper understanding or develop central mathematical competencies (Hiebert, 2003; Lithner, 2017). It is claimed that extensive use of algorithms is contra-productive (Hiebert, 2003; Niss, 2007), which indicates that the change from rote learning to learning for in-depth understanding is a very important task for the whole mathematics education community.

Research has shown that additional effort during practice will ease later retrieval of the practiced material (e.g. Pyc & Rawson, 2009; Wiklund-Hörnqvist et al., 2014). Solving tasks without solution templates (i.e., creating one's own solution strategies) is more efficient for later retrieval than solving tasks where solution templates are given (e.g. Fyfe & Rittle-Johnson, 2017; Jonsson et al., 2014; Norqvist, 2018), and this is the main rationale behind the LICR-program.

Results from the LICR-program

The main idea of this study is to understand more about possible explanations to previous results in the LICR-program. Jonsson et al. (2014) studied the eventual difference in task efficiency between tasks where the students were given a solution template or a formula (called AR-tasks, Algorithmic reasoning) and tasks where the students had to construct their own solution (called CMR-tasks, Creative mathematical reasoning). The efficiency was measured by post-test performance. The group of students who had practiced CMR-tasks outperformed the group practicing AR-tasks, and additionally there were indications that practicing CMR-tasks was especially beneficial for low-performing students. One critique to these results has been that similarities between practice tasks without solution templates and test tasks, so called transfer appropriate processing, could be the reason for the difference in test performance. Jonsson et al. (2016) studied how transfer appropriate processing and productive struggle influenced test scores for tasks with and without solution templates. Although there was a minor effect of transfer appropriate processing, the main reason for the observed difference in test scores was the increased effort needed during practice without solution templates. Variants of the study by Jonsson et al. (2014) illuminate different aspects of the situation. Norqvist (2018) argued that if an explanation of *why* the algorithm works was added to the AR-tasks, similar to textbooks or teacher instructions, the efficiency of these tasks could increase. The study did however show no such performance gain. CMR-tasks has also been shown to be beneficial when incorporating GeoGebra in mathematics education (Granberg & Olsson, 2015; Olsson, 2017; Olsson & Granberg, 2019). Olsson and Granberg (2019) confirmed that students that practiced with CMR-tasks showed a significantly higher post-test performance than the AR-practice group, although all students utilized GeoGebra during both practice and test. Eye-tracking has revealed that students that solve practice tasks with solution templates will focus on either a given formula or a solved example, while students that practice without solution template will have a larger focus on the illustration (Norqvist et al., 2019). However, Norqvist et al. (2019) showed that among the students that practice without solution templates there are a few students (six out of 25), with slightly lower cognitive abilities, that seemed to search for solution templates to imitate, focusing more on non-helpful information.

Earlier studies (Bergqvist et al., 2008; Liljekvist, 2014; Lithner, 2003) have shown that it is likely that tasks with solution templates will be solved by AR, while tasks without solution templates will mainly be solved by CMR. Several studies in the LICR-program also support the claim that when students are given an algorithm, they will use it (Jonsson et al., 2016; Lithner, 2017; Norqvist, 2018; Norqvist et al., 2019). This means that we will have a clear dividing line between the two types of tasks.

Aim and research questions

Previous studies have shown that CMR-practice is more efficient than AR-practice as measured by scores on a post-test (e.g. Jonsson et al., 2014; Norqvist, 2018). However, we do not know enough about the reasons for the difference in performance between students who practice on CMR-tasks and students who practice on AR-tasks.

The aim of this study is twofold, both to understand more about the students' choices of solution strategies during the practice and the test, and to gain insight into students' experiences from the practice and the test. Therefore, we will answer the following research questions:

RQ1: What types of solution strategies are used during practice and test and how do they relate to the two practice conditions (AR and CMR)?

RQ2: How do students experience AR and CMR tasks in the practice and test conditions?

Method

Participants

The study comprised a convenience choice of 10 pre-service teacher students at a university in Sweden, studying to become primary school teachers. A request to participate was given to a student group of 42 students. From this student group, 10 students volunteered to participate in the study, and were informed about their rights as participants in the study.

Design

The study utilizes a within-group design (i.e., all participants provide data for both conditions), which implies that eventual differences in results will be between conditions, not between participants. Hence, individual strengths or weaknesses among the informants will not affect the results since an important aspect of using a within-group design is that it compensates for individual competence. This design makes it possible to compare the results from this study with previous results in the LICR-program, despite eventual differences in the participants' mathematical level. We know from several studies (e.g. Bertilsson et al., 2017) that a within-group design works well enough for this type of study. The participants were met individually on two occasions, a practice session and a test session. During the practice session all ten participants solved 6 task-sets (see next section for a detailed description) with a suggested solution method (AR sub-tasks) and 6 task-sets without a solution method (CMR sub-tasks) (see figure 1). For five participants, task-sets 1–6 comprised AR sub-tasks and task-sets 7–12 comprised CMR sub-tasks. For the other half of the participants the task types were reversed (i.e., task-sets 1–6 were CMR sub-tasks and task-sets 7–12 AR sub-tasks). This was done to control for eventual differences in difficulty between task-sets.

A test session was conducted one week after the practice session. The main reason for delaying the test was that in previous studies in the LICR-program there was a one-week delay between practice and test. Such a delay is often the case in a school setting (i.e., the test takes place days or weeks after the individual lessons). Since we want to address previous results in the LICR-program, we wanted similar conditions in this study. During the test session, each participant worked on 24 sub-tasks (2 sub-tasks for each practiced task-set), all of them similar to the CMR-tasks they got during practice, but with different numerical values.

Tasks

The tasks mainly focused on algebraic thinking (e.g. recurring patterns) with different contexts. All practice sub-tasks provided basic description, an illustration, and a question. Additionally, the AR sub-tasks also provided a formula and a solved example (see figure 1). Since the AR sub-tasks provide a solution method, they can generally be completed faster (this was also evident from previous studies in the LICR-program, e.g. Jonsson et al., 2014). Therefore, each AR task-set included 5 sub-tasks, while CMR task-sets included 2 sub-tasks each. Sub-tasks within the same task-set were also designed with increasing difficulty, starting with answers that could be deduced from, or by slightly expanding the figure (for example, in the task-set presented in figure 1 the first sub-task asked for 6 squares in a row), while the following sub-tasks demanded calculation or a mathematical idea to be solved. Hence, the only thing that differed between sub-tasks within a task-set was the numerical value of the question (e.g. the number of squares in figure 1). Altogether, each participant worked on 42 sub-tasks during practice, comprising 30 AR sub-tasks and 12 CMR sub-tasks.

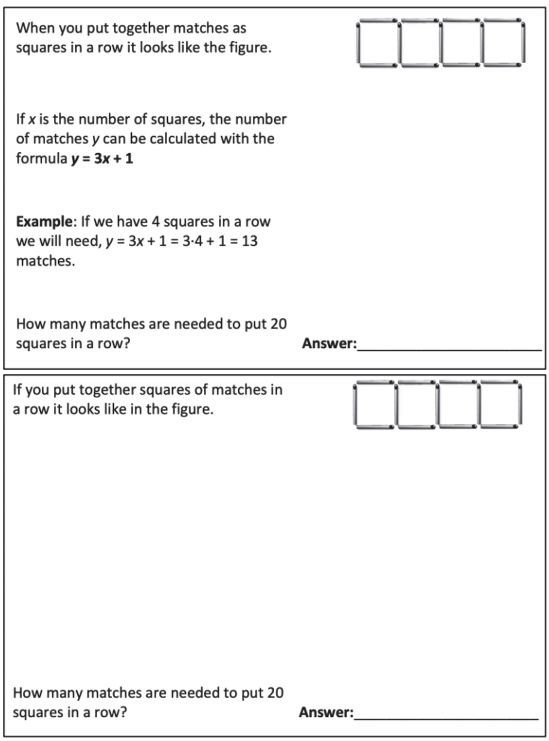

Figure 1. *Example of AR (above) and CMR (below) tasks.*
Note. For additional examples of tasks used in the LICR-program, see Norqvist et al. (2019).

Test tasks were identical to the CMR practice tasks, albeit with other numerical values. However, during the test the sub-tasks were presented with decreasing difficulty (i.e., the reverse order of CMR-practice). This design was made to identify two aspects. Firstly, to see if the participants could use the practiced solution method, or a mathematical idea, on the first sub-task where an eventual (re)construction was harder. Secondly, if they could (re)construct the solution method on the simpler second sub-task, if it was forgotten. The similarity between test sub-tasks and CMR-practice sub-tasks stems from the fact that tasks that occur on tests in school seldom have given solution methods (which also is a part of a CMR-design).

Data

Data comprised practice and test scores, papers with students' written solutions, video-recordings of practice and test, and an interview at the end of the recorded sessions. To try to eliminate an eventual effect of having an observer (one of the authors) in the room whilst solving tasks, we decided to observe and record half of the participants during practice and the other half during the test. In retrospect this might have been overly cautious since the participants did not seem to be bothered by the presence of a researcher and a video camera. Hence video-recordings and interviews did not include both sessions for all participants. To complement the lack of recordings we analyzed the written solutions from all students, both from the recorded and the non-recorded sessions.

During the recorded practice and test sessions, an observer used short comments and questions of the type ”what are you doing” to prompt the student to think-aloud. The sessions varied from 37 to 80 minutes in length, with an average of 55 minutes. At the end of each recorded session, the students were interviewed for about 10 minutes on their experiences. The interview questions consisted of both task specific questions and more general questions. A task specific question was when the observer had identified an interesting strategy choice or had failed to understand what the student was doing on a specific task: ”How did you think on the task with the matches?” or ”On task 3 you started with one idea, but then you changed your mind. Why?”. The more general questions (especially after the test session) considered the whole situation: ”What do you remember from the practice?” or ”Did you recognize the tasks?”. These general questions, or variants thereof, were asked of all recorded participants after the test session.

Method of analysis

The main analysis focuses on the choices of solution strategies. These strategies were identified using three types of data (i.e., written solutions, video recordings and interview answers). We searched both for previously identified strategies and other ways of solving the tasks. There were two previously identified strategies, found in earlier studies in the LICR-program (Lithner, 2017): i) understanding and completing the pattern, and ii) finding and analyzing the number series. For example, on the task presented in the framework-section above (figure 1), the first strategy would work like this:

– Realize that for each new square three new matches are needed

– Realize that you need do add an additional match to complete the pattern.

– This gives $3x + 1$

– 3 x 20 + 1 = **61**

The second strategy, to use a series of examples to find a number series. On the match-squares task this could work like this:

– 1 square = 4 matches

– 2 squares = 7 matches

– 3 squares = 10 matches

– 4 squares = 13 matches

- 4, 7, 10, 13, 16, 19, ... , 55, 58, **61**

Additional solution strategies were expected to be found in the analysis. We were open for many ways to solve the tasks, both correct and incorrect ways. Two such strategies were *guessing a formula* and *adapting the solution to the illustration*. By guessing a formula that solves a task it is possible to get a correct answer, but this is probably rare (none of the students who guessed a formula in this study got a correct answer). The strategy adapting the solution to the illustration means that a formula or relation that superficially matches the example given in the task is used for all situations (all answers created using this strategy were incorrect). For example, if 4 squares require 13 matches, then 20 squares will require five times as many: 5 x 13 = 65 matches.

The interview data was analyzed using thematic analysis (Braun & Clarke, 2012). In the first step we identified all relevant utterances regarding the research questions. That is, utterances about strategy choice, solution strategies, the training situation, the test situation, and the relation between training and test. There were also utterances about difficulty and the students' own performance that were relevant for the aim of the study. 51 utterances were identified and coded into 10 different codes, based on similarities related to solution strategies and student behavior. Most of these 51 utterances came from the five students interviewed after the test session. The codes were combined into four themes, see table 1, where three of them were connected to students' experiences and one regarded strategy choice. The themes will be presented more in detail in the results section.

Additionally, we also made a comparison of the differences between practice and test scores for each condition, as an indication of the efficiency of practice. With such a small sample, an inferential statistical analysis would be of small or no use and therefore we only present mean values for the difference between practice and test scores on a group level.

Table 1. *Examples of the thematic analysis*

Example of utterances	Codes	Themes	
If you cut the first there will be this many pieces. When you cut one more there will be this many extra.	Patterns	Solution strategies	Strategy choice
So, if there are 45 gray you will need 44 blacks.	Mathematical comments		
I remembered this one from last time.	Recognizing from training		
It was raised to and such stuff, I can't figure it out in my head. My brain doesn't work that way.	Can't solve this	Aspects of understanding	Student experiences
I noticed that it was easier to solve those I had solved before.	Easier when you have been thinking about it before		
Stupid that I can't remember the formula ... but I should take a time b and then take away something ...	Can't remember the solution	Memory	
I get so annoyed on that formula... you take that times that and that, I think it was like that, kind of, and then you took minus b times 2 or raised to 2 or something.	Remembering there is a formula		
Remembering stuff you have been taught takes much longer than something you have found out yourself.	Easy to remember what you have found out yourself		
Really, this was so much easier last time because that time we had a formula.	Easy with a formula	Formula	
I don't think formulas has helped me later, only in the moment	The formula helps locally		

Results

In this section we will present the results in relation to our research questions. First, we will present what strategies the students used on the tasks, both in practice and in test, along with practice and test scores, to see if solution strategies differ between the two practice conditions and in that case how. Secondly, we will present how the students commented on their experiences from the practice and test, to investigate how the two practice conditions might influence their strategy choices.

Strategy choices and their impact on scores on practice and test tasks

To answer the first research question, we categorized the students' strategy choices based on how they approached the tasks. All students made an initial strategy choice at the start of every task-set and clung to it throughout each sub-task within that task-set. Hence, we have chosen to present a total of 60 instances of strategy choice (6 task-sets x 10 students) for each condition, in total, 120 instances. Six solution strategies were observed during the practice and test sessions (see table 2).

Table 2. *Instances of observed solution strategies during practice and test*

Observed Strategies	Practice		Test	
	AR	CMR *	Solution strategy after AR practice	Solution strategy after CMR practice
Understand the pattern	-	27	39	32
Find a number series	-	1	5	7
Guess the formula	-	3	-	-
Adapt to illustration	-	5	-	1
Remember the formula	-	-	7	6
Use the given solution template	60	-	-	-
Not distinguishable **	-	22	9	14

Notes. A total of 60 possible instances in each column (6 task-sets x 10 students).
* Two task-sets were left unanswered
** Includes situations where only an answer was produced or other instances where the solution strategy could not be determined.

Within the AR practice task-sets, only one strategy was observed (i.e., utilizing the given formula to solve the task). When practicing with the CMR task-sets, four different strategies were observed. The most common by far was to try to understand the pattern and how this related to the question at hand (27 out of 60 instances). This included instances when students were either drawing new illustrations or discussing the pattern given in the sub-task. In nine instances the students used other strategies: finding a number series (1), guessing a formula (3), and adapting the solution to the illustration (5).

During the test the most common strategy, regardless of practice type, was again to understand the pattern (71 instances), followed by finding the number series (12 instances). There were also instances when students remembered, or tried to remember, the formula from practice (13 instances), and adapting the solution to the illustration occurred once.

In some situations (both in practice and test sessions) it was not possible to determine what strategy the students used. The reasons varied, but in most of these situations only an answer was given. In other situations, there were a few things scribbled on the paper, but no connection to the task could be identified. These were all put in the "not distinguishable" category. Where the sessions had been captured on video, some student comments or expressions were used to help us to classify their solution strategy. The interviews could provide indications of the sort "it is just to use the formula", or "half are white here [points to the illustration], if I take all tiles, we need 3774, then, technically, half of them should be red" or other expressions, indicating the student's strategy choices. These types of expressions were coded into the first theme, *solution strategies*, in the thematic analysis but used here to help determine students' strategy choice.

To get an indication of an eventual difference between practicing by AR and by CMR we also compared the difference between practice and test scores (see table 3). A mean difference between practice and test scores for the AR and CMR conditions shows that the tasks practiced by CMR yielded a 29 % higher score (from 44 to 57) on the test than during practice, while the performance on the tasks practiced by AR decreased by 35 % from practice to test (from 96 to 63). This could indicate that CMR-practice provides opportunities to practice lasting solution strategies which AR-practice does not. If only test scores for tasks where practice was successful are included, there is a clear difference between the two practice conditions in favor of CMR-practice (see table 3).

Table 3. *Mean solution frequency in percent in practice and test*

Condition	Practice	Test – all tasks	Test – successful practice
AR	96 (8)	63 (22)	64 (23)
CMR	44 (23)	57 (28)	83 (23)

Note. Numbers in brackets indicate standard deviation.

In summary, these results show that a given solution method, as in the AR-practice condition, will promote a utilization of the formula, with less regard given to the underlying mathematical principles. The CMR-practice condition seems to require the students to mainly use strategies that include some regard to the underlying mathematical principles (see table 2). However, there is no evident difference in strategy choice during the test between the two conditions. The results also indicate CMR-practice performance is somewhat stable between practice and test while the AR-condition shows a decline from practice to test, and that a successful practice is especially important for a good test performance in the CMR-condition.

Students' experiences of the AR and CMR tasks

To answer the second research question, a thematic analysis of student comments and answers to interview questions was conducted, which resulted in four themes: *Solution strategies* (presented in the section above), *aspects of understanding*, *memory* and *formula*. The themes were based on comments from students in relation to both practice conditions (i.e., reasoning type) and was not correlated to a particular solution strategy. For example, a student talking about the formula could either have been talking about a given formula (as in the AR-condition), or about a constructed formula (as in the CMR-condition). Student comments in the three latter themes were used to get information of how the students experienced the difference between the two conditions.

Student comments under the following three themes were stated during or after the test and were used to get a clearer picture of students' experiences of the two practice conditions. The notation by the quotes (AR or CMR) denote the practice condition for the specific task(s) that the students were referring to, and the number is an identification of the student (1–10).

Aspects of understanding

The students implicitly expressed different aspects of understanding, from when they did not grasp the mathematics or did not think that they would, to instances when they expressed that it was easier to work with the tasks when they had tried to construct solution methods during the practice (which also was a common comment after the test session). It's easier to work with the tasks that you have already been thinking about, and if you succeeded last time, you would remember how you dealt with the task.

> I can't do this one. I know that immediately, because it was powers and stuff, I can't figure that out in my head. My brain doesn't work that way. (AR, 7)

> I noticed that it was much easier to solve the tasks I had solved before. (CMR, 8)

The students also commented that if you had figured something out it is easier to recall what you have figured out than things you have read or been told.

> Remembering things that I have learned takes a lot longer than something I have figured out myself. (CMR, 4)

Some comments also showed that the students were aware of solution methods that were more efficient but could not recreate them.

> It feels like I should find an easier way, but I can't find an easier way right now. (AR, 4)

Memory

There were also comments directly connected to memory traces of the practice tasks. Some students expressed that they recognized the tasks from the CMR-practice, or that they had worked on a task for a long time. They also expressed that they remembered that there were formulas in AR-practice, but in most cases, they couldn't recall what they looked like.

> Oh no, I remember this one too. I worked on that one really long. (CMR, 8)

> I remembered on some that it should be squared, but I don't remember what parts. (AR, 7)

The students also commented that if you had figured something out it is easier to remember what you have figured out than things you have read or been told.

> Remembering things that I have learned takes a lot longer than something I have figured out myself. (CMR, 4)

Formulas

Parallel to the students' comments about memory, there were similar statements that focused more on the formulas themselves. They commented that a formula is very good for solving the tasks, but when you don't get the formula in the task the situation becomes difficult.

> It was so easy when you had the formula, but if I don't have it, that's tough! (AR, 3)

One student formulated very clearly what other students also hinted at, that formulas only help you in the practice session, not in the long run.

> I don't think that formulas have helped me later, only in the moment. (AR, 2)

One student also commented on the difference between constructing his or her own solution during the test compared to using a given formula during practice.

> It's maybe easier since I have been brooding more over them [the test tasks] than I did last time when there was a formula, so even if I made three of each then ... maybe that's it. At least if you're like me and don't get the formula but only use it, maybe. (AR, 7)

Summing up the students' comments gives two insights. The first is that the students express that the AR-practice hardly helps in the choice of solution strategy when taking the test since they used the formulas without reflection during practice, and to a large extent they cannot remember them. The second is that the students indicate that CMR-practice makes you a bit prepared since you have tried to solve similar tasks before and had to come up with solution strategies by yourself, which seems to be easier to recall.

Overall, the results from the two analyses indicate that CMR-practice, although more difficult, provides important experience, easier retrieval, and understanding, which AR-practice does not seem to do.

Discussion

Regarding solution strategies, our results show that the formula is used extensively during AR-practice, and is, as the students say, not really reflected upon. During CMR-practice most students used a strategy aiming at understanding the pattern, which also was the most common strategy during the test, regardless of practice condition. Hence, the results suggest that task design will influence students' choice of solution strategies in a way that given solution methods are used but not always understood. It also appears likely that CMR-practice prepare students for solution strategies needed during a test, which AR-practice does not do unless the students remember the formula. Students' experiences also suggest that the given solution methods during AR-practice are harder to remember than the solution methods they have constructed themselves, mostly since during AR-practice focus is on getting a correct answer and not on understanding why the formula works. The experience of having to construct a solution during CMR-practice was also expressed as helpful when put in a test situation.

During the practice on AR-tasks, the students made almost no mistakes, they used the given formulas and found the answer easily. Some also reflected that practice was so much easier since there was a formula present, showing that they were aware that such practice tasks are very easy to solve. An interesting situation was when one student looked on a formula (on one of the first sub-tasks) for some time and then asked, ”Do I need to understand the formula?”. When the researcher said that she was supposed to solve the task, she just used the formula and got a correct answer. This comes as no surprise, it is in line with how most tasks in mathematics textbooks function all over the world where 79 % of all tasks can be solved using a template or worked example (Jäder et al., 2020). This leads to a high competence at using templates or given formulas.

During CMR-practice, more errors are made, but the students indicate that the tasks are remembered better. One student said, ”Oh no, this is the one I didn't succeed ...”. That the students might remember the tasks better is in line with previous studies (e.g. Jonsson et al., 2014). One possible reason could be related to the idea of ”productive struggle” (Hiebert & Grouws, 2007). The students work harder during CMR-practice and the additional effort will make the retrieval of the practiced material easier (e.g. Pyc & Rawson, 2009; Wiklund-Hörnqvist et al., 2014). The fact that we saw a slight increase in test score compared to the practice score for the tasks that were practiced by CMR, also support the notion that retrieval might be easier after an initial struggle.

There were no major differences between the solution strategies used in the test for the two conditions. The strategy choice in the test situation after both AR-practice and CMR-practice was mainly *understanding the pattern*. However, when practicing by CMR, students in general use the same or similar solution strategies as they later used in the test situation, something that is not true when practicing by AR. It is apparent in our results that AR-practice, when a solution method is available, will consist of applying that method with little regard to other solution strategies. We also have indications (see table 3) that a successful CMR-practice more often will yield a correct solution on the corresponding test task, which again can be related to the benefits of productive struggle (Hiebert & Grouws, 2007) during the practice session. Some students also indicated this benefit in their comments on the CMR-tasks. A successful practice is not as important during AR-practice, in part because AR-practice has fewer obstacles, hence less struggle, and does not require the student to have a deeper understanding of why the formula works (Brousseau, 1997), and in part because they are tested on something that they have not been practicing (i.e., finding a solution method).

Conclusions

Our main result is that there is no real difference in how students address a task in the test depending on the type of practice. This was not completely expected since the difference in practice condition could have made a difference in strategy choice during the test. However, when students are taking tests on areas where they have practiced using solution templates, they are in most situations tested on something else than what they have practiced. If they are allowed to practice in the same way as they are tested, the students remember the tasks and can utilize similar strategies as in practice. The students must of course also be aware of, and accept, that during the practice you do not solve all tasks correctly without help. Formulas and algorithms are important tools in mathematics but are not primarily constructed to provide understanding (Brousseau, 1997). It is therefore of importance that the teacher's help does not include presenting a complete solution strategy, since that would increase the risk for students to utilize the strategy without reflection.

References

Andrews, P., Sunde, P. B., Nosrati, M., Petersson, J., Rosenqvist, E. et al. (2021). Computational estimation and mathematics education: a narrative literature review. *Journal of Mathematics Education*, 14 (1), 6–27.

Bergqvist, E. (2007). Types of reasoning required in university exams in mathematics. *Journal of Mathematical Behavior*, 26 (4), 348–370. doi: 10.1016/j.jmathb.2007.11.001

Bergqvist, T., Lithner, J. & Sumpter, L. (2008). Upper secondary students' task reasoning. *International Journal of Mathematical Education in Science and Technology*, 39 (1), 1–12. doi: 10.1080/00207390701464675

Bergqvist, T. & Lithner, J. (2012): Mathematical reasoning in teachers' presentations. *Journal of Mathematical Behavior*, 31 (2), 252–269. doi: 10.1016/j.jmathb.2011.12.002

Bertilsson, F., Wiklund-Hörnqvist, C., Stenlund, T. & Jonsson, B. (2017). The testing effect and its relation to working memory capacity and personality characteristics. *Journal of Cognitive Education and Psychology*, 16, 241–259. doi: .10.1891/1945-8959.16.3.241

Bjork, E. L. & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough & J. R. Pomerantz (Eds.), *Psychology and the real world: essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Braun, V. & Clarke, V. (2012). Thematic analysis. In H. E. Cooper, P. Camic, D. L. Long, A. T. Panter, D. E. Rindskopf & K. J. Sher (Eds.), *APA handbook of research methods in psychology. Vol 2: Research designs: quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Kluwer Academic.

Fan, L. & Bokhove, C. (2014). Rethinking the role of algorithms in school mathematics: a conceptual model with focus on cognitive development. *ZDM*, 46, 481– 492. doi: 10.1007/s11858-014-0590-2

Fyfe, E. R. & Rittle-Johnson, B. (2017). Mathematics practice without feedback: a desirable difficulty in a classroom setting. *Instructional Science*, 45 (2), 177–194. doi: 10.1007/s11251-016-9401-1

Granberg, C. & Olsson, J. (2015). ICT-supported problem solving and collaborative creative reasoning: exploring linear functions using dynamic mathematics software. *Journal of Mathematical Behavior*, 37, 48–62. doi: 10.1016/j.jmathb.2014.11.001

Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to Principles and standards for school mathematics* (pp. 5–23). NCTM.

Hiebert, J. & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age.

Jonsson, B., Kulaksiz, Y. C. & Lithner, J. (2016). Creative and algorithmic mathematical reasoning: effects of transfer-appropriate processing and effortful struggle. *International Journal of Mathematical Education in Science and Technology*, 47 (8), 1206–1225. doi: 10.1080/0020739x.2016.1192232

Jonsson, B., Norqvist, M., Liljekvist, Y. & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *The Journal of Mathematical Behavior*, 36, 20–32. doi: 10.1016/j.jmathb.2014.08.003

Jäder, J., Lithner, J. & Sidenvall, J. (2020). Mathematical problem solving in textbooks from twelve countries. *International Journal of Mathematical Education in Science and Technology*, 51 (7), 1120–1136. doi: 10.1080/0020739X.2019.1656826

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science: An International Journal of the Learning Sciences*, 38 (6), 523–550. doi: 10.1007/s11251-009-9093-x

Karlsson Wirebring, L., Lithner, J., Jonsson, B., Liljekvist, Y., Norqvist, M. & Nyberg, L. (2015). Learning mathematics without a suggested solution method: durable effects on performance and brain activity. *Trends in Neuroscience and Education*, 4 (1-2), 6–14. doi: 10.1010/j.tine.2015.03.002

Kilpatrick, J., Swafford, J. & Findell, B. (2001). *Adding it up: helping children learn mathematics*. National Academy Press.

Liljekvist, Y. (2014). *Lärande i matematik: om resonemang och matematikuppgifters egenskaper* [Learning mathematics: on the features of reasoning and mathematical tasks] (Ph. D. thesis). Karlstad University.
Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational Studies in Mathematics*, 52 (1), 29–55. doi: 10.1023/A:1023683716659
Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, 67 (3), 255–276. doi: 10.1007/s10649-007-9104-2
Lithner, J. (2017). Principles for designing mathematical tasks that enhance imitative and creative reasoning. *ZDM*, 49, 937–949. doi: 10.1007/s11858-017-0867-3
NCTM (1989). *Curriculum and evaluation standards for school mathematics*. National Council of Teachers of Mathematics.
NCTM (2011). *Common core state standards for mathematics*. National Council of Teachers of Mathematics.
Newton, D. & Newton, L. (2007). Could elementary mathematics textbooks help give attention to reasons in the classroom? *Educational Studies in Mathematics*, 64 (1), 69–84. doi: 10.1007/s10649-005-9015-z
Niss, M. & Jensen, T. H. (2002). *Kompetencer og matematiklaering* [Competencies and mathematical learning]. Undervisningsministeriet.
Niss, M. (2003). Mathematical competencies and the learning of mathematics: the Danish KOM project. In A. Gagtsis & Papastavridis (Eds.), *Third Mediterranean conference on mathematical education, 3–5 january 2003, Athens, Greece* (pp. 115–124). The Hellenic mathematical society.
Niss, M. (2007). Reactions on the state and trends in research on mathematics teaching and learning. From here to Utopia. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1293–1312). Information Age.
Norqvist, M. (2018). The effect of explanations on mathematical reasoning tasks. *International Journal of Mathematical Education in Science and Technology*, 49 (1), 15–30. doi: 10.1080/0020739X.2017.1340679
Norqvist, M., Jonsson, B., Lithner, J., Qwillbard, T. & Holm, L. (2019). Investigating algorithmic and creative reasoning strategies by eye tracking. *The Journal of Mathematical Behavior*, 55, 100701. doi: 10.1016/j.jmathb.2019.03.008
Norqvist, M., Jonsson, B. & Lithner, J. (2019). Eye-tracking data and mathematical tasks with focus on mathematical reasoning. *Data in Brief*, 25, 104216. doi: 10.1016/j.dib.2019.104216
Olsson, J. (2017). *GeoGebra, enhancing creative mathematical reasoning* (Ph. D. thesis). Umeå University.

Olsson, J. & Granberg, C. (2019). Dynamic software, task solving with or without guidelines, and learning outcomes. *Technology, Knowledge and Learning*, 24, 419–436. doi: 10.1007/s10758-018-9352-5

Pólya, G. (1945). *How to solve it*. Princeton University Press.

Pyc, M. A. & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, 60 (4), 437–447. doi: 10.1016/j.jml.2009.01.004

Schoenfeld, A. (1985). *Mathematical problem solving*. Academic Press.

Shield, M. & Dole, S. (2013). Assessing the potential of mathematics textbooks to promote deep learning. *Educational Studies in Mathematics*, 82 (2), 183–199. doi: 10.1007/s10649-012-9415-9

Sidenvall, J., Lithner, J. & Jäder, J. (2015). Students' reasoning in mathematics textbook task-solving. *International Journal of Mathematical Education in Science and Technology*, 46 (4), 533–552. doi: 10.1080/0020739X.2014.992986

Skolverket. (2011a). *Curriculum for the compulsary school, preschool class and school-age educare 2011*. Skolverket.

Skolverket. (2011b). *Curriculum for the upper secondary school 2011*. Skolverket.

Wiklund-Hornqvist, C., Jonsson, B. & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, 55 (1), 10–16. doi: 10.1111/sjop.12093

Tomas Bergqvist

Tomas Bergqvist is an associate professor in mathematics education at Umeå University in Sweden, and Chair of Umeå Mathematics Education Reseach Centre (UMERC). His research mainly concerns mathematical reasoning and argumentation.

tomas.bergqvist@umu.se

Mathias Norqvist

Mathias Norqvist is a Senior Lecturer in Mathematics Education at Umeå University. His main research interest concerns students' mathematical reasoning and how task design can influence different types of reasoning.

mathias.norqvist@umu.se