This is the published version of a paper presented at 42nd Annual Meeting of the International Group for the Psychology of Mathematics Education, Umeå, Sweden, July 3–8, 2018.

Citation for the original published paper:


N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-146866
COGNITIVE ABILITIES AND MATHEMATICAL REASONING IN PRACTICE AND TEST SITUATIONS

Mathias Norqvist

Department of Science and Mathematics Education, Umeå University, Sweden
Umeå Mathematics Education Research Centre, Umeå University, Sweden

Research studies have shown that to develop conceptual understanding of mathematics, practice needs to that focus this skill. In this study, the aim is to examine how different practice tasks, which promotes either imitative or creative mathematical reasoning, can influence which variables (i.e., cognitive abilities, mathematics grade, and gender) that are important for task completion. Two earlier studies show that cognitive abilities are more important in the test situation when students have practiced with imitative tasks. The result from this study indicate that although cognitive abilities are important when practicing with creative tasks, the influence of cognition is only implicit during the test. Since students often practice imitatively with given solution methods, this study suggests that a substantial part of what we test in school could be cognitive abilities rather than mathematics.

INTRODUCTION

Many studies have shown the inefficiency of rote-learning that transpires without understanding (e.g., Hiebert, 2003). Hiebert and Grouws (2007) argue that students need to struggle with important mathematical concepts or properties in order to get a deeper understanding of mathematics. This positive productive struggle could occur if the task involves some desirable difficulties that forces the student to regard the mathematical properties of the task. Bjork and Bjork (2011) argue that desirable difficulties are important whatever you are trying to learn, and can, while not as efficient at first, be more efficient in the long run (Fyfe & Rittle-Johnson, 2016). It is however important that the imposed difficulty should be surmountable and relate to the subject or skill you are about to learn (Bjork & Bjork, 2011). Otherwise it would be enough to just turn out the lights to create difficulty in the mathematics classroom. But since this has nothing to do with mathematics it would be an obstruction rather than a desirable difficulty. Desirable difficulties can however induce some amount of failure during task solving, but this might not necessarily be a bad thing. In a number of studies Kapur explored productive failure as an instructional design, and found that it can be effective for developing conceptual understanding of mathematics (e.g., Kapur, 2010; Kapur, 2015).

Studies of mathematics textbooks have shown that most textbook tasks lack the difficulties and struggle that Bjork and Bjork (2011) and Hiebert and Grouws (2007) are arguing for. Most textbook tasks are procedural and can mostly be solved by provided

solution methods or by looking at worked examples (e.g., Jäder, Lithner, & Sidenvall, 2015; Newton & Newton, 2007; Shield & Dole, 2013). Jäder et al. (2015) concluded in their cross-national study of textbooks from twelve countries that only 9% of the tasks required more extensive conceptual knowledge and justification, while 79% of the tasks could be solved completely by imitating or following given instructions. The textbooks also contain more tasks than what is reasonable for a student to solve during a course, and the more demanding tasks are commonly located to the last part of each section. This implies that many students have to select which tasks to solve. Since they tend to choose the basic tasks first they will often not reach the more demanding tasks at the end of the section (Sidenvall, Lithner, & Jäder, 2015). Bergqvist and Lithner (2012) observed that teacher presentations also are dominated by procedure and that most presentations consider how tasks should be solved rather than the concepts or properties behind the procedures. Conceptual instruction has proven to be more beneficial than procedural instruction when trying to promote a more thorough understanding of procedures and concepts (Rittle-Johnson, Fyfe, & Loehr 2016). Hence, more conceptual tasks seem to be needed in both teacher instruction and textbooks.

**FRAMEWORK**

Lithner (2008) proposed a research framework for mathematical reasoning where he concludes that there are two main types of reasoning that can occur while solving mathematical tasks, imitative and creative. Imitative algorithmic reasoning (AR) occurs where a solution method is already known or presented in close proximity to the task, so that the student can imitate or recall a solution method. An understanding of the concepts or mathematical properties is not imperative for this type of reasoning to solve the task at hand. The second type, creative mathematically founded reasoning (CMR), concerns student reasoning where no solution method is available. Not giving a solution method in advance force the students to consider mathematical properties when constructing a valid solution method. There is of course an effort involved in this process which is not necessary when solving a task imitatively and this effort, or struggle if you will, is close to what Hiebert and Grouws (2007) argued for. However, as most textbooks provide procedural solution methods, imitative reasoning is normal practice for most students in school.

From a theoretical perspective, Brousseau (1997) states that it is imperative for students to take responsibility for their own solution process. Brousseau argues that for this to happen, the teacher has to hand over responsibility to the students after constructing a well-designed task where the students can, with some work and arguments, construct the solution by themselves. When students work with tasks where they can imitate a given solution method, this argumentation and construction will not take place. Brousseau denotes the activity where the students work alone to solve the task an adidactical situation. In a study where students practice solution methods by either imitative or creative reasoning, Jonsson, Norqvist, Liljekvist, and Lithner (2014) utilized this adidactical situation to study the effectiveness of the different practice tasks (AR or CMR). The study showed that practicing with tasks that promote CMR is more
effective with regard to test scores than practicing with AR-tasks. Two additional studies have confirmed this result (Norqvist, 2018; Wirebring et al., 2015).

Studies have also shown that students’ cognitive abilities is important for mathematical achievement (e.g., Primi, Ferrão, & Almeida, 2010; Swanson & Alloway, 2012). For example, working memory (i.e., the ability to store information whilst processing other information) is highly correlated to mathematical achievement (Swanson & Alloway, 2012) and has been shown to be predictive of mathematics learning in the early school years (Passolunghi, Vercelloni & Schadee, 2007). Another cognitive ability that is closely related to mathematics is fluid reasoning, which is related to faster mathematical learning (Primi et al., 2010).

Jonsson et al. (2014) also showed that cognitive abilities (i.e., working memory and fluid reasoning) are important for test scores, especially for students that practice by AR. To rule out that this was not contributed to similarities between CMR-practice tasks and the test tasks that were used, a follow-up study was made. Here transfer-appropriate processing was contrasted to productive struggle to see if the higher test performance could be attributed to similarities in task design (Jonsson, Kulaksiz, & Lithner, 2016). The results showed that transfer appropriate processing accounted for only a minor part of the efficiency of the CMR-group. However, since Jonsson et al. (2014) and Norqvist (2018) did focus on test scores and the efficiency of AR and CMR, neither of the studies gave much notice to the practice scores and which variables that were important for the two practice conditions. A study of this could help us understand why CMR seems to be more efficient and at the same time give us a clue to why AR appears to be efficient during practice but not when it comes to the post-test.

AIM AND RESEARCH QUESTIONS

In the previous studies (Jonsson et al., 2014; Norqvist, 2018), there are strong indications that cognition play an important role for solving test tasks, especially for students that have practiced by AR. Jonsson et al. (2016) also showed that the difference in test-scores between the two practice groups (AR and CMR) was not attributed to similarities between practice- and test-task, so called transfer appropriate processing. However, the practice session could also help us understand why CMR-practice has proven to be more efficient as measured by test scores. The aim of this study is therefore to examine which of the measured variables (i.e., mathematics grade, gender, and cognitive abilities) influence students task solving during practice, depending on what type of reasoning the students utilize during practice. It is also interesting to examine if there is any difference between practice and test, regarding taxation on cognitive abilities, since this could be of importance for teaching.

1. How will the combined sample affect earlier results of the importance of cognitive abilities for the to test-scores?

2. How does the practice condition, AR or CMR, affect which variables (i.e., mathematics grade, gender, and cognitive abilities) that are most influential on students’ completion of the given practice tasks?
METHOD

Participants
The present study utilize data gathered in two earlier studies where the efficiency of different mathematics tasks was in focus. The samples in (Jonsson et al., 2014) and (Norqvist, 2018) comprised a total of 252 students in the natural science program in Swedish upper secondary school (16-17 y.o.). 44 students were excluded due to attrition. Before the analysis the practice and test data was scanned for outliers, and to compensate for eventual ceiling or floor effects participants that had the maximum score, or that scored lower than 10%, during practice where removed from the sample. This control excluded 13 participants from the sample. Also, in Norqvist (2018) 38 students were practicing with a third task type, and these students were also excluded from this study. Finally, 157 participants remained in the sample. In both studies, the participating students were divided into matched practice groups based on cognitive abilities (i.e., working memory and fluid reasoning), mathematics grade, and gender.

When squares are put in a row it looks like the figure to the right. 13 matches are needed for four squares.

If $x$ is the number of squares, then the number of matches $y$ can be calculated by the function

$$y = 3x + 1$$

Example: If 4 squares are put in a row, then

$y = 3x + 1 = 3 \cdot 4 + 1 = 13$ matches are needed.

How many matches are needed to get 20 squares in a row?

Figure 1: Example of an AR practice-task. The text written in italics is absent in the corresponding CMR-task.

Practice and test
The data collection was partly designed to mimic a common situation in school, where students often practice by solving textbook tasks of an AR-type, and often meet more complex tasks in a test. The students practiced 14 solution methods (i.e., formulas) by either AR- or CMR-tasks depending on the group (see example in Figure 1). During practice the teacher did not intervene and the students were instructed not to talk to each other. Since each AR-task is faster to complete than the corresponding CMR-task, the AR-group did more tasks for each solution method to get a comparable practice time. A computer software recorded solution frequencies and the time the students spent on each task.

One week later the students took a test where each solution method was tested with three tasks. The first asked for the formula used during practice while the second and
third task asked for a numerical answer. The first two test-tasks were limited to 30 seconds each so that no (re)construction could take place, while the last task was limited to 5 minutes. In this study, cognitive data, mathematics grade, gender, practice-scores (i.e., the proportion of correct answers to the practice-tasks), and test-scores (i.e., the proportion of correct answers to the test-tasks) from the mentioned studies are analyzed to find answers to the research questions.

**Analysis method**

Results from the cognitive tests (i.e., Operation span that tests working memory and Raven’s progressive matrices which is a test for fluid reasoning) were standardized and used to calculate a cognitive index. The practice-score was transformed to compensate for skewness (-3.66) and kurtosis (18.3) of the practice-score for the AR-group. This transformed practice-score was later used in the following analyses.

Two separate analyses were conducted on each of the practice-groups. First, a regression analysis, with test-score as the dependent variable and practice-score, CPI, gender, and mathematics grade as independent variables, was performed. This was done to control that the results presented in Jonsson et al. (2014) and Norqvist (2018) were valid for the combined sample as well (i.e., that cognitive abilities influenced the test-scores to a higher extent in the AR-group than in the CMR-group). Secondly, another regression analysis was performed, with practice-score as the dependent variable, to see to what degree the independent variables (i.e., cognitive index, mathematics grade, and gender) did influence students’ completion of the given practice-tasks. All statistical analyses were made in SPSS 24.

**RESULT**

<table>
<thead>
<tr>
<th>Variables</th>
<th>AR</th>
<th></th>
<th></th>
<th>CMR</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>β</td>
<td>B</td>
<td>SE B</td>
<td>β</td>
</tr>
<tr>
<td>Cognitive index</td>
<td>.504</td>
<td>.359</td>
<td>.359**</td>
<td>.117</td>
<td>.099</td>
<td>.104</td>
</tr>
<tr>
<td>Mathematics grade</td>
<td>.115</td>
<td>.308</td>
<td>.308**</td>
<td>.021</td>
<td>.027</td>
<td>.074</td>
</tr>
<tr>
<td>Practice score</td>
<td>.411</td>
<td>.175</td>
<td>.175**</td>
<td>.836</td>
<td>.127</td>
<td>.665*</td>
</tr>
<tr>
<td>Gender</td>
<td>.067</td>
<td>.046</td>
<td>.046</td>
<td>.048</td>
<td>.143</td>
<td>.027</td>
</tr>
<tr>
<td>$F$ total</td>
<td>8.057*</td>
<td></td>
<td></td>
<td>22.681*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.284</td>
<td></td>
<td></td>
<td>.543</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .001, **p < .01.

Table 1: Regression Analysis Summary for Variables Predicting Test Score.

The first regression analysis showed that the previous results, regarding the difference in relation between cognitive abilities and test-scores for the different practice groups, were valid for the combined sample. Cognitive index and mathematics grade were highly predictive of test-score in the AR-group while practice-score did predict the test-score in the CMR-group (see Table 1).
The second regression analysis showed that none of the included variables were predictive of the practice-score for the AR-group, while the CMR-practice score is highly dependent on cognitive index and mathematics grade (see Table 2).

<table>
<thead>
<tr>
<th>Variables</th>
<th>AR</th>
<th></th>
<th></th>
<th>CMR</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>β</td>
<td>B</td>
<td>SE B</td>
<td>β</td>
</tr>
<tr>
<td>Cognitive index</td>
<td>.201</td>
<td>.119</td>
<td>.208</td>
<td>.301</td>
<td>.086</td>
<td>.366*</td>
</tr>
<tr>
<td>Mathematics grade</td>
<td>.043</td>
<td>.032</td>
<td>.167</td>
<td>.107</td>
<td>.022</td>
<td>.468*</td>
</tr>
<tr>
<td>Gender</td>
<td>-.022</td>
<td>.188</td>
<td>-.014</td>
<td>.044</td>
<td>.134</td>
<td>.031</td>
</tr>
<tr>
<td>F total</td>
<td>2.432</td>
<td></td>
<td></td>
<td>14.555*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.057</td>
<td></td>
<td></td>
<td>.358</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .001.

Table 2: Regression Analysis Summary for Variables Predicting Practice Score.

**DISCUSSION AND CONCLUSION**

Introducing CMR-practice as an alternative has been shown to be more efficient than the procedural AR-tasks that constitutes the main part of textbook tasks (Jonsson et al., 2014; Norqvist, 2018). The results from the regression analyses confirm, as indicated in (Jonsson et al., 2014), that students that practice by AR are more dependent on cognitive abilities during the test situation than students practicing by CMR. This does not seem to be attributed to transfer-appropriate processing but rather to the productive struggle that the CMR-participants meet during practice (Jonsson et al., 2016). The results from the regression analysis on practice-scores show that CMR-practice is more taxing on cognitive abilities than AR-practice. This confirms that CMR-tasks are creating some struggle for the students and, as Hiebert and Grouws (2007) argued, this would then yield higher test scores. The desirable difficulties that CMR-tasks provides will not only help students to focus the important mathematical properties, but could actually force them to take these properties into account. AR-practice lacks the desirable difficulty that CMR-practice provides and the difficulty will therefore emerge during the test instead. Hence, the higher strain on cognitive abilities.

The importance of the practice-score for later test-performance in the CMR-group would indicate that successful CMR-practice leads to a deeper processing of the learning material, which in turn leads to easier retrieval during the test. The importance of a good practice-score for the CMR-students could of course be problematic for students with lower cognitive abilities or lesser mathematical skills, but this is where the teacher comes in. If the teacher (or maybe the author of the textbook) has designed the tasks well, so that the task will help students to consider the important mathematical properties, the teacher can support student learning by asking questions aimed at these properties. The difficulty could hereby be reduced but not removed completely, and it would still concern the important mathematics.
During AR-practice there are less difficulties for the students since they can use the given solution method to solve the task. Also, if a student would have difficulties with an AR-task, the focus would most likely be on how to use the given information (e.g., arithmetic difficulties when calculating the answer or how to exchange a variable with a number), and not mainly on the intrinsic mathematical properties in the task.

If the ability to solve novel problems or solve tasks that require transfer of knowledge is something that students are supposed to show during a test, they should have had the opportunity to practice these skills during lessons. Practicing with mainly AR-tasks gives an impression that tasks are standardized and short where the solution method is known (e.g., all tasks in the section can be solved by setting up and solving a linear equation). This can be preferable if the aim is to develop computational fluency but not if the aim is to develop conceptual understanding (Hiebert & Grouws, 2007). Students that only meet computational difficulties and never have to give any thought on why a solution method works will not be prepared to find new methods or solve novel tasks, neither during the upcoming test nor in their future working life. It would therefore be important to have textbooks and teacher presentations that include more tasks that compel students to consider mathematical properties to overcome the desirable difficulties of the task. Otherwise, there is a risk that the upcoming test will measure cognitive abilities rather than mathematical knowledge.

References


Norqvist


