university of groningen

MSC APPLIED COGNITIVE NEUROSCIENCE

MASTER THESIS

Towards reducing mathematical anxiety through adaptive learning

---

Author:

Stefania Denisa Iancu

Supervisors:

prof. dr . D.H. (Hedderik) van Rijn

Thomas Wilschut

October 1, 2022

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Abstract

Adaptive learning algorithms have gained popularity in educational contexts to effectively improve declarative fact learning. Such systems can improve learning by determining individual schedules of facts repetition, based on individual parameters, such as reaction time and accuracy. While studying declarative facts with an adaptive algorithm has been repeatedly proven to outperform less adaptive systems when using these performance parameters, it has never been examined when learning procedural knowledge, such as mathematical multiplications. Here, we are particularly interested in individuals with mathematical anxiety, their association with lower performance on mathematical tasks, and whether using adaptive methods might minimize their perceived mathematical anxiety. Our results indicate that differences between the two algorithms are small, with a slight advantage in accuracy and reaction time on test with the non-adaptive one. However, the adaptive algorithm shows a slight improvement in reaction time on learning. Additionally, our results point out that individuals with higher scores on mathematical anxiety tend to forget slightly faster when learning new facts. Contrary to previous results, we did not observe any association between performance parameters and mathematical anxiety. The current results help us better understand the differences in memory retention between declarative and procedural facts, and possible links between poor performance in math-related tasks and anxiety.

Keywords: adaptive learning, mathematical anxiety, procedural knowledge, memory

Towards reducing mathematical anxiety through adaptive learning

Mathematical knowledge is fundamental in many activities, ranging from high-technology jobs requiring data analysis, accounting and statistics, to everyday activities such as budgeting for our expenses. Lacking mathematical skills might affect our career choices, making us prone to choose college majors unrelated to mathematics (Ma and Johnson, 2008).

Studies looking into the cause of decreased mathematical skills pointed towards mathematical anxiety as an essential factor (Barroso et al., 2021). Meta-analyses have shown that mathematical anxiety, defined as fear and worry about our mathematical performance, has a moderate negative association with mathematical achievement (Barroso et al., 2021; Hembree, 1990; Ma, 1999). Studies show that students with increased anxiety about mathematics tend to avoid careers that imply mastering mathematical concepts or STEM (science, technology, engineering and math) related fields (Ahmed, 2018; Ashcraft, 2002). Some studies point out that females report higher levels than men (Devine et al., 2012), whereas others indicate that the difference is similar for both genders (Barroso et al., 2021). Therefore, it is important to acknowledge the possible link between mathematical anxiety and mathematical skills and find solutions for preventing or reducing it.

An FMRI study (Supekar et al., 2015) investigating whether one-to-one math tutoring programs can reduce math anxiety in third-grade children, shows that negative emotional response to mathematics is decreased by such intervention, pointing out the importance of math-tutoring programs in decreasing mathematical anxiety. As an alternative to tutoring, learning with computerized methods is increasing in popularity, due to studies indicating their efficacy in learning declarative facts (Van Rijn et al., 2009). By taking into account the learners' individual differences, such algorithms can optimize learning facts compared to non-adaptive algorithms (Sense et al., 2016). The existing literature points out that the learner's accuracy and reaction time are suitable performance parameters to approximate how well one can memorize an item (Anderson, 2009).

Whereas some previous studies have focused their attention on adaptive learning methods to improve educational achievement (Johnson et al., 2009; Liu et al., 2017), we want to extend the ongoing topic by comparing memory retention differences between an adaptive and non-adaptive algorithm. Therefore, an interesting topic of research is whether an adaptive algorithm could efficiently improve performance parameters on memorizing procedural knowledge while solving mathematical multiplications. Additionally, given the association between mathematical anxiety and poor mathematical skills, we were interested in measuring whether an adaptive algorithm is perceived to decrease the anxiety related to mathematics by improving one's performance on the task. Following, we present a review of the literature on such algorithms and possible mechanisms that could explain the link between mathematical anxiety and mathematical performance.

Literature review

Mathematical anxiety

Although the mechanism through which mathematical anxiety impairs mathematical performance is yet unknown, several theories point out that working memory might play a role (Ashcraft and Moore, 2009). In this view, the term *affective drop* indicates a drop in one's mathematical performance due to mathematical anxiety rather than capabilities or competencies (Ashcraft and Moore, 2009). Initially, limited working memory capacity was attributed to decreased mathematical performance because resources are shared between intrusive thoughts and mathematical problems. This framework is known as the processing efficiency theory (PET) (Eysenck and Calvo, 1992). Supporting evidence for the working memory and mathematical anxiety link comes from Ashcraft and Kirk (2001), who tested participants' working memory capacity while performing a complex mathematical task. In this study, participants stored words or digits in their working memory while processing arithmetic or verbal tasks. For the arithmetical tasks, participants had to solve simple additions, whereas for the verbal one, they had to answer simple questions about the content of a sentence. Participants then recalled the last words/numbers from the equation or sentence in each trial for both conditions. Results showed that the group with the highest score on the mathematical anxiety scale had a lower working memory span compared to the groups with lower scores, indicated by a lower number of numbers/words recalled. The authors hypothesized that due to a disruption in the central executive processes, there is an increase in demand during the online task for this particular group, without relating this to any intrusive thought. Additionally, results show that math anxiety depends on the complexity of the mathematical problems, with higher working memory demand for multi-digit additions than single-digit ones (Ashcraft and Faust, 1994). In short, there is some evidence suggesting that math performance is associated with mathematical anxiety, as both use limited working memory capacity.

Later on, Hopko et al. (2002) questioned the possibility of an attentional-control deficit among individuals with high mathematical anxiety. Results reported in their study (Hopko et al., 1998; 2002) suggested modifying the processing efficiency theory (Eysenck and Calvo, 1992). In their view, the link between mathematical anxiety and decreased mathematical performance is due to failure to inhibit one's intrusive thoughts and lack of attentional control. Since the two explanations lacked precision and explanatory power, the *attentional-control theory* reinterpreted PET (Suárez-Pellicioni et al., 2016; Eysenck et al., 2007). In agreement with the attentional-control theory, anxious individuals have an imbalance in attention by increasing their attention to the stimulus-driven system and decreasing it to the goal-oriented system. Therefore, individuals are more prone to distractions (bottom-up attentional intrusions) and are less focused on the ongoing task (top-down attention). Based on such evidence, anxiety impairs inhibition and attentional control, regardless of the source of stimuli (external or internal) (Suárez-Pellicioni et al., 2016; Eysenck et al., 2007). The literature overview above points out that working-memory load impacts mathematical anxiety and performance, even though the relation is not yet fully understood. Regardless of being an intentional-control deficit or a decrease in working-memory capacity, mathematical anxiety might reduce the mental resources available for solving mathematical problems.

Here, we aim to improve mathematical skills through scheduled memory retrieval training and improve performance for mathematical multiplications. By increasing mathematical skills through practice and at individual schedules, we aim to reduce at least some part of the source of the mathematical anxiety, by reducing mental resources used on the mathematical task. In this manner, as positive responses to math performance increase, the resources used for negative thoughts towards mathematics decrease.

Adaptive learning algorithms

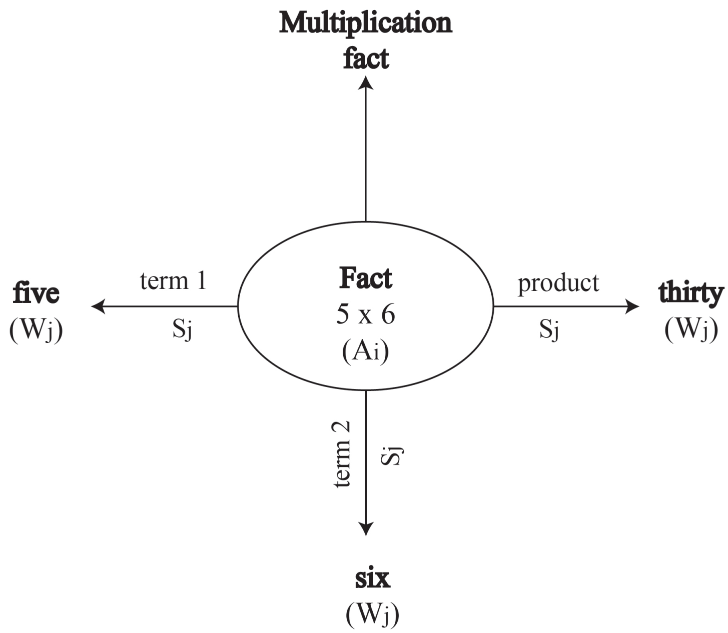
Adaptive learning algorithms have the capacity to give students feedback specific to their learning requirements, and improve their learning based on individual differences. The SlimStampen algorithm is an example of adaptive algorithm, a system that improves the learning efficacy by using an adaptive learning system, which results in better retention at the end of each study session (Sense et al., 2016; Van Rijn et al., 2009; Wilschut et al., 2021). Starting from Pavlik's and Anderson's (2005) ACT-R model of retention, which assumes that based on the difficulty of an item, there is an individual rate of forgetting, the adaptive algorithm can estimate individual differences in learning. Specifically, by determining one's reaction time and accuracy for an item response, the system estimates a parameter called rate of forgetting (RoF) for that particular item and aims to repeat it for retrieval at an individual schedule just before being estimated to be forgotten (Van Rijn et al., 2009). To determine RoF, an α parameter will be calculated to determine the decay at which one item should be presented from its first encounter (see subsection *The ACT-R model*). In the classic non-adaptive flashcard method, learners will initially be presented with all items to be retrieved, and all incorrect answers will be placed at the end of the set, to be repeated again in the order they have been mistaken. So far, the SlimpStampen adaptive learning algorithm has proved its efficacy in varied fields such as learning different declarative fact material (Sense et al., 2016) or learning modalities (Wilschut et al., 2021).

An open question is whether the SlimStampen adaptive learning system is efficient when solving mathematical problems. This study is the first to investigate whether an adaptive learning algorithm that takes into account individual differences in learners' reaction time and accuracy proves to be superior to a non-adaptive learning one when solving mathematical multiplications. A better understanding of the strategies used in mathematical calculations is required to explore this possibility. Along the strategies most used by learners, we count direct retrieval (i.e., remembering the solution of the problem) and procedural operations (i.e. actively calculating the answer to a question) (Polspoel et al., 2017). Anderson (1993) distinguishes two types of memories that might help us hypothesize the retrieval process in solving mathematical problems: declarative and procedural memory. According to the ACT-R theory (Anderson, 1993), the core theory behind the SlimStampen algorithm, a declarative memory will be stored in chunks, a product of the elements inside it, strengthened in time ($A_i = \sum W_j S_{ij}$) (see Figure 1). For example, when solving $5 \times 6 = 30$, the product will be retrieved by elements 5 and 6, reflecting the source activation of the chuck (W_j). In contrast, the degree of activation will determine the strength of activation (S_{ij}) between the two elements (Anderson, 1993, 1996).

On the other hand, procedural memory requires a production rule conditioned by the circumstance on which one can apply it, and at the specific time, it can be applied (Anderson, 1993, 1996). Roussel et al. (2002) suggest that in computing mathematical multiplications, the learner activates declarative knowledge by directly retrieving the answer of a calculation, compared to additions that rely on both types of memories. Therefore, when offering solutions for multiplication problems, the retrieval mechanism relies on declarative knowledge by activating facts within an associative network (Roussel et al., 2002). Similarly, Groen and Parkman (1972) argued that adults directly retrieve solutions for pairs of calculations from memory, relying on procedural rules only when they fail to. As the learner improves its arithmetical strategies, a shift from procedural memory to declarative one occurs, leading to faster reaction time in solving the mathematical problem, due to increased strength of association between the problem and the solution (Siegler and Shrager, 1984).

Figure 1

ACT-R theory declarative chunk.



Note. Adapted from Anderson et al., 1996 illustrates an example of a declarative chunk of the multiplicative fact $5 \times 6 = 30$ corresponding to the ACT-R theory.

Research purpose

In this view, the adaptive learning algorithm will determine an optimal schedule for retrieving the multiplications based on the particular reaction time and accuracy score while learning (Van Rijn et al., 2009; Sense et al., 2016). Therefore, there is a higher memory strength of association between the multiplication elements encountered more often, thus a shorter reaction time and better accuracy. If reaction times decrease and accuracy increases, we expect participants to improve their performance for the mathematical task. We are particularly interested in initial differences in reaction time and accuracy for participants who presented higher initial levels of mathematical anxiety and the associations between them. Because RoF is higher for participants who tend to forget declarative facts quicker, we are invested in its association with mathematical anxiety scores (Van Rijn et al., 2009). Additionally, we examine the individual preferences in using one of the two algorithms. Finally, by increasing one's performance on the mathematical task at an individual level, we expect participants to self-report a decrease in their mathematical anxiety by using the adaptive algorithm.

Therefore, the objectives of the investigation are:

1. To examine the differences in accuracy and reaction time between an adaptive learning algorithm and a non-adaptive one.

2. To observe individual differences in initial mathematical anxiety scores on accuracy and reaction time when solving mathematical multiplications.

3. To examine participants' preferences in learning with one of the two algorithms.

4. To examine the self-reported efficacy of an adaptive learning algorithm in improving mathematical skills and reducing mathematical anxiety.

We predict that the SlimStampen algorithm will result in higher accuracy and slower reaction time in solving multiplications than with the non-adaptive one. We hypothesize that initial higher mathematical anxiety scores will correlate negatively with accuracy and positively with reaction time and RoF. Finally, we hypothesize that participants will show a higher preference in learning with the SlimStampen algorithm compared to the non-adaptive one, and will indicate a decrease in their mathematical anxiety.

Method

Participants

Participants included 74 students from the University of Groningen, who voluntarily participated in the experiment in exchange for course credits. All participants completed the experiments conditions of the first session. Only 71 participants completed both sessions. There were no exclusion criteria before starting the experiment.

Materials

Measure for mathematical anxiety

To reduce administration time, mathematical anxiety was assessed with the brief version of the Mathematical Anxiety Rating Scale (MARS) (Richardson and Suinn, 1972), the reduced version of the 98-items initial scale. The computer-based scale consisted of all 30 items from the (MARS-SV) (Suinn and Winston, 2003), describing different situations encountering mathematical anxiety. Examples of items are *"Studying for a mathematics test"* or *"Figuring out your monthly budget"* that were rated on a Likert-type scale from 1 ("not at all anxious") to 5 ("very much anxious") (See Appendix A for additional sample items). Psychometric data indicate that there is high internal consistency with the initial scale, with an alpha Cronbach of .96, and test the reliability of .90 ($p < 0.0001$) (Suinn and Winston, 2003).

Measure for math skills

To assess mathematical skills, participants completed two unstandardized tests for the first session and one unstandardized test for the second session. An initial set of multiplications ranging from thirteen up to nineteen was selected. Multiplications of 1, 2, and 10 were subjectively assessed as too easy and removed from the experiment. The multiplication facts were presented by starting with the higher number as first factor (i.e. 13) and the lowest number as second factor (i.e. 3). To randomly allocate an equal number of items to one of the learning blocks, one more item was eliminated (15×5), with a remaining set of 48 items. All multiplications were shuffled and randomly selected for both learning sessions.

Measures of self-reported preferences and efficacy of the algorithms

We were interested in measuring self-reported preferences in learning with one of the algorithms, and therefore we asked participants to rate on a Likert-like scale from 1 ("not at all true") to 5 ("very much true") both statements: *"I liked learning with the first algorithm"* and *"I liked learning with the second algorithm".* At the end of the second session, participants self-reported the usefulness of the algorithm in reducing mathematical anxiety on a Likert-type scale from 1 ("not at all") to 10 ("very much likely") with the following statement *"How likely is the first learning method going to help you reduce your anxiety towards mathematics?".*

The ACT-R model

Based on Anderson (1996, 2009), a declarative memory equation can be defined under the ACT-R framework. Here, we are specifically interested in whether the following model can be used in solving mathematical multiplications. In this view, an activation value is defined for each mathematical item, which decays over time.

In the following equation,

$$A_i(t, n) = \sum_{j=1}^{n} (t - t_j)^{-d_j},$$

A_i is the activation of item i at time point t for the n items studied. Depending on the time t of the item encountered at j point in time, the activation A_i decreases or increases based on the recency of the encountered item. Here, the d_j parameter indicates how quickly a fact will decay over time.

In the second equation,

$$d_j = ce^{A_i(t)} + \alpha,$$

α represents the decay intercept that determines the decay value of an item from its first encounter. Lower decay values are attributed to items that have been encountered but for which presentation was long ago in time. If the estimated reaction time was longer than the actual reaction time, decay should have been longer.

Finally, the last formula:

$$RT_i(t) = Fe^{-A_i(t)} + \text{fixed time},$$

shows that we can determine an estimated reaction time by scaling the activation F of a particular item i at a point in time and summing it with a fixed time. The three formulas above-mentioned aim to optimize memory retention at an individual level. In the updated version of the algorithm (Van Rijn et al., 2009), the α parameter of a specific item compares the estimated reaction time with the observed reaction time. In this manner, the decay of an item's presentation is settled to optimize memory retention. Different from Pavlik's and Anderson's (2005) algorithm, the α parameter is updated for each item, showing individual differences in reaction time, accuracy and RoF.

To determine individual RoF, the item that dropped below the retrieval threshold will be the next to be presented to the participant at retrieval. Only when no item has dropped below the threshold will a new item be presented to be studied. Based on RoF, the algorithm checks which item needs to be presented for repetition.

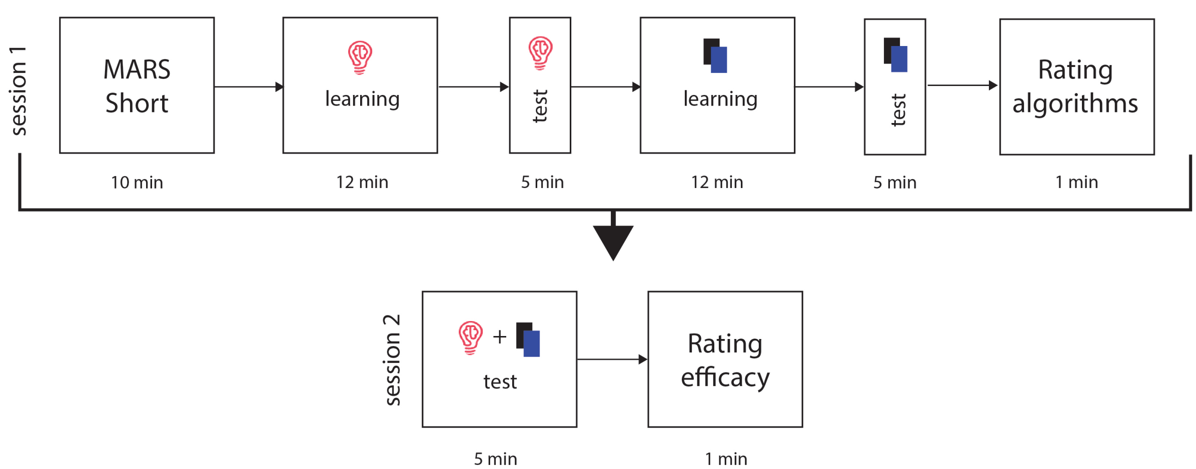
Procedure

In the present study, participants had to participate in two different sessions, one week apart, to observe if there is memory retention between the two-time occasions. In this experiment, there were two experimental conditions, one consisting of the adaptive learning algorithm and one of a non-adaptive one. In the first session, participants had to complete both a learning block and a test block for each experimental condition. In the second session, participants had to complete a test block. After completing informed consent, the experiment started with the first session and lasted approximately 40 min. The second session lasted around 10 min.

In the first session, participants had to complete the short mathematical anxiety scale (MARS-SV) (Suinn and Winston, 2003), the two mathematical skill learning blocks of 12 min each, one for each condition, followed by two short 5 min test blocks, one for each condition (see Figure 2). Afterwards, participants were asked about their preferences in learning with the two algorithms. In a single trial, participants were presented with a mathematical multiplication on the screen and asked to learn the answer. Then, participants are asked to type the answer to the next multiplication. Participants receive feedback ("correct"/"incorrect") for each response. There was a time limit of response set to 30 s per item. Participants started studying with the SlimStampen algorithm, to determine the same number of items for the non-adaptive learning one. Both number and order of items presented in the SlimStampen algorithm condition were determined based on RoF (see *The ACT-R model*), a parameter initially set at 0.3, and subsequently adjusted based on individual performances. For the non-adaptive algorithm, all items were shuffled over four stacks, and each mistaken item went back at the end of the corresponding stack. Then, the next item to be presented is the first mistaken one from the corresponding stack. The procedure continues until the learning block is over. After completing each learning block, participants were tested on the same items they studied.

Figure 2

Experimental design.



In the second session, participants had to complete a 5 min test comprising the previously studied items from both conditions from the week before. After being debriefed, participants were asked to rate the usefulness of reducing mathematical anxiety by using the SlimStampen algorithm.

Analysis

The experiment was realized using Java. The analysis of this study is made in R 4.2 and is based on the package lme4 (Bates et al., 2015) for the linear-mixed effects modeling, and ggplot2 for visualizing the data. To plot RoF and individual α's across participants, we used the SlimStampenRData package.

Results

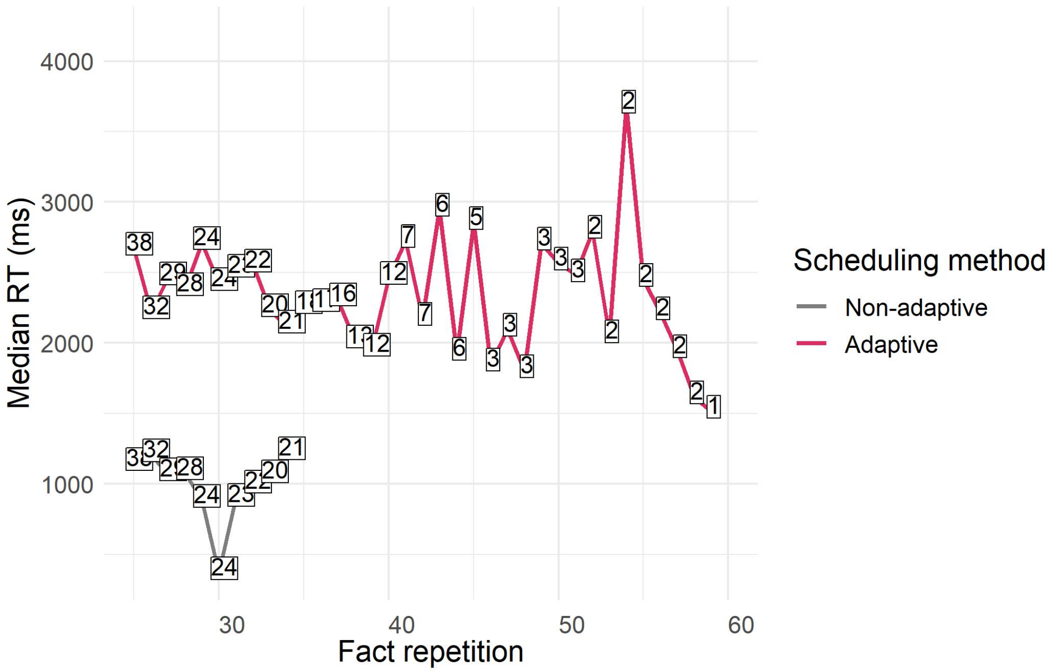
In the current study, we assessed whether an adaptive-learning algorithm outperforms a less adaptive flashcard-algorithm, in both reaction time and accuracy, while solving mathematical multiplications. Furthermore, we investigated whether there is any association between the scores on the mathematical anxiety scale and the reaction time, accuracy and RoF of participants. We then evaluated which algorithm was preferred, and whether participants reported the adaptive algorithm to efficiently reduce their mathematical anxiety.

Participants were presented with 10 items on average, ranging between 4 and 24 items. During the 12 min of a learning session, participants gave an average of 116 responses with the SlimStampen algorithm and 120 responses with the non-adaptive algorithm, with a mean of 4429.78 ms and 4542.73 ms, respectively. The average number of repetitions for each item is 18 and 16 for adaptive and non-adaptive algorithms respectively. However, by inspecting the unique number of repetitions over reaction time, we observe that the adaptive algorithm had more repetitions for an unique item compared to the non-adaptive one, with a maximum of 59 and 34 repetitions respectively (see Figure 3).

For a qualitative assessment of the relative difficulty of all items, Figure 4 shows that the easier items were “14 × 7”, “15 × 3” and “15 × 4” based on higher accuracy and “13 × 3”, “15 × 3” and “15 × 4” based on faster median reaction time. Additionally, the items that were less accurate were “17 × 8”, “18 × 3” and “19 × 8” with “18 × 7”, “19 × 3” and “19 × 4” having the slowest median reaction time.

Figure 3

Reaction times for unique items over repetition



Note. Numbers inside the box indicate the number of unique items repeated for both conditions.

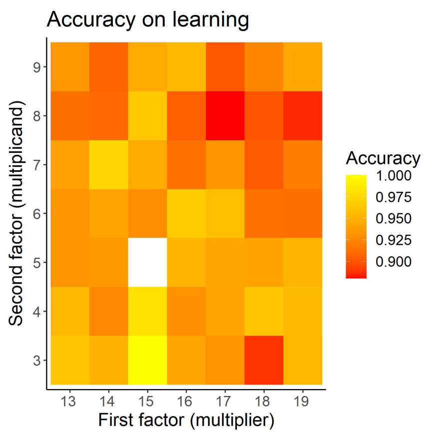
Differences between the adaptive and non-adaptive algorithm.

To answer our first main research question, we fitted mixed-effects regression models for both variables, accuracy and reaction time, and compared differences between the two conditions. We conducted four mixed-effects regression models for both the learning and test blocks, with reaction time and accuracy added as fixed effects, and with items and participants as random intercepts. We used contrast coding analysis on all mixed effects regression models.

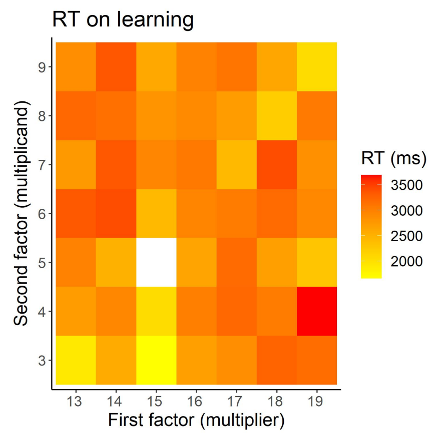
To determine the differences in accuracy during learning and test between the two algorithms, the two sessions and their interaction effects during learning, we fitted two logistic mixed effect regression models. Results show that the probability of giving a correct answer during the test block was 0.8 percentage points higher in the first session

Figure 4

Average accuracy differences for both sessions on learning (a) and on test (b) for the Adaptive algorithm and the Non-Adaptive one.



(a)



(b)

Note. Accuracy and median reaction time by multiplication item during the learning block.

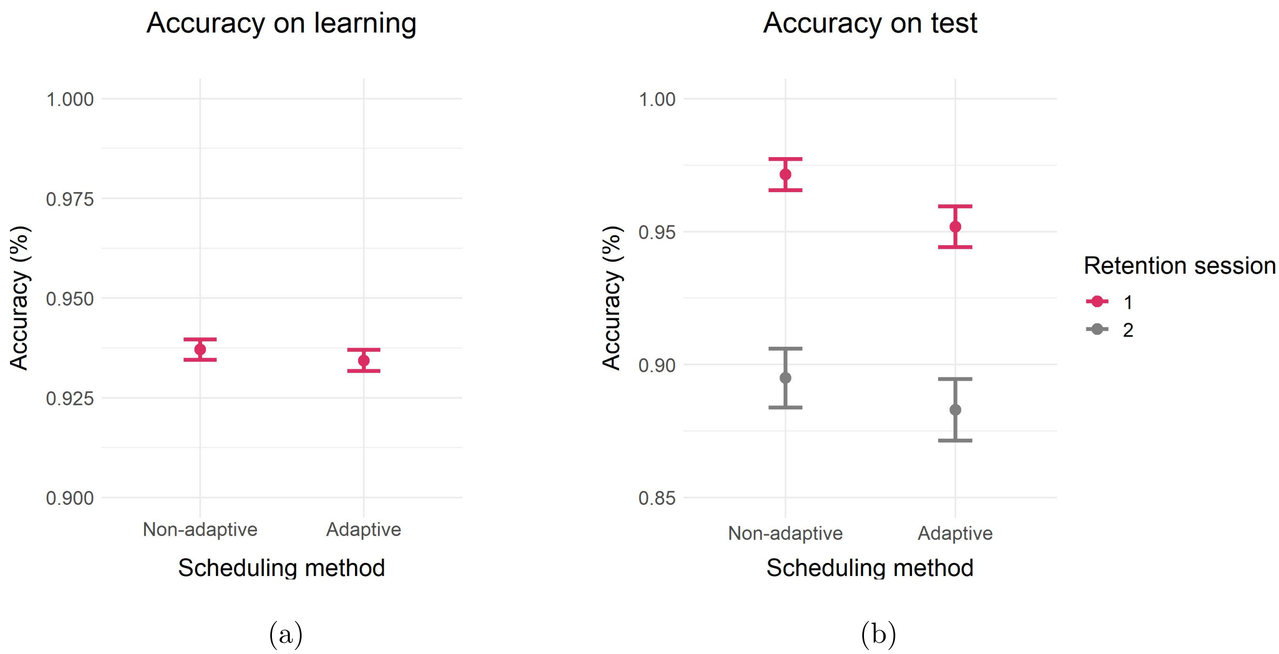
(see Table B1), and 2.6 percentage points higher for the second session by using the non-adaptive algorithm, in contrast to the adaptive one. The probability of giving a correct answer during the test block was 4.8 percentage points higher during the first session in contrast to the second one, with the non-adaptive learning, and 6.6 percentage points higher on the first session in contrast to the second one, with the SlimStampen algorithm.

To examine differences in accuracy during the learning block between the two algorithms, we fitted a second logistic regression mixed effect. According to the results, the probability of giving a correct answer by using the non-adaptive learning algorithm was 0.01 percentage points higher (see Table B1) in contrast to the adaptive one.

Subsequently, we fitted two regression mixed effect models to observe differences in reaction time between the two algorithms, the two sessions and their interaction effects during learning. As the reaction time distributions were highly skewed, we used the median reaction time for all analyses. Results, as seen in Table B2 indicate that, during the test block the non-adaptive algorithm reaction times are 336.2 ms faster in contrast to the

Figure 5

Average accuracy differences for both sessions on learning (a) and on test (b) for the Adaptive algorithm and the Non-Adaptive one.



Note. Standard errors are represented by the errors bars.

adaptive algorithm ($t(3040)$ = - 2.987, $p = 0.002$), and 1974.3 ms faster in the first session in contrast to the second one ($t(3028)$ = 17,739, $p < 0.001$), indicating that during test participants were slightly faster in the first session by using the non-adaptive algorithm. There was no significant interaction between the type of algorithm and session ($t(3009)$ = 1.943, $p = 0.0521$), indicating that reaction times did not change between conditions across sessions.

Finally, we examined differences in reaction time during learning by fitting one last mixed-effect model. There was, on average, a 236.3 ms advantage for the SlimStampen algorithm compared to the non-adaptive one (see Table B2). These results show that there is a small difference between the two algorithms, with a slight advantage for the non-adaptive algorithm for both accuracy and reaction time on the test block, and a small advantage on accuracy on the learning block. However, there is a slight advantage for the adaptive algorithm on the learning block in reaction time, compared to the non-adaptive one.

Figure 6

Median reaction time differences for both sessions on learning (a) and on test (b) for the Adaptive algorithm and the Non-Adaptive one.

(a) (b)

Note. Standard errors are represented by the errors bars.

Linear associations between mathematical performance and mathematical anxiety scores.

To observe differences in accuracy, reaction time, and RoF between participants with different scores on the mathematical anxiety scores, three simple Pearson's correlations were fitted, to assess the linear relationship between them. The distribution of the mathematical anxiety scores is normally distributed, as revealed by Q-Q analysis. However, there was a low number of participants for both high and low scores on mathematical anxiety, resulting in a low number of observations for the high mathematical anxious individuals. Therefore, no post-hoc analysis was realized to observe differences between the low and high mathematical anxiety group.

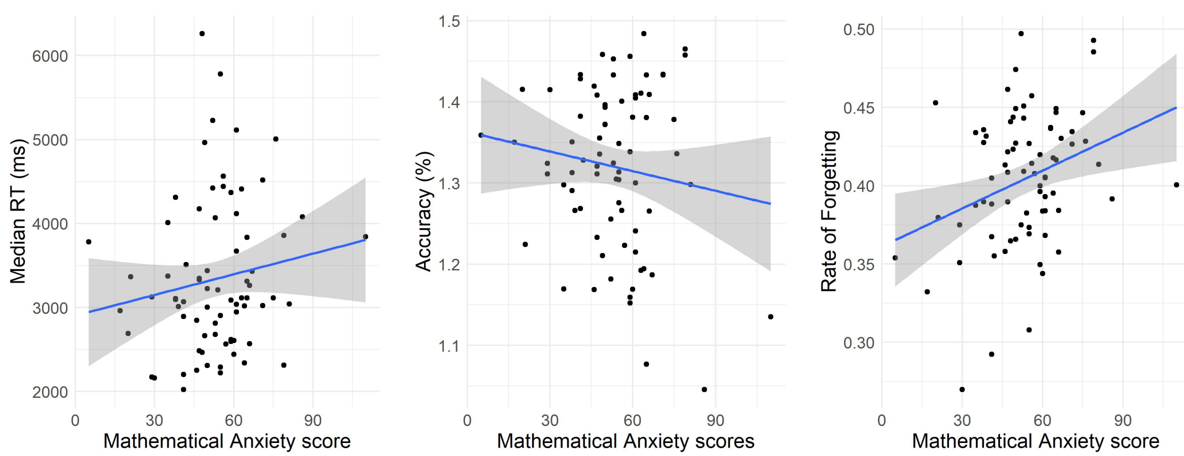
The first Pearson correlation computed was between the accuracy rate and mathematical anxiety scores. Arcsine square root transformations were performed for this data, with 0 : incorrect and 1 : correct. We expected accuracy scores to be lower for participants with higher scores on the mathematical anxiety scale. There was however a non-significant negative correlation between the mathematical anxiety scores and accuracy ($r(73)$ = - 0.13, $p = 0.257$).

Next, we computed a Pearson correlation between the relationship of the math anxiety scores and median reaction time. Results show a positive non-significant association between the two variables ($r(73) = 0.15$, $p = 0.199$).

Finally, we were interested in observing the linear relationships between RoF and mathematical anxiety scores. We expected that participants with higher RoF to show higher mathematical anxiety scores. Results indicate that there is a weak to moderate positive association between the two variables ($r(73) = 0.31$, $p = 0.007$).

Figure 7

Linear association between median RT (a), accuracy % (b), RoF (c) and mathematical anxiety scores.



(a) (b) (c)

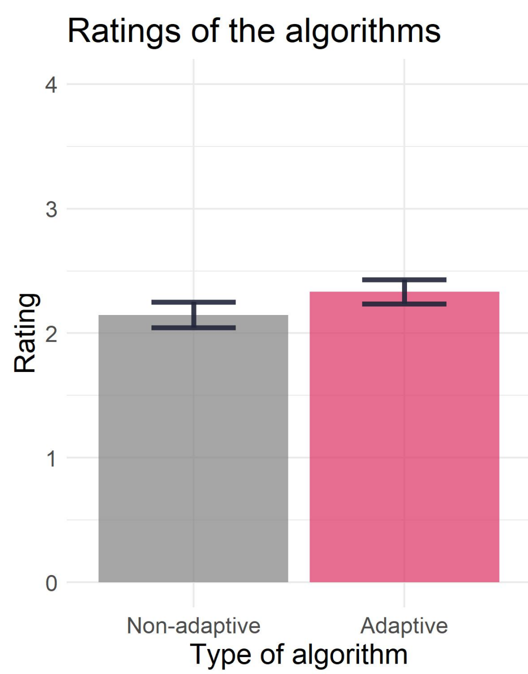
Altogether, these results show that, contrary to previous findings that point out negative associations between high scores on mathematical anxiety and weak math performance (Barroso et al., 2021; Hembree, 1990; Ma, 1999), mathematical anxiety is not a significant predictor of slower reaction time and lower accuracy, but is, however, a weak to moderate predictor for RoF, indicating that participants that have higher scores on the mathematical anxiety scale might forget material more quickly.

Algorithm self-reported efficacy ratings.

To approach our last two research questions regarding the self-reported preferences in learning with one of the algorithms and reported efficacy in reducing mathematical anxiety by using the adaptive algorithm, means and differences in means were computed.

Figure 8

Mean difference in self-reported preferences in learning with the SlimStampen algorithm and the Flashcard algorithm on a scale from 0 to 4 (0 = not at all true, 4 = very much true).



As shown in Figure 8 there is a small non-significant difference ($t(74)$ = -1.41, p = 0.163) between the mean of the participants' preferences in learning with the SlimStampen algorithm ($M = 2.33$, $SD = 0.89$) compared to the non-adaptive one ($M = 2.15$, $SD = 0.84$). On a scale from 0 to 9, participants rated on average 4.04 ($M = 4.04$, $SD = 2.17$), the degree to which the SlimStampen algorithm is likely to reduce their mathematical anxiety by using it when learning mathematical multiplications. Thus, no further interpretations based on the available data can be offered, participants do not express any particular preference for one of the algorithms, and their rating of the algorithm in decreasing their mathematical anxiety is moderate.

Discussion

To our knowledge, the present research was the first to investigate whether an adaptive learning algorithm that takes into account individual learner performance, such as reaction time and accuracy, can outperform a less adaptive one when solving mathematical multiplications.

The findings of this study are threefold, we will discuss each of these points in turn below. First, in contrast with previous results that show that an adaptive algorithm is more efficient in learning declarative knowledge facts (Van Rijn et al., 2009), our findings suggest that a non-adaptive algorithm might have a slight advantage over an adaptive one when learning and testing to calculate multiplications. This advantage can be observed while testing participants' accuracy and reaction time in solving mathematical multiplications, as well as their accuracy on learning, but not for their reaction time on learning. A possible explanation for our findings might be due to the fact that we did not ask participants explicitly to memorize the answer to the multiplications. According to the ACT-R theory, the information of a declarative fact is stored in chucks, and it is strengthened in time by repeated activation (Anderson et al., 1996). Here, participants had around 30 s to answer the multiplication question, which indicates that they might have calculated the answer and did not explicitly store it as a declarative fact. The procedural component might slow down the processing due to using "production rule" and different steps of cognition (Roussel et al., 2002). A possible explanation for the current results is that, given the fact that solving mathematical multiplications is procedural knowledge, one needs more time to shift from it to a declarative one (Siegler and Shrager, 1984). For example, if participants were to learn for longer study sessions (longer than 12 min) and across repeated sessions, they would need more time to shift from procedural to declarative knowledge.

Secondly, we hypothesized that participants with higher scores on mathematical anxiety show a slower reaction time and a lower rate of accuracy while learning mathematical multiplications. Additionally, we predicted that the individual RoF would be positively associated with high scores on the mathematical anxiety scale. However, our results suggest that only the RoF is weakly to moderately associated with participants' higher scores on mathematical anxiety, indicating that these participants tend to forget more easily while learning. The absence of a significant association between the performance parameters, such as reaction time and accuracy with mathematical anxiety might be due to the distribution of the participants' scores itself, with a relatively low number of participants that show both high and low anxiety towards mathematics. This might indicate that more observations of highly anxious participants for mathematics are needed to conclude results on the ongoing hypothesis. Another possible explanation of the current results could be due to the choice of performance parameters, such as accuracy and reaction time. There is a possibility that our results are inconsistent with previous findings because mathematical skills were previously measured with other performance parameters (i.e. scores on statistical tests, scores on standardized tests).

Finally, we were interested in observing whether participants show any particular preference in learning with one of the algorithms, and whether they subjectively indicate that the adaptive algorithm might be useful in decreasing their mathematical anxiety. Contrary to our expectations, results suggest that participants show no particular preference, reflected by no significant difference in preference, and their ratings on the usefulness of the adaptive algorithm are moderate. As a possible explanation, participants might need more interaction with both algorithms in order to observe differences between them, and therefore express a particular preference. Furthermore, using one informal direct question with no previous validity analysis is risky in concluding the following results. It might be that the question itself is too broad in ambiguity (i.e. participants do not understand what it means to decrease mathematical anxiety).

Future directions and limitations

Before mentioning the limitations of the current study, it is worth acknowledging one of its highest strengths. As already mentioned, the current study was the first to investigate whether an adaptive learning algorithm that takes into account individual parameters in learning, such as reaction time and accuracy performs better than a non adaptive one. We are therefore first in exploring this topic, and encourage more research on improving the adaptive algorithm for learning procedural knowledge. The lack of previous documentation might suggest that some experimental design errors have been influenced by this matter, and upcoming studies could aim to improve it and explore the possibility of different results.

A possible limitation of the current study is the mechanism of retention behind the adaptive algorithm, which might not be robust to procedural knowledge. As possible solutions, future research should focus on exploring the procedural memory equation behind the adaptive algorithm. Alternatively, with respect to the adaptive algorithm used in the current study, future studies should emphasize learning mathematical multiplication items as declarative facts. Therefore, by shortening the response time window (i.e. 5 s) and emphasizing memorizing the answer, rather than calculating it, it would be interesting if any different results arise. An alternative could be increasing the difficulty of the items, by including multiplications of two digits. In this manner, participants memorize the answer to the multiplication, instead of calculating it, given that the response window will be shorter.

Another possible limitation concerns the experimental design and the order of the experimental condition learning blocks. The number of items presented in both algorithms was determined based on RoF, suggesting that the first algorithm to start with was always

SlimStampen. This might indicate that the adaptive algorithm has been disadvantaged by being first in learning multiplications, because participants had time to rehearse the multiplication tables. It is possible that some multiplications have been previously rehearsed using the SlimStampen algorithm, facilitating the learning session with the non-adaptive one. It has been shown that retrieval of previous facts facilitates learning of new information (Pastötter and Bäuml, 2014). An interesting alternative could be setting four blocks of learning sessions, two for each experimental condition. By contrasting the two adaptive algorithms blocks we could further observe whether the difference could be explained by the disadvantage of the adaptive algorithm in starting first or not.

Perhaps the most contrasting difference between the two algorithms is the number of repetitions per unique item (see Figure 3). We observed that this was higher in the adaptive algorithm condition compared to the non-adaptive one, indicating that difficult items have been repeated more often than the easier one. By repeatedly solving the same multiplication for a prolonged time, participants might decrease their motivation. This interpretation is partially in line with previous findings, which indicate that there is a link between mental fatigue and task disengagement, which might influence the overall performance (Hopstaken et al., 2015). To avoid this, future research might set a limit on the maximum repetitions for each unique item for the adaptive algorithm.

With regards to the findings related to math anxiety, possible improvements are worth mentioning. Here, we measured decrease in anxiety with an informal question. However, an alternative to detect changes in math anxiety, is by measuring it pretest and posttest, after a longer period of learning (i.e. six weeks). Moreover, we reasoned that by adapting the learning experience to individuals' parameters we might observe a decrease in mathematical anxiety, or, other variables might be considered in future study, such as success rate. Jansen et al., (2013) study investigated the decrease in math anxiety for children who studied mathematics by using a computer-adapting program that takes into account individuals' ability level, and observed that there was a higher improvement in math performance associated with higher success rates. In their study, children solving more difficult problems had a higher success rate, by practicing more problems, resulting in an improvement in math performance. We suggest that future research pursue integrating an overall success rate after each learning session, where participants can see in real time their improvement on the learning path. Alternatively, it could be investigated whether there is a decrease in math anxiety between pretest and posttest associated with higher success rates due to learning with the SlimStampen adaptive algorithm compared to a non-adaptive one.

A potential problem with the external validity of the current study is both about the sample that included students from the University of Groningen and that it took place in a laboratory. It would be of great interest to replicate the current study in real educational contexts, such as classrooms, and in another sample, such as children from primary or secondary schools.

Conclusion

The following study helps us improve our understanding of how adaptive learning algorithms can personalize the educational experience in learning mathematics. To sum up, the current results indicate that when solving mathematical multiplications, an adaptive algorithm does not show any advantage in reaction time and accuracy compared to a less adaptive one during test, but does show a slight advantage in reaction time while learning. Here, we observed that participants with a higher RoF tend to score higher on the mathematical anxiety scale, which indicates that mathematical anxious participants tend to forget quicker when learning. Although our findings do not show any association between mathematical anxiety and performance parameters, it does provide a basis for future investigations. Importantly, the current results indicate that interventions with learning algorithms may be useful in the educational contexts, and that learning could be optimized in the future with their use, if slight adjustments to the current algorithms are made. Moreover, these findings can bring much-needed insight for educators who consider integrating adaptive learning algorithms in classrooms, to improve both efficiency and experience of learning.

References

Ahmed, W. (2018). Developmental trajectories of math anxiety during adolescence: Associations with stem career choice. *Journal of adolescence*, *67*, 158–166. https://doi.org/10.1016/j.adolescence.2018.06.010

Anderson, J. R. (1993). *Rules of the mind*. Psychology Press.

Anderson, J. R. (2009). *How can the human mind occur in the physical universe?* Oxford University Press.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive psychology*, *30*(3), 221–256. https://doi.org/10.1006/cogp.1996.0007

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current directions in psychological science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Ashcraft, M. H., & Faust, M. W. (1994). Mathematics anxiety and mental arithmetic performance: An exploratory investigation. *Cognition & Emotion*, *8*(2), 97–125. https://doi.org/10.1080/02699939408408931

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of experimental psychology: General*, *130*(2), 224–237. https://doi.org/10.1037/0096-3445.130.2.224

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational assessment*, *27*(3), 197–205. https://doi.org/10.1177/0734282908330580

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, *147*(2), 134. https://doi.org/10.1037/bul0000307

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4* [R package version 1.1-8]. http://CRAN.R-project.org/package=lme4

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and brain functions*, *8*(1), 1–9. https://doi.org/10.1186/1744-9081-8-33

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & emotion*, *6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336. https://doi.org/10.1037/1528-3542.7.2.336

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological review*, *79*(4), 329. https://doi.org/10.1037/h0032950

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for research in mathematics education*, *21*(1), 33–46. https://doi.org/10.5951/jresematheduc.21.1.0033

Hopko, D. R., Ashcraft, M. H., Gute, J., Ruggiero, K. J., & Lewis, C. (1998). Mathematics anxiety and working memory: Support for the existence of a deficient inhibition mechanism. *Journal of anxiety disorders*, *12*(4), 343–355. https://doi.org/10.1016/s0887-6185(98)00019-x

Hopko, D. R., McNeil, D. W., Gleason, P. J., & Rabalais, A. E. (2002). The emotional stroop paradigm: Performance as a function of stimulus properties and self-reported mathematics anxiety. *Cognitive Therapy and Research*, *26*(2), 157–166. https://doi-org.proxy-ub.rug.nl/10.1023/A:1014578218041

Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology*, *52*(3), 305–315. https://doi.org/10.1111/psyp.12339

Jansen, B. R., Louwerse, J., Straatemeier, M., Van der Ven, S. H., Klinkenberg, S., & Van der Maas, H. L. (2013). The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. *Learning and individual differences*, *24*, 190–197. https://doi.org/10.1016/j.lindif.2012.12.014

Johnson, B. G., Phillips, F., & Chase, L. G. (2009). An intelligent tutoring system for the accounting cycle: Enhancing textbook homework with artificial intelligence. *Journal of Accounting Education*, *27*(1), 30–39. https://doi.org/10.1016/j.jaccedu.2009.05.001

Liu, M., McKelroy, E., Corliss, S. B., & Carrigan, J. (2017). Investigating the effect of an adaptive learning intervention on students' learning. *Educational technology research and development*, *65*(6), 1605–1625. https://doi.org/10.1007/s11423-017-9542-1

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for research in mathematics education*, *30*(5), 520–540. https://doi.org/10.2307/749772

Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. https://doi-org.proxy-ub.rug.nl/10.1037/11706-002

Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in psychology*, 286. https://doi.org/10.3389/fpsyg.2014.00286

Pavlik Jr, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive science*, *29*(4), 559–586. https://doi.org/10.1207/s15516709cog0000_14

Polspoel, B., Peters, L., Vandermosten, M., & De Smedt, B. (2017). Strategy over operation: Neural activation in subtraction and multiplication during fact retrieval and procedural strategy use in children. *Human Brain Mapping*, *38*(9), 4657–4670. https://doi.org/10.1002/hbm.23691

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of counseling Psychology*, *19*(6), 551.

Roussel, J.-L., Fayol, M., & Barrouillet, P. (2002). Procedural vs. direct retrieval strategies in arithmetic: A comparison between additive and multiplicative problem solving. *European Journal of Cognitive Psychology*, *14*(1), 61–104. https://doi.org/10.1080/09541440042000115

Sense, F., Behrens, F., Meijer, R. R., & van Rijn, H. (2016). An individual's rate of forgetting is stable over time but differs across materials. *Topics in cognitive science*, *8*(1), 305–321. https://doi.org/10.1111/tops.12183

Siegler, R., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? *Origins of cognitive skills*.

Suárez-Pellicioni, M., Núñez-Peña, M. I., & Colomé, À. (2016). Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(1), 3–22. https://doi.org/10.3758/s13415-015-0370-7

Suinn, R. M., & Winston, E. H. (2003). The mathematics anxiety rating scale, a brief version: Psychometric data. *Psychological reports*, *92*(1), 167–173. https://doi.org/10.2466/pr0.2003.92.1.167

Supekar, K., Iuculano, T., Chen, L., & Menon, V. (2015). Remediation of childhood math anxiety and associated neural circuits through cognitive tutoring. *Journal of Neuroscience*, *35*(36), 12574–12583. https://doi.org/10.1523/JNEUROSCI.0786-15.2015

Van Rijn, H., van Maanen, L., & van Woudenberg, M. (2009). Passing the test: Improving learning gains by balancing spacing and testing effects. *Proceedings of the 9th International Conference of Cognitive Modeling*, *2*(1), 7–6.

Wilschut, T., Sense, F., van der Velde, M., Fountas, Z., Maaß, S. C., & Van Rijn, H. (2021). Benefits of adaptive learning transfer from typing-based learning to speech-based learning. *Frontiers in artificial intelligence*, *4*. https://doi.org/10.3389/frai.2021.780131

Appendix A

Instrument

The 30-item MARS-SV scale

1. Taking an examination (final) in a mathematics course.
2. Thinking about an upcoming mathematics test one week before
3. Thinking about an upcoming mathematics test one day before.
4. Thinking about an upcoming mathematics test one hour before.
5. Thinking about an upcoming mathematics test five minutes before.
6. Waiting to get a mathematics test returned in which you expected to do well.
7. Receiving your final mathematics grade in the mail.
8. Realizing that you have to take a number of mathematics classes to fulfill the requirements in your major.
9. Being given a pop quiz in a mathematics class.
10. Studying for a mathematics test.
11. Taking the mathematics section of a college entrance examination.
12. Taking an examination (quiz) in a mathematics course.
13. Picking up the mathematics textbook to begin working on a homework assignment.
14. Being given a homework assignment of many difficult problems which is due the next class meeting.
15. Getting ready to study for a mathematics test.
16. Dividing a five digit number by a two digit number in private with pencil and paper.

17. Adding 976 + 777 on paper.
18. Reading a cash register receipt.
19. Figuring the sales tax on a purchase that costs more than $ 1.00.
20. Figuring out your monthly budget.
21. Being given a set of numerical problems involving addition to solve on paper.
22. Having someone watch you as you total up a column of figures.
23. Totaling a dinner bill that you think overcharged you.
24. Being responsible for collecting dues for an organization and keeping track of the amount.
25. Studying for a driver's license test and memorizing the figures involved, such as the distances it takes to stop a car going at different speeds.
26. Totaling up the dues received and the expenses of a club you belong to.
27. Watching someone work with a calculator.
28. Being given a set of division problems to solve.
29. Being given a set of subtraction problems to solve.
30. Being given a set of multiplication problems to solve.

Appendix B

Table section

Table B1

Mixed-effects logistic regression models result from contrast coding for accuracy.

Accuracy on test	β	SE	z	p
Intercept	3.0598	0.1715	17.846	< 0.001***
Adaptive - NonA	0.3530	0.1642	2.149	0.0316
Sess 1 - Sess 2	-1.2572	0.1640	-7.668	< 0.001***
Interaction	-0.4541	0.3246	-1.399	0.1619
Accuracy on learning	β	SE	z	p
Intercept	2.9027	0.1097	26.459	< 0.001***
Adaptive - NonA	-0.0292	0.0636	-0.459	0.0316

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$

Note. The logistic regression coefficients are converted into probabilities by an inverse logit transformation with contrast coding.

Table B2

Mixed-effects regression models result from contrast coding for reaction time.

RT on test	β	SE	df	t	p
Intercept	4879.0	227.9	105	21.410	< 0.001***
Adaptive - NonA	-336.2	112.6	3040	-2.987	0.0028
Sess 1 - Sess 2	1974.3	0.1640	3028.3	17.639	< 0.001***
Interaction	431.1	0.3246	3009.2	1.943	0.0521
RT on learning	β	SE	df	t	p
Intercept	3907.57	145.10	113.36	26.459	< 0.001***
Adaptive - NonA	236.35	46.39	17634.63	5.095	< 0.001***

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$