Solving multiplication and division problems: latent variable modeling of students' solution strategies and performance

Fagginger Auer, M.F.

Author: Fagginger Auer, M.F.
Title: Solving multiplication and division problems: latent variable modeling of students' solution strategies and performance
Issue Date: 2016-06-15

— 6 —

Single-task versus mixed-task mathematics performance and strategy use: Switch costs and perseveration

Abstract

The generalization of educational research to educational practice often involves the generalization of results from a single-task setting to a mixed-task setting. Performance and strategy use could differ in these two settings because of task switching costs and strategy perseveration, which are both phenomena that have yet to be studied with more complex educational tasks. Therefore, the problem solving of 323 primary school students in a single-task and mixed-task condition was investigated. The tasks that students had to do were typical educational tasks from the domain of mathematics that are especially interesting with regard to strategy use: solving twelve multidigit division problems that were intended to be solved with written, algorithmic strategies, and twelve non-division mathematical problems that do not call for such strategies. The results indicated no condition differences in performance or strategy use. This suggests that generalization of problem solving in single-task setting to a mixed-task setting is not necessarily problematic.

6.1 Introduction

An important challenge for educational research is its generalization to educational practice. The present study addresses a possible issue in generalization that does

---

This chapter is currently submitted for publication as: Fagginger Auer, M. F. (submitted). *Single-task versus mixed-task mathematics performance and strategy use: Switch costs and perseveration.*

I would like to thank the schools and students for their participation in the experiment, Anton Béguin and Floor Scheltens for their assistance in conceptualizing the study, and the Dutch National Institute for Educational Measurement Cito for allowing use of the assessment items.

not appear to have been investigated so far: the generalization of single-task research to mixed-task practice. In the daily educational practice of lessons and tests, students generally do not work on one task exclusively, but switch between different tasks as they go from problem to problem: for example, a mathematics test usually does not concern only a single mathematical operation (e.g., multiplication), but consists of different types of problems that require different operations. Also at the higher level of evaluating educational achievement in (inter)national assessments, tasks are presented mixed with each other rather than in isolation (e.g., Mullis & Martin, 2014; Scheltens et al., 2013).

Yet, much of educational research consists of single-task experiments, such as multiplication (Siegler & Lemaire, 1997), addition (Torbeyns et al., 2005), or spelling (Rittle-Johnson & Siegler, 1999). Sometimes, single-task experiments are even used for explanation of results of mixed-task assessments (e.g., Hickendorff et al., 2010). The use of single-task designs for experiments is logical, given the nature of experiments: the evaluation of the effects of controlled manipulation of only one or a few factors at once. However, when using single-task designs, it is important to know to what extent this may limit the generalizability of results to educational practice. Therefore, in the present study two aspects of problem solving are considered that may differ for single-task versus mixed-task designs: performance and solution strategy use.

6.1.1 Possible causes of differences between single-task en mixed-task results

Two phenomena could play a role in creating differences in problem solving.

Switch costs

The first is the well-established phenomenon of task switching costs in terms of accuracy and speed. A long line of research has established in increasingly advanced experiments that switching between tasks incurs costs. Various explanations for this phenomenon have been proposed (Kiesel et al., 2010). One is that costs occur because of active preparation for the upcoming task, while another posits passive decay of the previous task. Another explanation is interference from the other task (that was previously performed or is expected to be performed) in performing the current task. The research on task switching usually concerns very simple tasks, such as determining whether a number is even or odd or whether a stimulus is a number or a letter, and describes switch costs in terms of milliseconds. In contrast, most tasks in education are much more complex, and therefore the extent to which switch costs will occur in an educational context is not self-evident and has yet to be investigated.

Strategy perseveration

The second phenomenon that could play a role is that of strategy perseveration. This topic has not been studied widely yet, but has received recent research attention (Lemaire & Lecacheur, 2010; Luwel, Schillemans, Onghena, & Verschaffel, 2009; Luwel, Torbeyns, Schillemans, & Verschaffel, 2009; Schillemans, Luwel, Bulté, Onghena, & Verschaffel, 2009; Schillemans, Luwel, Onghena, & Verschaffel, 2011a, 2011b). Strategy perseveration is the continuing use of the same strategy as in previous solutions, even though another strategy may be more suitable or efficient for the problem at hand. Schillemans (2011) has described several explanations for this perseveration. One is the Einstellung effect, which is individuals' tendency to become blinded to other strategies, even though they may be more suitable than the previously applied strategy. A second explanation is priming, where the strategy that was previously used is more highly activated and therefore more likely to be selected. A third explanation is strategy switch costs, which are the costs involved in switching between strategies (which may occur through similar mechanisms as task switching costs; Lemaire & Lecacheur, 2010).

Perseveration has been shown to occur in single-task settings (Lemaire & Lecacheur, 2010; Luwel, Schillemans, et al., 2009; Luwel, Torbeyns, et al., 2009; Schillemans et al., 2009, 2011a, 2011b), but what occurs in a mixed-task setting has yet to be investigated: the mixing would seem to prevent perseveration as the alternation of tasks makes it impossible to keep applying the same strategy, but possibly perseveration in a similar but not identical strategy could occur (e.g., an algorithmic approach on one task might increase the probability of a (different) algorithmic approach on a subsequent other task).

6.1.2 The present study

Given these possible and as of yet unknown effects of task switching costs and strategy perseveration in an educational setting, the present study compares performance and strategy use in a single-task versus a mixed-task condition. The task used is the solving of mathematical problems in the domain of multidigit division

(division with larger numbers or decimal numbers, such as $1536 \div 16$ or $31.2 \div 1.2$). This task is a typical educational task, making it suitable for the goal of investigating task switching and strategy perseveration in an educational context, and is also especially interesting with regard to the latter strategy phenomenon.

This is because multidigit division problems are traditionally associated with solution strategies that involve writing down calculations (especially algorithmic strategies), or even defined as problems that make such an approach necessary of desirable (J. Janssen et al., 2005; Scheltens et al., 2013). However, in mixed-task large-scale assessments only around half of students' solutions involve written (mostly algorithmic) strategies (Scheltens et al., 2013), even though these strategies are much more accurate than non-written strategies (Hickendorff et al., 2009). Possibly, the mixing of multidigit division problems with other problems that do not call for written, algorithmic strategies makes students persevere in using mental, non-algorithmic strategies, or conversely, prevents students from persevering in written algorithmic strategies on the division problems. The comparison of single-task and mixed-task division problem solving in the present study could shed light on the extent to which this is the case.

The division problems are contrasted with other mathematical problems that do not involve division and that were selected to elicit mental, or at least non-algorithmic strategy use. Rather than contrasting division with a single other task, non-division problems from (nearly) all regularly assessed mathematics domains were included, to more closely approximate educational practice. Division and non-division problems from the two most recent national large-scale assessments of mathematics at the end of primary school in the Netherlands were used, because they reflect typical problems in Dutch primary school mathematics and were rigorously pretested.

Research questions

The first research question addressed by this study was the following: to what extent does mathematical performance differ in single-task and mixed-task conditions? Given the well-established existence of switch costs, it was expected that in the case of any differences between conditions, performance (whether in accuracy or speed) would be worse in the mixed-task than in the single-task conditions. However, because the task of multidigit division problem solving is much more complex than the elementary tasks usually employed in task switching, it could be that so many facets are already involved in performing just the mathematics task, that additional costs in switching between different mathematical tasks are negligible. In that case, performance in both conditions would be comparable.

The second research question that was addressed was: to what extent does the occurrence of strategy perseveration differ in single-task and mixed-task conditions? Two types of perseveration could occur. One is perseveration in applying the mental, non-algorithmic strategies suitable for the non-division problems to the division problems in the mixed-task condition, where division problems always occur directly or shortly after non-division problems (which is not the case in the single-task condition). The other is perseveration in applying written, algorithmic strategies to the division problems when they are presented together in the single-task condition (which is not possible when the division problems are interspersed with non-division problems that cannot be solved with a division algorithm in the mixed-task condition).

6.2 Method

6.2.1 Participants

A total of 323 students at the end of primary school (sixth grade; 11-12-year-olds) from 15 different schools participated in the experiment, of whom 53 percent were girls and 47 percent were boys. Data on students' mathematical ability was available from standardized national tests that are administered at most Dutch primary schools (J. Janssen et al., 2010). Students were assigned to the single-task (50 percent of students) and mixed-task condition (the other 50 percent) according to a randomized block design (with blocking based on gender, ability quartile and school).

6.2.2 Materials

Students made a test consisting of twelve multidigit division problems and twelve problems of other types (see Table 6.1 for the problems). The problems came from the two most recent (2004 and 2011) national large-scale assessments of mathematics performance at the end of primary school (Scheltens et al., 2013; J. Janssen et al., 2005). All problems were open-ended, and all problems except $31 \div 1.2$ and $\frac{3}{8} + \frac{1}{4}$ were presented in a realistic problem solving context (such as determining how many bundles of 40 tulips can be made from 2500 tulips). The non-division problems were from (nearly) all mathematics domains investigated in the assessments except addition, subtraction, multiplication and division, as these problems were intended to evoke non-algorithmic, mental strategies.

The problems were printed in A4-booklets with two problems per page, so that there was ample space for writing down calculations. In the single-task condition, the first twelve problems in the test booklet were the division problems and the next twelve the non-division problems (or vice versa for half of the students that condition), whereas in the mixed-task condition, every time one or two non-division problems were followed by one or two division problems in an unpredictable way (see Table 6.1). The single task did not consist of solely division problems so that the total difficulty and time required for the test was the same in both conditions.

6.2.3 Procedure

Students made the tests in their classroom in the presence of the experimenter and had 45 minutes to do so. Students were instructed that if they wanted to write down calculations, they should do so in the test booklet. When a student had finished, the test completion time in minutes for that student was written down by the experimenter.

After students had made the test, their solutions were scored for accuracy and strategy use. For division problems, four categories of strategy use were discerned: the digit-based algorithm (a more traditional approach, where numbers are broken up into digits that can be handled without an appreciation of their magnitude in the whole number); the whole-number-based algorithm (a newer approach where every step towards obtaining the solution requires students to understand the magnitude of the numbers they are working with; Treffers, 1987a); non-algorithmic written solutions (such as only writing down intermediate steps); and no written work (see Table 6.2 for examples). For the non-division problems, the two algorithm categories were merged into one category, as whole-number-based algorithms are very infrequent for other operations than division (Buijs, 2008), and the other categories were the same.

6.2.4 Statistical analysis

Mixed models

The effects of condition (single-task or mixed-task) and student gender (boy or girl) and mathematical ability score and their interactions on speed and accuracy were investigated using mixed models: linear mixed models for test completion time and

Table 6.1: The twelve division and twelve other problems (order shown for the mixed condition).

type	item
surfaces	determining the surface of a triangle covering half of a 4×4 grid
division	$1536 \div 16 = 96$
tables	looking up the lesson taking place at a given time in a timetable
division	$872 \div 4 = 218$
division	*31.2 ÷ 1.2 = 26*
geometry	determining the number of windows based on a building scheme
money	determining the number of 20 cent coins in 80 euro
division	*6496 ÷ 14 = 464*
fractions	$\frac{3}{8} + \frac{1}{4}$
number line	? - 8 - 8.125 - 8.250 - 8.375 - 8.500
division	*544 ÷ 34 = 16*
division	$11585 \div 14 = 827.5$
length	converting 3.1 meters to centimeters
division	*47.25 ÷ 7 = 6.75*
division	$157.50 \div 7.50 = 21$
volume	reading off 1.5 liters from 2 liter container with 0.5 liter marks
time	determining the difference between 09:15 and 08:55
division	$2500 \div 40 = 62$
division	$1470 \div 12 = 122.50$
measurement	determining the height of a mentally rearranged tower of cubes
division	$736 \div 32 = 23$
weight	converting 3959 grams to kilograms
number line	2.06 - ? - 2.07
division	$16300 \div 420 = 39$

Note: Parallel versions of problems not yet released for publication are in italics.

Table 6.2: Examples for the different strategy coding categories for the division problem $544 \div 34$.

digit-based algorithm	whole-number-based algorithm	non-algorithmic strategies	no written work
34/544\16	544 : 34 =	$10 \times 34 = 340$	16
34	340 - 10×	$15 \times 34 = 510$	
204	204	$16 \times 34 = 544$	
204	102 - 3×		
0	102		
	102 - 3×+		
	0 16×		

logistic mixed models for accuracy (correct or incorrect). Both types of models included a random effect for schools, and the accuracy model also random effects for students and items (De Boeck, 2008) since it modeled data at the item level. The analyses were conducted using the package `lme4` in the statistical computing software `R` (Bates & Maechler, 2010).

Latent class analysis

Students' patterns of strategy use on the twelve division items were investigated using multilevel latent class analysis (MLCA). In LCA, individuals are classified in latent classes that are each characterized by a particular pattern of response probabilities for a set of items (Goodman, 1974; Hagenaars & McCutcheon, 2002). The multilevel aspect makes individuals' probability of being in latent classes dependent on the group they are in (in this study, the groups that are formed by the classes of the different teachers). Covariates can also be added to predict latent class membership. The multilevel latent class analysis was conducted with version 5.0 of the Latent GOLD program (Vermunt & Magidson, 2013). All twelve division strategy variables were entered as observed response variables and a teacher identifier variable as the grouping variable for a nonparametric multilevel effect. The optimal number of latent students and teacher classes was determined based on the Bayesian Information Criterion (BIC; Schwarz, 1978) and the effects of covariates were evaluated using Wald tests.

Table 6.3: Performance in the single and mixed task condition in terms of accuracy and speed.

	accuracy (percentage correct)		speed (minutes)
condition	non-division problems	division problems	whole test
single-task	69	44	37
mixed-task	70	45	36
total	70	45	36

Table 6.4: Strategy use in the single-task and mixed-task condition.

	non-division problems			division problems			
condition	A	NA	NW	DA	WA	NA	NW
single-task	2	7	91	13	36	20	32
mixed-task	3	7	90	17	34	19	30
total	2	7	91	14	35	19	31

Note: A=algorithm, NA=non-algorithmic, NW=no written work, DA=digit-based algorithm, WA=whole-number-based algorithm

6.3 Results

As can be seen from the performance descriptives in Table 6.3, students provided correct solutions to 70 percent of the non-division and 45 percent of the division problems, and completed the test in 36 minutes on average ($SD = 8$ minutes). Table 6.4 gives the frequencies of students' use of the different strategies. As intended, students almost never applied an algorithmic strategy to non-division problems (2 percent), and most often solved such problems without writing down any calculations (91 percent). For the division problems, students used an algorithmic strategy approximately half of the time: they applied the whole-number-based algorithm to 35 percent of the problems and the digit-based algorithm to 15 percent of the problems. Solutions without any written work were also frequent (31 percent), as were non-algorithmic written strategies (19 percent).

6.3.1 Task switching costs

To investigate whether the switching between division and non-division problems in the mixed-task condition incurred switch costs that did not occur in the single-task condition, accuracy and speed in the two conditions were compared.

Accuracy

Table 6.3 shows that the observed percentage of correct answers to division problems was nearly identical in the two conditions: 44 percent in the single-task and 45 percent in the mixed-task condition. A comparison of models using likelihood ratio tests confirmed a lack of differences between the conditions: the null model for the accuracy of division solutions (with only an intercept) was significantly improved by adding the student characteristics gender and ability (and their interaction) as predictors, $\chi^2(3) = 231.4$, $p < .001$, but adding a condition effect (and condition interactions with gender and ability) did not provide further improvement, $\chi^2(4) = 6.7$, $p = .15$. In the model with student characteristics, accuracy was found to be lower for boys than for girls, $z = -3.41$, $p < .001$, and accuracy was found to be positively related to ability score, $z = 10.25$, $p < .001$. The interaction between gender and ability was non-significant, $z = 1.75$, $p = .08$.

Speed

Table 6.3 also shows that average time in which students completed the whole test was nearly identical in the two conditions: 37 minutes in the single-task and 36 minutes in the mixed-task condition. Again, a comparison of models confirmed a lack of differences between the conditions: the null model for test completion time (with only an intercept) was significantly improved by adding student gender and ability, $\chi^2(3) = 27.3$, $p < .001$, but adding condition effects provided no further improvement, $\chi^2(4) = 4.4$, $p = .36$. In the model with student characteristics, boys were found to be faster than girls, $z = -5.23$, $p < .001$. Ability score did not have a significant effect, $z = -0.31$, $p = .38$, nor did the interaction between gender and ability, $z = 0.49$, $p = .31$.

6.3.2 Strategy perseveration

To investigate the effects of mixing division and non-division problems on strategy use, patterns of strategy use in the two conditions were compared. Table 6.4 shows that the overall percentage of division problems solved with each strategy was nearly identical in the single-task and mixed-task conditions: 13 and 17 percent respectively for the digit-based algorithm; 36 and 34 percent for the whole-number-based algorithm; 20 and 18 percent for non-algorithmic written strategies; and 32 and 30 percent for strategies without any written work.

A latent class analysis identified four different patterns of strategy use on the division problems (the BIC was lowest for a model with four latent student and three latent teacher classes): 44 percent of the students predominantly used the whole-number-based algorithm (mean probability of using that strategy on the different items of .72); 23 percent of students used mainly non-algorithmic written strategies (mean probability of .55) and answering without written work (mean probability of .28); 18 percent mostly answered without any written work (mean probability of .87); and 15 percent predominantly used the digit-based algorithm (mean probability of .71).

Again, adding student characteristics to the null model improved it, $\chi^2(9) = 58.4$, $p < .001$, while the subsequent addition of condition effects did not provide further improvement, $\chi^2(12) = 15.3$, $p = .23$. In the model with student characteristics, gender was significantly related to strategy use (the probability of the whole-number-based algorithm pattern was lower for boys than for girls, while the probability of the answering without any written work pattern was higher), $W^2 = 18.0$, $p < .001$. Ability score also had a significant effect (it was positively related to the probability of the algorithm patterns and negatively to the probability of the non-algorithmic written and no written work patterns), $W^2 = 10.4$, $p = .02$. The interaction between gender and ability was not significant, $W^2 = 6.9$, $p = .07$.

6.4 Discussion

As the generalization of educational research to educational practice often involves the generalization of results from a single-task setting to a mixed-task setting, the present study compared students' problem solving in these two conditions. The tasks that students had to do were typical educational tasks from the domain of mathematics that are especially interesting with regard to strategy use: solving multidigit division problems that are intended to be solved with written, algorithmic strategies, and non-division mathematical problems that do not call for such strategies. Differences in performance and strategy use on these tasks in the single-task and mixed-task conditions could occur because of task switching costs (both in terms of accuracy and speed) and because of strategy perseveration.

However, no differences between conditions were found: accuracy and speed did not differ, and though different patterns of strategy use were identified, students were equally likely to have those patterns in both conditions. There were gender and ability level differences in accuracy, speed and strategy use, but there were no significant interaction effects of these student characteristics with condition. Therefore, no support was found for task switching costs with these educational tasks. Possibly, the larger complexity of educational tasks compared to typical tasks from the task switching literature (such as deciding whether a number is odd or even) makes switching costs negligible, as the task itself already requires the switching between many different sub-tasks.

There was also no indication of strategy perseveration. There was a group of students who quite consistently answered without any written work on not only the non-division, but also the division problems. This might have indicated perseveration if this strategy choice pattern occurred more often in the mixed-task condition (where the division problems were preceded by non-division problems that elicited this strategy) than in the single-task condition, but this was not the case. There were also two groups of students who quite consistently used the digit-based or whole-number-based algorithms for division, which might have indicated perseveration if this pattern occurred more often in the single-task (where all division problems were presented in a row) than in the mixed-task condition, but this was also not the case.

All in all, the results of the present study therefore suggest that a generalization of performance and strategy use in a single-task setting to a mixed-task setting is not necessarily problematic.

6.4.1 Limitations

However, the detection of possible task switching costs and strategy perseveration may have been hindered by some limitations in the design of the present study.

Task switching costs

A limitation that may have prevented the finding of task switching costs is that a comparison of complete single-task blocks with mixed-task blocks is quite crude. A more refined comparison could be made between problems directly after a switch (switch trials) and problems not directly after a switch (repeat trials; Kiesel et al., 2010). This is not possible with the current data, however, as a fair comparison requires that each problem features as often in a switch as in a repeat trial (otherwise, type of trial and problem difficulty are confounded). This would necessitate extra versions of the problem set.

Another limitation is that switching costs in terms of speed were investigated using the amount of minutes it took students to solve all (division and non-division) problems combined, rather than the much more precise amount of seconds per problem. The latter could be assessed by making students do the test on a computer or individually in the presence of an experimenter who records the time, but both these situations are likely to affects students' performance and strategy use.

Another issue is the type of tasks that were used. The tasks in the present study may have been so complex that switching costs became negligible, but educational practice also involves simpler tasks where this may not be the case. In addition, division problems were mixed with many other tasks from the domain of mathematics, whereas in the task switching literature usually just two tasks are contrasted. Doing the latter could make differences between conditions more pronounced, though it would also reduce the similarity to educational practice.

Strategy perseveration

There are also two factors that may have prevented us from finding strategy perseveration effects. One is that strategy perseveration has only been demonstrated in the context of a single task for which different strategies are most appropriate depending on the characteristics of the problem at hand (Lemaire & Lecacheur, 2010; Luwel, Schillemans, et al., 2009; Schillemans et al., 2009, 2011b). In contrast, in the present study mental, non-algorithmic strategy use was elicited with non-division problems and the effect of this on strategy use on division problems was evaluated. Perseveration within the division problems could still have occurred in the single-task condition, but presumably in the form of repeated use of a written algorithm, and since this strategy is most accurate for this type of problem (Fagginger Auer et al., 2013) that would not constitute persevering in using a suboptimal strategy.

In addition, both Schillemans et al. (2009) and Lemaire and Lecacheur (2010) did not find perseveration or strategy switch costs generally, but only for problems with specific characteristics. Lemaire and Lecacheur (2010) found strategy switch costs particularly for easier problems, while the division problems in the present study were difficult (45 percent correct solutions), so strategy perseveration may be found with easier educational tasks.

6.4.2 Conclusion

It can be concluded that the results of the present study do not indicate particular problems for the generalization of performance and strategy use in single-task experiments to mixed-task educational practice. However, less complex tasks may induce more switch costs and strategy perseveration, and several adjustments to the experimental set-up would allow for a more thorough investigation (though possibly at the cost of similarity of the experiment to educational practice).