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

— 4 —

Solution strategies and adaptivity in multidigit division in a choice/no-choice experiment: Student and instructional factors

Abstract

Adaptive expertise in choosing when to apply which solution strategy is a central element of current day mathematics, but may not be attainable for all students in all mathematics domains. In the domain of multidigit division, the adaptivity of choices between mental and written strategies appears to be problematic. These solution strategies were investigated with a sample of 162 sixth graders in a choice/no-choice experiment. Children chose freely when to apply which strategy in the choice condition, but not in the no-choice conditions for mental and written calculation, so strategy performance could be assessed unbiasedly. Mental strategies were found to be less accurate but faster than written ones, and lower ability students made counter-adaptive choices between the two strategies. No teacher effects on strategy use were found. Implications for research on individual differences in adaptivity and the feasibility of adaptive expertise for lower ability students are discussed.

4.1 Introduction

Learning and problem solving are characterized by the use of a variety of strategies at every developmental stage (Siegler, 2007). Children's and adults' strategy use has been investigated for cognitive tasks concerning diverse topics such as class inclusion (Siegler & Svetina, 2006), analogical reasoning (Tunteler et al., 2008), and digital gaming (Ott & Pozzi, 2012). A well-studied area of investigation in

This chapter has been published as: Fagginger Auer, M. F., Hickendorff, M., & Van Putten, C. M. (2016). Solution strategies and adaptivity in multidigit division in a choice/no-choice experiment: Student and instructional factors. *Learning and Instruction, 41*, 52-59.

We would like to thank the schools and students for their participation in the experiment.

solution strategy research is strategy use for arithmetic problems. Many studies have been conducted on strategies in elementary addition, subtraction, multiplication and division (e.g., Barrouillet & Lépine, 2005; Imbo & Vandierendonck, 2007; Mulligan & Mitchelmore, 1997; Van der Ven et al., 2012), which concern operations in the number domain up to 100 that are taught in the lower grades of primary school. However, there is a notable scarcity of research on strategy use of higher grade students on more complex arithmetic problems (though not an absence; see for example Van Putten et al., 2005; Selter, 2001; Torbeyns, Ghesquière, & Verschaffel, 2009). This more advanced arithmetic is called multidigit arithmetic, as it involves larger numbers and decimal numbers. Multidigit arithmetic is particularly interesting with regard to strategy use, as the higher complexity of the problems allows for the use of a wider range of strategies.

4.1.1 Solution strategies and adaptivity

To chart strategy use for a given domain, Lemaire and Siegler (1995) proposed a general framework consisting of four aspects of strategic competence: strategy repertoire (which strategies are used); frequency (how often each strategy in that repertoire is chosen for use); efficiency (performance with use of each strategy); and adaptivity (the appropriateness of a choice for a strategy given its relative performance). While the first three aspects of the framework are quite straightforward, the aspect of adaptivity has been conceptualized in various ways by different researchers. Verschaffel et al. (2009) reviewed the existing literature on this topic and identified three factors that play central roles in the different conceptualizations.

First there is the role of task variables, which concern the adaptation of strategy choices to problem characteristics. For example, for a problem such as $62 - 29$ the adaptive strategy choice could be defined as compensation (Blöte, Van der Burg, & Klein, 2001): the problem can be greatly simplified by rounding the subtrahend 29 to 30, and then compensating for this after the subtraction ($62 - 30 + 1$). Second there is the role of subject variables, which concern the adaptation of strategy choices to strategies' relative performance for a particular individual (for a particular problem), such as in the Adaptive Strategy Choice Model (ASCM; Siegler & Shipley, 1995). Third there is the role of context variables, which can be both in the direct context of the task (such as time restrictions) and in the broader sociocultural context (such as the value placed on accuracy versus speed). Verschaffel et al. (2009) combine all three factors (calling for more research attention for context variables especially) in defining a strategy choice as adaptive when it is most appropriate for a particular problem for a particular individual, in a particular sociocultural context.

A second issue in determining adaptivity is that often there is not one unequivocal best performing strategy, as the most accurate strategy is not always also the fastest. This can be addressed by combining speed and accuracy in a definition of the best performing strategy as the one that leads to the correct solution the fastest (e.g., Luwel, Onghena, et al., 2009; Torbeyns, De Smedt, et al., 2009; Kerkman & Siegler, 1997). Yet, even with this definition, researchers tend to consider accuracy and speed separately in their statistical analyses in practice (with the exception of Torbeyns et al., 2005).

4.1.2 Adaptive expertise in mathematics education

Debates of its exact definition aside, adaptivity has become more and more important in the educational practice of primary school mathematics. Reforms in mathematics education have taken place in various countries over the past decades (Kilpatrick et al., 2001) and they have reshaped the didactics for multidigit arithmetic from prescribing a fixed algorithmic strategy per problem type to building on students' own strategic explorations (Gravemeijer, 1997). For students, this means that performing well now requires more than perfecting the execution of a limited set of algorithmic strategies, because choosing the best performing strategy for solving a problem is also necessary. Adaptive expertise has become a central element of education: students should have an array of strategies at their disposal, that they can use efficiently, flexibly and creatively when they solve problems (Verschaffel et al., 2009). Investigations differ in their findings of whether such adaptivity is attainable for everyone: some have found evidence of a general adaptivity of strategy choices (e.g., Siegler & Lemaire, 1997; Torbeyns et al., 2005), while others found it only for students with a high mathematical ability (e.g., Hickendorff et al., 2010; Torbeyns, Verschaffel, & Ghesquière, 2006), and some not at all (e.g., Torbeyns, De Smedt, et al., 2009).

In addition to providing more space for informal strategies, the reforms introduced new standardized approaches for the more complex multidigit problems. With traditional algorithms the large numbers in such problems are considered one or two digits at a time, without an appreciation of the magnitude of those digits in the whole number being necessary, while new approaches place more focus on the whole number (as such, the former approaches have been labeled 'digit-based' and the latter 'whole-number-based'; Van den Heuvel-Panhuizen et al., 2009). Espe-

Table 4.1: Examples of applications of the different strategies on 850 ÷ 25.

digit-based algorithm	whole-number-based algorithm		repeated addition or subtraction		simplifying strategies
25/850\34	850 : 25 =		4 ×	100	850 ÷ 25
75	250 -	10 ×			= 3400 ÷ 100
100	600		32 ×	800	= 34
100	500 -	20 ×	2×	50	
0	100		34×	850	
	100 -	4 ×			
	0	34 ×			

cially for multidigit division, digit-based algorithms (e.g., long division) have been de-emphasized or even abandoned in favor of whole-number-based approaches (e.g., partial quotients; Buijs, 2008; Scheltens et al., 2013). Table 4.1 provides examples of digit-based and whole-number-based approaches for division: while they both consist of standardized steps with a schematic notation, the digit-based algorithm breaks the dividend up into digits (e.g., in Table 1, the 85 part of 850 is considered separately when subtracting 75, and the rest of the dividend is only considered in a later step), whereas the whole-number-based algorithm considers the dividend as a whole (e.g., 250 is subtracted from 850).

However, dismissing a digit-based algorithm does not necessarily mean that a whole-number-based algorithm will be used instead; an increase in the use of more informal, non-algorithmic strategies is also possible, even though they may be less suited for challenging problems. For example, the decrease in the use of the digit-based division algorithm in Dutch national assessments from 1997 to 2004 was paired by an almost equal increase in answering problems without writing down any calculations (Van Putten, 2005), which should be interpreted as mental calculation (Hickendorff et al., 2010). This switch from written to mental calculation turned out to be very unfortunate, as the probability for a student to solve a division problem accurately was drastically lower with mental than with written calculation (Hickendorff et al., 2009), and the overall performance level on multidigit division decreased sharply from 1997 to 2004 (J. Janssen et al., 2005). This trend over time of an increasing percentage of students choosing an inaccurate strategy suggests that the reform goal of adaptive expertise may not be feasible for some domains of mathematics.

4.1.3 The present study

The present study therefore constitutes an in-depth experimental investigation of adaptivity in this domain of mathematics that was particularly affected by the reforms: multidigit division. An experimental approach is necessary, because performance estimates of strategies may be biased by so-called selection effects (Siegler & Lemaire, 1997): for example, though mental strategies produce a low percentage of correct solutions for multidigit division problems, this performance estimate may be biased because of the mathematical ability level of the students who choose to use this strategy or because of the difficulty of the problems it is applied to. If mental calculation were used equally by all types of students on all types of problems, then a different estimation of its performance could very well result. Hickendorff et al. (2010) experimentally compared a condition in which students freely chose when to write down calculations and one in which they had to write down calculations for every problem, and found that written calculation was at least as accurate or more accurate than mental calculation, especially for weak students. Mental calculation, however, was only observed in this study when spontaneously chosen and therefore performance estimates were biased by selection effects. In addition, only accuracy and not solution times were measured, so the role of speed in strategy choices and adaptivity remained unclear.

The present study addresses these two issues by experimentally investigating students' spontaneous strategy choices for multidigit division and both their accuracy and speed with required written and required mental calculation. The participants are sixth graders, because the radical changes in performance and strategy use were demonstrated for this age group in the aforementioned large-scale assessment. The aim of the present study is to systematically chart the four aspects of strategic competence of Lemaire and Siegler (1995) - repertoire, frequency, efficiency and adaptivity - with special attention to adaptivity, because of its high relevance to mathematics education and to multidigit division specifically. This was done using the choice/no-choice paradigm introduced by Siegler and Lemaire (1997) to allow for the unbiased assessment of strategy performance characteristics, that has since been applied in numerous solution strategy investigations (e.g., Imbo & Vandierendonck, 2007; Lemaire & Lecacheur, 2002; Torbeyns et al., 2005).

This design consists first of a choice phase in which participants freely choose between strategies in solving a set of problems. This phase provides information on strategy repertoire and the frequency with which strategies in that repertoire are chosen. The choice phase is followed by a no-choice (NC) phase, with a separate

NC-condition for each strategy under investigation, in which participants have to solve (parallel versions of) the problems from the choice phase with that strategy. This provides strategy efficiency estimates unbiased by selection effects, as every participant has to solve each problem with each strategy. Adaptivity can be judged based on the two phases combined: it can be evaluated whether the strategies that the participant chose in the choice phase were the most accurate and fastest in the no-choice phase for him or her.

Hypotheses

There were several hypotheses regarding the four aspects of students' strategic competence in multidigit division. As for strategy repertoire and frequency, previous research indicates that a majority of the students predominantly use written calculation for multidigit division, sometimes with mental calculation for particular problems, while around one third predominantly uses mental calculation (Hickendorff et al., 2009). Girls may use more written calculation than boys, as girls use more algorithmic strategies, while boys tend to use more intuitive, less formal strategies (Carr & Jessup, 1997; Davis & Carr, 2002; Hickendorff et al., 2009). As for strategy efficiency, mental calculation was expected to be less accurate than written calculation (see section 1.2). The fact that mental calculation is used frequently despite its apparent inaccuracy, suggests that it may offer advantages in terms of speed.

As for adaptivity, it was expected that counter-adaptive choices with regard to accuracy would be made for mental rather than written calculation, given the apparent role of increased mental calculation in the Dutch performance decline. Considering the previously described differences in adaptivity for different levels of mathematical ability, this counter-adaptivity may occur particularly in lower ability students. Adaptivity with regard to the sociocultural context was expected, given the large influence on strategy choices that the sociocultural context exerts by defining what choices are appropriate, as described in a review on this topic by Ellis (1997). Among other factors, Ellis (1997) describes cultural values regarding the use of mental strategies and the originality of employed strategies as influential. In the present study, these values (and values regarding the digit-based versus the whole-number-based algorithm) were measured in the students' teachers, and we expected students' strategy choices to be related to these cultural values of the teacher.

Table 4.2: The three versions of the eight problems in the division problem set.

problem							
1	2	3	4	5	6	7	8
$47 \div 2$	$93 \div 4$	$810 \div 30$	$850 \div 25$	$136 \div 32$	$308 \div 14$	$216 \div 6$	$861 \div 7$
$87 \div 2$	$77 \div 4$	$510 \div 30$	$675 \div 25$	$175 \div 28$	$414 \div 18$	$231 \div 7$	$732 \div 6$
$67 \div 2$	$85 \div 4$	$720 \div 30$	$925 \div 25$	$189 \div 36$	$336 \div 16$	$306 \div 9$	$976 \div 8$

4.2 Method

4.2.1 Sample

A sample of 162 sixth graders (11-12-year-olds) from 25 different primary schools participated, of whom 81 were boys (50 percent) and 81 were girls (50 percent). Seventy-two of these students had a mathematical ability score below the national median (44 percent) and the remaining 90 a score above the median (56 percent), as measured by standardized national tests that are administered at most Dutch primary schools (J. Janssen, Verhelst, Engelen, & Scheltens, 2010).

4.2.2 Materials

Division problems

Three comparable versions of a set of eight multidigit division problems were constructed (see Table 4.2). The characteristics of the dividends and divisors were varied systematically: there were two problems with a two-digit dividend and one-digit divisor (e.g., $93 \div 4$); two problems with a relatively easy combination of a three-digit dividend and two-digit divisor (e.g., $850 \div 25$); two problems with a more challenging combination of a three-digit dividend and two-digit divisor (e.g., $308 \div 14$); and two problems with a three-digit dividend and a one-digit divisor (e.g., $861 \div 7$).

Teacher questionnaire

A questionnaire for the students' teachers was constructed to assess the values regarding arithmetic in the sociocultural context formed by the teacher (see Table 4.3). Two questions in the questionnaire concerned teachers' values regarding the type of division algorithm (digit-based or whole-number-based). The rest of the questionnaire focused on two values described as influential by Ellis (1997): men-

Table 4.3: The questions from the values questionnaire for the students' teachers.

Whole-number-based or digit-based algorithm
Which division algorithm best reflects the practice in your class?
whole-number-based - both - digit-based
To what extent do you as a teacher prefer a division algorithm?
strong preference whole-number-based - digit-based (5-point scale)
Mental versus written calculation
What is important to you when your students solve multidigit problems?
that they try that with mental calculation - written calculation (5-point scale)
How important is the skill of writing down calculations to you?
not important - very important (5-point scale)
How often do your students write down their calculations?
very infrequently - infrequently - sometimes - regularly - often
How important is advising students to write down calculations to you?
How important is instructing students in writing down calculations to you?
very unimportant - very important (5-point scale)
Original strategy use
How important is teaching students multiple solution strategies to you?
How important is letting students choose their own solution strategies to you?
very unimportant - very important (5-point scale)
How often do you devote attention to convenient solution strategies?
How often do you devote attention to multiple strategies per problem type?
$<$ *1/month - 1×/month - 2×/month - 1×/two weeks -* $\geq$ *1/week*

Note: Response options are in italics under the question(s) they apply to.

tal (as opposed to written) calculation (five questions) and originality of strategies (four questions). The three scales were found to have adequate reliability (Cronbach's alphas of .75, .75 and .65 for the algorithm, mental calculation and originality scales respectively). Validity was not separately investigated, but previous research indicates that teachers' self-reports of instructional practice converge with classroom observations of independent observers and that teachers feel that self-report measures can capture how they teach (Mayer, 1999; Martinez, Borko, & Stecher, 2012).

4.2.3 Procedure

Students were tested individually in a quiet room. They solved the three different versions of the same set of multidigit division problems according to a choice/no-choice design (Siegler & Lemaire, 1997). The students solved the first set of problems in the choice condition, in which they were free to choose whether they wanted to write down calculations or not. The second and the third problem set were offered in two NC-conditions: one in which the entire set had to be solved without writing down any calculations (the NC mental calculation condition), and one in which calculations had to be written down for every problem in the set (the NC written calculation condition). Both the order in which the different versions of the problem set were presented and the order of the NC-conditions were counterbalanced.

The solution time for each problem was recorded by the experimenter using a stopwatch. Student's strategy use on the division problems was inferred from their written work, and when no calculations were written down for a problem, students were interviewed on their solution strategy. Five different strategy categories were discerned (both within mental and written calculation; see Table 4.1 for examples): the digit-based algorithm; the whole-number-based algorithm (both algorithms were discussed in section 1.2); non-algorithmic strategies that involve repeated addition (or subtraction) of multiples of the divisor; strategies that involve a simplification of the problem (such as the compensation strategy discussed in section 1.1); and remaining solution strategies (unclear strategies, misconceptions such as multiplying rather than dividing, and guessing).

The students' teachers filled out the questionnaire on the day that the experimenter was present at the school for testing, and also solved one of the sets of eight division problems so that their free strategy use and performance could be assessed.

4.2.4 Statistical analysis

Binary logistic mixed models (e.g., Molenberghs & Verbeke, 2006) were used for analyzing the accuracy scores for each problem (correct or incorrect), strategy choices on each problem (mental or written calculation), and students' overall strategy choices in the choice condition (at least once or never mental calculation). Linear mixed models were used for analyzing the proportion of correct solutions with each version of the problem set, and the time students took to obtain the solution to each problem. This solution time was log-transformed to normalize its strongly skewed distribution (as in Klein Entink, Fox, & Van der Linden, 2009).

For analyses at the problem level, random effects were added for the students and the schools, to account for the dependencies of problem solving within students and within schools. For analyses at the student level, only random school effects were added. All mixed model analyses were carried out using the `SAS` procedure `GLIMMIX` (Schabenberger, 2005). Ninety-five percent confidence intervals (95% CIs) are reported for the regression coefficient estimates (which equal the log of the odds ratio (OR) in the logistic models) and differences in estimated means for an indication of the magnitude of the effects. In addition, the standardized versions of these mean differences (SMDs) are reported as effect sizes for the linear models (where values of 0.2, 0.5 and 0.8 can be considered to reflect small, medium and large effects respectively; J. Cohen, 1988), and ORs for the logistic models (where values of 1.5, 3.5 and 9.0 can be considered small, medium and large respectively; J. Cohen, 1988).

4.3 Results

The difficulty of the three versions of the problem set (aggregated over all conditions) was comparable: students did not differ significantly in their proportion of correct solutions for the first (M = .62) and second version (M = .62) of the problem set, $z = -0.37$, $p = .71$, 95% CI [-0.04, 0.03], SMD $= -0.01$, or for the first and third version (M = .59), $z = -1.84$, $p = .07$, 95% CI [-0.07, -0.01], SMD $= -0.07$, (and given the intermediate difficulty of the second version, also not for the second and third version).

4.3.1 Strategy repertoire and frequency

Table 4.4 provides information on students' strategy repertoire and the frequency of use of strategies in that repertoire in the three conditions of the choice/no-choice experiment. In the choice condition, students solved 29 percent of the problems using mental calculation, but this varied both between problems (from 18 percent of mental calculation for problem 6 to 56 percent for problem 1) and between students: 40 percent of the students never used mental calculation in the choice condition, 30 percent used it at least once but for less than half of the problems, and 30 percent applied it to half of the problems or more. There were no significant differences between students who did and did not use any mental calculation in the choice condition in terms of gender, $z = 0.24$, $p = .81$, 95% CI [-0.86, 1.10], OR $= 1.13$, or mathematical ability level, $z = 1.22$, $p = .22$, 95% CI [-0.36, 1.54], OR $= 1.80$, or interaction between gender and ability, $z = 0.78$, $p = .43$, 95% CI [-0.82, 1.90], OR $= 1.72$.

Table 4.4: Strategy use in the choice, NC-mental and NC-written calculation condition.

condition	mental/written	dig. alg.	num. alg.	rep. +/−	simp.	rem.
choice	mental (.29)	.01	.14	.46	.20	.18
	written (.71)	.13	.55	.22	.08	.02
NC	mental	.03	.23	.43	.19	.13
	written	.11	.49	.24	.11	.04

Note: dig. alg. = digit-based algorithm; num. alg. = whole-number-based algorithm; rep. +/− = non-algorithmic repeated addition or subtraction; simp. = simplifying strategies; rem. = remaining strategies

In the free strategy choice condition, algorithms (both digit-based and whole-number-based) were used much more often in written than in mental solutions. In contrast, non-algorithmic repeated addition or subtraction and simplifying strategies (and also remaining strategies) were more frequent in mental solutions. The strategy use within mental and within written solutions was similar in the choice and NC-conditions.

4.3.2 Strategy efficiency

We investigated the relative efficiency of mental and written calculation strategies by comparing students' performance in the NC-mental and NC-written calculation conditions (see Table 4.5 for accuracy and speed averages per condition). Students had a higher probability of solving a problem correctly in the NC-written calculation condition (probability of a correct solution (P) of .70) than in the NC-mental calculation condition (P = .54), $z = -7.48$, $p < .001$, 95% CI [-1.57, -0.92], OR = 3.48. For below median ability students, the difference in accuracy between NC-written and NC-mental calculation was much larger (ΔP = .26) than for above median ability students (ΔP = .06), z = 3.72, $p < .001$, 95% CI [0.34, 1.08], OR = 2.03. The accuracy difference did not depend significantly on student gender, z = 1.80, p = .07, 95% CI [-0.03, 0.71], OR = 1.40.

As for speed: students solved problems faster in the NC-mental calculation condition (estimated mean problem solving time of 38 s) than in the NC-written calculation condition (M = 50 s), $z = -5.52$, $p < .001$, 95% CI [-0.29, -0.14], SMD = -0.39. This speed difference was larger for boys (ΔM = 15 s) than for girls (ΔM = 11 s), $z = -2.47$, p = .01, 95% CI [-0.19, -0.02], SMD = -0.20, but did not depend significantly on students' ability level, $z = -1.15$, p = .25, 95% CI [-0.14,

Table 4.5: Efficiency of required mental and written calculation in the respective no-choice conditions.

		accuracy (proportion correct)		speed (problem solving time (s))	
		NC-mental	NC-written	NC-mental	NC-written
gender	girls	.51	.67	48	51
	boys	.56	.65	44	53
ability	below median	.36	.56	49	58
	above median	.68	.74	44	48
total		.53	.66	46	52

0.04], SMD = −0.09.

4.3.3 Strategy adaptivity

Student-level correlations

A first indication of adaptivity is a positive relation between the frequency with which a student chooses to use a strategy and the relative performance of that strategy for that student (e.g., more choices for written calculation by students for whom this strategy is generally more accurate and faster than mental calculation). To investigate whether such an adaptive association between strategy choices and performance exists, the total number of choices for written calculation by students was correlated with their relative accuracy with written calculation (number correct with NC-written minus that with NC-mental) and their relative speed (average solution time with NC-mental minus that with NC-written) using Spearman's rho. Students were found to adaptively choose more written calculation when it was relatively more accurate for them than mental calculation, $\rho = .35$, $df = 156$, $p < .001$, but not significantly so when it was relatively faster, $\rho = .08$, $df = 154$, $p = .32$ (though adaptivity with regard to speed was shown by the subgroup of higher ability students, $\rho_{above} = .24$, $df = 87$, $p = .03$).

Problem level adaptivity scores

However, such correlation analyses - though common in adaptivity investigations (e.g., Kerkman & Siegler, 1997; Siegler & Lemaire, 1997; Torbeyns, Ghesquière, & Verschaffel, 2009; Torbeyns, De Smedt, et al., 2009; Torbeyns et al., 2006) - only reveal general trends at the student level and do not utilize the information that is available at the problem level in a choice/no-choice experiment, where comparisons can be made between the different strategy conditions in which parallel versions of a single problem are presented. In addition, correlation analyses consider accuracy and speed in isolation, while it is more educationally relevant to consider them simultaneously and define a choice as adaptive when it is for the strategy that produces the correct solution the fastest (as in Luwel, Onghena, et al., 2009; Torbeyns, De Smedt, et al., 2009; Kerkman & Siegler, 1997). Following this definition, the following problem-level adaptivity judgments can be made: when one no-choice strategy was accurate and the other no-choice strategy inaccurate on parallel versions of the same problem for a student (e.g., NC-written correct and NC-mental incorrect on two of the versions of problem 5), a choice for the accurate strategy (in this example, written) on the other version of the problem by that student in the choice condition was defined as adaptive, and a choice for the inaccurate strategy (in this example, mental) as counter-adaptive. When both strategies were accurate, a choice for the faster strategy was defined as adaptive and a choice for the slower strategy as counter-adaptive. The case of two incorrect NC-solutions is undetermined, as then there is no 'best' choice to speak of.

Disregarding the undetermined trials (34 percent of all trials), 62 percent of choices were found to be adaptive using these criteria (of which 67 percent were for written strategies) and 38 percent counter-adaptive. This considerable percentage of counter-adaptive strategy choices was found to hardly vary over gender and ability subgroups (between 61 to 66 percent), though the percentage of undetermined trials was considerably higher in lower (40 percent) compared to higher ability students (29 percent) because of the larger proportion of incorrect answers in the lower ability students.

Relative performance with free strategy choice and required written calculation

In the introduction, it was suggested that requiring students to write down calculations might improve their performance. Therefore, we also investigated the adaptivity of students' strategy choices at the problem level in the following way: we examined whether students performed better in solving a problem when they were required to write down calculations, compared to when they were free to choose whether they wanted to. Table 4.6 shows students' accuracy and speed in the choice and NC-written conditions, separately for mental and written strategy choices in the choice condition. There was no general significant effect of condition on accuracy, $z = 1.26$, $p = .21$, 95% CI [-0.15, 0.67], OR = 1.30, but condition did interact with students' strategy choice in the choice condition, $z = 2.15$, $p = .03$, 95% CI [0.07, 1.53], OR = 2.23. There was also an interaction of condition, strategy choice and ability level, $z = -1.95$, $p = .05$, 95% CI [-1.62, 0.00], OR = 2.25: when below median ability students chose mental calculation, their accuracy improved with NC-written calculation (increase in probability of a correct solution of .14), which was not the case for students with an above median ability ($\Delta P = -.02$). When students chose written calculation in the choice condition, accuracy was largely unaffected by condition, both for below median ability students ($\Delta P = .01$) and above median ability students ($\Delta P = .03$). Condition, strategy choice and gender did not interact significantly, $z = -1.33$, $p = .18$, 95% CI [-1.36, 0.26], OR = 1.73.

Requiring written calculation affected speed, $z = -2.42$, $p = .02$, 95% CI [-0.19, -0.02], SMD = -0.22, and this condition effect interacted with strategy choice, $z = 7.51$, $p < .001$, 95% CI [0.43, 0.74], SMD = 1.23: when students chose mental calculation in the choice condition they were slower in the NC-written condition ($\Delta M = 19$ s), which did not hold when students chose written calculation ($\Delta M = -2$ s). This slowing effect of NC-written calculation when students chose mental calculation was stronger for higher ability students ($\Delta M = 21$ s) than for lower ability students ($\Delta M = 17$ s), $z = 2.59$, $p = .01$, 95% CI [0.05, 0.39], SMD = 0.46. Condition, strategy choice and gender did not interact significantly , $z = -0.65$, $p = .51$, 95% CI [-0.22, 0.11], SMD = -0.11.

Teachers' effects on strategy choices

No significant teacher effects on students' choices between mental and written calculation in the choice condition were found. Firstly, there were no significant effects of teacher's responses on the teacher questionnaire. To investigate this, mean scores were calculated for the responses per category (with one question transformed to a five-point scale). For the questions on the whole-number-based versus digit-based algorithm, these mean scores showed that teachers were on average more oriented towards the whole-number-based approach (M = 2.20, SD = 1.33), but these scores had no significant effect on students' use of mental calculation, $z = -0.42$, $p = .68$, 95% CI [-0.53, 0.34], OR = 1.10. Mean scores for the questions on mental versus written computation showed that teachers on average considered written computation more important (M = 4.27, SD = .49), but these scores also had no significant effect, $z = 0.26$, $p = .80$, 95% CI [-0.96, 1.25], OR = 1.16. Mean scores for the questions on originality showed that teachers on average found orig-

Table 4.6: Performance in terms of accuracy and speed with free strategy choice and NC-written calculation, split by strategy choice in the choice condition.

		accuracy (proportion correct)			
		mental choice		written choice	
		choice	NC-written	choice	NC-written
gender	girls	.52	.64	.68	.68
	boys	.66	.59	.68	.68
ability	lower	.40	.50	.57	.58
	higher	.73	.69	.73	.76
total		.60	.62	.66	.68
		speed (problem solving time (s))			
		mental choice		written choice	
		choice	NC-written	choice	NC-written
gender	girls	29	48	60	52
	boys	25	46	59	57
ability	lower	33	51	64	60
	higher	23	44	55	49
total		27	47	59	54

inality important (M = 4.04, SD = .76), but these scores also had no significant effect, $z = -0.21$, $p = .84$, 95% CI [-0.95, 0.76], OR = 1.09. Secondly, there were no significant effects of how the teachers solved the eight problems: neither for the number of times a teacher used mental calculation (M = 2.13, SD = 2.03), $z = 0.10$, $p = .92$, 95% CI [-0.27, 0.30], OR = 1.01; nor for the number of correctly solved problems (M = 6.61, SD = 1.12), $z = -0.99$, $p = .33$, 95% CI [-0.79, 0.26], OR = 1.30.

4.4 Discussion

In this study, students' mental and written solution strategies for multidigit division problems were investigated. Using the choice/no-choice paradigm, the four dimensions of strategy use proposed by Lemaire and Siegler (1995) were charted: repertoire, frequency, efficiency, and adaptivity. The repertoire that students demonstrated contained mental strategies for more than half of the students, half of whom applied it to a majority of the problems. In line with the more informal nature of mental strategies (Blöte, Klein, & Beishuizen, 2000), mental strategies were found to be non-algorithmic and simplifying more often than written strategies. As expected, mental strategies were found to be faster but less accurate than written strategies, and earlier estimates of the inaccuracy of mental strategies (Van Putten, 2005) were probably even still too optimistic because of selection effects (Siegler & Lemaire, 1997): the percentage correct difference between mental and written strategies was smaller in the free strategy choice condition (6 percentage points) than in the unbiased NC-conditions (13 points).

We first investigated adaptivity by evaluating the degree to which students adapted their strategy choices to their relative performance with these strategies. Using student-level correlations, students were found to adaptively choose written strategies more when these were relatively more accurate for them than mental strategies, and above median ability students also when written strategies were relatively faster. However, using problem-level adaptivity scores that labeled a strategy choice as adaptive when it was for the fastest accurate strategy, we found that a considerable portion of the strategy choices was counter-adaptive (around a third), also for higher ability students. Particular counter-adaptivity was indicated for lower ability students who chose mental calculation, as their accuracy improved when they were required to write down calculations - which was also found by Hickendorff et al. (2010), though they did not find the effect to depend on ability level.

Following the suggestion of Verschaffel et al. (2009) of high importance of the sociocultural context, we also devoted attention to adaptivity in the sense of adaptation of solution strategies to that context in the form of the students' teachers attitudes towards various aspects of strategy use and teachers' own strategy application. However, we found no significant effects, suggesting that students' division strategy use may not be very sensitive to that context, at least to the extent that that context is shaped by their current teacher. Other studies of sociocultural context effects which operationalized that context more broadly did find effects, for example by including parents in addition to teachers (Carr & Jessup, 1997), or contrasting vastly different contexts in which Brazilian children functioned as a street vendor or as a student (Nunes, Schliemann, & Carraher, 1993). Sociocultural effects on mental division strategy use might therefore be found by taking broader approaches such as also including teachers from earlier stages of mathematics learning instead of only the teacher from the final year of primary school, or by also including less formal sociocultural influences such as parents and peers. Contrasting distinct contexts could be achieved by comparing mental strategy use in different countries.

Several interesting individual differences were found. Mental strategies offered boys a larger speed advantage relative to written strategies than they did for girls, which could contribute to the finding of Hickendorff et al. (2009) that boys use mental strategies more than girls (though we did not replicate that finding). As for ability level, while the rate of choices for mental strategies did not differ significantly between levels, the accuracy advantage of written compared to mental strategies was larger for lower than for higher ability students, and lower ability students demonstrated less adaptivity (as in several other studies, e.g., Foxman & Beishuizen, 2003; Hickendorff et al., 2010; Torbeyns et al., 2006). These results indicate that mental strategies are especially risky for lower ability students: not only are these strategies especially inaccurate for this group, these weaker students also appear to have problems with determining when they should and should not be applied. What makes this especially worrisome, is that lower ability students nonetheless appear to use mental strategies as often as higher ability students (or even more often, as found by Hickendorff et al., 2009).

The finding that lower ability students benefit from being required to write down calculations while higher ability students do not (who instead are slowed down more) is in line with the expertise reversal effect in cognitive load theory, which states that instructional techniques can have differential (and even reversed) effects on cognitive load (and thereby, performance) depending on the expertise of the learner (Kalyuga, Ayres, Chandler, & Sweller, 2003). In low-expertise students, writing down calculations may free working memory resources for division problems that otherwise pose a cognitive load that is too high, whereas in high-expertise students, writing down calculations may be a redundant activity that places an unnecessary extra load on those resources. Such an expertise reversed effect implies that this technique of requiring writing down calculations should only be used for expertise levels for which it is effective: lower ability students.

4.4.1 Methodological considerations

Two aspects of the methodology of the current investigation warrant further attention. The first is the strategies evaluated in the choice/no-choice experiments. The choice/no-choice paradigm is often employed to compare specific strategies such as direct subtraction and indirect addition (Torbeyns, Ghesquière, & Verschaffel, 2009), but in our case broader categories of strategies are compared. As criticized by Luwel, Onghena, et al. (2009), such broad categories can in turn consists of several strategies, which is indeed the case here (both written and mental strategies are further classified into five categories). However, we argue that our comparison of mental and written strategies is very meaningful in light of their large performance difference and the apparently important role of this difference in performance level changes (as discussed in the introduction), and note that in their introduction of the choice/no-choice paradigm, Siegler and Lemaire (1997) also compared mental and written strategies. In addition, comparing more specific division strategies in the Dutch situation is complicated, as the division strategies that Dutch students are taught differ and therefore not all students can be expected to be able to execute particular strategies.

The second methodological aspect is the statistical conceptualization of adaptivity. In this study different approaches were taken, that each shed their own light on the degree of adaptivity displayed by (subgroups of) students. A first consideration is the level at which adaptivity is evaluated. Performance on individuals problems may be unreliable, making aggregating over problems necessary for stable results Luwel, Onghena, et al. (2009), but this discards a lot of information and treats problems as if they are interchangeable, which even within a domain as specific as multidigit division seems unreasonable (e.g., see Table 4.2). The ASCM posits that strategy choices are based on a weighted combination of global data averaged over all problems, featural data per particular structural problem feature, and local data per particular problem, and that the weighing depends on the familiarity of the problem (Siegler & Shipley, 1995). Therefore, it might be argued that aggregating over all problems is more suitable when problems are relatively unfamiliar, and less so when problems or problem features are more familiar. The present study appears to lie somewhere in between, as multidigit division should by a very familiar domain for students, but particular problems are typically not repeatedly encountered, and both problem and individual approaches were taken.

A second consideration in determining adaptivity - also touched upon in the introduction - is whether to consider accuracy and speed in isolation or to combine them. In this study both approaches were taken, and for the combination speed was only considered when both strategies were accurate, defining choices for the fastest accurate strategy as adaptive (as in Luwel, Onghena, et al., 2009; Torbeyns, De Smedt, et al., 2009; Kerkman & Siegler, 1997). Trials where both NC-solutions were inaccurate remained undetermined, which is not the case when speed is considered in isolation, but one could question whether speed differences for inaccurate strategies are as relevant as those for accurate strategies. All in all, we urge investigators of adaptivity to be aware that different choices with regard to analysis level and accuracy and speed can highlight different aspects of adaptivity.

4.4.2 Implications

The findings of this study have implications for cognitive psychological research on solution strategies and for educational practice. As for cognitive research, we found that written strategies appear to be chosen more for their accuracy, while mental strategies appear to be chosen more for their speed. These considerations did not play in equal measure in everyone: the strength of the accuracy effect depended on mathematical ability level and the strength of the speed effect on gender, and differences in adaptivity indicated that students differ in the extent to which they adapt their strategy choices to strategies' accuracy and speed. Choices between mental and written calculation may therefore in part be determined by individual differences in the relative value assigned to accuracy and speed, and therefore in part reflect students' speed-accuracy tradeoff (MacKay, 1982). Factors that may play a role in the relative favoring of accuracy and speed are traits which are traditionally associated with academic success (Bembenutty, 2009): academic delay of gratification (which is generally higher in girls) parallels sacrificing speed for accuracy, and self-efficacy (higher in boys) could determine the speed that students allow themselves while still feeling confident about their accuracy. Future research on adaptivity could extend existing models such as the ASCM (Siegler & Shipley, 1995) to accommodate individual differences in preferences for accuracy and speed, and provide more insight into the sources of these individual differences by relating them to other factors such as the ones discussed.

As for educational practice, results suggest that for some students it may be too ambitious to strive for what is a central element of mathematics reforms: adaptive expertise in choosing from an array of formal and informal strategies, rather than mastery of a limited set of algorithmic strategies. We found that lower ability students appear to use mental strategies as often as higher ability students, while mental strategies are especially inaccurate for them and adaptivity in choosing when to apply these strategies appears problematic. Lower ability students' performance may therefore be improved by providing them with more direction in their strategy choices. The present study provided support for beneficial effects of doing this directly by simply requiring students to write down calculations, and a broader change in strategy behavior might be accomplished by targeting the sociocultural context (Verschaffel et al., 2009).

As described in a review by Ellis (1997), cultural values in this context concerning various aspects of problem solving exert an important influence on children's strategy choices. She discusses values regarding speed and accuracy, mental strategies, originality, and independent performance, and contrasts these values in different cultures (such as Western cultures as apposed to Navajo, Asian and aborigine cultures). Given the results of the present study, values for speed and accuracy and mental strategies appear especially relevant to performance in multidigit arithmetic. The suboptimal choices for mental strategies that we have observed may be related to typical Western values in these areas: the favoring of fast performance (rather than error-free performance, as for example in the Navajo culture), and of solutions constructed in the head without any external aids. Therefore, performance might be improved by making efforts to adjust these norms so that accuracy is more important than speed, and that solutions constructed in the head are not more desirable than those constructed with the external aids of paper and pencil. The results of the present study suggest that such an adjustment may require a broader approach of sociocultural effects than just the students' current teacher. All in all, we feel that it would be highly relevant for mathematics education to devote more research efforts to investigating the feasibility of the educational goal of adaptive expertise for lower ability students, and evaluating sociocultural influences more broadly to see how strategy choices may be favorably influenced.