Explanatory latent variable modeling of mathematical ability in primary school : crossing the border between psychometrics and psychology

Hickendorff, M.

CHAPTER 5

Solution strategies and adaptivity in complex division: A choice/no-choice study

This chapter is co-authored by Marije F. Fagginger Auer.

ABSTRACT

The current study systematically investigated mental and written solution strategies for solving complex division problems (e.g., 306÷17), with the main focus on strategy adaptivity. Eighty-six Dutch 12-year-olds were tested using the choice/no-choice design. They first solved division problems in the free strategy choice condition, and consecutively with forced mental and forced written computation in the two respective no-choice conditions. Strategy choice and strategy performance (accuracy and speed) were recorded. Findings showed that mental computation was usually chosen for reasons of speed, while choices for written computation were fit to accuracy characteristics. Moreover, there were group differences regarding gender and mathematics achievement level of the student in the relative preference for accuracy and speed in choosing between mental and written strategies.

5.1 INTRODUCTION

Solution strategies for solving cognitive tasks have been an important psychological research topic. Especially solution strategies for mathematics problems have received considerable attention, since they are interesting both from a cognitive psychological perspective and from the viewpoint of mathematics education. Until recently, these studies were mostly limited to elementary addition, subtraction, and multiplication in the number domain up to 100. In contrast, complex arithmetic that is part of the curriculum of higher grades of primary school (i.e., operations with multidigit numbers for which one may use a written procedure) – and particularly complex division – has not received much research attention. However, systematic studies in complex division are needed, particularly in the Netherlands. Dutch national assessments showed a descending achievement trend on complex arithmetic in general, and on complex division in particular (J. Janssen et al., 2005), and this trend appears to be related to a shift in strategy use from written to mental strategies (Hickendorff et al., 2009b).

Therefore, the present study aims at a systematic investigation of the characteristics of mental and written solution strategies Dutch children at the end of primary school use to solve complex division problems, with a special focus on adaptivity: to what extent do the children choose the strategy (mental or written) with which they perform best? In the remainder of this section, we discuss research into solution strategies including strategy adaptivity, and previous research in the domain of complex division. We end this section with the design and aims of the current study.

5.1.1 Solution strategies

Children and adults know and use multiple strategies to solve cognitive tasks, including mathematics problems. The many studies into solution strategies for elementary addition and subtraction (e.g., Carr & Jessup, 1997; Carr & Davis, 2001; Torbeyns et al., 2002, 2004a, 2005), elementary multiplication (e.g., Anghileri, 1989; Imbo & Vandierendonck, 2007; Lemaire & Siegler, 1995; Mabbott & Bisanz, 2003; Mulligan & Mitchelmore, 1997), and mental multidigit addition and subtraction (e.g., Beishuizen, 1993; Beishuizen et al., 1997; Blöte et al., 2001; Torbeyns et al., 2006) have resulted in near consensus on the strategies used and the characteristics thereof for these mathematical domains.

Research into strategies for solving mathematics problems has been carried out in the field of cognitive psychology and in the field of mathematics education. Cognitive psychology acknowledges that arithmetic performance depends on the type of strategies that a subject uses (Lemaire, 2010). Within the cognitive psychological framework, the work of Siegler and his colleagues has been very influential (e.g., Lemaire & Siegler, 1995; Shrager & Siegler, 1998; Siegler, 1988a, 1988b). Lemaire and Siegler distinguished four dimensions of strategic competence on which individuals may differ: their strategy repertoire (which strategies are used), their strategy distribution (the frequency with which the strategies are used), their strategy efficiency or performance (strategy speed and/or accuracy), and their strategy selection or adaptivity (how strategies are chosen, related to problem and individual strategy characteristics). These four dimensions are central to the current study, with the main focus on the last one: strategy adaptivity.

Cognitive models of the underlying structures and mechanisms of strategy choice or adaptivity have been developed (Shrager & Siegler, 1998; Siegler & Shipley, 1995). In these models, an individual's strategy choice on a particular problem is for the largest part determined by the individual's strategy performance characteristics for that problem. According to Siegler and Lemaire (1997), people tend to choose their strategies adaptively: they choose the fastest and most accurate strategy for a given problem out of their strategy repertoire. Strategy speed and accuracy on a particular task may vary from individual to individual. However, not all research findings support this cognitive claim on the adaptivity of strategy choices. For example, suboptimal strategy choices have been observed in 2-digit addition and subtraction (Torbeyns, De Smedt, et al., 2009b) and in complex division (Hickendorff et al., 2010). Moreover, cognitive models on strategy choice have been argued to ignore the influence of sociocultural context variables such as sociomathematical norms (Ellis, 1997; Luwel, Onghena, Torbeyns, Schillemans, & Verschaffel, 2009; Verschaffel et al., 2009).

From the perspective of mathematics education, solution strategies are important in the international reform movement (e.g., Kilpatrick et al., 2001) for at least two reasons. First, the didactics for solving complex arithmetic problems have changed, from instructing standard written algorithms to building on children's informal strategies (Freudenthal, 1973; Treffers, 1987, 1993), and mental computation has become very important (Blöte et al., 2001). Second, mathematics education reform aims at attaining adaptive expertise instead of routine expertise: instruction should foster the ability to solve mathematics problems efficiently, creatively, and flexibly, with a diversity of strategies (Baroody & Dowker, 2003; Torbeyns, De Smedt, et al., 2009b). It is worth mentioning at this point that the terms 'adaptivity' and 'flexibility' are used with different meanings by different authors (for a discussion, see Heinze, Star, & Verschaffel, 2009, and Verschaffel et al., 2009). In the present study, adaptivity is defined with respect to both individual strategy performance characteristics and task characteristics, in the following way: to what extent does a child choose the strategy that is the most appropriate or efficient for him or her on a given problem?

Although this conceptualization of strategy adaptivity is used in the literature (e.g., Heinze, Star, & Verschaffel, 2009; Star & Newton, 2009; Torbeyns, De Smedt, et al., 2009b), it is not particularly well-defined, because what constitutes 'appropriate' or 'efficient' is ambiguous. These terms usually refer to the performance of a strategy, but there are at least two components to strategy performance: accuracy and speed. Problems arise when the most accurate strategy on a problem is not the fastest. For example, backup-strategies are slower but can be more accurate than retrieval (e.g., Kerkman & Siegler, 1997; Lemaire & Siegler, 1995; Siegler & Lemaire, 1997), and on complex arithmetic problems it has been suggested that written strategies are more accurate but slower than mental strategies (Hickendorff et al., 2010). In these instances, an adaptive strategy choice is not univocally defined. Some researchers leave the relative importance of accuracy and speed rather unspecified by defining the most efficient strategy as *the fastest and most accurate* (e.g., Lemaire & Callies, 2009; Siegler & Lemaire, 1997). Obviously, such a definition does not accommodate for the situations where one strategy is faster, but another strategy is more accurate. Other researchers defined the most efficient strategy as the one *leading fastest to the correct answer* (e.g., Kerkman & Siegler, 1997; Luwel et al., 2009; Torbeyns, De

Smedt, et al., 2009b; Torbeyns et al., 2004a, 2005, 2006) or the strategy that *produces the most beneficial combination of speed and accuracy* (Verschaffel et al., 2009). Although the latter definitions combine accuracy and speed, in the operationalizing analyses by these researchers accuracy and speed were generally not considered simultaneously but separately instead (an exception is the study by Torbeyns et al., 2005).

So, the relative importance of accuracy and speed plays a role in situations where the most accurate strategy is not the fastest. Moreover, individuals may differ in their relative favoring of accuracy and speed: they may have different speed-accuracy preferences (Ellis, 1997; Phillips & Rabbitt, 1995). In other words, they may differ in which combination of speed and accuracy they find most beneficial. Such considerations have not received much research attention in the research on strategy adaptivity. For elementary cognitive tasks, Siegler (1988a) discusses individual differences in the strength of the confidence criterion (i.e., the certainty required for stating an answer from retrieval), which relates to the individual differences in motivation to make few errors. In such elementary cognitive tasks backup strategies are clearly slower but evenly or more accurate than retrieval. Such clear performance differences do not necessarily exist in more complex cognitive tasks, in which relative strategy accuracy and speed may differ from individual to individual. In the current study, we try to gain insight into different patterns in the relative favoring of strategy accuracy and strategy speed in complex division problem solving. Such insight may have important educational implications, since instruction may be adapted to these individual differences. For instance, students who favor speed over accuracy may be encouraged to work slower but with fewer errors.

5.1.2 Complex division

In the present paper, complex division is defined as division problems in which the quotient is a multidigit number (e.g., $872 \div 4 = 218$)[1], and the divisor may be multidigit too (e.g., $306 \div 17 = 18$). This contrasts with simple division (division problems from the multiplication tables), in which the quotient is a single digit (e.g., $48 \div 8 = 6$; Robinson et al., 2006). Compared to addition, subtraction, and multiplication, the domain of (complex) division is understudied thus far. However, systematic studies on complex division are needed, particularly at the end of primary school in the Netherlands, for at least two reasons.

1 We consider a decimal number (e.g., $34 \div 4 = 8.5$) as multidigit too.

First, the most recent Dutch national assessment showed a large decline in sixth graders (12-year-olds) performance on complex division problems over a period of two decades (J. Janssen et al., 2005). Second, mathematics education reform has had a considerable impact on instruction in complex division. Under the influence of Realistic Mathematics Education (RME) the traditional long division algorithm has disappeared from mathematics textbooks, and has been replaced by more informal strategies based on repeatedly adding or subtracting multiples of the divisor (Freudenthal, 1973; Treffers, 1987). Figure 5.1 presents examples of such repeated addition and subtraction strategies, that differ in their level of abbreviation (i.e., the number of steps taken), see also Hickendorff et al. (2010) and Van Putten et al. (2005). Moreover, the traditional long division algorithm (and its notational form in the Netherlands and the US) is also presented. In addition to this shift in instruction in written strategies, another characteristic of the reform is that mental arithmetic plays a central role in mathematics education (Blöte et al., 2001). In 2004, nearly all Dutch primary schools used mathematics textbooks based on RME principles (J. Janssen et al., 2005), although a return to more traditionally oriented mathematics textbooks has been observed recently (KNAW, 2009).

In a recent study, secondary analyses on the student materials of the two most recent national assessments of 1997 and 2004 were carried out, aiming to relate the achievement decline on complex division to (changes in) the solution strategies used (Hickendorff et al., 2009a, 2009b). Results showed that two changes appeared to have contributed to the decline. First, strategy use had shifted: use of written procedures decreased (attributable to a decrease in the use of the traditional long division algorithm), while an increasing percentage of the students (more boys than girls) predominantly answered without calculations written down on scrap paper. This strategy shift was unfortunate, since answering these problems with a nonwritten strategy was less accurate than using a written strategy. Second, each of the solution strategies yielded less correct answers in 2004 than in 1997, with approximately the same amount of accuracy decrease per strategy. So, the performance decline over time on complex division in the Netherlands seems to be related to a change in strategy choice – in particular to a decrease in the use of written strategies – and to a general decrease in strategy accuracy as well.

These strategy change results of the national assessment data were descriptive by nature and therefore limited in several aspects, among which possible selection effects (cf. Siegler & Lemaire, 1997). That is, because strategy choice is probably influenced by the ability of the student and/or difficulty of the item, the strategy accuracy estimates

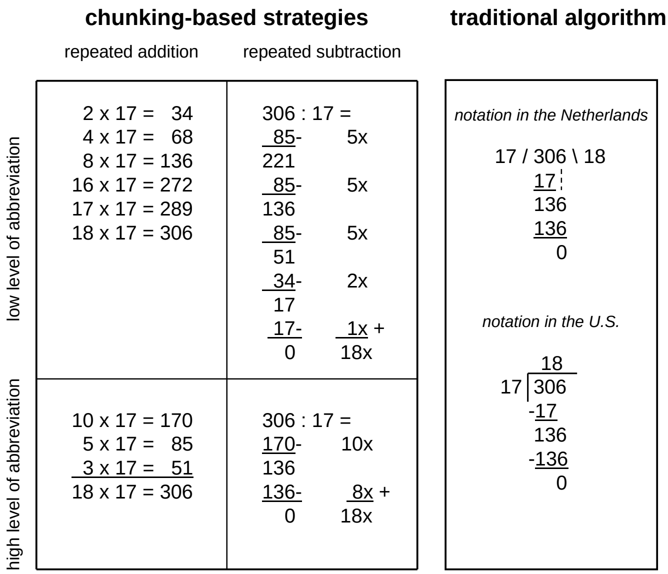

FIGURE 5.1 *Examples of solution strategies for the problem 306 ÷ 17.*

may have been biased by these student and item selection effects. To overcome this, Hickendorff et al. (2010) studied strategy choice in an experimental test design, in which sixth graders had to solve division problems under two conditions: free choice between mental and written computation, and forced written computation. One of the main findings was that accuracy of students who used mental calculation on a particular item in the free choice condition, improved by requiring the use of a written strategy in the forced written condition. So, these findings suggest that these choices for mental strategies were counter-adaptive with regard to accuracy. However, the methodology of this study hampered drawing conclusions on adaptivity rigorously for two reasons. First, data were collected on strategy accuracy but not on strategy speed, so only one aspect of strategy performance could be accounted for. Second, unbiased strategy characteristics (i.e., accuracies) were gathered only for written strategies and not for mental strategies, since it was deemed to be too demanding for a large number of students to solve complex division problems with large numbers with obligatory mental calculation. As a result, unbiased strategy characteristics of only one of the two strategies could be used in assessing adaptivity.

The current study extends the finding of Hickendorff et al. (2010) in a follow-up experiment in which these two main methodological limitations are overcome. Specifically, we also included speed measures in addition to accuracy data, as well as a condition in which students were forced to use mental computation in addition to a forced-written strategy use condition. In order to prevent students becoming frustrated from having to solve quite difficult problems in their heads, the presents study's division problems were designed to be somewhat less cognitively demanding compared to the ones used in Hickendorff et al. (2010).

5.1.3 The current study

The present study's aim is to systematically investigate the four dimensions of strategic competence in the domain of complex division problem solving (repertoire, distribution, performance, and adaptivity), distinguishing between mental and written computation strategies. Particularly the fourth dimension, adaptivity of the strategy choices, received special attention: to what extent do individual strategy performance characteristics (accuracy and speed) predict the choice of a strategy? We expected that different patterns in preference for accuracy and speed would be present, giving rise to different patterns of strategy adaptivity. Such findings may have implications relevant for educational practices, since students favoring speed over accuracy may require another instructional approach than students favoring accuracy over speed.

The main focus regarding solution strategies was on the distinction between written and mental computation, because secondary analyses on Dutch national assessments at the end of primary school showed that this was a very relevant distinction with respect to the observed decrease of performance over time: use of mental strategies increased over time, but their success rates lagged far behind those of written strategies (Hickendorff et al., 2009b). In the present study, mental computation was defined as carrying out arithmetical operations without the use of any recording devices such as pen and paper, similar to the definitions of for example Hickendorff et al. (2010), Ruthven (1998), Siegler and Lemaire (1997), and Timmermans et al. (2007), but unlike other studies (e.g., Beishuizen et al., 1997; Blöte et al., 2001; Torbeyns, De Smedt, et al., 2009b). Written strategies included all forms of calculations in which some part of the solution process was written down on paper, ranging from only recording intermediate solution steps to written algorithmic procedures. In order to be able to extend conclusions of the present experiment to the earlier studies on strategy use on complex division problems, participants and problems were chosen to resemble those of the Dutch national assessments. That is, students at the end of primary school (sixth graders) were selected, and we devised division problems presented in a realistic context involving multidigit quotient and/or divisor.

To estimate strategy characteristics in an unbiased manner, we used Siegler and Lemaire's (1997) *choice/no-choice* methodology. Each participant solved three parallel series of division problems under three different conditions. They first solved a series of division problems in the free strategy choice condition, and consecutively solved the two parallel series with forced mental and forced written calculation in the two respective no-choice conditions. From these two no-choice conditions, individual accuracy and speed characteristics of written and mental computation strategies were assessed without selection effects. Several studies in mathematics have applied the choice/no-choice methodology to study solution strategies (for an overview, see Luwel et al., 2009). An important feature of this design is that the adaptivity of strategy choices can be evaluated on an individual level, i.e., whether a subject chooses that strategy that *for him or her* is most efficient.

In addition to assessment of the four dimensions of strategic competence in general, we searched for effects of the student characteristics gender and general mathematics level. Gender differences in strategy use have been reported frequently. For example, girls have been found to have a larger tendency than boys to (quite consistently) use algorithmic strategies instead of using more intuitive, less structured strategies (Carr & Davis, 2001; Carr & Jessup, 1997; Hickendorff et al., 2009b, 2010; Gallagher et al., 2000; Timmermans et al., 2007). In contrast, gender differences in strategy adaptivity not been studied often, thus far. However, the findings of Hickendorff et al. (2010) showed boys making less adaptive strategy choices than girls, at least regarding accuracy, and it was suggested that girls and boys may weigh the importance of accuracy and speed differently. Regarding mathematics achievement level, it has been frequently (but not uniformly, see Torbeyns et al., 2005) reported that students of higher mathematical ability choose more adaptively between strategies than students of low mathematical ability (Foxman & Beishuizen, 2003; Hickendorff et al., 2010; Torbeyns, De Smedt, et al., 2009b; Torbeyns et al., 2002, 2006). So, we expect to find the same pattern in the current study.

5.2 METHOD

5.2.1 Participants

The participants were 86 students in the sixth grade (12-year-olds). They originated from 9 Dutch primary schools located either in the city or in a more rural area. All schools used a mathematics textbook based on RME principles, but they did not use the same textbook. In the sample, there were 43 girls and 43 boys. Also, information about the general mathematical level of the students was obtained: the students most recent level on CITO's Student Monitoring System mathematics test, a national standardized measurement instrument (in which speed of performance is not important) yielding a norm-referenced mathematics score, that we categorized into 2 levels: above the average of the norm group and below the average of the norm group.

5.2.2 Material

Three parallel sets of four complex division problems each were constructed, resulting in a total of 12 problems (see Appendix 5.A). These problems were designed to resemble those that students encounter in their classroom and testing practices. Each problem was presented within a realistic context: a situation that described a hypothetical real life mathematical problem. For each item, three parallel versions were constructed that were as similar as possible to each other with respect to number characteristics and realistic context, but that at the same time would not be perceived as identical problems to prevent practice effects. The 3 parallel sets were counterbalanced over the 3 conditions (choice, no-choice mental, and no-choice written; see below).

The number characteristics of each item set were as follows. In the first item set, the outcome was below 10, but students had to deal with a remainder. In the second item set, numbers were such that a compensation approach (rounding the dividend; e.g., $1089 \div 11 = 1100 \div 11 - 1 = 100$ - 1 = 99) would be efficient. In the third item set, the dividend and divisor were decimal numbers (while the outcome was not). Finally, in the fourth item set a 3-digit number had to be divided by a 2-digit number, with outcome also a 2-digit number.

5.2.3 Procedure

Per school, six to twelve students were randomly selected for participation. The students were tested individually in a quiet room outside their classroom. They were told that they would be given twelve division problems to solve. Each student was first tested in the free choice condition (C) and then in two no-choice conditions: forced mental calculation (NC-M) and forced written calculation (NC-W). The order of the 2 no-choice conditions was counterbalanced over students. All problems were presented one by one, and solution times were collected with a stopwatch on a trial-by-trial basis. The students received the following instruction: *'With this stopwatch, I will register what time you need to solve the problems, but you can take as much time as you need on each problem.'*

5.2.4 Conditions

The first four division problems were presented in the free choice condition. On these problems, students were free to choose whether they solved them by mental or written calculation. There was a pencil available for the student to use and space for writing down calculations in the booklet. At the end of this first set of problems, the children were asked to report verbally on the strategies they used on the problems that they solved by mental calculation.

In the no-choice mental condition, another parallel set of four problems was presented. The procedure was similar to the choice condition, except for the fact the students could not use paper and pencil in doing their calculations and thus were forced to use mental calculation. In addition, the students were asked to report on their calculation strategy verbally after each problem was solved.

In the no-choice written condition, the final set of four problems was presented. In this condition, students had to write down their calculation procedure and were thus forced to use written calculation.

5.2.5 Responses

For each trial, the following responses were coded: (a) the accuracy of the answer given, (b) the solution time (ST), (c) the main strategy used, mental or written calculation (only in choice condition), and (d) the type of written or mental solution strategy used.

The type of written or mental strategy used was coded to get more insight into the rather broad categorization into mental and written strategies. The types distinguished were (a) repeated addition or subtraction of multiples of the divisor (see left part of Figure 5.1), (b) traditional algorithm for long division (see right part of Figure 5.1), (c) wrong procedure: e.g., multiplication of dividend and divisor, numerical estimation and splitting up the divisor (e.g., solving $306 \div 17$ by $306 \div 10 + 306 \div 7$), and (d) unclear procedure or student did not remember. The type of written strategy was inferred from the solution steps that were written down. The type of mental strategy was inferred from the verbal reports.

5.2.6 *Statistical Analyses*

Because item responses were nested in individual students, observations were not independent. To account for these correlated responses we used (generalized) linear mixed (also called hierarchical, multilevel or random effects) regression models (e.g., Hedeker & Gibbons, 2006; Snijders & Bosker, 1999). All estimated models were random intercepts models, in which individual differences were accounted for by the intercept being random over students. The continuous dependent variable 'solution time' was analyzed with linear mixed models using the SAS procedure MIXED. Because solution times deviated from normality (the smallest z-value of skewness of solution times on the 12 problems was 4.32), solution times were log-transformed before entering the analyses (cf. Klein Entink, Fox, & Van der Linden, 2009). The binary dependent variables strategy choice (mental/written) and accuracy (incorrect/correct) were analyzed by mixed binary logistic regression models using the SAS procedure NLMIXED. The statistical significance of predictor effects was tested using a likelihood ratio (LR) test. The LR-test statistic is computed as two times the difference between the log-likelihood of the model with and the model without the predictor effect, and is asymptotically χ^2 distributed with df the number of parameters associated with the predictor effect.

5.3 RESULTS

Preliminary mixed linear (speed) and logistic (strategy choice and accuracy) regression analyses showed that that the three parallel item sets A, B, and C did not differ in proportion mental calculation (choice condition only; LR $= 1.0$, $df = 2, p > .05$) nor in average accuracy (LR $= .0$, $df = 2, p > .05$), but the effect of item set on speed did just reach significance (LR $= 6.1$, $df = 2, p = .048$). Furthermore, the order of the no-choice conditions did not affect accuracy (LR $= .1$, $df = 1, p > .05$) or speed (LR $= .6$,

$df = 1, p > .05$). In the main analyses, data were grouped over the versions of the item set and order of no-choice conditions, but in the speed analyses we statistically controlled for the item set version.

The main results are discussed in three sections: (a) repertoire and distribution of strategies in the choice condition, (b) strategy performance data (accuracy and speed) from the choice as well as from the two no-choice conditions, and (c) results on adaptivity of strategy choices.

5.3.1 Strategy repertoire and distribution in choice condition

Almost half of the students (42 students, 49%) used mental calculation and written calculation at least once. The remaining students used either written calculation on all items (33 students, 38%) or mental calculation on all items (11 students, 13%).

On 67% of all trials a written strategy was chosen. Girls chose a written strategy ($M = 80\%$) significantly more often than boys ($M = 53\%$), LR = 13.0, $df = 1, p < .001$. The difference between below-average ($M = 70\%$ written strategies) and above average mathematics achievers ($M = 64\%$ written strategies) was not significant, LR = .5, $df = 1, p > .05$. Finally, the four items significantly differed in percentage of written strategies, LR = 34.3, $df = 3, p = .001$. Post-hoc pairwise comparisons showed that on item 4 ($M = 84\%$ written strategies) significantly more written strategies were used than on item 3 ($M = 67\%$, $t(85) = 3.20$, $p = .002$), item 2 ($M = 59\%$, $t(85) = 4.40$, $p < .001$), and item 1 ($M = 57\%$, $t(85) = 4.55$, $p < .001$), respectively. All other pairwise comparisons were not significant.

Table 5.1 shows the distribution of each type of strategy used in the choice condition averaged over the four items. The distribution of repeated addition/subtraction strategies and the traditional algorithm on the one hand, and applying the wrong procedure or unclear strategy on the other hand, was significantly different for mental and written strategies (LR = 15.8, $df = 1, p < .001$). Within written as well as within mental strategies, repeated addition/subtraction strategies were dominant with 75% and 84% respectively. The traditional algorithm was used very infrequently (on 5% of all trials), and if it was used it was only within written strategies. Furthermore, executing the wrong procedure was more prevalent in mental strategies (17%) than in written strategies (5%). The same pattern holds for unclear strategies: 8% of the mental strategies, and 3% of the written strategies. Another interesting result (not presented in Table 5.1) was on the prevalence

TABLE 5.1 *Distribution of type of strategies used in choice condition.*

type of strategy	within mental strategies		within written strategies		total	
repeated addition/subtraction	86	(75%)	192	(84%)	278	(81%)
traditional algorithm	0	(0%)	18	(8%)	18	(5%)
wrong procedure	19	(17%)	12	(5%)	31	(9%)
unclear	9	(8%)	6	(3%)	15	(4%)
total number of trials	114	(100%)	228	(100%)	342	(100%)

Note. On two trials the student did not give an answer, so the strategy used could not be determined. As a result, the total number of trials equals 342.

of the compensation strategy on item 2 (a strategy in which the dividend was rounded; this was a specific form of the repeated addition/subtraction category): this shortcut strategy constituted 63% of all 35 mental strategies on this item, but only 12% of the 54 written strategies, a significant difference (z-test for proportions $= 4.68$, $p < .001$).

5.3.2 Strategy performance

Choice condition

Table 5.2 shows strategy performance data (accuracy and speed[2]) in the choice condition, by gender and general mathematics level. On average, the accuracy difference between mental strategies and written strategies was border significant (LR $= 3.8$, $df = 1, p = .051$). The speed difference was highly significant (LR $= 108.8$, $df = 1, p < .001$), with mental strategies being faster than written strategies.

Boys and girls did not differ significantly in average total accuracy in the choice condition (LR $= .2$, $df = 1, p > .05$), nor in accuracy within mental or within written strategies (LR $= 3.5$, $df = 1, p > .05$). Boys were significantly faster on average (LR $=11.3$, $df = 1, p < .001$), but within each strategy choice the gender difference in speed was not significant anymore (LR $= 2.1$, $df = 1, p > .05$). So, difference in strategy choice

2 All speed data presented in the Results are based on all trials (correct and incorrect ones), because we argue that this presents a more complete picture than presenting only speed of correctly executed strategies. However, we also analyzed speed data based on only the correct trials. Results were very similar, with the exception that correct responses were faster.

TABLE 5.2 *Strategy performance in the choice condition, by gender and general mathematics level.*

	accuracy (P (correct))			speed (ST in seconds)		
strategy choice	mental	written	total	mental	written	total
girl	.41	.57	.54	44.0	105.6	93.3
boy	.56	.68	.63	47.4	84.8	67.4
< average math level	.23	.48	.40	66.4	114.1	99.6
> average math level	.77	.77	.77	29.1	78.1	60.3
total	.52	.61	.58	46.4	97.2	80.3

accounted for gender differences in speed in the choice condition: boys were faster on average because they chose fast mental calculation more often than girls.

The effect of general mathematics level of the student was highly significant on average accuracy (LR $= 29.0$, $df = 1, p < .001$) as well as on average speed (LR $= 18.6$, $df = 1, p < .001$). Moreover, accuracy differences within mental and written strategy choices were also significant (LR $= 32.0$, $df = 1, p < .001$), as were speed differences within the two strategies (LR $= 20.5$, $df = 1, p < .001$). Below-average achievers had a lower proportion correct and were slower than above-average achievers within the two strategies as well as totaled over the strategy choices. The difference in performance between below-average and above-average achievers was the same in mental strategies as in written strategies, since the interaction between general mathematics level and strategy choice was not significant on either accuracy (LR $= 3.3$, $df = 1, p > .05$) or speed (LR $= 2.9$, $df = 1, p > .05$).

No-choice conditions

Table 5.3 shows strategy performance data (accuracy and speed) from the two no-choice conditions, by gender and general mathematics level. No-choice condition had a significant effect on accuracy (LR $= 11.8$, $df = 1, p < .001$) as well as on speed (LR $= 46.9$, $df = 1, p < .001$). These unbiased strategy performance data thus showed that forced mental strategies were less accurate but faster than forced written strategies.

The accuracy difference between boys and girls within each condition was not significant (LR $= 2.8$, $df = 1, p > .05$). By contrast, gender did have a significant effect

Table 5.3 *Strategy performance in the no-choice conditions, by gender and general mathematics level.*

	accuracy P (correct)		speed ST in seconds	
condition	NC-M	NC-W	NC-M	NC-W
girl	.46	.55	92.4	92.5
boy	.55	.67	67.2	90.4
< average math level	.25	.43	102.8	115.1
> average math level	.76	.81	55.8	66.9
total	.50	.61	79.6	91.4

on speed (LR $= 4.7$, $df = 1, p = .031$) with boys being faster than girls. Moreover, the interaction between gender and no-choice condition (mental versus written) was also significant (LR $= 6.6$, $df = 1, p = .010$). Post-hoc contrasts showed that boys were significantly faster than girls in the no-choice mental condition ($t(85) = 3.06$, $p = .003$), but that the gender difference in speed in the no-choice written condition was not significant ($t(85) = .96$, $p > .05$). Moreover, for boys ($t(85) = 6.97$, $p < .001$) as well as for girls ($t(85) = 3.11$, $p = .003$) forced mental strategies were significantly faster than forced written strategies.[3]

The effect of general mathematics level of the student was highly significant on accuracy (LR $= 63.6$, $df = 1, p < .001$) as well as on speed (LR $= 27.1$, $df = 1, p < .001$). Students with below-average mathematics level had a significantly lower proportion correct than above-average achievers, in both no-choice conditions (interaction between mathematics level and no-choice condition on accuracy not significant; LR $= 2.3$, $df = 1, p > .05$). Regarding speed, below-average achievers were significantly slower than above-average achievers, regardless of the strategy they had to use (interaction between mathematics level and no-choice condition on speed not significant; LR $= .9$, $df = 1, p > .05$).

3 Although for girls the mean solution times in the two no-choice conditions (92.4s. and 92.5s.) did not seem to differ, these means were influenced by the skewness of the distribution of raw solution times. Log-transformed STs were not affected by skewness, and the mean in the no-choice mental condition was significantly lower (4.14) than the mean in the no-choice written condition (4.36).

5.3.3 Strategy adaptivity

Comparing the strategy performance data from the no-choice conditions with the strategy choice made in the choice condition gives information on the adaptivity of the strategy choice. To what extent was the most appropriate strategy chosen, as evidenced from the individual strategy performance data from the no-choice conditions?

The issue of strategy adaptivity is approached in three ways. In the first two approaches, analyses were done on the item level (aggregating over students), and in the third approach they were done on the student level (aggregating over items) (cf. Luwel et al., 2009). In these latter analyses, group differences with respect to gender and mathematics achievement level were studied as well.

Adaptivity at the item level

In analyzing adaptivity at the item level, we ask the following question: Is the performance difference between forced mental and forced written strategy on an item in accordance with the strategy choice made in the choice condition on the parallel item? First, accuracy and speed are dealt with separately.

For the mental strategy choices in the choice condition, there was no difference in accuracy rates between forced mental ($M = .57$) and forced written computation ($M = .57$) on (the parallel versions of) that item ($t(85) = .00$, $p > .05$). So, regarding accuracy, these mental strategy choices were neither adaptive nor counter-adaptive. In contrast, these mental strategy choices were adaptive to speed: for instances in which a mental strategy was chosen in the choice condition, forced mental strategies were significantly faster ($M = 53.3$ seconds) than forced written strategies ($M = 84.7$ seconds), $t(85) = 8.00$, $p < .001$.

For written strategy choices in the choice condition, the forced mental strategy was significantly less accurate ($M = .48$) than forced written computation ($M = .64$) on the parallel items in the no-choice conditions, $t(85) = 4.07$, $p < .001$. So, the choices for a written strategy were adaptive regarding accuracy. In contrast, these choices were counter-adaptive to speed: for instances in which a written strategy was chosen in the choice condition, forced mental strategies ($M = 93.2$ seconds) were significantly faster than forced written strategies ($M = 94.8$ seconds), $t(85) = -3.31, p = .001$.

In short, these separate analyses on accuracy and speed suggest that on average, mental strategy choices seem adaptive to speed considerations as evidenced from the solution time differences between the no-choice conditions (mental strategies being faster), while there were no differences in unbiased accuracy characteristics. In contrast, written strategy choices seem adaptive to accuracy, but counter-adaptive to speed.

In the next approach, accuracy and speed were combined on a trial-level basis. Following Lemaire and Siegler (1995), an adaptive strategy choice was defined as choosing the strategy that leads the individual fastest to an accurate answer. To operationalize this definition, for each trial in the choice condition we combined accuracy and speed information from the two no-choice conditions, and compared it to the strategy choice made in the choice condition, similar to Imbo and LeFevre (2009).

There were three possible categories. First, a strategy choice was coded as *adaptive* either when (a) a correct answer was obtained on an item with both forced mental and forced written calculation and the fastest of these two strategy was chosen, or (b) when a correct answer was obtained in only one of the no-choice conditions and the strategy that had yielded the correct answer was chosen. For example, for a student who obtained the correct answer in the NC-W condition but an incorrect answer in the NC-M condition, choosing a written strategy in the choice condition was coded as an adaptive strategy choice. For these latter trials, potential differences in speed of the two strategies did not play a role in coding adaptivity: accuracy was deemed more important and therefore decisive. Second, a strategy choice was coded as *counter-adaptive* if (a) a correct answer on an item was obtained with both forced mental and forced written calculation and the strategy that was slowest was chosen, or (b) when the correct answer was obtained in only one of the no-choice conditions and the strategy that had yielded the incorrect answer was chosen. Finally, a strategy choice was coded *indeterminate* when a student answered the item in both no-choice conditions incorrectly. In these instances it is hard to think of adaptivity, since neither of the strategies yielded a correct answer, and hence could never lead to an adaptive choice. Results showed that strategy choices were adaptive on 43% of the items and counter-adaptive on 30% of the trials. In addition, 28% of the strategy choices were indeterminate with respect to adaptivity. Moreover, for each individual student summing the adaptivity scores over the 4 trials in the choice condition, showed that there was substantial variation between students. To illustrate, 51 students (59%) made an adaptive as well as a counter-adaptive choice at least once.

Adaptivity at the student level

In the preceding section results were aggregated over students, obscuring individual variations in accuracy differences and speed differences. In this section, we took these individual differences into account by analyzing adaptivity at the student level. Accuracy and speed were treated separately, by computing the correlation between the frequency of mental calculation of a student in the choice condition and the differences in accuracy (total number correct) and in speed (total log-transformed solution time) between the two strategies from the no-choice conditions (cf. Torbeyns, De Smedt, et al., 2009b).

Spearman's ρ correlation between frequency of mental calculation (choice condition) and the difference in the total number correct between no-choice mental and no-choice written conditions was positive and significant ($\rho = .28$, $df = 84, p = .009$), indicating that students took into account which of the two strategies was most accurate for them. Gender seemed to affect this correlation ($\rho_{\text{girls}} = .40$, $df = 41, p = .008$; $\rho_{\text{boys}} = .26$, $df = 41, p > .05$) but the difference was not significant ($z = .68, p > .05$). Mathematics level did have a significant effect on the correlation ($z = 2.85$, $p = .004$), $\rho_{\text{below}} = .00$, $df = 42, p > .05$; $\rho_{\text{above}} = .56$, $df = 40, p < .001$.

With respect to speed, Spearman's ρ correlation between frequency of mental calculation and differences in solution time between forced mental and written strategies was also significant ($\rho = -.32$, $df = 84, p = .002$). Note that the correlation is negative because solution times are inversely related to speed, so that this result indicates that students took into account their individual strategy speed characteristics. This time, gender had a significant effect on this relation ($z = 2.29$, $p = .022$): $\rho_{\text{girls}} = -.07$, $df = 41, p > .05$; $\rho_{\text{boys}} = -.52$, $df = 41, p < .001$. Although the effect of mathematics achievement level on the size of this correlation was not significant ($z = 1.62, p > .05$), the correlation was not significantly different from zero for below-average achievers ($\rho_{\text{below}} = -.15$, $df = 42, p > .05$), while it was for above-average achievers ($\rho_{\text{above}} = -.47$, $df = 40, p = .002$).

So, for the sample as whole, students seemed to adapt their strategy choices to their individual accuracy and speed characteristics of the two strategies. However, interesting gender differences were found. Girls appeared to fit their strategy choices to accuracy characteristics, ignoring speed characteristics. In contrast, boys showed the opposite pattern, by choosing adaptively regarding speed, and thereby paying less attention to accuracy (although the – nonsignificant – correlation with accuracy difference was positive and not significantly different from the correlation for girls). Moreover, there were also important differences with respect to mathematics level. Below-average achievers did not significantly fit their strategy choice to either accuracy or speed, while above-average achievers chose significantly adaptive regarding both accuracy and speed.

5.4 Discussion

In this study, mental and written solution strategies on complex division problems were investigated using the choice/no-choice paradigm. Students successively solved three parallel sets of four division problems with free strategy choice, forced mental, and forced written calculation, respectively. Besides assessment of strategy repertoire, distribution, and efficiency, an important focus was on strategy adaptivity. We will first discuss the results on the four dimensions of strategy competence, and then focus on gender differences and mathematics achievement level effects. We will end by discussing cognitive psychological and educational implications of the findings of the current study.

5.4.1 Strategy repertoire, distribution, efficiency, and adaptivity

Concerning strategy repertoire, findings showed that approximately one half of the students used written as well as mental strategies in the choice condition. The majority of the other half used only written strategies, and a small part used only mental strategies. Regarding strategy distribution, the relative frequencies of mental and written strategies in the choice condition showed that each item was solved most frequently by written calculation, but that there were differences in this respect between items. Analysis of the specific types of strategies showed that both written and mental strategies predominantly comprised repeated addition/subtraction. Item 2 deserves special attention because of the possibility of using a compensation strategy (a special case of the repeated addition/ subtraction strategy, in which it is possible to take advantage of the closeness of the dividend to a hundredfold of the divisor). The majority of mental strategies involved compensation, while this was not very frequent within written strategies. So, choosing between mental and written computation on that item mainly reflected using the compensation strategy or not: a similar finding to Hickendorff et al. (2010).

Strategy performance in the choice condition showed that, for the sample as a whole, freely chosen mental strategies were evenly accurate but faster than freely chosen written strategies. However, these accuracy data were probably affected by selection effects (cf.

Siegler & Lemaire, 1997), because the accuracy difference between the two no-choice conditions was significant: forced written strategies were on average more accurate than forced mental strategies. Speed differences between the no-choice conditions were congruent with those from the choice condition, with mental strategies being significantly faster than written strategies. Thus, unbiased strategy performance data showed that mental strategies were less accurate but faster than written strategies.

The main focus was on strategy adaptivity, which we approached in three ways. In each approach, we assessed whether the strategy selected in the choice condition was the most 'appropriate' one, as evidenced by the unbiased strategy performance information from the no-choice conditions. First, item-level analysis showed that, on average, mental strategy choices were adaptive to speed, and indifferent to accuracy. In contrast, written strategy choices were fit to accuracy differences, while being counter-adaptive to speed. Second, we combined accuracy and speed on a trial-basis by operationalizing the definition of the 'best' strategy as being the one leading fastest to an accurate answer. On average, on 43% of the trials the 'best' strategy was chosen, and on 30% the best strategy was *not* chosen. On the remaining 28% of the trials there was no correct answer obtained in either of the two no-choice conditions, so the strategy choice in the choice condition could not be scored with respect to adaptivity. Interestingly, more than half of the students made at least one adaptive and one counter-adaptive strategy choice on the 4 trials. Third, student-level analyses showed that the correlations between frequency of use of mental computation on the one hand and unbiased accuracy and speed differences on the other hand were significant and in the expected direction. So, in general, strategy choices seem adaptive both to accuracy and speed. In sum, we found that mental strategies were chosen in trials where they were faster but equally accurate according to the unbiased strategy performance data, while written strategies were chosen on trials on which it was the more accurate (albeit slower) strategy. Combining these findings resulted in the pattern that, on average, students chose adaptively both to accuracy and speed. However, when accuracy and speed were combined to define the optimal strategy, we found that students made a suboptimal strategy choice on a substantial percentage of items (30%).

5.4.2 Gender differences and mathematics achievement level effects

We found interesting gender differences in the dimensions of strategic competence. Regarding strategy choice, boys were more inclined to mental computation than girls, a finding resembling earlier research findings that girls favor structured, algorithmic strategies, while boys tend to use less structured, more intuitive strategies (Carr & Davis, 2001; Carr & Jessup, 1997; Gallagher et al., 2000; Hickendorff et al., 2009b, 2010; Timmermans et al., 2007). In the current study, there were no significant gender differences in strategy accuracy. Regarding speed, girls were slower than boys when they were forced to use mental computation. In all other conditions, gender differences in speed were not significant. Consequently, for boys the speed gains of choosing mental strategies over written ones was larger than for girls, which may partially account for boys' larger inclination of choosing mental strategies. In addition, boys and girls appeared to have different speed-accuracy preferences. Girls appeared to fit their strategy choices to accuracy considerations, ignoring speed, while boys had a preference for speed over accuracy. This may be related to individual differences in the confidence criterion that have been reported in children (Siegler, 1988a, 1988b) and in adults (Hecht, 2006).

Moreover, these found gender differences in strategy choice and strategy adaptivity may be related to the consistent finding that girls have lower levels of confidence with mathematics (Mullis et al., 2008; Timmermans et al., 2007; Vermeer et al., 2000), so they may perform more cautiously than boys and therefore choose accuracy over speed. In line with this reasoning, girls have been found to be less inclined to intellectual risk-taking than boys (Byrnes et al., 1999). In addition, girls tend to be more inclined to (academic) delay of gratification (Bembenutty, 2009; Silverman, 2003), which might partially explain that boys more often choose fast mental calculation over slower but more accurate written computation.

In addition to gender differences, there were mathematics achievement level effects. Below-average and above-average achievers chose mental computation equally often. Performance differences were as could have been expected, with below-average achievers less accurate and slower than above-average achievers, regardless of the strategy they chose (choice condition) or the strategy they had to use (no-choice conditions). Interesting differences were found in strategy adaptivity: below-average achievers did not take either accuracy or speed into account in their strategy choices, while above-average achievers fitted their strategy choices to both components of performance. Therefore, we argue that the present study's results implies that strategy adaptivity in complex division is currently only attained by the better achieving students, resembling findings of Foxman and Beishuizen (2003), Hickendorff et al. (2010), Torbeyns, De Smedt, et al. (2009b), and Torbeyns et al. (2002, 2006). Further research is needed into whether this important goal of mathematics education reform is feasible or desirable for the below-average achievers. Like Geary (2003), Torbeyns et al. (2006), and Verschaffel et al. (2009), we plead for more research-based evidence for striving for adaptive expertise in mathematics education, especially for average and weaker students.

5.4.3 Methodological considerations

Several methodological considerations of the current study merit attention. First, we focused on distinguishing between between mental strategies, defined as nothing written down on paper, and written strategies in which something was written down on paper, ranging from intermediate answers to procedural algorithms. Note that this is a similar categorization of strategies as in Siegler and Lemaire's (1997) original choice/no-choice study, in which they distinguished between using a calculator, using mental arithmetic, and using pencil and paper (experiment 3). Although this is arguably a rough classification, and other categorizations (for example with respect to the number of solution steps) are thinkable, we chose this strategy split for two reasons. First, earlier studies into strategy use on complex division by Dutch sixth graders showed that both strategy types were used, and that they had large predictive power of the accuracy of solutions (Hickendorff et al., 2009b, 2010). Second, the didactical practice in the Netherlands – with the disappearance of the traditional algorithm and many different informal strategies – leads to obstacles in studying the characteristics of different strategies in a choice/no-choice design. That is, if students are forced to use a particular strategy in the no-choice conditions, they should have those strategies in their repertoire, and many students did not get instruction in for example the traditional algorithm.

Second, the number of items (4 items times 3 conditions) was small, mainly based on practical considerations. Because these kind of complex division problems are quite demanding to solve for sixth graders (as also comes forward from the long solution times, on average 83.7 seconds with peaks to over 500 seconds), we believed it was not practically feasible to administer more than 12 problems to a student in a session. Given this limited number of items, we were unfortunately not able to rigorously analyze the effect of item features on strategy choice. However, we argue that we were able to study adaptivity, because we did found substantial variation in strategy choice, with half of the students using both strategies on the set of 4 problems in the choice condition. Moreover, the single strategy users were also spread over the two strategies: two-thirds of them consistently used written strategies, while one third consistently used mental strategies. Therefore, even on this small number of problems there were clear differences in strategy choice, that we tried to predict by strategy performance characteristics in order to investigate the issue of adaptivity.

5.4.4 Implications

In conclusion, we found several indications that there are different patterns of adaptivity to strategy accuracy and speed. If we look at the general pattern for the whole sample, it would seem that mental strategies are chosen for reasons of speed, while written strategies are chosen for reasons of accuracy. However, when we look at different subgroups of students (gender, mathematics level), we find that there are different adaptivity patterns. Interestingly, there are students who prefer accuracy over speed, while there are also students showing the opposite pattern. Moreover, the majority of students made both optimal and suboptimal strategy choices on the 4 items.

Individual differences in preference for accuracy and speed are important from a theoretical as well as from practical point of view. Theoretically, they have not been specifically addressed in the cognitive models of strategy choice and adaptivity (Shrager & Siegler, 1998; Siegler & Shipley, 1995), although the concept of different confidence criterion individuals may hold is related. Moreover, future research may investigate whether other individual differences constructs can (partly) account for the found accuracy-speed preferences. The cognitive style of impulsivity-reflection (Kagan, 1966) may very well be related (cf. Siegler, 1988a), with reflectives being slow but accurate, but impulsives being fast but with more errors (Phillips & Rabbitt, 1995). In addition, concepts such as academic delay of gratification (Bembenutty, 2009) and academic risk-taking (Byrnes et al., 1999) may be associated as well.

In addition to task and subject variables, sociocultural context variables may also affect strategy choices (Luwel et al., 2009; Verschaffel et al., 2009). Ellis (1997) pointed out the possibility of (sub)cultural differences in the weights assigned to speed versus accuracy of performance, and the value placed on solutions constructed in the head versus by means of external aids. For instance, classroom socio-mathematical norms and practices valuing speed over accuracy and/or mental strategies over written ones, may result in students overusing mental strategies at the cost of accuracy. Besides the implication that cognitive models for strategy adaptivity are limited in this respect (see also Ellis, 1997; Verschaffel et al., 2009, there are also important educational consequences. Most educators will agree that being correct is more important than being fast in learning mathematics. However, not all students (particularly boys) seem to reason in that way, and teachers should be aware of that. So, teachers may create a classroom environment in which accuracy is preferred over speed and using an external aid (paper and pencil) is not necessarily less valuable than working in the head. Furthermore, they may explicitly encourage consistently mentally calculation students (especially those with low mathematical ability) to write down their solution procedures, as this would improve their accuracy. Finally, it would be interesting to conduct a similar study in another educational climate with a larger focus on written arithmetic than in the Netherlands and compare the results.

Appendix 5.A Complete item set

Table 5.4 presents the number characteristics of the three parallel sets (A, B, and C) of 4 items each. The realistic contexts in set A were (translated from Dutch):

1. Four children go to an amusement park together. For admission, they have to pay 34 euro in all. How much is that per child?
2. A bookseller has earned 1,089 euro. He sold all his books for 11 euro each. How many books did he sell?
3. Robert is making a fence, which will have a length of 31.2 meters. The planks he uses are 1.2 m long. How many planks will he need for the entire fence?[4]
4. Anne has 304 biscuits. She divides the biscuits over 19 jars. How many biscuits are there in each jar?

The contexts of the items in set B and set C were comparable.

Table 5.4 *Number characteristics of the items.*

	item set		
item nr.	A	B	C
1	34 ÷ 4	52 ÷ 8	45 ÷ 6
2	1089 ÷ 11	2450 ÷ 25	1980 ÷ 20
3	31.2 ÷ 1.2	30.8 ÷ 1.1	32.2 ÷ 1.4
4	304 ÷ 19	306 ÷ 17	221 ÷ 13

4 NB. Original item included illustration, making clear that the height of the fence was the height of one plank.