Journal of Experimental Child Psychology 201 (2021) 104987

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Developmental changes in size effects for simple tie and non-tie addition problems in 6- to 12-year-old children and adults

Jeanne Bagnoud [a,*], Jasinta Dewi [a], Caroline Castel [b], Romain Mathieu [b], Catherine Thevenot [a,*]

[a] *Institute of Psychology, University of Lausanne, CH-1015 Lausanne, Switzerland*
[b] *Faculty of Psychology and Educational Sciences, University of Geneva, CH-1205 Geneva, Switzerland*

A R T I C L E I N F O

Article history:
Received 27 February 2020
Revised 10 August 2020
Available online 17 September 2020

Keywords:
Mental arithmetic
Mathematics
Operations
Strategies
Memory retrieval
Automatization

A B S T R A C T

In the domain of cognitive arithmetic, the size effect corresponds to an increase in solution times as a function of the size of the operands involved in the problems. In this study, we tracked the evolution of size effects associated with tie and non-tie addition problems across development. We scrutinized the progression of solution times for very small problems involving operands from 2 to 4, larger problems, and 1-problems (problems involving 1 as one of the operands) in children from Grade 1 to Grade 5 and adults. For the first time, we document the presence of a size effect for tie problems with a sum up to 8 in Grade 1 children. In contrast, from Grade 3 until adulthood, this size effect could not be evidenced. Crucially, for non-tie problems, whereas a general size effect is observed when contrasting small one-digit additions with large additions, we show that, from Grade 1 until adulthood, a continuous size effect as a function of the sum of the problems is not observed. In fact, for all age groups, medium problems with sums of 8, 9, and 10 do not present a size effect at all. Given that the problem size effect is sometimes referred to as one of the most robust and reliable effects in the numerical cognition literature, our results necessarily challenge its theoretical interpretation.

* Corresponding authors.
E-mail addresses: jeanne.bagnoud@unil.ch (J. Bagnoud), catherine.thevenot@unil.ch (C. Thevenot).

https://doi.org/10.1016/j.jecp.2020.104987
0022-0965/

Introduction

Researchers have studied strategies in the domain of mental arithmetic for a long time (see Campbell, 2015, for a review). Despite the impressive amount of research devoted to this important theoretical and educational topic, it is surprising that some basic assumptions in this domain are still under debate (Baroody, 2018; Chen & Campbell, 2018). Nowadays, one of the most debated effects is the size effect, according to which addition problems involving small operands such as 4 + 2 are solved more quickly than problems involving larger operands such as 7 + 6 (Campbell, 1995; Campbell & Xue, 2001; Chen, Loehr, & Campbell, 2019; LeFevre, Sadesky, & Bisanz, 1996; Thevenot, Barrouillet, Castel, & Uittenhove, 2016; Uittenhove, Thevenot, & Barrouillet, 2016; see Ashcraft & Guillaume, 2009, for a review). Describing the evolution of the size effect for tie and non-tie problems across development was the objective of the current article. Through a longitudinal and cross-sectional approach, we aimed at furthering understanding of the source of the size effect.

The most common explanations of size effects are given within a retrieval framework (Ashcraft, 1992; Campbell, 1995; Siegler & Shrager, 1984). In this framework, the strategy used to solve arithmetic problems shifts with practice from counting to direct retrieval of answers from long-term memory. Within this theory, problem size effects are explained by frequency and interference effects. Frequency effects correspond to the fact that smaller problems are more often solved by participants and therefore are solved more quickly because of better mental access (Hamann & Ashcraft, 1986). Interference effects correspond to the fact that problems with larger operands share their sums with more other problems than smaller problems, which produces larger fan effects and again more difficulties in accessing answers for larger problems. Moreover, mental representations of larger numbers overlap more than smaller numbers with mental representations of their neighbors (Campbell, 1995; Campbell & Graham, 1985; Campbell & Timm, 2000). More precisely, whereas mental representations of relatively small numbers are very distinct from each other, the representations of larger numbers become fuzzier due to shared representations (Pinel, Piazza, Le Bihan, & Dehaene, 2004). The memory association between large problems and their answers would also be weaker than that for small problems because large problems have been associated with more wrong answers than small problems during the course of development (Siegler & Shrager, 1984). Indeed, more counting steps in larger problems necessarily increase the likelihood of mistakes. Another explanation of the problem size effect is that reversal to less mature strategies than retrieval, such as counting and decomposition, occurs for some trials (e.g., LeFevre et al., 1996). Given that strategies derived from counting are usually considered slower than retrieval, and given that larger problems often require more counting steps than smaller problems, larger problems take longer to solve than smaller problems. Consequently, when occasional reversal to nonretrieval strategies occurs, solution times are longer for larger problems (Groen & Parkman, 1972).

Interestingly, for addition, the systematic observation of a size effect is specific to non-tie problems, which correspond to problems with different operands. For tie problems, which are constructed with repeated operands (e.g., 4 + 4), the effect is not always observed, at least in adults (LeFevre, Shanahan, & DeStefano, 2004). When observed, size effects are always smaller for tie problems than for non-tie problems (e.g., Campbell & Gunter, 2002). The fact that the retrieval process of tie problems is not subjected to the same effects as non-tie problems was interpreted by LeFevre, Smith-Chant, Hiscock, Daley, and Morris (2003), who suggested that non-tie problems are less often retrieved than tie problems. Therefore, more frequent resort to counting procedures in the case of non-tie problems than of tie problems would necessarily result in a larger size effect for the former than for the latter. Nevertheless, even when only problems where participants report retrieval are considered, size effects are still observed for the two types of problems and, according to Campbell and Gunter (2002), are still larger for non-tie problems than for tie problems. Therefore, additional interpretations were needed because it was necessary to explain why these two categories of problems do not suffer from the same interference and frequency effects. One of the interpretations is that non-tie problems activate two families of answers, each associated with one operand. Because tie problems contain only one operand, only one family is activated, and therefore they suffer from less interference (Graham & Campbell, 1992). Another interpretation is that tie problems could be represented in a memory network partially separated from that of non-tie problems. Tie and non-tie problems therefore would constitute two different categories of problems, with weak interferences across categories. In that case, tie problems would suffer from less interference effect because interferences within the small tie problem network would necessarily be weaker than those within the larger non-tie problem network (Campbell & Oliphant, 1992; Graham & Campbell, 1992). These interpretations lead to the idea that memory access would be more efficient for tie problems than for non-tie problems (Ashcraft & Battaglia, 1978; Campbell & Gunter, 2002; LeFevre et al., 2004). This memory access hypothesis would also explain why, irrespective of problem size effects, tie problems are solved more quickly than non-tie problems. Even though an interpretation in terms of faster encoding was initially given to explain this tie advantage (Blankenberger, 2001), all researchers now agree that this explanation is not sufficient (Campbell & Gunter, 2002; LeFevre et al., 2004).

Therefore, we know that tie problems constitute a special category of problem with shorter solution times than non-tie problems and little to no size effect. However, we do not know whether these particularities are observable from the beginning of learning or are acquired through learning. One approach to answering this question is to precisely analyze the pattern of solution times for tie problems during development. This was the first goal of our article. If tie problems present null or little size effects from the beginning of acquisition, this would imply that they are never solved through counting procedures and therefore that they are learned by rote in early education. Alternatively, counting procedures can initially be applied to tie problems, and associations between operands and answers could be constructed with practice. In this case, a substantial size effect should be observed at the beginning, but not at the end, of acquisition. For non-tie problems, substantial size effects will be observed at the beginning of learning because of the use of counting strategies. Such effects should decrease across development until only residual size effects due to interference and frequency effects within retrieval networks are observed. Nevertheless, such size effects could be limited to very small problems because it has been documented in adults that non-tie addition problems with a sum from 7 to 10 do not present a size effect (Uittenhove et al., 2016). Exactly as for tie problems, we do not know whether such a plateau is observable from the beginning of learning or is acquired through learning. Within non-tie problems, 1-problems (problems involving 1 as one of the operands) are also supposed to constitute a special category. They are indeed often viewed as solved by a rule consisting in uttering the next number after the operand that is not equal to 1 in the numerical sequence (i.e., the number-after rule; e.g., Baroody, 1995; Baroody, Eiland, Purpura, & Reid, 2012). To the best of our knowledge, a precise description of 1-problem solution times depending on the size of the operands has never been provided, and therefore it is impossible to know whether problem size effects are also observable for this category of problems from the beginning of learning and throughout development. Tracking the evolution of non-tie problems for each of these categories (i.e., small problems, sum 7 to 10 problems, and 1-problems) constituted the second goal of our article.

To answer our research questions, we tested 133 children (from the beginning of Grade 1 until Grade 5) and 34 adults. More precisely, we tested two different groups of children at the beginning and at the end of Grade 1, whereas we followed the same children from Grade 3 to 5. In addition to the cross-sectional approach, this longitudinal approach over 3 years allowed us to precisely track the developmental changes around the critical point where strategy is supposed to shift from counting to retrieval (Ashcraft & Fierman, 1982). Participants were asked to solve simple addition problems involving operands from 1 to 9. All participants solved the 81 problems constructed with these operands except first graders, for whom problems with a sum higher than 10 were too difficult. Consequently, they solved only the 45 problems with operands from 1 to 9 with a sum up to 10. We measured solution times and, as already explained, were especially interested in potential differences across development between tie and non-tie problems and between the different categories of problems mentioned above.

Method

Participants

In total, 133 children from the beginning of first grade to fifth grade and 34 adults were involved in this study. The data of 4 children were discarded because these children presented too many missing data. Cross-sectional analyses involved 37 first graders (13 female) tested three months after the beginning of the year (mean age = 6 years 10 months, *SD* = 5 months), 41 first graders (19 female) tested three months before the end of the year (mean age = 7 years 3 months, *SD* = 5 months), 51 fourth graders (26 female; mean age = 9 years 9 months, *SD* = 5 months), and 34 undergraduate students (22 females; mean age = 21 years 5 months, *SD* = 3 years 10 months). The 51 children who were assessed in fourth grade were also tested in third grade (mean age = 8 years 9 months, *SD* = 5 months) and fifth grade (mean age = 10 years 9 months, *SD* = 5 months). We tested children individually in French-speaking public schools in Switzerland. Approval from the local ethics committee of the psychology department at the University of Lausanne and parental consent were obtained before starting the experiment. The undergraduate students all were in their first year of psychology studies at the University of Lausanne and received course credit for participation. In accordance with the canton de Vaud policy, approval from the ethics committee was not needed because the experiment involved adult participants without a lack of discernment.

Materials and procedure

We tested participants individually on different tasks, but here we focus on the main arithmetic task. In this task, we asked participants to solve simple additions and to give their responses orally. The additions involved operands from 1 to 9. This task was designed using the DMDX software (Forster & Forster, 2003). Each trial began with a ready signal ("$$$") displayed at the center of a computer screen for 500 ms. Then, two operands separated by a "+" sign were presented simultaneously and remained at the center of the screen until the response was given orally. Finally, a blank screen was displayed for 500 ms before the start of the next trial. Children were presented with the problems across two experimental sessions of 30 min. Six blocks of problems, each containing all 81 problems presented once in random order, were constructed for children from Grade 3 to Grade 5. For a given participant, therefore, one problem could not be presented more often than another. Because it was not always possible to present all the blocks within the experimental time frame, some children were presented with only four or five blocks. To avoid excessive fatigue, only three blocks were presented to younger children. Adults were also presented with three blocks. Before the experimental phase, participants needed to solve 5 to 10 training additions, which allowed them to become familiarized with the task and allowed the experimenter to test the voice key sensitivity. Training problems were chosen randomly from the pool of problems presented to participants and were always the same for each age group.

The voice key stopped the timer when participants gave their response, but to correct voice key imprecision, solution times were manually adjusted for each response at the onset of the answer vocal signal using CheckVocal (Protopapas, 2007). These numerous imprecisions were due to environmental noise and failure of detection. In 6.73% of the trials, the computer did not record the answer at all, so we removed these trials from the analyses.

Problems were classified into different categories. We considered four categories for non-tie problems. Following Barrouillet and Thevenot (2013), small problems were constructed with operands smaller than or equal to 4, medium problems contained at least one operand larger than 4 and their sum was smaller than or equal to 10, and large problems had a sum larger than 10. One peculiarity of this classification is that two problems having the same sum belong to two different size categories. Indeed, whereas 3 + 4 belongs to the small category, 5 + 2 belongs to the medium category. We considered 1-problems as one of the operands as the fourth category of non-tie problems. For tie problems, this categorization was not retained because 1-problems and medium problems would have been represented by only one problem each (1 + 1 and 5 + 5, respectively). Therefore, following Wilson, Revkin, Cohen, Cohen, and Dehaene (2006), we considered only two subcategories of problems: small tie problems with a sum smaller than 10 and large tie problems with a sum larger than 10.

Analytical and statistical approach

To begin with, we conducted overall analyses between tie and non-tie problem solution times to assess whether a tie advantage could be observed throughout development. Then, we studied the evolution of the problem size effect for tie and non-tie problems separately, focusing on the different categories of problems.

To better understand problem size effects, we capitalized on our longitudinal approach and aimed at determining whether tie and non-tie problem size effects in Grade 5 could be predicted by children's problem-solving speed 2 years earlier. Finally, we looked at individual differences to determine whether our general pattern of results was driven by the majority of our participants.

For solution times, which presented a skewed distribution, we calculated median values for each participant and each problem. In contrast, we calculated mean values for the percentages of correct responses. We performed statistical analyses on R (R Core Team, 2017). The data containing the averaged solution times for each problem and each age group are available on the Open Science Framework (OSF) (doi:https://doi.org//10.17605/osf.io/etgbs).

Because parts of our data are not cross sectional but rather longitudinal, we performed analyses on three different sets of data for general analyses of variance (ANOVAs). The first set on which we conducted an ANOVA was the cross-sectional data set of Grade 1 children (beginning and end of Grade 1), the second was the longitudinal data set (Grade 3, Grade 4, and Grade 5), and the last was the data set from the adult group. When needed, results were corrected for violation of the sphericity assumption using the Greenhouse–Geisser correction. All the results were Holm corrected.

Results

Accuracy

A precise description of the mean percentages of correct answers for each category of problems in each age group can be found in Table 1. A noticeable result is that, on average, accuracy was higher for tie problems than for non-tie problems.

Solution times

Tie versus non-tie problems

Mean solution times depending on the problem sums for tie and non-tie problems are represented in Fig. 1 for each age group. The similarity in the distribution of solution times across age groups is striking for problems with a sum up to 10. We confirmed this similarity through a series of Holm-corrected correlational analyses conducted on solution times depending on the sum for tie and non-tie problems altogether. These analyses, which are reported in detail in the online supplementary material, revealed correlations ranging from .95 to .99 (Table S1).

To analyze the difference between tie and non-tie problems, we conducted a series of ANOVAs on solution times. Size effects were not statistically considered in these analyses but are analyzed in the following parts of this Results section. For Grade 1 children, we conducted an ANOVA with age group (beginning or end) as a between-factor variable and problem type (tie or non-tie) as a within-factor variable. There was a significant main effect of problem type, $F(1, 76) = 288.98$, $p < .001$, $\eta_p^2 = .79$, with shorter solution times for tie problems (1932 ms) compared with non-tie problems (3414 ms), as well

Table 1
Mean percentages of correct answers for each problem type, problem category, and age group.

Age group	Non-tie problems				Tie problems	
	1-Problems	Small	Medium	Large	Small	Large
Beginning of Grade 1	96.94	89.49	82.58		95.47	
End of Grade 1	98.67	95.12	92.93		96.83	
Grade 3	97.39	96.64	96.50	91.62	98.99	91.70
Grade 4	97.98	97.29	96.95	93.84	98.56	95.89
Grade 5	97.46	97.09	96.84	93.00	98.56	97.17
Adults	99.33	97.55	98.09	94.83	99.02	99.02

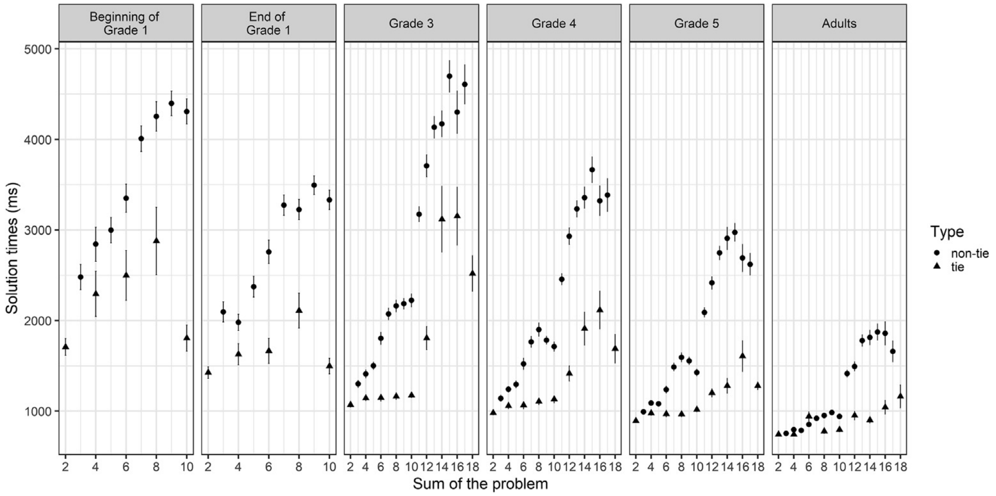

Fig. 1. Mean solution time (with standard errors) for tie and non-tie problems depending on their sums for each age group.

as a main effect of age group, $F(1, 76) = 11.85$, $p < .001$, $\eta_p^2 = .13$, with longer solution times at the beginning of Grade 1(3655 ms) than at the end of Grade 1 (2893 ms). The interaction between the two variables was not significant, $F(1, 76) = 1.65$, $p = .20$.

For the longitudinal data set, we conducted an ANOVA with age group (Grade 3, Grade 4, or Grade 5) and problem type (tie or non-tie) as within-factor variables. The analysis indicated a main effect of problem type, $F(1, 50) = 115.64$, $p < .001$, $\eta_p^2 = .70$, with shorter solution times for tie problems (1439 ms) than for non-tie problems (2345 ms). The results also revealed a main effect of age group, $F(1.41, 70.70) = 123.58$, $p < .001$, $\eta_p^2 = .71$. Contrasts indicated that solution times in Grade 3 (2730 ms) were significantly longer than those in Grade 4 (2179 ms), $t(100) = 9.55$, $p < .001$, which in turn were longer than solution times in Grade 5 (1825 ms), $t(100) = 6.04$, $p < .001$. The interaction between age group and problem type was significant, $F(1.39, 69.61) = 5.48$, $p = .01$, $\eta_p^2 = .10$, showing that the difference between tie and non-tie solution times decreased with age. However, a series of Holm-corrected contrasts revealed that the difference was significant for all age groups: Grade 3, $t(81) = 10.87$, $p < .001$; Grade 4, $t(81) = 9.29$, $p < .001$; Grade 5, $t(81) = 8.12$, $p < .001$.

For adults, we conducted a one-way ANOVA with problem type (tie or non-tie) as a within-factor variable, which indicated a significant difference between tie problems (895 ms) and non-tie problems (1235 ms), $F(1, 33) = 72.83$, $p < .001$, $\eta_p^2 = .69$.

These results revealed that, in each age group, solution times were shorter for tie problems than for non-tie problems. Moreover, as observable in Fig. 1, tie and non-tie problems seem to present different developmental patterns, and thus we considered them separately in the following analyses.

Non-tie problems

We conducted Holm-corrected linear regression analyses to determine which predictors could account for the best for non-tie problem solution time distributions. To determine the best predictor to characterize the problem size effect, we employed the variables commonly used to describe it to fit non-tie solution times for all age groups together. When large problems were excluded, analyses with the first operand (O1), second operand (O2), largest operand (maximum), or smallest operand (minimum), or with the sum of the squared operands ($O1^2 + O2^2$), as the predictor were not significant: minimum, $F(1, 2) = 10.74$, $p = .41$, and Fs < 1 for the other predictors. In contrast, the product of the operands, the sum of the operands squared ($[O1 + O2]^2$), and the sum of the operands all were significant predictors, $R^2_{adj} = .75$, $F(1, 15) = 50.16$, $p < .001$; $R^2_{adj} = .78$, $F(1, 6) = 25.52$, $p = .01$; and $R^2_{adj} = .87$, $F(1, 6) = 47.86$, $p = .003$, respectively. As can be seen, the sum of the operands best fit solution times. When we considered each age group separately, the sum of the problem was still the best significant predictor for each group (see Table S2 in supplementary material). The slopes with respect to the sum were 298 ms at the beginning of Grade 1, 229 ms at the end of Grade 1, 150 ms in Grade 3, 104 ms in Grade 4, 85 ms in Grade 5, and 34 ms for adults. With large problems, which were not solved by first graders, the sum of the problem remained the best predictor of solution times for all age groups together: sum, R^2_{adj} = .93, $F(1, 13)$ = 194.10, p < .001; sum squared, R^2_{adj} = .88, $F(1, 13)$ = 100.33, p < .001; minimum, R^2_{adj} = .82, $F(1, 6)$ = 31.84, p = .005; maximum, R^2_{adj} = .73, $F(1, 6)$ = 20.18, p = .01; product, R^2_{adj} = .65, $F(1, 29)$ = 56.95, p < .001; second operand, R^2_{adj} = .55, $F(1, 7)$ = 10.97, p = .03; first operand, R^2_{adj} = .47, $F(1, 7)$ = 7.98, p = .03; sum of the squared operands, R^2_{adj} = .38, $F(1, 32)$ = 20.88, p < .001. When we considered each age group separately, solution times were best predicted by the sum for Grade 3 and Grade 4 children, by the maximum for Grade 5 children, and by the second operand for adults (see Table S3 in supplementary material). Even though the sum was not the best predictor for Grade 5 and adult solution times, it was still a good fit with a difference of adjusted R^2, with the best predictor of maximum .03. Because the sum of the problem best fit the solution time in the vast majority of cases, we used it to characterize the problem size effect.

Even though the sum of the problem significantly predicted solution times for each age group, a close look at solution time distributions according to the problem sum showed that, already from Grade 1, solution times did not linearly increase with the sum of the problems (Fig. 1). To characterize the problem size effect more precisely, we focused on solution time distribution according to problem category. When solution times were rescaled for each age group (Fig. 2), impressive similarities in these patterns for each category of problem were observed across development.

Small non-tie problems. Because small non-tie solution times follow a linear pattern throughout development, it was possible to fit them with a linear regression for each participant, with the problem sum as the predictor. The slopes were extracted and Holm-corrected *t*-tests were performed to assess whether they were statistically greater than 0 for each age group. This was indeed the case, with $t(36)$ = 6.00, p < .001 for beginning of Grade 1; $t(40)$ = 5.62, p < .001 for end of Grade 1; $t(50)$ = 6.74, p < .001 for Grade 3; $t(50)$ = 6.20, p < .001 for Grade 4; $t(50)$ = 6.23, p < .001 for Grade 5; and $t(33)$ = 4.52, p < .001 for adults.

Medium non-tie problems. Concerning medium problems, and in sharp contrast to small problems, solution time distributions were not linear. This was true for each age group. In fact, medium problems barely presented a problem size effect, especially for sums to 8, 9, and 10. To scrutinize this plateau, we performed an ANOVA on solution times for each data set (Grade 1, longitudinal, or adult) with age group and problem sum (sum to 8, sum to 9, or sum to 10) as variables (see Table S4 in supplementary material for complete results of analysis). Interestingly, the significant differences in solution times with respect to problem sum always contradicted classical size effects, with longer solution times for smaller problems compared with larger problems. More precisely, solution times were significantly different between sum to 8 and sum to 10 problems for Grade 5, $t(265)$ = 2.95, p = .03, and Grade 4, $t(265)$ = 3.59, p = 003, and between sum to 9 and sum to 10 problems for adults, $t(66)$ = 2.57, p = .04.

Furthermore, it is striking that, for each age group, problems with a sum to 7 were solved quickly when they belonged to the medium category rather than the small category. We confirmed this through a series of Holm-corrected *t*-tests, showing that this difference was significant in each age group: $t(32)$ = 2.99, p = .02 for beginning of Grade 1; $t(39)$ = 5.04, p < .001 for end of Grade 1; $t(50)$ = 4.04, p = .001 for Grade 3; $t(50)$ = 2.84, p = .02 for Grade 4; $t(50)$ = 2.75, p = .02 for Grade 5; and $t(33)$ = 3.51, p = .005 for adults. Note that children who presented missing data were removed from this analysis.

1-Problems. Finally, 1-problems presented a limited size effect throughout development and an M shape with systematic decrease in solution times around 4 + 1 and 5 + 1 problems. A closer examination of solution times (Fig. 3) confirmed the M−shaped distribution with a decrease at sum to 5 and sum to 6 problems. This pattern of distribution was particularly puzzling because it appeared from the beginning of Grade 1, where solution times were about 2000 to 3000 ms, and persisted until adulthood, where solution times were about 700 to 900 ms.

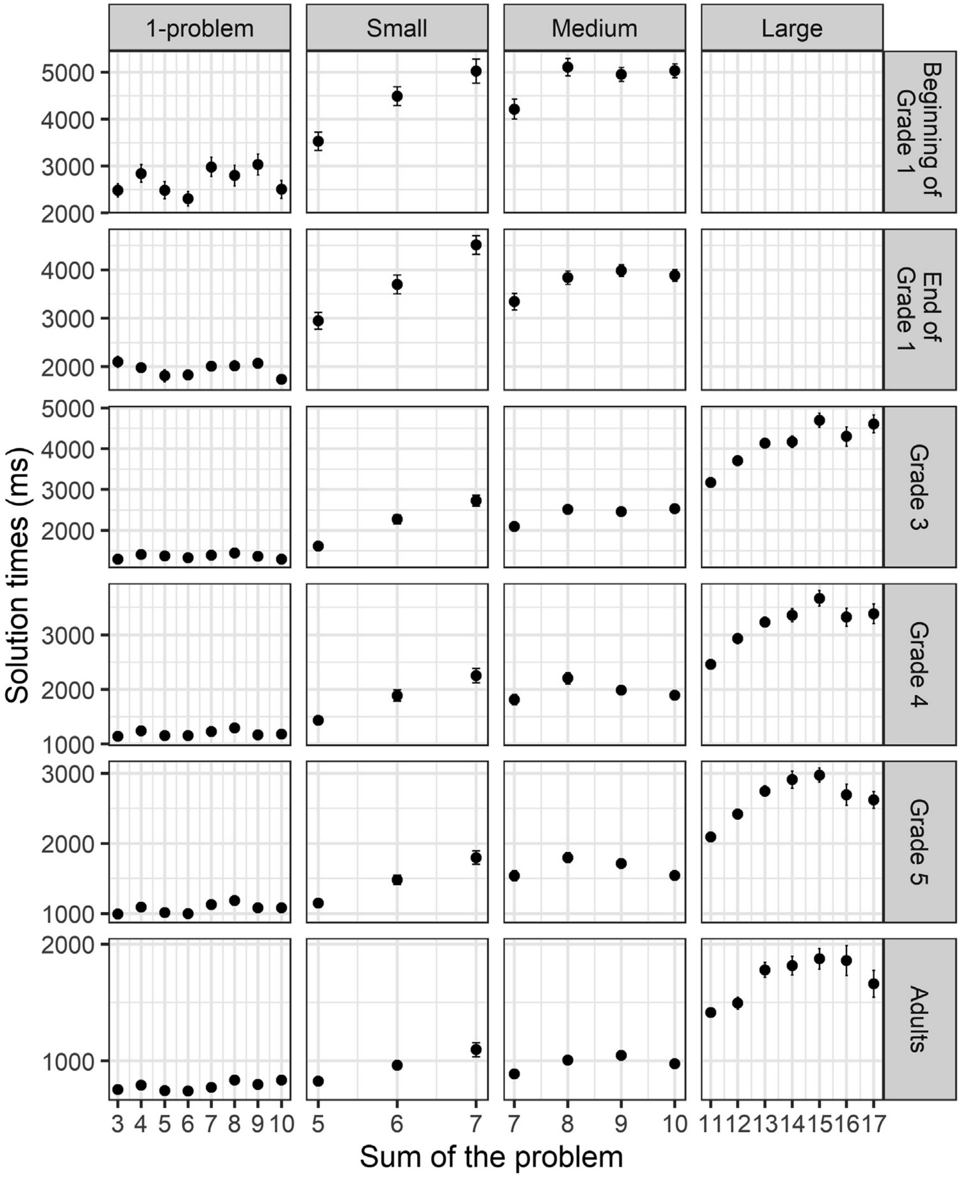

Fig. 2. Mean non-tie problem solution times (with standard errors) for each age group and each problem category according to the sum of the problem.

Large non-tie problems. For large problems, and as observable in Fig. 2, there was an increase in solution times between sum to 11 and sum to 13 problems for each age group. For sum to 13 problems, the problem size effect was less clear. For sum to 15 problems and regardless of age group, solution times did not monotonically increase with respect to the sum.

Tie problems

For a given tie problem, the first (O1), second (O2), largest (maximum), and smallest (minimum) operands all are equal to the same value. Consequently, among the classical predictors of solution

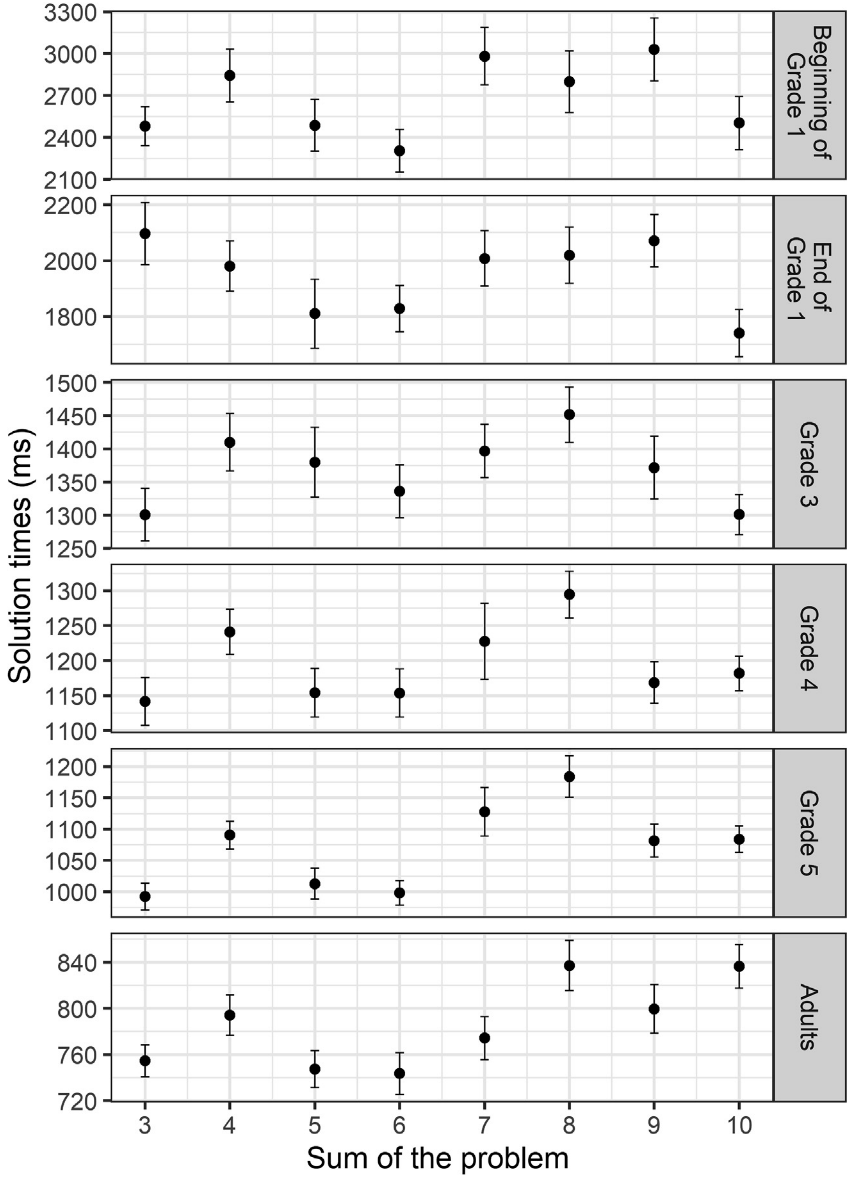

Fig. 3. Mean non-tie 1-problem solution times (with standard errors) for each age group according to the sum of the problem.

times, it is possible to consider only the sum and the sum of the operands squared (i.e., $[O1 + O2]^2$). Our analyses revealed that the two variables predicted equally well tie problem solution times (see Table S5 in supplementary material for detailed results of analyses). Therefore, and for the sake of clarity and simplicity, we represented size effects for tie problems using the sum of the problems.

Small tie problems. In Grade 3, solution times for problems with sums up to 10 were relatively constant across sums (Fig. 4). In contrast, in Grade 1, and especially at the beginning of Grade 1, solution times monotonically increased with the sums of the problems until the sum to 8 problems, and they abruptly decreased for the sum to 10 problems. When we removed 5 + 5 problems from the analyses,

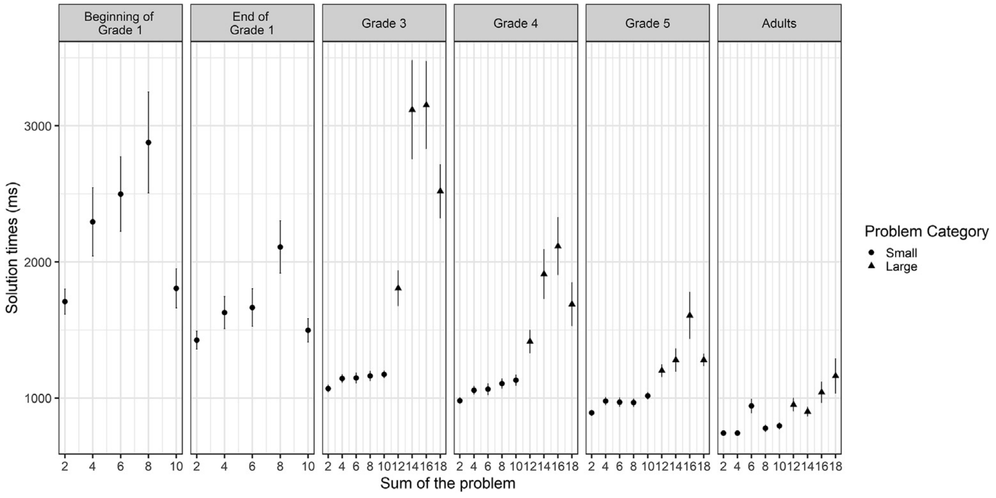

Fig. 4. Mean tie problem solution time (with standard errors) according to the sum of the problem for each age group and each problem category.

solution times in Grade 1 were predicted by the sums of the problems [beginning of Grade 1, R^2_{adj} = .94, $F(1, 2)$ = 52.29, p = .04, slope of 186 ms; end of Grade 1, R^2_{adj} = .81, $F(1, 2)$ = 13.94, p = .07, slope of 104 ms] and by the sums of the problems squared [beginning of Grade 1, sum squared, R^2_{adj} = .84, $F(1, 2)$ = 16.51, p = .06, slope of 18 ms; end of Grade 1, R^2_{adj} = .89, $F(1, 2)$ = 25.80, p = .07, slope of 11 ms]. This demonstrated the existence of a size effect for small problems with sums up to 8 in Grade 1 children. By contrast, and concerning the same problems, none of the predictors accounted for solution time distributions from Grade 3 [for the sums: Grade 3, $F(1, 2)$ = 6.73, p = .24, slope of 14 ms; Grade 4, $F(1, 2)$ = 17.71, p = .10, slope of 19 ms; Grade 5, $F(1, 2)$ = 1.87, p = .61, slope of 11 ms; adults, $F < 1$, slope of 15 ms; for the sums squared: Grade 3, $F(1, 2)$ = 3.39, p = .24; Grade 4, $F(1, 2)$ = 8.38, p = .10; Grade 5, $F < 1$; adults, $F < 1$, slopes of 1 ms for all age groups from Grade 3]. Therefore, no evidence existed for an increase in solution times for tie problems with sums up to 8 from Grade 3 to adulthood.

Large tie problems. For each age group, large problems presented longer solution times than small problems (Fig. 4). Nevertheless, solution time patterns for large tie problems changed throughout development, with shorter solution times for 9 + 9 problems than for 7 + 7 and 8 + 8 problems until Grade 5 but longer solution times for 9 + 9 problems than for smaller problems in adults.

Longitudinal profiles

We used our longitudinal approach to determine whether the problem solution speed in Grade 3 could predict the problem size effects in Grade 5. We calculated the problem size effects only for small problems because they followed a general linear trend for both tie and non-tie problems, contrary to the other categories of problems that we studied. More precisely, for fifth graders, we fitted solution times with a linear regression for small tie and non-tie problems separately using the problem sum as the predictor. We then extracted slopes from these regressions to characterize the tie and non-tie problem size effect. Finally, we performed linear regressions to assess whether Grade 5 slopes could be predicted by the mean solution times of children in Grade 3 (i.e., overall problem-solving speed). The results indicated that the problem-solving speed in Grade 3 significantly predicted small non-tie problem slopes, R^2_{adj} = .44, $F(1, 49)$ = 39.52, $p < .001$, with larger slopes for slower children. In contrast, problem-solving speed did not predict small tie problem slopes, $F(1, 49)$ = 1.77, p = .19.

Individual differences

For non-tie problems, we found a classical problem size effect pattern in 53.67% of our participants for small problems (i.e., sum to 7 > sum to 6 > sum to 5) but in only 16.67% of our participants for medium problems (i.e., sum to 10 > sum to 9 > sum to 8). For small tie problems, we found a classical problem size effect pattern in 20.38% of our participants (i.e., sum to 8 > sum to 6 > sum to 4). We performed chi-square for goodness of fit tests to determine whether the percentage of participants showing the classical size effect pattern differed across groups. The analyses revealed nonsignificant results for small non-tie problems, $\chi^2(5, N = 139) = 1.20$, $p = .94$, and medium non-tie problems, $\chi^2(5, N = 44) = 3.11$, $p = .68$, as well as for small tie problems, $\chi^2(5, N = 53) = 8.58$, $p = .13$ (see the "Individual differences" section of the supplementary material for more information).

Discussion

In this study, we tracked the evolution of one-digit addition problem solving through development by considering specific categories of problems. We observed solution times associated with 1-problems and small, medium, and large additions in children from Grade 1 to Grade 5 and in adults. We were especially interested in the progression of size effects in tie and non-tie problems within the categories mentioned above. We thought that this approach could be valuable for studying the puzzling effects of the current literature in more depth. Among these effects, we focused our attention on tie problems, which present null or very limited size effects compared with non-tie problems (e.g., LeFevre et al., 2004). We also focused on medium problems with sums of 7, 8, 9, and 10, which were also reported as not presenting size effects (Uittenhove et al., 2016). In addition, we focused on problems involving 1 (i.e., 1-problems), which are sometimes considered to be part of a special category and are solved through the use of the number-after rule (e.g., Baroody, 1995).

We showed that, from the beginning of Grade 1, 1-problems are solved more quickly than any other problems whatever the size of the other operand might be (from 3 to 9). Moreover, they never present the classical size effect because their solution times are relatively constant across problems despite an M-shaped distribution. This M-shaped distribution is created via a decrease in the solution times for 4 + 1 and 5 + 1 problems. These peculiarities (i.e., short and relatively constant solution times across problems and the M-shaped distribution) give credit to the hypothesis that 1-problems are processed differently from other problems, starting at the beginning of learning. The idea that the number-after rule is used to solve them is coherent with the general pattern that we observed even though, to our knowledge, existing theories do not account for shorter solution times for 4 + 1 and 5 + 1 problems. Note that the conclusion that a rule is used for very small problems stands in opposition to Uittenhove et al.'s (2016) assumption that very small 1-problems all are processed through automatic counting. This discrepancy between conclusions stems from 1-problems involving 4, which showed a decrease in solution times compared with smaller problems for all age groups in the current study. Uittenhove et al. did not observe such a decrease in adults.

Concerning tie problems, we showed for the first time that size effects can be documented for small problems in young children. Indeed, we observed steep slopes from problems ranging from 1 + 1 to 4 + 4. This result suggests that children do not systematically learn the answers associated with tie problems by rote at the beginning of schooling. In fact, small tie problems behave exactly as retrieval models could predict. The shift from counting to retrieval is indeed perfectly operationalized in a shift from steep slopes associating solution times and sums in the early stage of development to flat slopes from Grade 3. In fact, from Grade 3 to adulthood, 83% of our population showed no size effect for small tie problems (i.e., sum up to 10). Note, however, that this consistency in the behavior of our population from Grade 3 was not as obvious in Grade 1, where we observed the size effect in only 23% of children for problems ranging from 2 + 2 to 4 + 4. However, the problem of 5 + 5 constitutes an exception to the model just described. In Grade 1, solution times for this problem are shorter than they are for smaller problems. This is probably due to the fact that children have already memorized the answer 10 because they know that they have 10 fingers, with 5 fingers on each hand (e.g., McLennan, 2019). Still, from Grade 3, this early special status of 5 + 5 no longer affects solution times. From this developmental point, a difference in solution times between this problem and smaller problems is no longer observed. Another problem that presents a peculiarity in our data set is 3 + 3 in adults. Unexpectedly, this problem led to longer solution times than what would have been expected. We must admit that we do not have any plausible explanation for this phenomenon, but among the bulk of our results it might be deemed a negligible result.

Large tie one-digit problems, which we studied only from Grade 3, led to longer solution times than smaller problems, and their distribution was quite hectic and changed across age groups. Some interpretations given in the past for non-tie problems can be put forward here. First, they might lead to a mixture of retrieval and procedure strategies (LeFevre et al., 1996). Second, the procedures used might strongly depend on the nature of the problem (Chen & Campbell, 2014) and on individuals' habits and preferences (e.g., 9 + 9 can be solved by adding 10 and 10 and subtracting 2 or by rounding up only one of the operands, 10 + 9–1; e.g., Lemaire & Lecacheur, 2011).

Before concluding concerning tie problems, we need to explore the reason behind their relatively short solution times compared with non-tie problems from the beginning of learning. Several interpretations are provided in the literature. As already noted in the Introduction, tie problems could benefit from an encoding advantage over non-tie problems because repeated operands are perceptively apprehended more quickly than different operands (Blankenberger, 2001). Moreover, tie problems could be accessed more easily from memory than non-tie problems (LeFevre et al., 2004). Finally, answers to non-tie problems could be better accessed when the larger operand is presented first (i.e., Max + Min rather than Min + Max; Groen & Parkman, 1972). During the solving process, a stage consisting of the comparison of the sizes of the operands therefore would be needed to determine whether operands are presented in the preferred order. This stage is obviously not necessary for tie problems, hence the existence of shorter solution times for this category of problems than for non-tie problems (Butterworth, Zorzi, Girelli, & Jonckheere, 2001).

Concerning small non-tie addition problems involving operands from 2 to 4, their developmental pattern in solution times does not follow the same trajectory as that of tie problems. We observed a significant size effect created by a monotonic increase in solution times in each age group from the beginning of learning until adulthood. This result pattern was observed for the majority of our participants. The only change over the course of development was the size of the slope associated with the sums of the problems and the solution times, which decreased with age and practice. Two main explanations drawn from two concurrent theoretical models of the literature can be provided here. First, it is possible that the shift between counting to retrieval never occurs. This is coherent with some theories, such as the automatized counting theory, suggesting that even adults rely on procedural strategies to solve very simple problems. This theory assumes that the development of expertise for non-tie problems consists of the acceleration of one-unit step counting until automatization rather than a shift from counting to retrieval (Fayol & Thevenot, 2012; Mathieu, Epinat-Duclos, Léone, et al., 2018; Mathieu, Epinat-Duclos, Sigovan, et al., 2018; Mathieu, Gourjon, Couderc, Thevenot, & Prado, 2016; Uittenhove et al., 2016). Within this framework, only the answers of tie problems are eventually retrieved from memory, and this is the reason why no size effect was observed for this category of problems. On the contrary, linear increases in solution times, depending on the sizes of problems, would reflect counting. Interestingly, Svenson (1985) already presented the idea that very simple addition problems involving the operands of 1 and 2 are solved using one-unit step counting procedures. The use of reasoning strategies such as two-step counting could also explain the pattern of results obtained in the current study for small non-tie problems (Baroody & Coslick, 1998; Purpura, Baroody, Eiland, & Reid, 2016). Second, and alternatively, residual size effects in older children and adults could reflect the increasing interference and decreasing frequency and strength of associations between answers and operands with an increase in the magnitude of the operands within the retrieval networks (Campbell, 1995; Hamann & Ashcraft, 1986; Siegler & Shrager, 1984). Nevertheless, and as addressed in the next paragraph, these latter explanations are difficult to defend in light of the results that we obtained for the medium problems.

Concerning non-tie medium problems—that is, problems with sums up to 10 and at least one operand larger than 4—it is striking to note that no size effect can be observed whatever the age group of the participants may be. Indeed, solution times for problems with sums of 8, 9, and 10 do not increase. This result cannot be due to a lack of power in our analyses because, as can be seen in Fig. 2, a pattern wherein solution times increase monotonically across the three sums never occurs on average in our age groups. In fact, when solution times vary across the sums of 8, 9, and 10, inverse size effects can be observed. Moreover, individual difference analyses indicate that only 17% of our participants presented a classical size effect for such problems. This lack of size effect could be due to the fact that sum to 10 problems are often viewed as having a special status and could be retrieved more easily than other problems (e.g., Aiken & Williams, 1973), probably because they often serve as anchors for problem decompositions (e.g., 4 + 7 is 7 + 3 = 10 + 1; Chen & Campbell, 2018). However, our results showed a plateau between problems with sums from 8 to 10, not just a decrease in the solution times for sum to 10 problems. Still, Chen and Campbell (2018) argued that sum to 10 and sum to 9 problems artificially drive the lack of size effect for medium problems, which could benefit from the sum to 10 problem memory advantage. In light of the current results, we think that this line of reasoning does not hold because the absence of size effect for medium problems is observable even in young children, who are unanimously viewed as using counting procedures to solve such problems. Even if it could be considered that sum to 10 problems already have a special memory status at an early age, extensive memory practice should be necessary before sum to 9 problems could benefit from such a sum to 10 advantage. In addition, whatever the age of participants, solution times for problems with sums to 7 within the medium category are shorter than those within the small category, which is in contradiction with the existence of classical size effects. Once the lack of size effect for medium problems is admitted, an interpretation of size effects for small non-tie problems in terms of retrieval time variations is difficult to defend. Undeniably, medium problems should not be immune from the interference, frequency, and strength of association effects and therefore should present size effects in continuity with small problems. However, we think that the automatized counting theory is not defeated by the fact that size effects are observed only for the small category of problems. Indeed, counting theories suggesting the use of quick and automatized procedures are often limited to very small problems involving operands from 2 to 4 (Uittenhove et al., 2016) or even 1 or 2 (Svenson, 1985). The rationale behind this limitation is that too many steps due to large operands cannot be executed quickly, and individuals might prefer other strategies, including retrieval, for such larger problems (Thevenot, Dewi, Bagnoud, Uittenhove, & Castel, 2020). Nevertheless, nowadays this lack of variations for problems with sums to 8, 9, and 10 comes with no definite theoretical explanation. Further investigation might address the possibility that extreme variability in the use of arithmetic strategies or in the application of heuristics and rules can affect average solution times to such an extent that no size effect is observable. This idea stems from the fact that a similar plateau is observed for large problems with a sum ranging from about 13 to 17, and the results of such large problems are usually not considered to be retrieved from memory (e.g., in adults: Chen & Campbell, 2014; in 10-year-old children: Fanget, Thevenot, Castel, & Fayol, 2011).

To conclude, an important aspect of our data set is that it provides support for a theory advocating that the developmental pattern for small tie problem consists of a shift from counting to retrieval, whereas the developmental pattern for small non-tie problems consists of the acceleration of counting procedures until automatization. Another argument for this conclusion might be found in the results concerning the individual profiles that we established from our longitudinal approach. We showed that the overall speed of arithmetic problem solving in Grade 3 predicts the magnitude of size effects for small non-tie problems but not for small tie problems in Grade 5. This might suggest that the same cognitive mechanisms do not drive tie and non-tie problem-solving processes. This is coherent with the idea that the counting speed in Grade 3 predicts the counting speed in Grade 5 for non-tie problems. Counting speed would not logically predict the retrieval speed for tie problems in Grade 5. Nevertheless, it is also possible to advocate that non-tie problems are not all solved through retrieval in Grade 5. Rather, massively resorting to this strategy will occur later during development. As already explained, this is not what our results suggest, but further longitudinal studies on a broader age range might be useful for addressing these questions.

To sum up, we showed in the current research that 1-problems constitute a special category of problems from the beginning of learning because they are solved more quickly than any other problems, whatever the size of the problem, and because they never present classical size effects. Without providing direct evidence for the use of a specific strategy, our results are coherent with the idea that a 1-problem is solved using a rule according to which its sum is the number after the other operand in the count sequence (e.g., Baroody, 1995). In contrast to 1-problems, small tie problems present classical size effects at the beginning of learning that quickly disappear during the course of development. The evolution of small tie problem solving therefore is very likely to consist of a shift from counting strategies to retrieval. A drastically different process appears to account for the evolution of small non-tie problem solving, which seems to consist of the acceleration of counting procedures until automatization. This conclusion is based on the observation of clear classical size effects throughout development whatever the age of individuals may be. Retrieval time variations fail to account for these effects given that we showed here that solution times do not continuously increase with the size of problems. Therefore, our research questions the existence of one of the supposedly most robust and reliable effects described in the numerical cognition literature, namely the size effect.

A final point that deserves to be discussed is related to educational practices and the question of the generalization of our results. In the French-speaking part of Switzerland, where testing took place in the current study, children are supposed to have memorized all simple addition facts by the end of Grade 2. However, teachers can choose the concrete educational tools adopted to reach this goal. Overpracticing counting (e.g., 5 + 3 is 6, 7, 8), promoting finger counting, and implementing rote learning, especially for tie and sum to 10 problems, are among these tools. Nevertheless, informal interviews with teachers in the French-speaking part of Switzerland revealed that addition problems are never learned by rote as systematically as multiplication tables. Investigating the evolution of addition size effects in countries where rote learning is practiced more intensively, such as the Dutch-speaking part of Belgium (De Smedt, 2016), therefore might be very interesting for pursuing our line of research. Note, however, that education does not always seem to influence individuals' performance and strategies in arithmetic. Campbell and Xue (2001) showed that in Asian and Canadian adults, culture, more than education, was the key determinant in the rate of retrieval in mental arithmetic. Investigating the role of children's home numeracy environments in addition size effects might therefore also constitute an interesting future line of research.

Acknowledgment

This work was funded by the Swiss National Science Foundation (100014_152543 and 100019_169208).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2020.104987.

References

Aiken, L. R., & Williams, E. N. (1973). Response times in adding and multiplying single-digit numbers. *Perceptual and Motor Skills, 37*, 3–13.
Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75–106.
Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 527–538.
Ashcraft, M. H., & Fierman, B. A. (1982). Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology, 33*, 216–234.
Ashcraft, M. H., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. In B. H. Ross (Ed.), *Psychology of learning and motivation* (pp. 121–151). San Diego: Academic Press.
Baroody, A. J. (1995). The role of the number-after rule in the invention of computational shortcuts. *Cognition and Instruction, 13*, 189–219.
Baroody, A. J. (2018). A commentary on Chen and Campbell (2017): Is there a clear case for addition fact recall? *Psychonomic Bulletin & Review, 25*, 2398–2405.
Baroody, A. J., & Coslick, R. T. (1998). *Fostering children's mathematical power*. Mahwah, NJ: Lawrence Erlbaum.
Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2012). Fostering at-risk kindergarten children's number sense. *Cognition and Instruction, 30*, 435–470.
Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account?. *Cognition, 128*, 35–44.
Blankenberger, S. (2001). The arithmetic tie effect is mainly encoding-based. *Cognition, 82*, B15–B24.
Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. R. (2001). Storage and retrieval of addition facts: The role of number comparison. *Quarterly Journal of Experimental Psychology A, 54*, 1005–1029.
Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121–164.

Campbell, J. I. D. (2015). How abstract is arithmetic?. In R. C. Kadosh & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 140–157). Oxford, UK: Oxford University Press.
Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 39*, 338–366.
Campbell, J. I. D., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition, 86*, 71–96.
Campbell, J. I. D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In J. I. D. Campbell (Ed.), *Advances in Psychology, vol. 91: The nature and origins of mathematical skills*. Amsterdam: North-Holland, pp. 331–364.
Campbell, J. I. D., & Timm, J. C. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin & Review, 7*, 692–699.
Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*, 299–315.
Chen, Y., & Campbell, J. I. D. (2014). Generalization effects in Canadian and Chinese adults' simple addition. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 68*, 152–157.
Chen, Y., & Campbell, J. I. D. (2018). "Compacted" procedures for adults' simple addition: A review and critique of the evidence. *Psychonomic Bulletin & Review, 25*, 739–753.
Chen, Y., Loehr, J. D., & Campbell, J. I. D. (2019). Does the min-counting strategy for simple addition become automatized in educated adults? A behavioural and ERP study of the size congruency effect. *Neuropsychologia, 124*, 311–321.
Fanget, M., Thevenot, C., Castel, C., & Fayol, M. (2011). Retrieval from memory or use of procedures for addition in children: The use of the operand-recognition paradigm in 10-year-old children. *Swiss Journal of Psychology, 70*, 35–39.
De Smedt, B. (2016). Individual differences in arithmetic fact retrieval. In D. B. Berch, D. C. Geary, & K. Mann-Koepke (Eds.), *Development of mathematical cognition: Neural substrates and genetic influences* (pp. 219–243). San Diego, CA: Elsevier Academic Press.
Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition, 123*, 392–403.
Forster, K. I., & Forster, J. C. (2003). DMDX: A window display program with millisecond accuracy. *Behavior, Research Methods, Instruments, & Computers, 35*, 116–124.
Graham, D. J., & Campbell, J. I. D. (1992). Network interference and number-fact retrieval: Evidence from children's alphaplication. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 46*, 65–91.
Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329–343.
Hamann, M. S., & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction, 3*, 173–192.
LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 216–230.
LeFevre, J. A., Shanahan, T., & DeStefano, D. (2004). The tie effect in simple arithmetic: An access-based account. *Memory & Cognition, 32*, 1019–1031.
LeFevre, J. A., Smith-Chant, B. L., Hiscock, K., Daley, K. E., & Morris, J. (2003). Young adults' strategic choices in simple arithmetic: Implications for the development of mathematical representations. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 203–228). Mahwah, NJ: Lawrence Erlbaum.
Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26*, 282–294.
Mathieu, R., Epinat-Duclos, J., Léone, J., Fayol, M., Thevenot, C., & Prado, J. (2018). Hippocampal spatial mechanisms relate to the development of arithmetic symbol processing in children. *Developmental Cognitive Neuroscience, 30*, 324–332.
Mathieu, R., Epinat-Duclos, J., Sigovan, M., Breton, A., Cheylus, A., Fayol, M., ... Prado, J. (2018). What's behind a "+" sign? Perceiving an arithmetic operator recruits brain circuits for spatial orienting. *Cerebral Cortex, 28*, 1673–1684.
Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., & Prado, J. (2016). Running the number line: Rapid shifts of attention in single-digit arithmetic. *Cognition, 146*, 229–239.
McLennan, D. M. P. (2019). Joyful number talks in kindergarten. *Journal of Teaching and Learning, 13*, 43–54.
Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron, 41*, 983–993.
Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*, 859–862.
Purpura, D. J., Baroody, A. J., Eiland, M. D., & Reid, E. E. (2016). Fostering first graders' reasoning strategies with basic sums: The value of guided instruction. *Elementary School Journal, 117*, 72–100.
R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL: https://www.R-project.org/.
Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Lawrence Erlbaum.
Svenson, O. (1985). Memory retrieval of answers of simple additions as reflected in response latencies. *Acta Psychologica, 59*, 285–304.
Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: A counting model account. *Cognition, 146*, 48–57.
Thevenot, C., Dewi, J., Bagnoud, J., Uittenhove, K., & Castel, C. (2020). Scrutinizing patterns of solution times in alphabet-arithmetic tasks favors counting over retrieval models. *Cognition, 200*. 104272.
Uittenhove, K., Thevenot, C., & Barrouillet, P. (2016). Fast automated counting procedures in addition problem solving: When are they used and why are they mistaken for retrieval? *Cognition, 146*, 289–303.
Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*, 20.