

“Compacted” procedures for adults’ simple addition: A review and critique of the evidence

Yalin Chen¹ · Jamie I. D. Campbell¹

Published online: 9 June 2017
© Psychonomic Society, Inc. 2017

Abstract We review recent empirical findings and arguments proffered as evidence that educated adults solve elementary addition problems ($3 + 2$, $4 + 1$) using so-called compacted procedures (e.g., unconscious, automatic counting); a conclusion that could have significant pedagogical implications. We begin with the large-sample experiment reported by Uittenhove, Thevenot and Barrouillet (2016, *Cognition*, 146, 289–303), which tested 90 adults on the 81 single-digit addition problems from $1 + 1$ to $9 + 9$. They identified the 12 very-small addition problems with different operands both ≤ 4 (e.g., $4 + 3$) as a distinct subgroup of problems solved by unconscious, automatic counting: These items yielded a near-perfectly linear increase in answer response time (RT) yoked to the sum of the operands. Using the data reported in the article, however, we show that there are clear violations of the sum-counting model’s predictions among the very-small addition problems, and that there is no real RT boundary associated with addends ≤ 4 . Furthermore, we show that a well-known associative retrieval model of addition facts—the network interference theory (Campbell, 1995)—predicts the results observed for these problems with high precision. We also review the other types of evidence adduced for the compacted procedure theory of simple addition and conclude that these findings are unconvincing in their own right and only distantly consistent with automatic counting. We conclude that the cumulative evidence for fast compacted procedures for adults’

simple addition does not justify revision of the long-standing assumption that direct memory retrieval is ultimately the most efficient process of simple addition for nonzero problems, let alone sufficient to recommend significant changes to basic addition pedagogy.

Keywords Addition · Procedures · Counting · Retrieval interference

A recent series of prominent articles (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012; Mathieu, Gourjon, Couderc, Thevenot, & Prado, 2016; Thevenot, Barrouillet, Castel & Uittenhove, 2016; Uittenhove, Thevenot, & Barrouillet, 2016) have presented experimental evidence and arguments that educated adults perform simple addition (e.g., $4 + 3 = ?$) using fast procedural algorithms (e.g., automatic counting). This conclusion stands in contrast to the long-held view that adults’ solving of simple addition problems normally evolves from relatively slow procedural strategies in childhood, such as deliberate counting, to fast direct retrieval of answers from an associative network of memorized addition facts, especially for addition problems with a sum ≤ 10 (e.g., Ashcraft & Guillaume, 2009; Barrouillet & Fayol, 1998; Campbell, 1995; Siegler & Shrager, 1984). If, in fact, the development of fast, compacted counting procedures for simple arithmetic is pervasive, it would not only overturn long-standing cognitive theory about simple addition but also related pedagogical practices. Fayol and Thevenot (2012) concluded that this new view would have a “massive impact in the domain of numerical cognition,” including potential effects on national standards for mathematics education (p. 401; see also Thevenot et al., 2016, p. 55). In this review, however, we critique the evidence proffered for a compacted counting theory of adults’ single-digit addition, showing that there are clear violations of

✉ Yalin Chen
yalin.chen@usask.ca

✉ Jamie I. D. Campbell
jamie.campbell@usask.ca

¹ Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, SK, Canada S7N 5A5

the counting model's predictions in the presented data. We also show that the critical data are also predicted by a well-known model of direct memory retrieval for addition facts, the network interference model (Campbell, 1995). We conclude that the evidence for a compacted procedure theory of fast addition performance does not constitute a coherent body of convincing evidence and is far from sufficient to recommend significant revision to modern theory or educational practices for elementary addition.

Compacted counting for simple addition by skilled adults

We begin with what might be perceived as the most direct evidence for fast addition procedures. Barrouillet and Thevenot (2013) conducted new analyses of previously published data (Barrouillet, Lépine, & Camos, 2008) that included adults' response times (RTs) for very-small additions involving the numbers 1 through 4. These were the only addition problems tested, but there were other cognitive tests, including a working memory measure. The large number of observations (92 participants \times 6 repetitions) collected for each of the 16 small addition problems afforded precise estimates of mean RTs across problems. Barrouillet and Thevenot (2013) found that RT for correct answers to these very-small addition problems increased linearly and monotonically with the sum of the operands. Furthermore, the slope for problems composed of two different addends (i.e., the nonties such as $1 + 2$ and $3 + 4$) was 20 ms, about twice the 11 ms slope observed for the tie problems with a repeated operand (i.e., $1 + 1$, $2 + 2$, $3 + 3$, $4 + 4$). Participants with a high working memory span were faster and also presented a shallower RT problem-size slope with the sum of nontie operands (8 ms) compared to low span individuals (26 ms).

Barrouillet and Thevenot (2013) proposed that the very-small addition nonties activated a compacted counting procedure that gives rise to a relatively steep linear problem size effect, whereas the ties were solved by direct fact retrieval from memory, which they assumed is relatively insensitive to problem size. With respect to the differences between individuals with low and high working memory spans, Barrouillet and Thevenot suggested this could reflect differences in the efficiency of a "basic general-purpose resource that affects each atomic step of cognition," such that the "faster responses and flatter slopes observed with higher working memory capacities could result from a capacity to perform more quickly each step of the procedure" (p. 43).

Uittenhove et al. (2016) undertook to replicate the findings of Barrouillet and Thevenot (2013), but tested all 81 addition problems between $1 + 1$ and $9 + 9$ and collected strategy self-reports for each problem (e.g., just remembered the answer, counted up). The purpose of the strategy reports was to

identify specific problems for which reconstructive procedure use was consciously experienced and reported. The conscious procedural strategy trials were expected to be relatively slow (e.g., Campbell & Xue, 2001; LeFevre, Sadesky, & Bisanz, 1996), whereas the compacted counting procedure proposed by Barrouillet and Thevenot was assumed to be very fast and not accessible to conscious experience and therefore likely to be reported as direct memory retrieval (Uittenhove et al., 2016, p. 299). It was important to exclude problems solved by slow, conscious procedures, such as intentional counting or reconstructive strategies, because these could contribute to a problem-size effect for the very-small problems and thereby contaminate measurement of a problem-size effect owed to an automated counting procedure.¹

The theoretically critical analyses of the Uittenhove et al. (2016) experiment were those based on the subset of 51 participants among the 90 tested who were identified as frequent retrievers for small problems. These participants reported no less than 97% retrieval for problems with a sum ≤ 10 and reported 100% retrieval for the very-small nontie additions with both operands ≤ 4 . The size of the sum was used to predict participants' mean problem RT for four types of problems. Very-small nontie problems (both operands ≤ 4) had a mean slope of 46 ms ($SD = 37$) per increment in the sum ($p < .0001$ against the null hypothesis of 0 ms slope), for tie problems the slope was 8 ms ($SD = 7$, $p < .001$), $n + 1$ with $n > 4$ problems had a slope of 7 ms ($SD = 22$, $p < .05$), but there was no problem-size effect (-5 ms, $SD = 44$, $p > .20$) for medium-small nontie problems with sums from 7 to 10 and at least one operand > 4 . The positive RT slopes for very-small nonties, ties, and $n + 1$ problems cannot be attributed to differences in sum articulation time because participants' mean RT for each problem was corrected for differences in time to verbally produce the sum (Uittenhove et al., 2016, p. 293).

The key result for Uittenhove et al. (2016), however, was the substantial problem-size effect for the 12 very-small addition problems with both operands ≤ 4 , which replicated Barrouillet and Thevenot's (2013) results for the same problems. A very similar pattern of mean RTs for these problems was also observed in a group of 10-year-old children (Thevenot et al., 2016), although of course the children were substantially slower, with a steeper problem-size slope than the adults. The relatively small, albeit statistically significant, problem-size effect for ties and $n + 1$ problems found were assumed to be negligible and their sources unclear (Thevenot et al., 2016, p. 295). The absence of a problem-size effect for the 18 medium-small problems, however, suggested a

¹ Thevenot et al. (2016, p. 55) expressed doubt about whether "the term 'counting' used to qualify those automated procedures [for addition] is the most appropriate" and that "a more general term such as 'reconstructive' procedures might be more suitable." Uittenhove et al. (2016), however, used the term *reconstructive strategies* to refer to conscious addition procedures that were measured through participants' verbal self-reports.

boundary condition when small addition problems had at least one operand greater than 4. Moreover, for the 51 frequent retrievers in the Uittenhove et al. study, the RT slope associated with the sum for very small problems was negatively related to working memory capacity ($r = -.36, p < .01$) so that lower capacities were associated with a steeper slope, again replicating Barrouillet and Thevenot; but this association between RT slope and working memory span was not observed for ties, $n + 1$, or medium-small problems.

Following Barrouillet and Thevenot (2013), Uittenhove et al. (2016) speculated about the cognitive implementation of an automated counting procedure for very-small addition problems and illustrated the model applied to solving $3 + 2$ in their Fig. 9 (p. 299). According to the model, the two operands are sequentially encoded in working memory in an analogical form. In their Fig. 9, the encoding of the operand 3 in working memory is depicted by a horizontal string of three dots or tokens. This representation triggers a recursive algorithm that successively maps each working memory token to the sequential values of a number list in long-term memory, which is represented in Fig. 9 as a list of successive Arabic numbers starting with 1. This next-token-next-value production cycle continues until the third token is mapped to the value 3. Then, the second operand (2) is encoded in working memory, triggering the algorithm to successively map the two tokens generated to the two values succeeding 3 in the number list (i.e., 4 and 5). The counting algorithm terminates, and the final value reached in the number list is expressed by the participant as the spoken word *five*. Variation across problems in time to complete the procedure is determined by the number of production cycles, which corresponds to the sum of the two operands. The time for each cycle is assumed to be a few 10s of milliseconds. The applicability of the procedure “is limited to small quantities that can be represented analogically in a single focus of attention (i.e., no more than four elements)” and the procedure “is so fast that a subject is only aware of its product, hence the subjective experience of a direct retrieval from memory” (Uittenhove et al., 2016, p. 299). Despite these elaborate theoretical and empirical arguments, we argue here that the automatic counting theory, some of its basic assumptions, and the supporting evidence are all questionable. It is important to raise these counterarguments against automatic counting theory because unchallenged acceptance of the theory and evidence, even in this relatively limited application, would establish a *prima facie* case for working-memory-mediated sequential cognition that can be automatized and operate outside awareness.

Challenges for the automatic counting theory and its evidential basis

The counting model outlined by Uittenhove et al. (2016) corresponds to an automated version of the sum strategy for

simple addition that is initially used by children when they count out one by one the cumulative values of both addends, perhaps on their fingers. As Geary, Hoard, Byrd-Craven, and Desoto (2004) observed, however, “improvements in children’s conceptual understanding of counting . . . is reflected in a gradual shift from frequent use of the sum procedure to the min procedure” (p. 122; see also, e.g., Geary, Bow-Thomas & Yao, 1992; Jones & VanLehn, 1994; Siegler, 1987; Siegler & Jenkins, 1989; Svenson, 1985). The min strategy, which starts from the value of the larger operand (the max) and proceeds by counting up a number of times equal to the min, saves counting steps and is generally more efficient than the sum strategy. The transition toward the min strategy is not monolithic, and children continue to use multiple strategies even for the same problem as skill develops (Siegler & Jenkins, 1989; Siegler & Shipley, 1995), but the min strategy eventually becomes the dominant counting procedure used by children for single-digit addition (e.g., Shrager & Siegler, 1998, Fig. 1).² Given this common progression in children’s counting strategies, it is surprising that the addition procedure assumed to become automated by Uittenhove et al. is the more-primitive sum strategy rather than the min strategy that commonly replaces it. Even if we admitted the sum strategy model as plausible, however, the Uittenhove et al. results do not fit it neatly, as we explain next.

Violations of the sum-counting model’s RT predictions for the very-small problems

The major conclusions of Uittenhove et al. (2016) rest on the assumption that “there is a size effect intrinsically related with the [nontie] problems involving operands that do not exceed 4 (i.e., the very small problems) that is significantly stronger than the size effect that can affect any other type of small problems” (pp. 295–296); that is, it rests on the assumption that this category of very-small addition problems exists as a statistically and functionally distinct subset of problems. This conclusion is supported statistically in their paper by an analysis showing that the RT slope of the six $n + 1$ problems (e.g., $3 + 1$ or $1 + 3$), with n between 2 and 4, was steeper than when n was greater than 4 (28 ms vs. 7 ms; p. 295). Thus, within the $n + 1$ problems, there appears to be a statistical boundary in the problem-size effect at $n \leq 4$. When the $n + 1$ very-small problems were excluded and only the other six very-small problems were included for analysis (i.e., $2 + 3, 3 + 2, 2 + 4, 4 + 2, 3 + 4, 4 + 3$), there remained a statistically robust RT slope related to the sum (34 ms; p. 295). Additionally, the RT for medium-small problems (sums from 7 to 10, with both addends greater than 1 and at least one operand greater than 4)

² It is important to emphasize that all of these researchers proposed that direct memory retrieval ultimately becomes the dominant strategy over any procedural strategy for small simple addition problems.

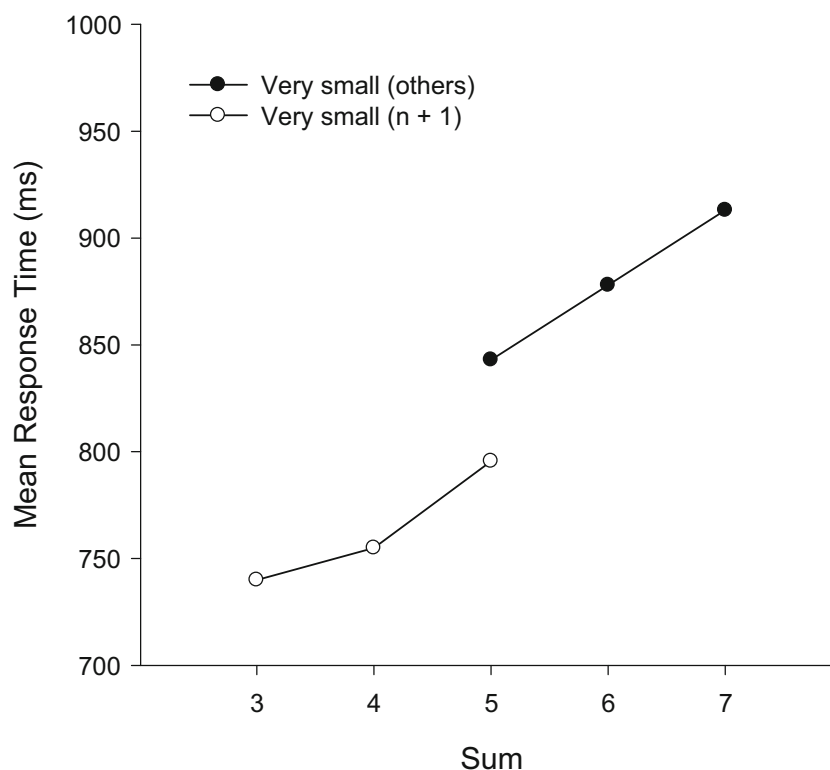


Fig. 1 Mean RT (ms) for the 12 very small problems averaged over operand order from the Uittenhove et al. (2016) frequent-retriever analysis (derived from their Fig. 6, $n = 51$)

did not correlate significantly with the sum (p. 295). Thus, again, there appeared to be a boundary at a maximum operand of 4 in the problem-size effect for small problems. Taken together, these results seem to identify the set of very-small nontie problems with operands ≤ 4 as a statistically distinct category of problems with a robust problem-size effect linked to the sum of the operands.

In our view, however, the analyses reported neglect important features of the data and also neglect known RT phenomena of simple addition that would contribute to their result. First, there is very clear statistical evidence in the results reported that there is an RT boundary within the 12 very-small problems between the six problems involving 1 and the other six very-small problems. Figure 1 depicts the mean RT for the 12 very-small problems averaged across operand order from the Uittenhove et al. (2016) frequent-retriever analysis derived from their Fig. 6 (which is also depicted in the right panel of our Fig. 3). According to the automatic counting model proposed by Uittenhove et al., the pattern of means in our Fig. 1 is determined by a sum-counting algorithm that applies uniformly across this special group of problems. We conducted a regression analysis of the 12 RT means with two predictors: the problem's sum and whether or not the problem contained 1 as an operand (the six $n + 1$ problems were coded with a one, and the other six very-small problems were coded as zero). The resultant regression model had an adjusted R^2 of .948,

$F(2, 9) = 100.7, p < .0001$. The sum had a slope of 31.9 ms ($\beta = .639, t = 5.86, p < .001$), and the problem-type factor (i.e., $n + 1$ vs. others) entered the model with a slope of -51.1 ms ($\beta = .396, t = -3.64, p = .005$). In other words, once variation in mean RT associated with the sum was accounted for, the $n + 1$ problems were 51 ms faster than the remaining very-small problems. This indicates that there was a statistically robust RT discontinuity between these two subsets of the very-small problems and raises doubts that the very-small problem set, as defined, is a valid category.

Furthermore, $1 + 4$ and $4 + 1$ were faster than $2 + 3$ and $3 + 2$, although both operand pairs (i.e., $1, 4$ and $2, 3$) have sums of 5 and should have the same RT according to the sum-counting model. We used the values presented in Uittenhove et al. (2016) Table 1 (p. 294) for all 90 participants to compare the averaged mean RT (817 ms) and standard deviation (125.5) for $1 + 4$ and $4 + 1$ to the averaged mean RT (885.51) and standard deviation (177) for $2 + 3$ and $3 + 2$, treating operand pair ($1, 4$ vs. $2, 3$) as an independent groups factor. The 68.5 ms RT advantage for adding 1 and 4 relative to adding 2 and 3 was significant, $t(178) = 2.995, p = .003, SE = 22.871, \eta^2 = .05$. This analysis, however, includes the 15% of trials for which participants reported reconstructive strategies for these items (Uittenhove et al., 2016, Fig. 3, p. 295). Reconstructive strategy trials might have contributed to this difference. Nonetheless, the frequent retrievers ($n = 51$)

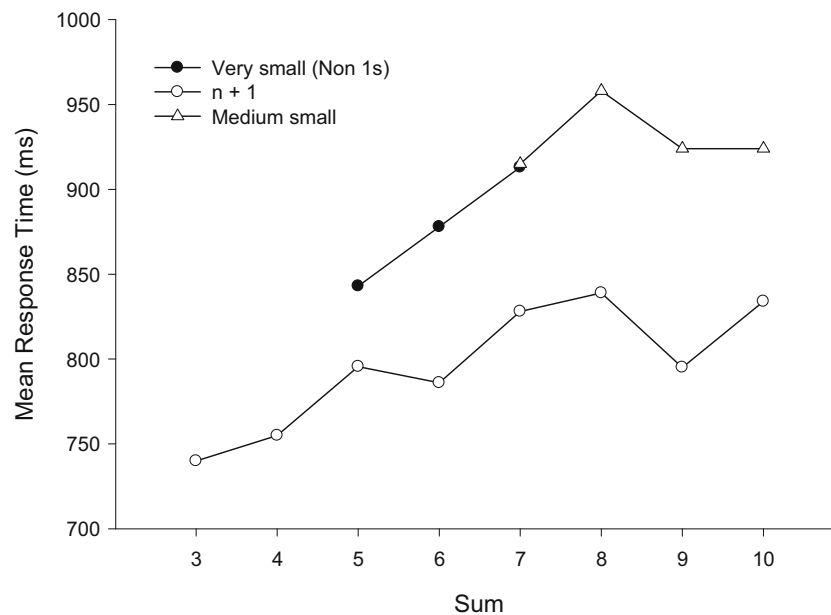


Fig. 2 Mean RT (ms) for $n + 1$, very small and medium small addition problems as a function of the sum based on Uittenhove et al. (2016) Figs. 5 and 6 ($n = 51$ frequent retrievers)

presented a 48-ms RT advantage for adding 1 and 4 (795 ms) relative to adding 2 and 3 (843 ms) that was statistically significant, $t(50) = 2.099$, $p = .04$, $SE = 22.871$, $\eta^2 = .08$, using the standard error from the preceding analysis. Thus, the unpredicted RT advantage to add 1 and 4 compared to adding 2 and 3 was also observed in the frequent-retriever data.

The very-small addition category

There are other challenges for the validity of the very-small problem category identified by Uittenhove et al. (2016). These are apparent in our Fig. 2, derived from Figs. 5 and 6 in the Uittenhove et al. paper. This figure presents mean RT for the very-small problems, the medium-small problems, and the $n + 1$ problems as a function of the problem sum averaged over operand order. In Fig. 2, the average RT for the $n + 1$ operand pairs $4 + 1 = 5$ and $1 + 4 = 5$ is plotted among the $n + 1$ problems rather than counted among the very-small problems as assumed by Uittenhove et al. In the very-small problem plot, the sum of 5 is represented by the average of the other sum-to-five pair $3 + 2$ and $2 + 3$; as we have already shown, and in contradiction to the Uittenhove et al. model, adding the operands 2 and 3 was slower than adding 1 and 4 in this data set.

With respect to our Fig. 2, we draw attention to two features of the pattern of means derived from the data presented by Uittenhove et al. (2016). First, for the $n + 1$ problems, the correlation between the sum and mean RT in Fig. 2 is $.966$ (slope = 20 ms) for n from 2 to 7, but there is a marked drop in RT for the sum of 9 (i.e., $8 + 1$). The very strong RT linearity for the $n + 1$ problems up to

a sum of 8 indicates there is no reason to believe that there is an important RT boundary at $n = 4$ for the $n + 1$ problems. Similarly, for the very-small problems (excluding $n + 1$) and medium-small problems, there is little reason to believe that there is an important RT boundary at $n = 4$ because the RT function is almost perfectly linear for sums 5 through 8 with no evidence of a boundary at $n = 4$; specifically, $r = .995$ (slope = 45 ms) between mean RT in Fig. 2 and the sum of the operands. Thus, the Uittenhove et al. analyses defining the boundary for very-small and medium-small problems between the sums 7 and 8 is not justified by the data, because the RT function is perfectly linear across the 7–8 sum boundary and includes problems with $n > 4$ (e.g., $2 + 5$, $2 + 6$, $3 + 5$). As with the $n + 1$ problems, again there is a marked RT drop at a sum of 9 for the other small problems.

The results in our Fig. 2 can thus be summarized as follows: In the Uittenhove et al. (2016) data, there is a very strong linear relationship between RT and the sum of the operands up to a sum of 8, with $n + 1$ problems presenting a shallower slope than the others (20 ms vs. 45 ms). The mean RT for the problems with sums of 9 or 10 are lower than their sum would predict based on the linear RT pattern up to a sum of 8. With respect to the sum-to-10 problems, it has been observed for decades that the simple addition combinations that sum to 10 are fast and accurate relative to their magnitude (Aiken & Williams, 1973; Campbell, 1995). LeFevre et al. (1996; see also Geary, Bow-Thomas, Liu, & Siegler, 1996) noted that the majority of self-reported transformation strategies for simple addition involved the use of facts that summed

to 10 (e.g., 4 + 7 transformed to 7 + 3 + 1), suggesting that the sum-to-10 facts are salient and highly accessible, often serving as anchors in people's addition strategy repertoire. The sum-to-nine problems could benefit from proximity to the very high memory strength sum-to-10 problems. Indeed, Fig. 3 in Uittenhove et al. (2016, p. 295) shows that sum-to-10 problems had the highest rate of self-reported memory retrieval among the nontie problems in their experiment. None of the analyses or conclusions presented by Uittenhove et al., however, take into account their own or previous evidence that sum-to-10 problems have special memory status in simple addition skills.

Our Fig. 2 shows that an RT boundary associated with the addend 4 in the Uittenhove et al. (2016) data was more apparent than real. Uittenhove et al. carefully selected their analyses to create the appearance of a RT boundary at 4. Thus, it is a theoretical boundary and not an unambiguous empirical boundary. Neither Barrouillet and Thevenot (2013) nor Thevenot et al. (2016) provided direct evidence for an RT boundary at 4 because they tested only problems with both addends ≤ 4 . Furthermore, the theoretical argument supporting the proposed boundary is weak. There is substantial evidence that the span of visual spatial attention is about four items (e.g., Cowan, 2001) and that encoding the numerosities of up to three or four visual objects can be accomplished by an automatic subitizing process (e.g., Mandler & Shebo, 1982). There is, however, no evidence that we are aware of that the Arabic digits up to four (i.e., 1, 2, 3, 4) automatically activate a corresponding number of tokens in spatial working memory, but the Arabic digits for five or higher do not, as the Uittenhove et al. model assumes. Thus, there is no evidence for this central and crucial assumption of the automatic counting theory.

Measurement of reconstructive strategy trials

Another reason to doubt that the results of Uittenhove et al. (2016) support an automatic counting theory is the likely contamination of the frequent-retriever analysis by inclusion of reconstructive strategy trials. Successful excision of reconstructive strategies based on participants' strategy self-reports was critical for the following reasons. As their own data indicate (see also Campbell & Xue, 2001; LeFevre et al. 1996), the rate of self-reported use of reconstructive strategies increases linearly with problem size even for small problems with sums ≤ 10 (see Fig. 3, Uittenhove et al., 2016, p. 295). Furthermore, their Fig. 2 (Uittenhove et al., 2016, p. 294) shows that the increasing proportion of reconstructive strategies as problem size increases inflates the problem-size effect on RT for these problems. This occurs because reconstructive strategies require more, or more difficult, intermediate steps as problem size increases (Campbell & Xue, 2001; LeFevre

et al., 1996). Consequently, even low rates of contamination by reconstructive strategies would generate the problem-size effect observed for the frequent retrievers on very small problems. The extent of such contamination to produce a small-problem-size effect correlated with the sum for the very-small problems (about 50 ms per increment in Uittenhove et al., 2016; about 20 ms in Barrouillet & Thevenot, 2013) would be modest because a small proportion of slow procedure-based trials mixed with faster retrieval trials can substantially inflate the problem-size effect (Campbell & Xue, 2001, p. 311). For example, Fig. 3 in Uittenhove et al. shows a rate of about 22% reconstructive strategies for nontie problems with a sum of 7, and their Fig. 2 shows that this translated into a RT cost of approximately 250 ms.

Such contamination of the retrieval analysis by reconstructive strategy trials is highly probable because Uittenhove et al. (2016) did not measure strategies during the trials actually used for their RT analysis. The addition task data used for the main analyses were collected in Session 1, in which participants completed six blocks of the 81 addition problems without strategy reports. In Session 2, participants received a final addition block of the 81 problems and provided a strategy self-report for each problem (p. 292). The strategies reported in this final block in Session 2 were assumed to represent the strategies used in the first six blocks in Session 1 and used as the basis to exclude participants who reported reconstructive strategies in the final block. The flaw with this design is that the six blocks of Session 1 provided multiple practice trials with each problem and self-reported retrieval for simple arithmetic increases across blocks in which the problems are repeated (Campbell & Timm, 2000; Imbo & Vandierendonck, 2008). Consequently, participants who used reconstructive strategies for various problems early in Session 1 had multiple opportunities to strengthen associative memory for those problems and increase the probability of using direct memory retrieval before strategy use was measured. If retrieval was reported in the final addition block, however, then all the trials from the addition task in Session 1 were included in the RT analysis, although the strategies used on these trials were not measured. Given the relatively high rate of reconstructive strategies reported in Session 2, even after six recent practice trials with each problem (Fig. 3, Uittenhove et al., 2016, p. 295), there can be little doubt that the analysis of the frequent retrievers based on the Session 1 data was contaminated by reconstructive strategies.

Although the precise extent of this contamination can only be guessed, contamination of the frequent-retriever RT analysis by reconstructive strategies would explain the surprisingly long mean RTs reported. For purposes of comparison, we combined data from two recent publications that reported simple addition experiments using

similar methodologies. Campbell and Beech (2014) recruited 64 student volunteers from the University of Saskatchewan Psychology participant pool, and Chen and Campbell (2014) tested 36 Canadian and 36 Chinese adults recruited through the participant pool or by online advertisements. We excluded the Chinese sample from this analysis because of their known superior simple arithmetic skills (e.g., Campbell & Xue, 2001). The problem sets tested included single-digit plus single-digit addition problems including $0 + N$, $1 + N$, $N + N$ (ties), other small nonties with sums ≤ 10 , and large nonties. Each experiment consisted of two blocks of 48 trials, and each problem was tested in each block. Half of the 100 participants received nonties in min-left order (e.g., $1 + 2$), and the other half received nonties in max-left order ($2 + 1$). As is discussed later, the purpose of these experiments was to investigate transfer of practice, but here we use the results for the very-small addition problems to illustrate the relative slowness of the participants in Uittenhove et al. (2016). Considering only the 12 very-small problems, the average RT for frequent retrievers (derived from Fig. 6) in Uittenhove et al. was substantially slower (821 ms, $SD = 67.4$) than the Canadian participants for these items (729 ms, $SD = 42.7$), $t(11) = 7.61$, $p < .001$, $SE = 12.1$, even though we did not collect strategy reports and therefore had no means to excise relatively slow reconstructive strategies, if they occurred. The Uittenhove et al. results were based on six repetitions of each problem compared to only two repetitions per problem in our sample. We would normally expect faster, not slower, RTs with more repetitions. Similarly, the Uittenhove et al. participants were significantly slower to answer very small additions compared to the participants tested by Barrouillet and Thevenot (2013; 716 ms, $SD = 30.6$), $t(11) = 8.20$, $p < .001$, $SE = 12.8$, and again, the latter authors did not collect strategy reports and therefore had no means to exclude reconstructive strategies.³ The unusually long mean RTs reported for the frequent-retriever analysis would result from contamination by inadvertent inclusion of reconstructive strategies.

Such contamination, even if the extent of it was relatively small, would generate or inflate a problem-size effect on RT observed in the frequent retriever analysis because the rate of contamination would be roughly proportional to the rate of use of the relatively slow reconstructive strategies, which generally increases with problem size (Uittenhove et al., 2016, Fig. 3, p.

295). Indeed, for nontie problems, the correlation was .95 between the percentage of reconstructive strategies reported as a function of the sum (Uittenhove et al., 2016, Fig. 3) and mean RT for direct retrieval as a function of the sum (Uittenhove et al., 2016, Fig. 2). In other words, after ostensible exclusion of reconstructive strategies, mean RT as a function of the sum for the remaining trials (ostensibly direct retrieval) was still very well predicted by the percentage of reconstructive strategies reported. This is exactly as would be expected if the retrieval analysis was contaminated by inclusion of reconstructive strategies. Furthermore, use of reconstructive strategies requires storage and processing of intermediate results that tap working memory resources. Consequently, contamination of the frequent-retriever analysis by reconstructive strategies would also explain the observation that the slope of the problem-size effect for very-small problems was steeper for frequent retrievers with low working memory span compared to participants with a higher working memory span. The absence of a significant correlation between sum-related RT slope and working memory span for ties, $n + 1$, and medium-small problems may simply reflect a smaller effect size and consequent low power to detect this relation.

A retrieval-based problem-size effect for very-small addition problems

Regardless of contamination of the frequent-retriever analysis by inclusion of reconstructive strategies, the Uittenhove et al. (2016) claim that the sum-related increase in RT for very-small addition problems is difficult to reconcile with a retrieval model (pp. 297–298) is not correct. To the contrary, the network interference model and computer simulation proposed by Campbell (1995; see also Campbell & Oliphant, 1992; Whalen, 1997) predict precisely this finding. This is a widely cited theory of arithmetic fact retrieval (although it was not specifically referenced by Uittenhove et al., 2016, or Barrouillet & Thevenot, 2013) and its core assumption of similarity-based interproblem interference has substantial converging evidence (e.g., Campbell, 1987; Campbell & Thompson, 2012; Galfano, Rusconi, & Umiltà, 2003; Griffiths & Kalish, 2002; Phenix & Campbell, 2004; Whalen, 1997). According to this view, a problem-size effect on RT arises in both simple addition and multiplication fact retrieval because interference from competing arithmetic facts increases with problem size (see also Whalen, 1997). When a problem is presented, all problem nodes in an associative network of arithmetic facts receive similarity-based excitatory input and compete by way of mutual inhibition until one node reaches the activation threshold for retrieval. Similarity between the target fact and other arithmetic facts is based on feature overlap (e.g., common operands) and the magnitude similarity of stored answers. In the model, magnitude similarity is calculated using Welford's (1960) function $\text{LOG}[L/(L-S)]$, where S stands for

³ Neither Uittenhove et al. (2016) nor Barrouillet and Thevenot (2013) provided information about their participants' training or background in math or science, but we are doubtful in any case that differences in this regard between their samples could account for the relatively slow addition performance of the Uittenhove et al. participants. Chen and Campbell (2016) found that a sample of 36 engineering and computer science students did not demonstrate superior performance on a standardized test of basic arithmetic fluency.

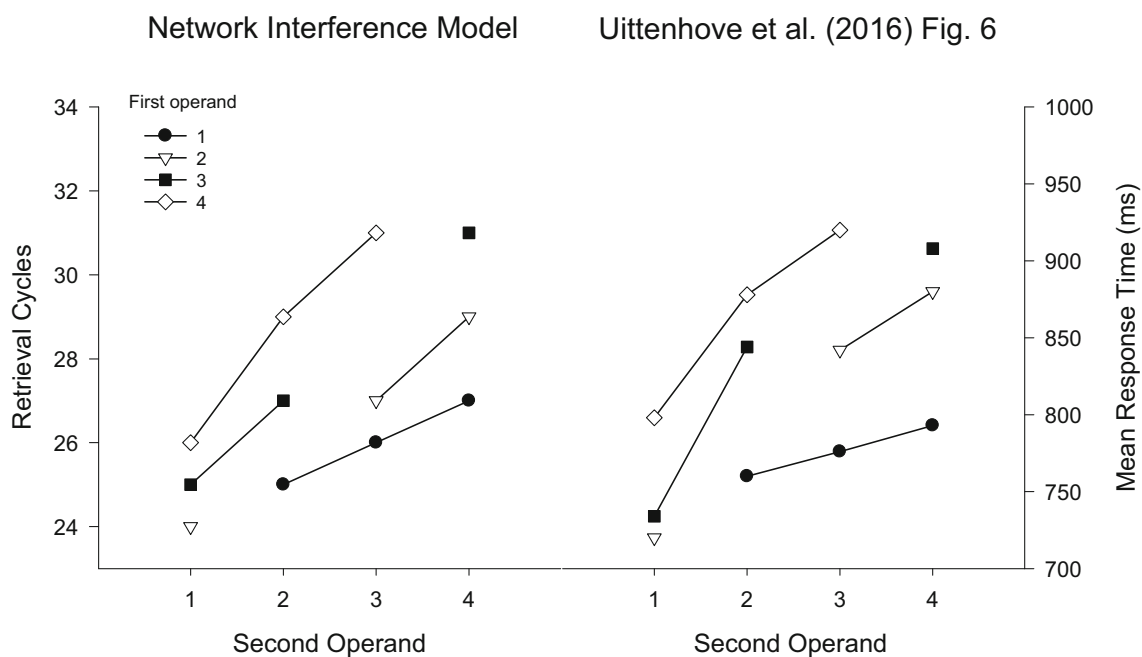


Fig. 3 Mean RT (ms) for very small problem based on Uittenhove et al. (2016) Fig. 6 ($n = 51$) and predicted RT (retrieval cycles) from the network interference model

the smaller and L for the larger of the magnitudes. This implements the long-held and widely held view that the scale of number similarity is compressed as number magnitude increases (e.g., Dehaene, 1989; Dehaene, Dupoux, & Mehler, 1990; Moyer & Landauer, 1967). One consequence of this magnitude-related compression in the model is that, as answer size increases (i.e., the sum for addition), the strength of inhibitory retrieval competition increases and slows the rate of activation of the target fact toward retrieval criterion. RT in the model corresponds to the number of excitation-inhibition retrieval cycles required to reach criterion. These basic assumptions of the model (Campbell, 1995) account for a large number of salient and subtle features of RT and error characteristics in both simple addition and multiplication. Furthermore, recent findings in both normal and impaired arithmetic development have identified susceptibility to memory interference as an important factor in individual differences in arithmetic learning (De Visscher & Noël, 2013, 2014a, 2014b; De Visscher, Noël, & De Smedt, 2016).

The original network interference model for addition was implemented for the so-called standard set of problems composed of the operands from 2 to 9 (i.e., $2 + 2$ to $9 + 9$). The $n + 1$ problems were not implemented at that time because we had no normative data for them, but it was a simple matter to modify the model to include these items and obtain RT predictions for the very-small addition problems identified by Uittenhove et al. (2016). The implementation tested was essentially the same as the original model (Campbell, 1995), and there was no kind of special treatment in connection with the very-small addition problems. Figure 3 presents the number of model retrieval cycles to criterion for the very-small problems

(the correct sum was produced in every case) and a reproduction of the corresponding Fig. 6 from Uittenhove et al. (2016). Obviously, the network interference retrieval model captures major features of the RT pattern across the very-small addition problems in the Uittenhove et al. data. The correlation between problem RT and the problem's sum in their experimental data was .945, and in the model it was .967. The correlation between the experimental and modelled data across the 12 very-small additions was .964 (adjusted $r^2 = .921$), with a regression slope of 28 ms per retrieval cycle in the model.⁴ The correlation between the model answer times and the mean RTs for the 12 very-small problems by the low span individuals in Barrouillet and Thevenot (2013) was .881. Thus, the network interference model of addition fact retrieval provides a very good fit to their experimental data for these items.⁵ Furthermore, because retrieval of simple addition problems

⁴ Despite the very high correlation between the data and model, the model did not produce the overall RT advantage for $n + 1$ problems relative to other very small additions that was observed in the data. The Campbell (1995) network interference model produces faster RTs for ties and sum-to-10 problems by treating these as distinctive subcategories of problems that activate within-category items more than extracategory items. This reduces retrieval interference and speeds RT and yields a shallower problem-size effect for these items. A similar mechanism could apply to the $n + 1$ problems.

⁵ Barrouillet and Thevenot (2013) used a retrieval interference index that they called overlap (the number of different problems sharing the same sum), but this index did not correlate significantly with RT across the 12 very-small nontie additions. This is not surprising from the standpoint of the network interference model. In the model, the strongest retrieval competitors are neighbouring problems that share a common operand (e.g., $3 + 4$, $3 + 5$, $4 + 5$). Addition problems with a common sum necessarily involve different operand pairs (e.g., $1 + 4 = 5$, $2 + 3 = 5$).

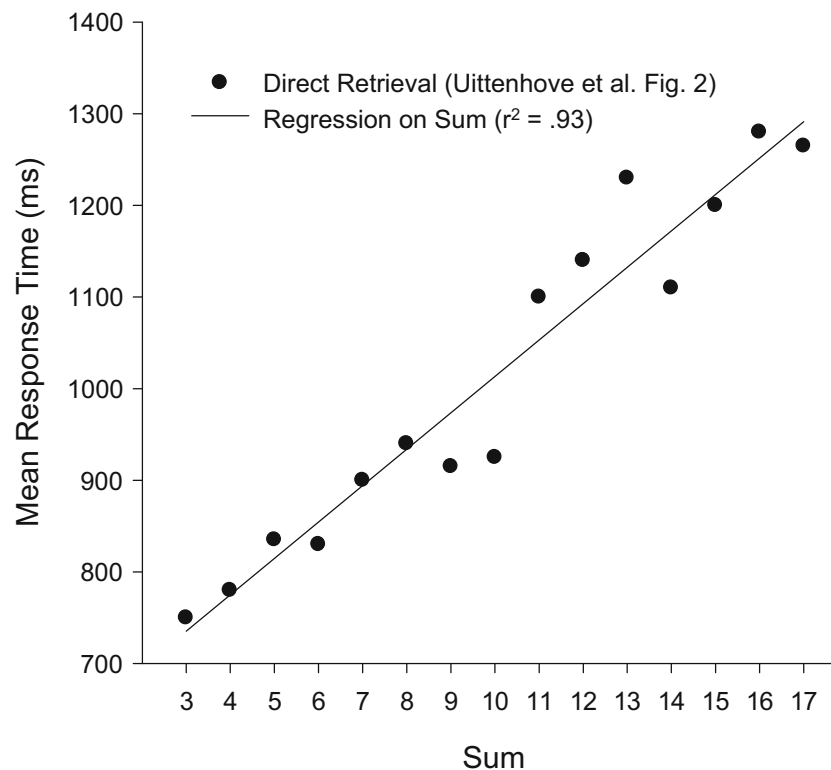


Fig. 4 Mean RT (ms) for retrieval trials as a function of the addition sum (after Uittenhove et al., 2016, Fig. 2)

loads on the central executive of working memory (Hubber, Gilmore, & Cragg, 2014; Imbo & Vandierendonck, 2007; see also Barrouillet, Benardin, & Camos, 2004; Barrouillet et al., 2008; Uittenhove et al., 2016, pp. 292, 299), the observed relation between performance of the very-small problems and working-memory capacity is consistent with a fact-retrieval process. Indeed, Uittenhove et al. (2016, p. 296) reported a correlation of $-.44$ ($p < .01$) between their working-memory-span measure and mean RT for small-tie-addition problems (i.e., higher span, faster tie RTs), which they assumed are solved by direct memory retrieval.

Proponents of the automatic counting theory might point out that the network interference model (or any retrieval model) fails to predict the null problem-size effect Uittenhove et al. (2016) observed for the medium-small problems and that it is precisely this null effect that rules out a retrieval model of their results. We have shown, however, that the dichotomy of very-small versus medium-small problems is not supported by an empirical RT boundary related to the sum in the Uittenhove et al. data (see our Fig. 2). For reported retrieval trials, Fig. 2 in Uittenhove et al. (2016, p. 294) showed a very strong linear relationship ($r^2 = .93$, slope = 40 ms) between mean RT and the sum for nontie problems across the full range of sums from 3 to 17 as predicted by the network interference model (see our Fig. 4). The greater variability in RT as problem size increases (e.g., the RT means and *SD* values in Uittenhove

et al. Table 1, were correlated $.97$) is also predicted by the network interference model (Campbell, 1995, p. 143).⁶ Thus, the model provides very good prediction of the Uittenhove et al. self-reported retrieval RT data. There are deviations from linearity in their Fig. 2 (we have already noted that problems summing to 9 and 10 were faster than predicted by the sum), but there could be idiosyncratic features of this population's learning history for addition that contributes to these (e.g., using sum-to-10 facts as the basis to construct solutions to other addition problems). Given this, we should be careful not to reify the Uittenhove et al. data or assume that their results necessarily generalize widely. For example, for the 100 (mostly Canadian) participants in the reanalysis of the generalization studies described previously, the sum predicted only 70% of the RT variability across the 12 very-small problems compared to greater than 90% in the Uittenhove et al. data. Also, our participants did not show a mean RT advantage for sum-to-nine problems (784 ms) compared to sum-to-eight problems (775 ms) as observed by Uittenhove et al., but sum-

⁶ Table 1 included all 90 participants. Uittenhove et al. (2016) did not provide a comparable table of means and standard deviations for the 51 frequent retrievers, but we would expect the correlation across problems between RT mean and standard deviation to be high here as well. A very strong correlation between mean RT and standard deviation across problems is also observed in adults' simple multiplication, $r(62) = .96$, in Campbell (1995), which is widely believed to primarily involve direct memory retrieval.

to-10 problems were faster (765 ms) as expected. Although Uittenhove et al. tested 90 participants on each problem six times, the standard deviations for many of the larger nontie problems were very large relative to the means (for example, for $8 + 6$ the mean was 1,698 ms with an *SD* of 732 ms in their Table 1). Given such variability, it would be difficult for any model to account precisely for the pattern of problem means across the full range of simple addition problems.

It is important to clarify, though, that we are not trying to reify the Campbell (1995) network interference model or suggest that interference is the only factor that contributes to problem-size effects in simple addition. Although the network interference model's basic assumptions remain plausible, a large amount of relevant empirical and theoretical work has occurred since its publication that would inform an updated version of the theory. Instead, our point was to demonstrate that a linear problem-size effect related to the sum in simple additions RT, even for the very-small problems, does not necessarily imply a counting process and is compatible with a retrieval-based account.

No generalization of addition practice

The validity of fast procedures for adults' simple addition is also challenged by several failures to observe generalization of practice for nonzero simple addition. Repeatedly practicing a procedure results in its speeding up (Singley & Anderson, 1989). As a result, speed-up with practice should transfer or generalize to different, unpracticed problems that use that procedure. To investigate the possibility of fast procedures in adults' simple addition, Campbell and Beech (2014) examined generalization of practice in students at the University of Saskatchewan, Canada. They argued that if simple addition problems were based on a fast procedure, then practicing a subset of problems (e.g., $4 + 3$) should produce speed-up in subsequent performance of similar, unpracticed problems (e.g., $3 + 2$). The results showed that there was no generalization of practice for nonzero simple addition problems, but the procedure-based $n + 0 = n$ problems presented clear evidence of generalization (i.e., practicing a subset of $n + 0$ problems facilitated a different subset of $n + 0$ problems). If automatic counting for simple addition existed, it might be most likely for the $n + 1$ problems because these require only a single counting increment; but generalization was not observed for $n + 1$ problems. Generalization for $n + 0$ problems, but no generalization for nonzero simple addition problems, has been repeatedly replicated (Campbell & Beech, 2014; Campbell, Dufour, & Chen, 2015; Campbell & Theriault, 2013; Chen & Campbell, 2014, 2016, 2017). Campbell, Chen, Allen, and Beech (2016) demonstrated robust generalization of practice in a counting-based alphabet addition task (e.g., $B + 5 = G$), thereby confirming generalization of counting-based procedures. Furthermore, one cannot argue that simple addition skills for small problems are so overlearned that they would not benefit from generalization or practice. Even $n + 1$ problems present

robust speed-up when tested a second time (Campbell & Beech, 2014; Campbell et al. 2015), indicating that there is ample potential for these problems to show effects of learning and transfer.⁷

The null generalization results of Campbell and Beech (2014) and subsequent studies listed in the previous paragraph cast doubt on the general applicability of the theory of fast procedures for simple addition. Uittenhove et al. (2016) did not cite Campbell and Beech or any of the other articles that demonstrated no generalization for nonzero simple addition. Thevenot et al. (2016), however, mentioned the finding of no generalization for $n + 1$ problems in a footnote, dismissing it by noting that “Campbell and Beech's conclusions have in turn been challenged by Baroody, Eiland, Purpura, and Reid (2014) who noted that generalization effects should have also been observed for addition problems involving 1, which are known to be solved by procedural rules” (p. 49). Baroody et al. (2014) did not actually refer to the Campbell and Beech experiment, but they did discuss young children's use of what they call the “number-after rule, which specifies that the sum of 1 and a number n is the number after n in the counting sequence” (p. 160). Baroody et al. (2014) did not provide evidence of rule-based performance of $n + 1$ problems, but Baroody, Purpura, Eiland, and Reid (2015) found that training of $n + 1$ problems transferred to novel $n + 1$ items in young children not initially fluent with the $n + 1 =$ successor of n relation. Whether educated adults rely on knowledge of the relation between counting and addition to answer $n + 1$ problems when n is a small number (≤ 9), which are the $n + 1$ items potentially relevant to the fast procedure theory, or alternatively have memorized this set of common addition facts, is precisely the issue in question and not to be taken for granted. The generalization results indicate they are solved, at least in the adult samples we have tested, by using an item-specific mechanism that does not generalize to unpracticed $n + 1$ problems with small values of n .⁸ Fact retrieval from

⁷ In the generalization studies, transfer of practice has never been observed for large (sum > 10) simple addition problems, although Canadian samples report some procedure use for these items; typically 25% to 50% of trials (see Campbell & Xue, 2001). Generalization would be difficult to measure for these items, however, because there is large RT variability for reconstructive strategies, and there are a variety of idiosyncratic procedures used for large additions (e.g., $6 + 8 = 6 + 6 + 2 = 12 + 2 = 14$; $6 + 9 = 6 + 10 - 1 = 16 - 1 = 15$; see Appendix A in LeFevre et al., 1996) rather than a common omnibus procedure.

⁸ There is further converging evidence the $n + 1$ problems are solved by an item-specific process. Practicing small multiplications (3×4) or additions ($3 + 4$) interferes with (i.e., slows performance of) the cross-operation counterparts ($3 + 4$ or 3×4), a form of retrieval-induced forgetting (RIF; Campbell, Dufour & Chen, 2015). Practicing rule-based problems (i.e., $n + 0 = n$, $n \times 0 = 0$, and $n \times 1 = n$), however, does not induce this RIF effect on their cross-operation counterparts, presumably because these problems are solved by retrieving a rule (rather than separately represented in memory as individual facts) and this does not interfere with individual cross-operation counterparts. In contrast, practicing $n + 1 = ?$ problems does produce item-specific interference with corresponding $n \times 1 = n$ items (Campbell et al., 2015). This is additional strong evidence that $n + 1 = ?$ ($n \leq 9$) problems are represented and solved as individual facts.

associative memory is a plausible candidate for this mechanism. Indeed, the substantial problem-size effect for $n + 1$ problems (see Fig. 2; see also Butterworth, Zorzi, Girelli, & Jonckheere, 2001), a result predicted by the network interference model, is difficult to reconcile with a rule-based account (Butterworth et al., 2001).

Other potential evidence for compacted procedures for simple addition

Operator priming effects

Fayol and Thevenot (2012) initiated the recent interest in the possibility of fast procedures for simple addition. They reported two experiments using an operator-priming paradigm (see also Roussel, Fayol, & Barrouillet, 2002; Sohn & Carlson, 1998). Experiment 1 tested Swiss engineering students in blocks of mixed simple addition, multiplication, and subtraction problems. When the operation sign (+, −, or ×) appeared 150 ms before the problem operands, mean response time (RT) for addition and subtraction problems was 40–60 ms faster relative to simultaneous presentation; but there was no operator preview effect for multiplication. Experiment 2 used the same operator priming paradigm but tested the 100 simple addition and multiplication combinations. Only participants who scored above a minimum criterion on a standardized test of arithmetic fluency were included for the reported analyses (the effect of interest was not observed in the complete sample; Fayol & Thevenot, 2012, p. 397). For addition, significant operator priming was observed for special problems that involved a zero or one (24 ms), small nontie problems with a sum ≤ 10 (31 ms), and large nonties with a sum > 10 (32 ms), but the 29-ms difference observed for the addition ties (3 + 3, 8 + 8, etc.), was not significant. For multiplication, no problem type showed significant operator priming. Fayol and Thevenot concluded that, in highly skilled performers, nontie addition problems were solved using a fast compacted procedure that could be primed by a preview of the plus sign, whereas multiplication involved direct retrieval of individual facts and was not subject to operator priming. The absence of significant operator priming for the addition ties (although the observed effect size of 29 ms for ties was equal to the average of 29 ms observed for the other three problem types) was taken as evidence that addition ties, like single-digit multiplications, were solved by direct memory retrieval.

To link the operator priming results to their automatic counting theory, Barrouillet and Thevenot (2013) noted that “such a compacted procedure could correspond to the reconstructive strategy primed by the anticipated presentation of the additive sign in Fayol and Thevenot’s (2012) study” (p. 43), although Barrouillet and Thevenot only examined the very-small nontie and tie problems with $n \leq 4$, whereas Fayol and

Thevenot also observed operator-preview facilitation for the larger nontie addition problems (sum > 10). Furthermore, Fayol and Thevenot emphasized the importance of high levels of arithmetic skill for operator priming, whereas Barrouillet and Thevenot (see also Uittenhove et al., 2016) argued that their effect of interest (i.e., the sum-related RT slope for very-small nontie additions) was stronger in the less skilled (i.e., slower) performers. Thus, the two phenomena do not converge neatly and there is little reason to believe they reflect a common underlying mechanism.

Additionally, Chen and Campbell (2015) found results using the operator priming paradigm that question the generality of the Fayol and Thevenot (2012) findings and conclusions. Chen and Campbell tested Chinese and Canadian adults ($n = 144$) in the operator priming paradigm used by Fayol and Thevenot (2012, Experiment 2), testing ties and small and large nontie addition and multiplication problems. In contrast to Fayol and Thevenot, Chen and Campbell observed robust operator priming for both addition and multiplication for both the Canadian and Chinese samples. They also observed a robust 26 ms ($SE = 9$) operator preview facilitation effect for addition ties ($p < .005$), which was equivalent to the effect observed for large nonties (24 ms, $SE = 11$) but smaller than that observed for the small non-ties (46 ms, $SE = 8$). The 24 ms operator priming observed for addition ties by Chen and Campbell was statistically robust whereas the nominally larger 29-ms effect for ties in Fayol and Thevenot was not significant, indicating relatively low power to detect an effect of this magnitude. As the addition ties are widely assumed to be solved by direct retrieval (e.g., Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012; Uittenhove et al., 2016), the finding of operator priming for addition ties by Chen and Campbell strongly suggests that operator priming is not necessarily a signature of procedure use in simple addition.

Furthermore, given that operator priming did not differ between addition and multiplication or between highly skilled Chinese and less skilled Canadians in the Chen and Campbell (2015) study, there was no evidence that operator priming discriminated between operations or depended on arithmetic skill level. Thus, these results, based on a large sample size, did not reinforce the major assumptions, findings, or conclusions of Fayol and Thevenot (2012). Specifically, if we assume, as did Fayol and Thevenot, that most single-digit additions are solved by procedures but that their multiplication counterparts are solved by fact retrieval, then the results of Chen and Campbell imply that operator priming does not reliably discriminate procedure use from fact retrieval.

This conclusion was reinforced by the findings of Chen and Campbell (2016), who applied the operator-priming paradigm to addition and multiplication identity-rule problems ($n + 0 = n$, $n \times 1 = n$, $n \times 0 = 0$) and $n + 1$ problems with n ranging from zero to 9. Chen and Campbell found that all three identity-rules demonstrated both operator-preview facilitation and

generalization of practice (e.g., practicing $0 + 3$ sped up unpracticed $0 + 8$), the latter being a signature of procedure use as explained previously. Chen and Campbell, however, also found operator-preview facilitation for $n + 1$ problems in the absence of generalization, which implies that the $n + 1$ problems were solved by an item-specific process (e.g., fact retrieval) but nonetheless were facilitated by an operator preview. Thus, the operator-priming paradigm is not reliably diagnostic of use of procedures versus use of direct memory retrieval as assumed by Fayol and Thevenot (2012).

RIF of small nontie additions for Canadian but not Chinese adults

Campbell, Chen, and Maslany (2013) provided another type of potential evidence that small addition problems may be solved by fast procedures, at least in highly skilled individuals. They examined Canadian and Chinese adults' performance in an arithmetic retrieval-induced forgetting (RIF) paradigm. Practicing small multiplication problems (e.g., 2×3) slowed RT to answer addition counterparts ($2 + 3$) for the Canadian group but not for the Chinese group. As this retrieval-induced forgetting effect in addition had previously been shown to be induced by number-fact retrieval practice but not by practice of arithmetic procedures (Campbell & Theriault, 2013), Campbell et al. proposed that the arithmetically superior Chinese participants might solve small addition problems by fast procedures whereas the Canadians used number-fact retrieval. Chen and Campbell (in press) replicated the finding of RIF for addition ties and not for small nonties in a study of 48 Chinese adults. Therefore, we believe this is a genuine effect, at least in the local Chinese population available to our research in Saskatchewan, Canada. Nonetheless, the absence of RIF for small nonties is not a definitive signature of procedure use. Indeed, Campbell et al. explicitly acknowledged that their RIF group differences did not provide direct evidence of automatic procedures for addition, given that RIF is subject to several boundary conditions (e.g., Storm & Levy, 2012). For example, RIF is competition dependent, occurring only when the target memory has strong competitors that require inhibition. Unlike North American children, Chinese children often learn their nontie multiplication facts in a preferred operand order (smaller \times larger), but this practice is not applied to the learning of addition facts (LeFevre & Liu, 1997). As a consequence of this, both problem encoding processes and the organization of the memory networks in Chinese adults might evolve quite differently for multiplication and addition. As a result, retrieving a nontie multiplication problem may not strongly activate the addition counterpart, which consequently does not attract inhibition and RIF because it is not a strong retrieval competitor. For the tie problems, however, order is irrelevant (e.g., 3×3 , 4×4), so the encoding process for multiplication and addition

would be the same resulting in strong activation of addition counterparts (e.g., $3 + 3$, $4 + 4$) and RIF for these items.

Rightward spatial shifts of attention with single-digit addition

Finally, Mathieu et al. (2016) tested adult participants on single-digit addition, subtraction, and multiplication problems with the first operand, operator (+, −, or \times), and second operand displayed visually in sequence. The second operand was displaced to the left or right of the central screen location where the first operand and operator appeared. Addition problems were solved faster when the second operand appeared to the right side compared to the left side, whereas subtraction problems were answered faster with the second operand displayed to the left compared to right side. There was no effect of second operand position for multiplication. Mathieu et al. concluded that simple addition and subtraction, respectively, entail rightward and leftward horizontal shifts of attention (see also Masson & Pesenti, 2014), but multiplication does not induce such a spatial bias.

To link these findings to the compacted counting theory, Mathieu et al. (2016) proposed that “it is plausible that some algorithmic procedures (e.g., step-by-step internal counting) explicitly used by children when learning arithmetic are progressively internalized into rapid left–right attentional movement of the [mental number line] MNL”; but these authors also acknowledge that their “data do not directly speak to the question of whether single-digit addition and subtraction problems are solved by means of procedural or retrieval strategies” (p. 237). Indeed, Masson and Pesenti (2014) concluded that it is “unclear whether attentional shifts are necessary, or even useful, in arithmetic processes” (p. 1524), although they may have a functional role when addition requires a carry operation (Masson, Pesenti, & Dormal, 2016). The link between compacted counting theory and the spatial bias phenomena is tenuous also because the current versions of the counting model is assumed to apply only to addition problems with both operands ≤ 4 (Barrouillet & Thevenot, 2013; Uittenhove et al., 2016), whereas the spatial bias effects reported by Mathieu et al. were observed for both small and large simple addition problems. This discontinuity questions the theoretical coherence of directly linking the two phenomena.

Efficient strategies in adults' elementary arithmetic

While we are doubtful about the automatic counting theory of small addition problems, there are other phenomena that do indicate an important role for relational knowledge in adults' arithmetic strategy repertoire. For example, there is evidence that adults can solve subtraction problems presented in an addition format ($8 = 2 + ?$) just as quickly or faster than when

presented in standard subtraction format ($8 - 2 = ?$). Similarly, simple divisions ($8 \div 2 = ?$) can be solved as quickly or faster when presented in a multiplication format ($8 = 2 \times ?$) (Campbell & Alberts, 2010; Mauro, LeFevre, & Morris, 2003). Adults can also very efficiently generate the factors associated with multiplication products (e.g., given 30, generate 5 and 6) despite relatively little factoring experience (Campbell & Robert, 2008; Rickard, 2005). These findings and others (e.g., Campbell, 1999, 2008; Campbell & Agnew, 2009) suggest that adults often solve simple subtraction and division problems by reference to their addition and multiplication counterparts. The efficiency of these operations implies flexible memory representations for addition and multiplication facts that afford fast procedures to extract different elements from the retrieval structure depending on the required operation. There is no evidence that we are aware of, however, that these are compacted procedures (i.e., automatic and unconscious); in fact, their use often appears in participants' self-reported strategies for subtraction and division (Campbell & Agnew, 2009; Campbell & Alberts, 2009).

Another case in which relational knowledge is exploited concerns the commutativity property of addition ($a + b = b + a$) and multiplication ($a \times b = b \times a$). The commutativity property is incorporated in adults' basic addition and multiplication fact retrieval, in as much as both orders are not represented redundantly in long-term memory. Operand order (e.g., 3×2 vs. 2×3) has little or no systematic effect on adults' RT to answer simple addition or multiplication problems (but see LeFevre & Liu, 1997), and positive transfer from practice (i.e., RT gains) between identical pairs (e.g., practice 2×3 and test 2×3) is practically equivalent to that between commuted pairs (e.g., practice 3×2 and test 2×3 ; Campbell, Fuchs-Lacelle, & Phenix, 2006; Rickard & Bourne, 1996). This is evidence that the two orders of commuted pairs are efficiently referred to a common long-term memory representation, probably because commuted problems are composed of identical elements (Rickard, 2005; Rickard & Bourne, 1996; but see Campbell, 1995; Campbell & Agnew, 2009); however, the specific perceptual or cognitive process by which North American adults rapidly map visually presented commuted pairs to a common memory representation remains unknown (Robert & Campbell, 2008).

Conclusions

When Fayol and Thevenot (2012) rekindled the debate about fast procedures for adults' simple addition, they believed this new view would have a "massive impact in the domain of numerical cognition" (p. 401), including possible revision of educational practices for elementary addition (see also Thevenot et al., 2016). The subsequent series of studies published over several years attempting to develop this potentially

important view, however, has not yielded a coherent or convincing set of findings or theory. In this review, with respect to the compacted counting theory, we have shown that there are clear violations of the predictions of the Uittenhove et al. (2016) sum-strategy counting model in their own data for the very-small addition problems. The network interference model of addition fact retrieval (Campbell, 1995) predicts the RT results observed for these problems at least as well as the counting theory. Our alternative graphic depiction of the means for small problems (see Fig. 2) shows that, in fact, there is not an unambiguous RT boundary for problems with both operands ≤ 4 ; so the proposed category (see also Barrouillet & Thevenot, 2013) of very-small addition problems is a statistical artefact of the particular analyses focussed on by Uittenhove et al. and has no other supporting evidence. Furthermore, the deferred strategy measurement procedure used by Uittenhove et al. to identify and excise reconstructive strategy trials from the data was seriously flawed and the inadvertent inclusion of reconstructive strategies would contribute to the RT problem-size effects and unusually long mean RTs reported for putative retrieval performance.

Finally, we reviewed the other types of evidence adduced for the compacted procedure theory of addition and concluded that these findings are unconvincing in their own right and, at best, only distantly consistent with the fast counting theory. Given also that there appears to be little other evidence to support the idea that mediational strategies become proceduralized with practice (see Bajic & Rickard, 2009; Kole & Healy, 2013; Rickard, Lau, & Pashler, 2008), we conclude that the cumulative evidence for fast compacted procedures for simple addition in the articles reviewed is not convincing and does not justify significant revision of the long-standing assumption in cognitive science that direct memory retrieval is ultimately the most efficient process of simple addition for nonzero problems, let alone sufficient to recommend significant changes to basic addition pedagogy. Operator priming, generalization effects, and attentional shifts induced while performing arithmetic are important but still poorly understood phenomena. Integrating them within the wider arithmetic literature remains an important theoretical objective.

Acknowledgments This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to Jamie Campbell.

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., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. In B. Ross (Ed.), *The psychology of*

- learning and motivation* (Vol. 51, pp. 121–151). Burlington: Academic Press.
- Bajic, D., & Rickard, T. C. (2009). The temporal dynamics of strategy execution in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 113–121.
- Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2014). Fostering first graders' fluency with basic subtraction and larger addition combinations via computer assisted instruction. *Cognition and Instruction*, *32*, 159–197.
- Baroody, A. J., Purpura, D. J., Eiland, M. D., & Reid, E. E. (2015). The impact of highly and minimally guided discovery instruction on promoting the learning of reasoning strategies for basic add-1 and doubles combinations. *Early Childhood Research Quarterly*, *30*, 93–105. doi:10.1016/j.ecresq.2014.09.003
- Barrouillet, P., Benardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory span. *Journal of Experimental Psychology: General*, *133*, 83–100.
- Barrouillet, P., & Fayol, M. (1998). From algorithmic computing to direct retrieval: Evidence from number and alphabetic arithmetic in children and adults. *Memory & Cognition*, *26*, 355–368.
- Barrouillet, P., Lépine, R., & Camos, V. (2008). Is the influence of working memory capacity on high level cognition mediated by complexity or resource-dependent elementary processes? *Psychological Bulletin & Review*, *15*, 528–534.
- 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.
- Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. R. (2001). Storage and retrieval of addition facts: The role of number comparison. *The Quarterly Journal of Experimental Psychology*, *54A*, 1005–1029.
- Campbell, J. I. D. (1987). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 109–123.
- 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. (1999). Division by multiplication. *Memory & Cognition*, *27*, 245–275.
- Campbell, J. I. D. (2008). Subtraction by addition. *Memory & Cognition*, *36*, 1094–1102. doi:10.3758/MC.36.6.1094
- Campbell, J. I. D., & Agnew, H. (2009). Retrieval savings with non-identical elements: The case of simple addition and subtraction. *Psychonomic Bulletin & Review*, *16*, 934–944. doi:10.3758/PBR.16.5.938
- Campbell, J. I. D., & Alberts, N. M. (2009). Operation-specific effects of numerical surface form on arithmetic strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 999–1011.
- Campbell, J. I. D., & Alberts, N. M. (2010). Inverse reference in adults' elementary arithmetic. *Canadian Journal of Experimental Psychology*, *64*, 77–85. doi:10.1037/a0015720
- Campbell, J. I. D., & Beech, L. C. (2014). No generalization of practice for non-zero simple addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1766–1771. doi:10.1037/xlm0000003
- Campbell, J. I. D., Chen, Y., Allen, K., & Beech, L. (2016). Transfer of training in alphabet arithmetic. *Memory & Cognition*. doi:10.3758/s13421-016-0631-x
- Campbell, J. I. D., Chen, Y., & Maslany, A. J. (2013). Retrieval-induced forgetting of arithmetic facts across cultures. *Journal of Cognitive Psychology*, *25*, 759–773.
- Campbell, J. I. D., Dufour, K. D., & Chen, Y. (2015). Retrieval-induced forgetting of multiplication facts and identity rule. *Memory & Cognition*, *43*, 672–680. doi:10.3758/s13421-014-0483-1
- Campbell, J. I. D., Fuchs-Lacelle, S., & Phenix, T. L. (2006). Identical elements model of arithmetic memory: Extension to addition and subtraction. *Memory & Cognition*, *34*, 633–647. doi:10.3758/BF03193585
- Campbell, J. I. D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference theory and simulation. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 331–364). Amsterdam: Elsevier.
- Campbell, J. I. D., & Robert, N. D. (2008). Bidirectional associations in multiplication memory: Conditions of negative and positive transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 546–555. doi:10.1037/0278-7393.34.3.546
- Campbell, J. I. D., & Theriault, N. H. (2013). Retrieval-induced forgetting of arithmetic facts but not rules. *Journal of Cognitive Psychology*, *25*, 717–724.
- Campbell, J. I. D., & Thompson, V. A. T. (2012). Retrieval-induced forgetting of arithmetic facts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 118–129. doi:10.1037/a0025056
- 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*, *68*, 152–157.
- Chen, Y., & Campbell, J. I. D. (2015). Operator and operand preview effects in simple addition and multiplication: A comparison of Canadian and Chinese adults. *Journal of Cognitive Psychology*, *27*, 326–334.
- Chen, Y., & Campbell, J. I. D. (2016). Operator priming and generalization of practice in adults' simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 627–635. doi:10.1037/xlm0000196
- Chen, Y., & Campbell, J. I. D. (2017). Transfer of training in simple addition. *The Quarterly Journal of Experimental Psychology*. doi:10.1080/17470218.2017.1318295
- Chen, Y., & Campbell, J. I. D. (in press). An evaluation of sex and cultural differences in arithmetic retrieval-induced forgetting. *Journal of Cognitive Psychology*.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.
- De Visscher, A., & Noël, M.-P. (2013). A case study of arithmetic facts dyscalculia caused by a hypersensitivity-to-interference in memory. *Cortex*, *49*, 50–70.
- De Visscher, A., & Noël, M.-P. (2014a). Arithmetic facts storage deficit: The hypersensitivity-to-interference in memory hypothesis. *Developmental Science*, *17*, 434–442.
- De Visscher, A., & Noël, M.-P. (2014b). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*, 2380–2400.
- De Visscher, A., Noël, M. P., & De Smedt, B. (2016). The role of physical digit representation and numerical magnitude representation in children's multiplication fact retrieval. *Journal of Experimental Child Psychology*, *152*, 41–53. doi:10.1016/j.jecp.2016.06.014
- Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception and Cognition*, *45*, 557–566.
- Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 626–641.

- Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, *123*, 392–403.
- Galfano, G., Rusconi, E., & Umiltà, C. (2003). Automatic activation of multiplication facts: Evidence from the nodes adjacent to the product. *Quarterly Journal of Experimental Psychology*, *56A*, 31–61.
- Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development*, *67*, 2022–2044.
- Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, *54*, 372–391.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*, 121–151.
- Griffiths, T. L., & Kalish, M. K. (2002). A multidimensional scaling approach to mental multiplication. *Memory & Cognition*, *30*, 97–106.
- Hubber, P., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *Quarterly Journal of Experimental Psychology*, *67*, 936–954. doi:10.1080/17470218.2013.838590
- Imbo, I., & Vandierendonck, A. (2007). The role of phonological and executive working memory resources in simple arithmetic strategies. *European Journal of Cognitive Psychology*, *19*, 910–933. doi:10.1080/09541440601051571
- Imbo, I., & Vandierendonck, A. (2008). Practice effects on strategy selection and strategy efficiency in simple mental arithmetic. *Psychological Research*, *72*, 528–541. doi:10.1007/s00426-007-0128-0
- Jones, R. M., & Van Lehn, K. (1994). Acquisition of children's addition strategies: A model of impasse-free, knowledge-level learning. *Machine Learning*, *16*, 11–36. doi:10.1023/A:1022626701243
- Kole, J. A., & Healy, A. F. (2013). Is retrieval mediated after repeated testing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 462–472.
- LeFevre, J., & Liu, J. (1997). The role of experience in numerical skill: Multiplication performance in adults from Canada and China. *Mathematical Cognition*, *3*, 31–62.
- 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*, 216–223.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22.
- Masson, N., & Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1514–1526.
- Masson, N., Pesenti, M., & Dormal, V. (2016). Impact of optokinetic stimulation on mental arithmetic. *Psychological Research*, *80*(4). Advance online publication. doi:10.1007/s00426-016-0784-z
- 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.
- Mauro, D. G., LeFevre, J., & Morris, J. (2003). Effects of problem format on division and multiplication performance: Division facts are mediated via multiplication-based representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 163–170.
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.
- Phenix, T. L., & Campbell, J. I. D. (2004). Effects of multiplication practice on product verification: Integrated structures model or retrieval induced forgetting? *Memory & Cognition*, *32*, 324–335. doi:10.3758/BF03196862
- Rickard, T. C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 250–257. doi:10.1037/0278-7393.31.2.250
- Rickard, T. C., & Bourne, L. E., Jr. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1281–1295.
- Rickard, T. C., Lau, J. S., & Pashler, H. (2008). Spacing and the transition from calculation to retrieval. *Psychonomic Bulletin & Review*, *15*, 656–661.
- Robert, N. D., & Campbell, J. I. D. (2008). Simple addition and multiplication: No comparison. *European Journal of Cognitive Psychology*, *20*, 123–138. doi:10.1080/09541440701275823
- Roussel, J. L., Fayol, M., & Barrouillet, P. (2002). Procedural vs. direct retrieval strategies in arithmetic: A comparison between additive and multiplicative problem solving. *European Journal of Cognitive Psychology*, *14*, 61–104. doi:10.1080/09541440042000115
- Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405–410.
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264.
- Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale: Erlbaum.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale: Erlbaum.
- Siegler, R. S., & Shrager, J. (1984). Strategy choice addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale: Erlbaum.
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge: Harvard University Press.
- Sohn, M.-H., & Carlson, R. A. (1998). Procedural frameworks for simple procedural skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1052–1067.
- Storm, B., & Levy, B. (2012). A progress report on the inhibition account of retrieval-induced forgetting. *Memory & Cognition*, *40*, 827–843. doi:10.3758/s13421-012-0211-7
- 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*, 289–303.
- 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.
- Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics*, *3*, 189–230.
- Whalen, J. (1997). The influence of semantic magnitude representations on arithmetic: Theory, data, and simulation. *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society, USA*, *19*, 814–819.