Cognition, 44 (1992) 75–106

Cognitive arithmetic: A review of data and theory

Mark H. Ashcraft*

Department of Psychology, Cleveland State University, Cleveland, OH 44115, USA

Abstract

Ashcraft, M.H., 1992. Cognitive arithmetic: A review of data and theory. Cognition, 44: 75–106.

The area of cognitive arithmetic is concerned with the mental representation of number and arithmetic, and the processes and procedures that access and use this knowledge. In this article, I provide a tutorial review of the area, first discussing the four basic empirical effects that characterize the evidence on cognitive arithmetic: the effects of problem size or difficulty, errors, relatedness, and strategies of processing. I then review three current models of simple arithmetic processing and the empirical reports that support or challenge their explanations. The third section of the review discusses the relationship between basic fact retrieval and a rule-based component or system, and considers current evidence and proposals on the overall architecture of the cognitive arithmetic system. The review concludes with a final set of speculations about the all-pervasive problem difficulty effect, still a central puzzle in the field.

Introduction

Twenty or so years ago, mainstream cognitive psychology viewed mental arithmetic not as an area of research interest, but as a tool to be used for other purposes. Putting it bluntly, what we knew was that counting backwards by 3s provided an excellent short-term memory distractor task. Even Newell and Simon's (1972) extensive work on cryptarithmetic problem solving, an apparent exception to this generalization, showed very little interest in arithmetic knowledge itself, compared with the problem-solving steps revealed by that knowledge.

Correspondence to: Mark H. Ashcraft, Department of Psychology, Cleveland State University, Cleveland, OH 44115, USA; email: r0599@csuohio.bitnet.

*I wish to thank Jamie Campbell, David Geary, Michael McCloskey, Robert Siegler, and several anonymous reviewers for their helpful comments on earlier drafts of the manuscript.

The publication of Groen and Parkman's (1972) *Psychological Review* paper, however, signalled an important change in this attitude, a genuine interest in number and arithmetic *per se*, and in the mental processes that deal with that knowledge. Within the past decade in particular, this interest has flourished. The area of research and theory known as *metal* or *cognitive arithmetic* is now an active field, with a substantial empirical base, a variety of theoretical perspectives, and a growing core of investigators. Recent extensions of the research to different populations of subjects, for example, learning disabled children, indicate that the field can now productively apply its basic knowledge to new situations.

Given these developments, it is appropriate that the area of cognitive arithmetic be reviewed, that its empirical effects be described, and that its theoretical claims and disagreements be portrayed to the general community of cognitive psychologists. This is the first and most basic purpose of the present paper. A second purpose is to explore several issues of concern to other areas of cognitive psychology that relate to the area of cognitive arithmetic in important ways, for example, automaticity of processing, modularity. Finally, a few of the effects that characterize number and mathematics processing address other areas in cognitive psychology, implying that the study of cognitive arithmetic can advance cognitive psychology as a whole. Stated simply, then, the purposes of the present paper are to review the existing area of cognitive arithmetic, and to explore areas and issues of reciprocal influence between cognitive arithmetic and the general field of cognitive psychology.

The paper is divided into three major sections. First, I present a discussion of four basic empirical effects that form the boundaries of models dealing with simple mental arithmetic. The first of these effects is also used to illustrate both the rise and fall of the ancestral model of mental arithmetic, Groen and Parkman's *min* model (1972). I then discuss three current models of simple arithmetic processing: my own *network retrieval model* (e.g., Ashcraft, 1982, 1987), Siegler's *distribution of associations model* (e.g., Siegler, 1988b; Siegler & Jenkins, 1989; Siegler & Shrager, 1984), and Campbell's *network interference model* (e.g., Campbell, 1987a; Campbell & Clark, 1989). Following this, I consider several papers that address the "rules of arithmetic" (e.g., Brown & Burton, 1978; Widaman, Geary, Cormier, & Little, 1989), a discussion that leads naturally to questions of the overall architecture of the cognitive arithmetic system (e.g., Campbell & Clark, 1988; McCloskey, Caramazza, & Basili, 1985) The final section concludes with a few notes on unresolved issues and theoretical advances in the area. Although the paper speculates briefly on an integrated theory of mental arithmetic performance, its primary purpose is a tutorial portrayal of research and theory in the area – a progress report, as it were, on the first twenty years of the field.

One final introductory note deserves mention here. Because of the close relationships between arithmetic performance on the one hand, and the intertwined effects of age and educational level on the other, research and models in this area tend to be more responsive to developmental concerns than is typical in other specialties within cognitive psychology: see, for example, the accompanying article on children's crucial experiences in learning to count (Gallistel & Gelman, this issue) and the impact this has on early arithmetic performance, reviewed below. Thus, an adequate model of mental arithmetic must not only explain the variety of effects found in normal adults but must also, at the minimum, be plausible in its account of age-related changes in performance.[1] Indeed, the criterion of developmental plausibility itself, and the additional leverage and insight provided by developmental data, have been critically important to the field of cognitive arithmetic. They would probably be beneficial to other specialty areas as well.

Basic effects

The area of cognitive or mental arithmetic asks a seemingly simple question: how do people do arithmetic in their heads? That is, how does an educated person, one who has studied and used arithmetic and mathematics since the earliest years of formal schooling, accomplish routine acts such as adding or multiplying single-digit numbers, to say nothing of more complex mathematical operations and processes, or naturally occurring math (e.g., Lave, 1988)?

Modern cognitive psychology is not the first discipline to ask such questions, of course. Important work on the relative difficulty of the number combinations, and on educational practice designed to overcome that difficulty, dates at least to the early years of this century (e.g., Brownell, 1928; Clapp, 1924; Thorndike, 1922; Washburne & Vogel, 1928; see Resnick & Ford, 1981, Chapter 2, for an excellent discussion). And yet much of this work is either theoretically or empirically tangential to contemporary concerns. Put simply, our current concerns involve the two basic questions of structure and process: how is a person's knowledge of number and mathematics organized in memory, and what are the processes by which this knowledge is accessed and applied in various settings? Four important empirical effects illustrate the depth of these questions, and have determined the overall shape or architecture of the three most current models of arithmetic performance.

The problem size/difficulty effect

Consider the simple, single-digit addition problems from $0 + 0$ up through $9 + 9$.

[1]McCloskey (this issue; see also McCloskey et al., 1985) argues that the models must also be consistent with the growing body of data on disruptions of arithmetic processing among brain-damaged individuals.

and their companion multiplication problems, the "basic facts" or "basic number combinations." The most widely reported result in the literature on mental arithmetic is the problem size or problem difficulty effect, which concerns just these basic facts. The effect is simply that reaction times (RTs) are longer for the larger facts; that is, that problems with larger addends or multipliers, and hence with larger answers, are in some fashion more difficult to solve than those with smaller numbers and answers. The effect is akin to such effects as word frequency and semantic relatedness in studies of lexical and semantic representation, both because it is an overwhelmingly common result and because it motivates the basic structural and processing assumptions about the memory system.

Groen and Parkman's (1972) paper reported the first thorough examination of the problem size effect, and the first serious attempt to explain it in terms of postulated mental structures and processes. In one of their experiments, first graders were timed as they gave answers to addition problems with sums less than or equal to 9, in a production task (state the answer to $3+5$). In two other studies, college adults responded to the entire set of basic addition (Parkman & Groen, 1971) and multiplication (Parkman, 1972) facts in a verification task (make a true/false decision to $3+2=5$). In all three studies, RT increased directly as a function of problem size.

Analyses of the children's latencies evaluated the fit of five different models, all of which attributed changes in RT across problems to some internal counting mechanism. The best-fitting model in this set of candidates was termed with *min* model, because the minimum, that is, smaller, addend in the problem yielded the highest R^2 value. The mechanism underlying this variable was a counting or incrementing process, according to Groen and Parkman, in which the numerical value held in memory, the 4 in $4+3$, was incremented by 1s repeatedly until the number of increments equalled the *min*, 3 in this example (this pattern is also widely known as "counting-on"). Thus, first graders were said to perform simple addition by means of a simple incrementing process. The 400-ms slope of the regression line was interpreted as an estimate of their rate of mental counting, 400 ms per increment.

Some of the limits of counting models became apparent when Groen and Parkman examined adult performance, however. Although adults' addition RTs were well predicted by either the min or the sum, the 20-ms slope of the regression function across min was judged an implausibly rapid rate for the hypothesized incrementing process. Instead, Groen and Parkman suggested that adult performance to addition problems might be accomplished by some direct access retrieval process, with counting use as a backup in the event of retrieval failure. If such failure occurred on 5% of all trials, they reasoned, the obtained slope could then be interpreted as an artifact of averaging across the two trial types: direct access trials which required only a constant amount of time, and trials on which backup counting had taken place at the rate of 400 ms/increment.

Striking similarities between adults' addition and multiplication performance, however, were reported by Parkman (1972); that is multiplication RTs were well predicted by the *min* or the *sum* of the two multipliers. The difficulty with this finding was that a straightforward extension of the *min* counting model to multiplication had to claim that subjects multiplied by counting-on, using a unit size specified by the larger multiplier; for example, for 5×4, counting-on by 5s for 4 increments. Aside from the cumbersome nature of such a strategy, Miller, Perlmutter, and Keating (1984) noted the internal inconsistency of the extension: if multiplication problems are solved via counting-on by 5s (or any other unit size), how could the *min* model restrict addition to incrementing by 1s? Parkman's discussion both acknowledged the limitations of the counting approach, and suggested that a retrieval process of unknown character might be responsible for the obtained problem size effect.

Two important points require emphasis here: first, the robustness of the problem size effect, and second, the implications of this effect for the questions of mental representation and processing. Concerning the robustness of the effect, suffice it to say that every published report that has examined RTs on a problem-by-problem basis has reported the problem size effect. The effect holds for simple addition (e.g., Ashcraft & Battaglia, 1978; Parkman & Groen, 1971), subtraction (e.g., Siegler, 1987b; Woods, Resnick, & Groen, 1975), multiplication (e.g., Campbell & Graham, 1985; Parkman, 1972), and division (Campbell, 1985). It is apparent using both RT and error rates as the dependent variable (e.g., Miller et al., 1984; Siegler, 1988b). It is obtained both in the production and the verification tasks (Geary, Widaman, & Little, 1986; Miller et al., 1984). And, it characterizes performance across the entire span of ages, from kindergarten through college adults (e.g., Hamann & Ashcraft, 1985; Koshmider & Ashcraft, 1991; Siegler, 1987a) and the elderly (Geary & Wiley, 1991).

Concerning the second point, note that the nature of the problem size effect, its shape and temporal characteristics, is taken as evidence about the nature of the underlying processes of mental arithmetic. The systematic increase in RT as a function of problem size was initially, and parsimoniously, proposed to reflect mental counting in children's mental addition (Groen & Parkman, 1972). This counting-on model was judged empirically inconsistent with adults' rapid performance to the addition facts, however, and was thus rejected in favor of a direct access model. A variety of empirical reports subsequently provided evidence against this model as well. For instance, Ashcraft and Battaglia (1978) found an exponential increasing problem size effect in adults. The non-linearity of the effect argued against both the *min* and direct access models, as did Ashcraft and Stazyk's (1981) results, in which RTs matched the direct access model's predictions only at the implausible retrieval failure rate of 50%. Instead, both papers argued for a retrieval model of performance, one consistent with predictions drawn from network representation models of semantic memory. Stazyk, Ashcraft, and Hamann (1982), as well as Miller et al. (1984), made comparable arguments for multiplication performance. In short, the robust problem size effect was judged inconsistent with purely counting-based explanations, but supportive of a memory-based model of arithmetic knowledge.

By virtually all current accounts, as described below, this model involves retrieval from an organized long-term memory representation of fact knowledge: the characteristics of that organization remain under debate, of course (but cf. Gallistel & Gelman, this issue). Moreover, because of the emphasis on retrieval from memory, the term *problem size* is now generally considered a misnomer. That is, it misleadingly attributes the increasing RT patterns to a structural characteristic of the problems, such as the size or magnitude of the *min*, sum, product, etc. Several authors have argued, however, that these variables are only coincidentally found to be significant predictors of performance, because of *their* relationship to a more central variable, *problem difficulty* (e.g., Campbell & Graham, 1985; Stazyk et al., 1982). Measures of problem difficulty have been taken from normative studies (Stazyk et al., 1982), from speeded assessments of performance (Siegler & Shrager, 1984), and from analyses of children's classroom performance and textbooks (e.g., Clapp, 1924; Hamann & Ashcraft, 1986; Wheeler, 1939). In general, they provide a better fit to RT and error data than do structural variables (e.g., Koshmider & Ashcraft, 1991; Siegler, 1987a, 1988b).

The clear advantages of the term *problem difficulty* are both theoretical and empirical. First, suggesting that problem size is the operative variable implies, wrongly, that something inherent in larger numbers makes them more difficult to process. Strictly speaking, however, this should only be true if an individual is counting. More importantly, researchers have always acknowledged a major exception to the relationship with problem size, that *tie* problems like $2+2$ and $7+7$ are routinely found to be more easily processed than the size of their addends or sums would predict. Even Groen and Parkman (1972) suggested that first graders must be retrieving the answers to tie problems from memory, given the relatively flat RT function found for these problems. When problem difficulty is the underlying variable, however, then tie problems are no longer exceptions to the processing explanation offered for other problems. Instead, their rapid processing is a function of their strength in memory – in other words, their low level of difficulty.

Finally, conceptualizing the effect in terms of problem difficulty argues for close connections between mental representations of number, as postulated in the arithmetic models, and representations of lexical and semantic information, in which words' and concepts' accessibility is conceptualized in terms of variables such as word frequency, strength of storage, typicality, meaning dominance, and the like (e.g., Collins & Loftus, 1975; Kintsch, 1974; Simpson, 1984). The processing advantage for frequent words, for instance, stems from their stronger or more accessible representations in memory (e.g., Allen, McNeal, & Kvak, in press). We advance a parallel argument in arithmetic, that problems vary in difficulty due to differential experience, beginning at least with early formal education if not sooner (but cf. Campbell & Oliphant, 1992). This line of reasoning is amplified below, exploring ways of escaping the apparent tautology that more difficult problems are processed more slowly and inaccurately.

Error effects

An early canon of the information-processing approach involved errors in performance. As emphasized in Sternberg's (1969) classic work on short-term memory scanning and Pachella's (1974) treatise on speed–accuracy trade-off functions, meaningful interpretation of RT effects is only possible if error rates are uniformly low and constant across treatment conditions. A consequence of this canon was, in retrospect, rather superficial consideration of errors in performance. If speed–accuracy trade-offs could be ruled out as an explanation for RT effects, then little more in the way of error analyses was customarily provided.

While several investigators departed from this orientation (e.g., Resnick & Ford, 1981; Siegler & Robinson, 1982), work by Campbell and Graham (1985), especially, brought the importance of error patterns to the attention of researchers on adults' arithmetic processes. Campbell and Graham (see also Campbell, 1987b; Miller et al., 1984) tested subjects' multiplication performance in the production task, and tabulated the kinds of errors that were committed. For adults, more than 90% of the errors were "table errors," that is, answers to another basic number combination in multiplication, rather than merely values falling within the correct range (0–81). Even more revealing, multiples of the problems' operands (e.g., 32 for the problem 4 × 6), which are variously called table-related or operand-related answers, far outnumbered errors that merely stated a legal but unrelated answer (a table error, e.g., 35 for the same problem), by 79.1% versus 13.5%. Essentially the same patterns, but at lower levels, were found even for second and third-grade students, relative newcomers to formal multiplication (see also Cooney, Swanson, & Ladd, 1988).

As described below in connection with Campbell's model, such error patterns have strong implications for the nature of the memory representation of arithmetic facts. For now, the important point is that errors in arithmetic performance have a more important status than the above comments on the problem difficulty effect would imply. In particular, their regularity and lawfulness provide additional evidence about the representation of arithmetic knowledge and the overall architecture of the arithmetic processing system.

Relatedness effects

Ashcraft and Battaglia (1978) reported that several of their false addition stimuli were, by accident, correct problems under subtraction (e.g., $6 + 1 = 5$). They speculated that the abnormally slow and variable performance to these problems might be due to a relatedness effect among the arithmetic problems and operations. In fact, a similar effect had been reported several years earlier. In Winkelman and Schmidt (1974), addition problems with answers correct under multiplication ($5 + 4 = 20$) were processed more slowly and displayed higher error rates than problems lacking this relationship (see also Zbrodoff & Logan, 1986). Further investigation (e.g., Stazyk et al., 1982) confirmed that the effect is not limited to so-called "cross-operation confusions," but is also apparent within a single operation. In their study, Stazyk et al. found that confusion problems in multiplication (e.g., $7 \times 4 = 21$ or 35) were significantly slowed relative to non-confusion problems ($7 \times 4 = 18$), by up to 200 ms or more for large problems.

In important ways, these confusion effects tap the same phenomenon as the error effects described above – they display a general relatedness effect in performance. That is, Campbell and Graham's (1985) table-related errors, identified in the production task, are precisely those values found by Stazyk et al. to induce the confusion effect in verification – multiplies of the problems' operands. In fact, the strong tendency is for table-related answers, and for effective "confusion" answers, to be adjacent or "near" neighbors of the correct value, that is, answers to a problem differing by ± 1 in one of the operands. Likewise, cross-operation confusions occur with some frequency in production tasks (Miller et al., 1984), and indicate a degree of relatedness between the memory representations for (at least) addition and multiplication. Campbell in particular has demonstrated the importance of these relatedness effects (e.g., Campbell, 1987a; Campbell & Clark, 1989; Campbell & Graham, 1985), and, as described below, has made this a central feature of his model of arithmetic processing.

The similarity of these effects is important in another way as well. Stazyk et al. suggested that their obtained confusion effect stemmed from relatedness in the memory representation of multiplication facts. From the standpoint of processing, however, they suggested that relatedness disrupted the yes/no decision stage mechanism. Such a stage is a necessary component of verification performance, of course, but is presumably not necessary in the production task (but cf. Reder's (1982) discussion of simultaneous fact retrieval and plausibility judgments). The issue raises the question of whether or not the two tasks are directly comparable. Early indications that the tasks are similar in many respects (e.g., Ashcraft, Fierman, & Bartolotta, 1984) have given way to explicit evidence of task differences (Campbell, 1987b; Zbrodoff & Logan, 1990). In essence, these later studies show that the stated answer in a verification problem alters the course, and possibly the outcome, of retrieval.

But the relatedness issue is broader than that. It also implies, for example, that evidence on decision stage processes should be reconsidered in light of recent arguments. In particular, our early evidence concerning decision mechanisms involved the *split* effect (Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981), that RT to false trials declines as the incorrect answer differs more and more from the correct answer. But if the error and relatedness effects are both generated by a retrieval stage phenomenon, then it is not clear that split effects in addition, using the verification task, constitute separate evidence for a decision stage mechanism. Instead, the split effect may be the same as the near neighbor effect found in multiplication errors (note, though, that addition errors cannot be categorized as table-related, table-unrelated, and so forth, so addition data cannot disentangle these two effects). This question cannot be resolved with currently available data.

Strategies of processing

A clear implication in Groen and Parkman's (1972) original results is the notion that mental arithmetic operations change across ages, from operations that rely heavily on counting during the first years of elementary school, to retrieval operations later on (see for instance Ashcraft & Fierman, 1982; Cooney et al., 1988). There can be no doubt of the centrality of counting to the child's earliest experiences with number, as Gelman and Gallistel (1978) documented. And, as Siegler (1987b) has noted, counting provides the child a basis for understanding and computing simple arithmetic facts early in schooling, despite teachers' widespread discouragement of counting. Thus, the evidence indicates that both reproductive and reconstructive processes, both retrieval and counting, coexist at even the earliest stages of formal education.

Several studies have shown the need for two such components in models of arithmetic processing: one devoted to basic fact retrieval, the other to procedures like counting or other rule-based performance. For instance, Ashcraft and Stazyk (1981, Experiment 2) found that performance to more complex addition problems like $14 + 9$ involved both the basic fact retrieval processes for $4 + 9$ as well as a discernible carry operation (see the description below of work by Widaman et al., 1989, who present a model of such processing). Further, the unusual RT and error characteristics found to problems like $N \times 0 = 0$ (Parkman, 1972; Sokol, McCloskey, Cohen, & Aliminosa, 1991; Stazyk et al., 1982) suggested that such problems were being processed by a rule-based component or strategy, rather than normal retrieval.

Siegler's work (1987a, 1987b; 1988b), however, provides the most systematic examination of this strategic component of arithmetic processing. In his studies, children's strategies are assessed as they complete the arithmetic tests, and the chronometric relationships between strategies and performance are then explored. In one study (1987a), 99% of the sample of children in grades K–2 reported using at least two strategies, and 62% of the sample reported using three or more. When performance to individual problems was scored, 44% of the first graders' trials exhibited basic retrieval from memory, 38% of their trials showed *min* counting, 9% showed a decomposition strategy, and only 1% showed the inefficient count-all strategy (counting by 1s up to the first addend, then continuing with the second addend). Just as interesting, each strategy had a different time course and accuracy profile. Retrieval trials averaged 2.1 s for first graders, with a 4% error rate, compared with 6.9 s, and 17% errors, for *min* counting, 4.1 s and 8% errors for decomposition, and 16.3 s with 50% errors for count-all.

Summary

The most basic empirical effect in studies of mental arithmetic is the problem size/difficulty effect: as problems grow larger, they become more difficult to process, as shown by RTs and error rates. The pattern of these increases, furthermore, rules out counting as the single mental process by which adults perform the basic facts of addition and multiplication.[2] Instead, such performance is attributed to retrieval processes operating on an organized, long-term memory network of fact knowledge. An important degree of relatedness among the basic facts of arithmetic has been documented, both in the form of interference and errors in production tasks. These effects are not only important empirical demonstrations, but also constrain the nature of possible models. Furthermore, there are several consequences of processing a false answer, as typically tested in the verification task. Split effects show that an answer wrong by a small amount is more difficult to reject than an answer wrong by a larger amount. It is not yet clear whether this effect is due solely to a decision stage mechanism, interference during a retrieval stage, or a combination of the two. Finally, strategies other than retrieval (if retrieval can indeed be considered a strategy; see Ashcraft, 1990; Bisanz & LeFevre, 1990) are amply demonstrated among young school children, as are the contributions of other procedural or rule-based performance in adults (e.g., Widaman et al., 1989).

[2]In fact, in Siegler's view (1987a, 1988b), it is erroneous to search for any single process that characterizes all of a sample's or even subject's performance, that averaging across subjects will invariably and misleadingly combine across multiple strategies with different time characteristics. Note, however, that beyond approximately the 3rd or 4th grade level, the bulk of processing for the basic facts of addition and multiplication is generally retrieval (e.g., Ashcraft & Fierman, 1982; Cooney et al., 1988). Indeed, in Siegler's model, there is exactly this trend toward greater and greater reliance on retrieval.

Three current models

I turn now to a discussion of three current models of simple arithmetic processing. At a fairly macroscopic level. these models share several basic assumptions: performance on simple arithmetic facts depends on retrieval from long-term memory; the memory representation is organized and structured in terms of the strength of individual connections. and reflects varying degrees of relatedness among the elements; and the strength with which the elements are stored. hence the probability or speed of retrieving information. depends critically on experience, especially acquisition, rather than on numerical characteristics inherent in the information itself.

The models vary considerably in how they deal with the several empirical effects reviewed above, however, and they naturally differ in emphasis and focus; for example, Siegler's model is the most developmentally thorough of the three, but the least concerned with processing effects observed in adults. The models also differ in their complexity. Indeed, it is not inaccurate to view them, in their chronological order, as successive revisions and elaborations. with each model accommodating newer empirical evidence as it became available. What follows is an exposition of these models, along with a description of their most prominent contributions and weaknesses. Table 2, at the end of this discussion, provides a thumbnail sketch and summary of each model.

Ashcraft's network retrieval model

As discussed above, my early research (Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981; Stazyk et al., 1982) revealed strong evidence against counting-based models of adults' performance. Instead, the results indicated that basic addition and multiplication facts were represented in memory in an organized network of information, accessed and retrieved from the network via a process of spreading activation.

As proposed in Ashcraft (1982, 1987), the two most important structural aspects of the network involved the concepts of strength and relatedness among nodes. In the network, each problem-to-answer association was represented in terms of strength or degree of accessibility. Furthermore, the network also coded the degree of relatedness among problems and answers, in that adjacent, "near neighbor" nodes were more strongly interlinked than more distant, non-adjacent nodes. Retrieval from this network involved a spread of activation triggered by three sources: the problem addends, the answer stated in the problem, and those nodes in the network which had themselves been activated during retrieval. Activation from these sources spread throughout the network in parallel, with nodes accumulating various levels of activation depending on their strength and relatedness. The most highly activated node was the one selected as the answer to the problem at the end of retrieval, and elapsed retrieval time was a direct function of the accrued activation at that node.

Thus, as an example, the answer 24 to the problem 8×3 was the node in the network standing at the intersection of the "parent nodes" 8 and 3. This answer node was stored with a particular strength value, and with multiple pathways of differing strength to its neighbors, essentially the answers to the other "8 times" and "3 times" facts. Retrieval time, described metaphorically as the "distance" traversed in the network until an intersection was found, was predicted to be a function of the strength variable, because strength was the governing value in the accumulation of activation.

In the 1982 description of the model, there was no explicit discussion devoted to the source of the varying strengths among nodes in the network. The paper described the analogy between problem difficulty and the general semantic distance effect, and suggested that either subjective ratings of difficulty or actual problem size were indices of the underlying structure of the network. Thus, the original speculation about a table-based memory representation (Ashcraft & Battaglia, 1978) was abandoned, but not yet replaced with a specific proposal on the source of problem strength or difficulty.

Within three years, however, an explicit rationale was advanced (Ashcraft, 1985; Hamann & Ashcraft, 1986), that the strength with which nodes were stored and interconnected was a function of frequency of occurrence and practice, especially in early education. Computationally, for its basic memory strength values, the simulation (Ashcraft, 1987) used the Siegler and Shrager (1984) estimates of associative strength between problems and correct answers, converted to percentages. This pattern of differential strengths was corroborated by data taken from elementary school textbooks, which showed problem size and difficulty to be a direct function of both order and frequency of presentation (Hamann & Ashcraft, 1986). In this tabulation, smaller problems appeared earlier in instruction, and with far greater frequency, than larger problems. The only apparent exceptions here were addition problems with addends of 0 and 1, which occurred as infrequently as the large facts. Accordingly, the strength values in the model were computed as a function of two factors: initial strength and relative frequency of occurrence.

The most recent version of the model (1987) therefore asserted that the nature of instruction on arithmetic during early schooling, among other influences, has a direct effect on the strength of problem representation in long-term memory. This effect remains in evidence across the developmental span according to the model, given the assumption that the differences in problem frequency, apparent in early grade school, are maintained across later experience.

In short, the model asserted a straightforward frequency-to-strength hypothesis to account for problem difficulty effects during retrieval. Note that if some other

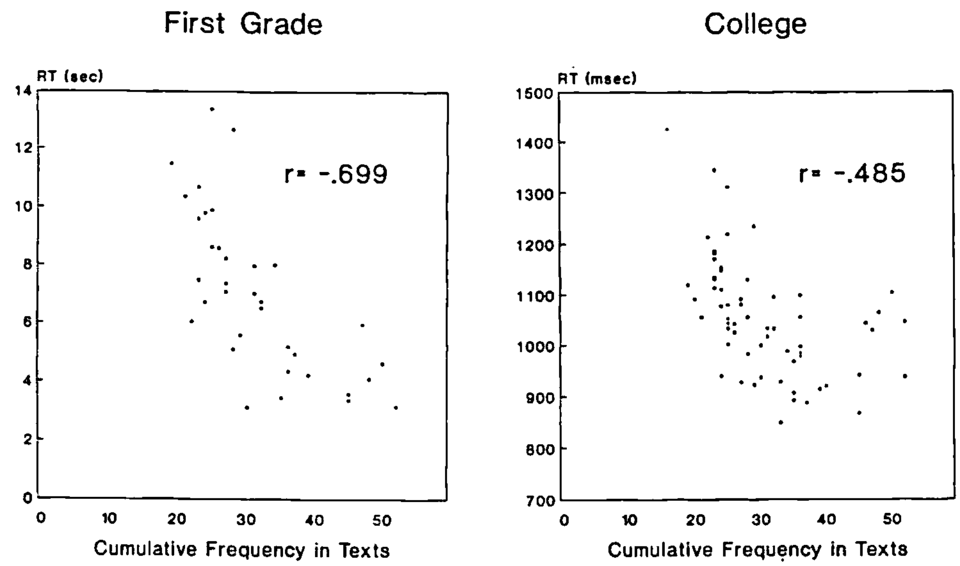

Figure 1. *Scatterplots of first-grade (left panel) and college (right panel) reaction times as a function of the frequency with which problems appeared in elementary texts, grades K–3. Data taken from Hamann and Ashcraft (1986).*

factor were responsible for the RT patterns, then the correlations between frequency of occurrence and RTs should not only be low, but in particular they should be lower than the intercorrelations of RTs across grade levels. Instead, the frequency–RT correlations are quite strong (see Figure 1), and are uniformly higher across grades than the RT intercorrelations, as shown in Table 1 (all data

Table 1. *Correlation and regression values between reaction time (RT) and problem frequency*[a]

Grade	Slope	Intercept	F_{reg}	Correlation of RT to frequency	Correlation of RT to College RT
1[b]	−228.4	13986	32.63	−.699	.441
4	−42.0	3960	14.52	−.547	.405
7	−22.3	2190	21.64	−.624	.505
10	−21.3	1952	23.74	−.641	.555
College	−6.4	1249	19.10	−.485	–

[a]All data from Hamann and Ashcraft (1986). Slopes and intercepts in milliseconds. The college F is based on 1, 62 df, all others on 1, 34. All Fs and rs are significant beyond 0.05.

[b]All analyses are based on the cumulative frequency of problems in grades K–3. The corresponding regression values for first graders' RTs correlated with problem frequency in the first-grade texts only are:

−283.2	10965	39.95	−.735	.441.

from Hamann & Ashcraft, 1986).[3] Despite evidence favorable to the model such as this, however, the assumption of differential problem frequency across the entire developmental span is apparently controversial (e.g., McCloskey, Harley, & Sokol, 1991). Further discussion of this point is provided in the conclusion of the present paper.

The model also made predictions about priming effects and automaticity of retrieval, by means of its rather straightforward assumptions about the spread and decay of activation. In particular, once a set of nodes is activated by a problem, the level of activation decays across some relatively short period of time. During this period, however, RT to a subsequent problem that accesses any of these nodes should be altered. This is an identical prediction to, for example, priming effects in semantic memory (e.g., Ashcraft, 1976; Loftus & Loftus, 1974). By the same priming mechanism, advance information about an upcoming problem was predicted to alter RT performance to the target problem. Several studies have now confirmed these general predictions (e.g., Campbell, 1987b, 1991; LeFevre, Bisanz, & Mrkonjic, 1988; Stazyk et al., 1982, Experiment 3). Koshmider and Ashcraft (1991), further, have shown how automatic and conscious priming effects in retrieval vary by age and problem difficulty.

Likewise, the model accounted for split and confusion effects on the basis of activation. Consider the false problem $8 \times 3 = 32$. If x and y are, respectively, the answer stated in the problem (32) and the answer retrieved from memory (24), then the decision stage mechanism has to discriminate between these values, based on their respective levels of activation. Neighbor nodes that received especially high levels of activation during search, for example 32, should disrupt decision processing to a greater degree, given that the discrimination is more difficult as the levels of activation of x and y converge. Again, a variety of studies have presented evidence of the split and confusion effects (e.g., Ashcraft & Stazyk, 1981; Stazyk et al., 1982; Winkelman & Schmidt, 1974; Zbrodoff & Logan, 1986), although a purely retrieval-based explanation has also been proposed (Campbell, 1987a, 1987b).

Finally, and especially relevant for younger subjects, the entire retrieval and decision process was said to occur in parallel with a procedural-based solution attempt, in which counting or other reconstructive processes began simultaneously with retrieval. Only among younger subjects, however, would retrieval processing be expected to be slow or to fail with any regularity, conditions under which the "horse race" between declarative (retrieval) and procedural routes might indeed be won by procedural processing (as noted above, procedural-based performance was also expected for problems with addends equal to 0 or 1).

[3]I wish to thank an anonymous reviewer for suggesting this line of reasoning and evidence. Note that these correlations and scatterplots exclude problems with addends equal to 0 and 1, given that these problems are generally conceded to be performed via rules rather than network retrieval.

Siegler's distribution of associations model

Two shortcomings of the network retrieval model, as well as a theoretically motivated dissatisfaction with prevailing methods of data analysis, were prominent among Siegler's reasons for proposing his distribution of associations model (Siegler, 1988b; Siegler & Shrager, 1984). The first shortcoming involved the relatively underdeveloped procedural knowledge component of the network retrieval model. Because retrieval was its primary focus, the model described in Ashcraft (1987) made explicit reference only to *min* counting as the method by which problems could be reconstructed (but cf. Hamann & Ashcraft, 1985). Siegler's empirical work, however, demonstrated clearly that children possess a variety of solution strategies. He described their "uncanny ability," further, to rely on one of these overt strategies in just those cases where overt strategies are most helpful, that is, on problems especially prone to error, on which simpler retrieval is especially difficult, and the like. Thus, Siegler's distribution of associations model contains explicit machinery to invoke these overt solution strategies. Notice that strategic solutions were proposed to occur after a failed retrieval attempt in Siegler's original model, instead of in parallel with retrieval as is the case in Ashcraft (1987).

The second shortcoming involved the errors that subjects make. In its focus on retrieval, against the then prevailing preference for counting approaches, the network retrieval model failed to address the error effects necessary for a complete understanding of performance. Indeed, it had to appeal, in post hoc fashion, to the pattern of activation after retrieval in order to suggest that incorrect answers would be essentially those nodes with elevated levels of activation.

In contrast, Siegler and Shrager (1984) proposed that the memory representation of arithmetic facts contains both correct and incorrect answers. In this scheme, associations between problems and answers (and, later, multipliers, etc.) are formed each time a child encounters an arithmetic problem, regardless of the setting, and regardless of the correctness or incorrectness of the answer. In other words, if a child mistakenly computes (say via counting-on) 7 as the answer to $4 + 2$, then an association is formed between that problem and the mistaken answer. For reasons related to highly practiced counting associations, children may mistakenly offer 5 as the answer to $3 + 4$, which would again store that association in memory. An explanation of retrieval errors, thus, was quite natural within the model; either the retrieval process accessed one of the incorrect associations stored in memory, or the overt strategy that governed performance operated inaccurately.

The interesting consequence of the varied learning history of different problems and answers is that associations vary considerably with respect to their strength (and, obviously, their accuracy). That is, problems like $2 + 2$ are rarely solved incorrectly, so relatively few associations to incorrect answers are stored in memory. The distribution of associations here is said to be "peaked," in that strength to the correct answer 4 is quite high, and strengths of associations to other answers, if there are any, are uniformly low. In such a situation, retrieval occurs rapidly and smoothly. Few if any interfering associations disrupt the retrieval process, and the retrieval probability of the correct association, the ratio of its strength to the sum of associative strengths for all answers to the problem, seldom falls below the child's confidence criterion for responding. A problem like 5 + 9, however, experiences far more counting-related errors during acquisition, and hence yields a relatively "flat" distribution of associations. In this case, the associative strength to the correct answer 14 may only be slightly higher than strengths to other, incorrect answers. Retrieval here is far more likely to access an incorrect answer, and/or to yield a strength well below the confidence criterion. In such cases, retrieval may be reattempted and/or abandoned in favor of some overt strategy.

Finally, the Siegler model is quite explicit about the matter of strategy choice. In Siegler's view, the decision to abandon a retrieval attempt and initiate a counting process is not a matter of metacognitive, conscious choice. Instead, the model predicts this choice to be a rather mechanistic one, based on two adjustable internal values: the confidence criterion and search length parameters. The confidence criterion is an internal threshold above which retrieval probability must fall before the child offers an answer. As indicated above, this probability is the strength of the answer relative to the strengths of all answers associated with the problem. Thus, if a child sets the confidence criterion at a value like 0.50, and then retrieves an answer with a retrieval probability of 0.80, the answer will be stated. If the retrieved answer's probability value falls below the criterion, then repeated attempts are made at retrieval, but only if the subject's search length (number of searches) has not been exceeded. Once the number of retrieval attempts exhausts the search length parameter, then the problem is reformulated; that is, its representation is elaborated, say, by representing the addends physically with one's fingers. Then, some counting process is initiated, and the result of the count is stated as the answer to the problem.

The major change in Siegler's recent revision (Siegler & Jenkins, 1989) involves a strategy choice mechanism. In the revision, retrieval is no longer the default strategy for a first attempt at solution. Instead, the model first selects a strategy based on the distribution of "strategy strengths" stored individually with each problem, and then attempts to execute that strategy. The strength of a strategy reflects both its accuracy and its speed of execution.[4] Thus, problems with

[4]The selection mechanism for a strategy is computed in the same fashion as a retrieval probability, that is, the strategy's strength divided by the summed strengths of all strategies. A difference, however, is that retrieval probability is specific to a single problem, whereas a strategy's probability of selection is based on its speed and accuracy for that problem and all others with similar characteristics, for example, same or similar first addend. This necessitates considerable "record keeping" on the part of the memory representation, since the stored strengths must be continually updated based on the speed and accuracy outcomes of strategies, both for specific problems and for problems in general.

flat distributions of associations will presumably be attempted with a counting or reconstructive procedure; that is the strength of the retrieval strategy for such problems would probably be lower than the strengths of other strategies. Problems with higher strengths for the retrieval strategy, and with more peaked distributions, are commonly retrieved, subject to the same search length parameter as before. Despite the modifications in the model, it is still the case that retrieval eventually dominates processing. With experience, the strength of retrieval and the strength of correct answer associations tend to overpower other strategies and associations.

A major advantage of Siegler's model is that it provides a more plausible explanation of errors in processing than the network retrieval model, not only because of the absence of error mechanisms in the latter, but also because of the learning mechanism Siegler discusses: *any* erroneous solution – and there are many of these in childhood – has an effect on the memory representation, with repeated errors or relatively infrequent correct solutions yielding a flat distribution of associative strength, hence slower and more error-prone performance. Furthermore, the model does an admirable job in predicting the various latency, error, and solution strategy results obtained in children's addition (Siegler & Shrager, 1984), subtraction (Siegler, 1987b), and multiplication (Siegler, 1988b). It has also fared well in discussing strategy choices in other domains, for example, spelling (Siegler, 1986). And finally, it is now being applied rather successfully to the performance of different subgroups of children, notably Siegler's "perfectionists" (1988a), Geary's (1990) mathematics-disabled children, and Geary, Fan, and Bow-Thomas's (in press) Chinese and American first graders. As an example, Geary (see also Geary, Brown, & Samaranayake, 1991) reported that normal children exhibited a strategy shift from counting-on to retrieval during first grade, as well as an improvement in their rate of strategy execution. A sample of math-disabled children who showed no academic improvement across the year displayed no tendency toward more frequent retrieval, and no improvement in their rate of counting.

Despite this support, however, there are still some difficulties with Siegler's model as a general explanation of mental arithmetic performance. It is not entirely clear, for example, that the distribution of associations model deals plausibly with the problem difficulty effect beyond the early stages of elementary education. That is, the most apparent explanation for the problem difficulty effect in children is the averaging of performance across problems and subjects, which in Siegler's (1987a) analysis of young children's performance clearly confounds problem difficulty with multiple strategy use. This averaging, in fact, was the source of Siegler's dissatisfaction with then prevailing methods of data analysis and resultant theorizing.

But for older children and adults, the continued significance of the problem difficulty effect is less explainable in the model, because their performance is admittedly due to retrieval the vast majority of the time. For instance, Siegler and

Jenkins' (1989) simulation reaches the point at which 99% of the trials are accomplished by retrieval; Geary and Wiley (1991) found that retrieval accounted for 88% of their college students' performance. It is safe to assume that the associative strengths for correct answers are, for the typical adult, large enough to exceed the confidence criterion, even though those strengths tend to be lower for larger, more difficult problems. And yet, the model does not predict RT to be a function of associative strength. Instead, the solution time on retrieval is predicted to be "proportional to . . . the number of searches on each problem before an answer is stated" (Siegler & Shager, 1984, p. 250). A straightforward reading of the model, then, suggests that since associative strength will be relatively high, the need for repeated searches would be correspondingly low. If so, then most retrieval trials would require only one search, predicting no RT increase at all. And even at the highest reasonable values for the search length parameter, say 4 or 5, it is not clear that this variation alone could predict the full problem difficulty effect. Thus, the model's account for problem difficulty effects may be somewhat implausible beyond the ages at which overt, reconstructive strategy use is common.

Note that Siegler's model does not explicitly discuss the verification task, or how the model predicts performance on that task. As such, the interference nature of verification effects is not generated directly by his model. Instead, the model's adequacy in dealing with the relatedness effects must be considered from a different perspective: would it generate the same kinds of errors in simulating the production task that have been shown to disrupt processing in verification performance? The answer to this question is yes. Siegler's model provides an explicit basis for the relatedness effect, in terms of each problem's associations to both correct and (potentially many) incorrect answers. There is thus a strong likelihood that a related answer, say 24 for 4×8, will match one of the stored, incorrect answers in memory. Thus, a fairly minor addition to the model, involving some sort of procedure for checking the stated answer against stored associations, might generate the interference effects that have been reported. The same mechanism would also predict the split effects described above, since incorrect addition answers wrong by a small amount are far more likely to have been stored in memory than answers wrong by a larger amount. In fact, if this additional mechanism does not violate the spirit of the model, the variations in associative strength to incorrect answers would provide a more detailed basis for prediction than simple levels of split. As an example, the incorrect association $3 + 4 = 5$ has considerably more strength than $3 + 4 = 9$, though both problems have a split equal to 2. The model should therefore predict greater disruption of performance to the former problem. No such specific tests have been reported, however.

Campbell's network interference model

Campbell's model of arithmetic performance (1987a: Campbell & Graham, 1985) represents an important modification of both Ashcraft's and Siegler's frameworks, although it is not a wholesale departure from those approaches. For example, it shares the important notions of spreading activation and strength of stored nodes with Ashcraft's model. It shares with Siegler's the conviction that the nature of errors committed by subjects is critically informative about the underlying memory representation; that is, errors provide an index of the relatedness among stored elements. Unlike the other two approaches, Campbell's unique contributions are his focus on interference as an unavoidable part of the retrieval process, and his elaboration of the kinds of stored connections that are potent influences on adults' RT performance.

The telling pattern of errors in multiplication, discussed previously, and its implication for the structure of the memory representation, form the core of Campbell's new perspective on mental arithmetic. There is no doubt, of course, that retrieval is "operand driven," that the mechanism of accessing stored information is activated and shaped by the operands in the problem. In other words, Campbell's model makes the uncontroversial claim that the values in a problem trigger the retrieval process. As in the simpler network retrieval model, this process is conceived as a spread of activation through the network, with nodes activated to different degrees by virtue of their different associative strengths.

Campbell notes, however, that retrieval cannot be influenced solely by the two separate operands in a problem. If only the operands trigger the spread of activation, then the system will retrieve both 24 and 32 as answers to 8×4, since both answers have associations to both operands (e.g., Campbell & Graham, 1985, p. 352); this is the case under Campbell's assumption that answers are not stored redundantly in the network (i.e., there is only one 32 node; but cf. Ashcraft, 1987). Thus, in Campbell's proposed network, there are not only associations or pathways from individual operands to answers, but also associations from whole problems to answers. Thus, in a straightforward production task, the outcome of retrieval will depend on two distinct kinds of parallel activation spreads, one originating from the individual operands, and one from the problem as a whole.

Because the spread of activation from operands will activate not only the correct answer but also the entire set of multiples from each operand, Campbell's model speaks of the "candidate set" of answers, that is, the set of nodes activated during retrieval. For the most part, the candidate set will contain table-related (operand-related) values. Because of the cumulative effects of errors across development, however, there will also be some spurious and incorrect associations coded in the network. Each such connection, if activated during retrieval, will "promote" another value into the candidate set, at a level dictated by the strength of that connection.

The pervasive influence of interference during routine retrieval is quite apparent in this scheme. Each act of retrieval activates a large number of associations, and a set of candidate responses. The retrieval process here must therefore discriminate among the elements of the candidate set, selecting the most strongly activated value from the entire set. Thus, reminiscent of a pandemonium-like system (Selfridge, 1959), accurate performance rests on greater activation of the correct answer than any of the remaining candidates. Interference effects, of course, result from retrieving an erroneous but related node, with a probability depending on the strength of that erroneous association and hence its level of activation.

Campbell's scheme has been criticized recently as less parsimonious than might be desirable, in its profusion of the different types of associations embedded into the network (e.g., McCloskey et al., 1991). It is indeed difficult to ascertain in the predictive sense what the exact chronometric or error characteristics of performance will be, since independent assessments of the strength of the various associations are not provided in the model (but see Campbell & Oliphant's (1992) description of their simulation model). On the other hand, Campbell's empirical demonstrations, especially the error priming effect, argue strongly that such a "tangle" of associations is justified.[5] In error priming, activation stemming from one trial influences the speed and accuracy of a subsequent trial. Thus, Campbell (1987a, 1991) found that retrieving an answer that is also a strong but false associate to a subsequent problem will in fact increase the likelihood that the false association governs the subsequent retrieval. For example, 56 is a frequent error to the problem 7×9. When it is correctly retrieved for the problem 7×8, that retrieval of 56 increases the probability that subjects will erroneously respond "56" to the problem 7×9 on a subsequent trial. In effect, the 56 has been primed by the earlier retrieval, and its residual activation interferes with the subsequent retrieval for 7×9.

Note that this error priming effect parallels inter-trial effects found elsewhere, for example in the semantic memory literature (e.g., Loftus & Loftus, 1974). As such, Campbell's demonstrations (1987a; Campbell & Clark, 1989) are important for two reasons. First, they expand the scope of the relatedness effect. In particular, relatedness can lead to either facilitation or disruption of performance, and not only within a particular trial (e.g., the confusion effect), but also across trials; at short lags between related trials, recently retrieved answers are inhibited, but then are promoted as errors at longer lags (Campbell, 1990, 1991). And second, they indicate the suitability, or even necessity, of spreading activation models in arithmetic performance. That is, these inter-trial influences on performance imply that a *general* carry-over mechanism is insufficient. The carry-over is specific to the information accessed in the previous trials, and hence implicates the continued influence of that specific information. Precisely such

[5]Campbell and Graham (1985) also draw an analogy between learning the basic arithmetic facts and paired-associate learning, with the parallel expectation that the similarities among arithmetic facts would generate substantial interference during acquisition.

patterns are characteristic of spreading activation models (as indeed are data indicating an influence of the answer in a verification task; e.g., Campbell, 1987b; Zbrodoff & Logan, 1990).

The network interference model holds an advantage over Siegler's framework because of its elaborated assumptions about network storage and retrieval processes; automatic retrieval effects are expected in Campbell's model, but not discussed directly by Siegler. Campbell postulates a network structure in which spreading activation accumulates and persists across time, but nothing in Siegler's model suggests a role for spreading activation, or predicts trial-to-trial effects. The role of answers in influencing RT is also clear in Campbell's model. As stated above, Siegler does not discuss how verification performance might operate within his model, and the Ashcraft model was completely silent on error effects. Conversely, Campbell's model is essentially silent on the issue of strategies other than retrieval, and their influence on processing. Just as Siegler's special focus has been on children's strategies, Campbell's has been on adults' RT effects.

Towards an integrated model

As indicated in the review, and in Table 2, the three current models show broad agreement on several fundamental principles: arithmetic facts are stored in an interrelated memory representation; the stored associations differ in their strength; retrieval has a central role in performance, even for young children; and strategy-based processing, especially common among younger children, yields substantially to retrieval across development. While not minimizing the discrepancies, it would seem that a judicious borrowing of features across models could integrate the models. What follows is a sketch of how such borrowing might modify the network retrieval model; for obvious reasons, I do not suggest modifications of others' models.

Consider the Ashcraft (1987) network retrieval model with the following two changes and elaborations:

(a) Each operand has stored pathways or associations to a variety of answers, correct as well as incorrect; that is, Siegler's and Campbell's assumption. These associations, which will vary in strength as a function of experience, will tend to reflect the error, confusion, and possibly the split characteristics described above.

(b) The procedural knowledge component is elaborated to capture the variety of strategies observed among children, as well as the more idiosyncratic strategies that are sometimes invented (e.g., the variety of specific methods that fall under the "decomposition" or "solve from known facts" strategies; see Hamann & Ashcraft, 1985; Siegler, 1987a). Notice that, as in the original model, some associations in the network will be of very low strength, such that functionally the

Table 2. *Summary of arithmetic models*

Model	Mental representation of facts	Hypothesized mental processes	Effects predicted	Effects not predicted
A. Groen and Parkman's *min* model (1972)	N/A	Counting; increment by 1 s	Counting; linear RT increase	Retrieval; errors; relatedness; mixture of strategies, priming
B. Groen and Parkman's direct access	Unspecified	Direct memory retrieval plus backup counting	Retrieval and counting; linear RT increase	Errors; relatedness; mixture of strategies; priming
C. Ashcraft's network retrieval (1982, 1987)	Interrelated network	Retrieval via spreading activation	Problem difficulty; relatedness; retrieval and counting; priming; automaticity	Errors; non-counting strategies
D. Siegler and Jenkins distribution of associations (1989)	Problem and strategy associations	Retrieval and strategies via associations, confidence criterion, search length	Problem difficulty in children; mixture of strategies; errors	Problem difficulty on retrieval trials; priming, automaticity
E. Cambell (1987) network interference	Interrelated network	Retrieval via spreading activation	Problem difficulty; relatedness; priming; automaticity; errors	Mixture of strategies

relevant problem may always be solved by a rule rather than by retrieval (e.g., $N \times 0$). Siegler's strategy strength mechanism may in fact operate in exactly this fashion.

Such a revision, note, maintains the notion that retrieval and strategy solutions are triggered in parallel, with the faster route governing performance. This may be most easily thought of in terms of Siegler's strategy associations, in which individual problems also have associative connections, of varying strengths, to strategies that have proved successful in previous circumstances. Note, however, that this proposal is still fundamentally different from Siegler's, in which only one association is operative at any given moment. Instead, this revision claims that strategy associations are accessed via spreading activation, as are answer associations, according to their individual strengths. Thus, all associations, regardless of their strength, receive activation during processing, but of course the degree of activation will depend on association strength. This is virtually identical to conventional network assumptions regarding semantic relatedness, typicality, etc. (e.g., Collins & Loftus, 1975).

One prediction of the parallel operation of retrieval and procedural processes might be that trials being processed via a strategy may be disrupted by retrieval-based interference, and vice versa. This issue, among others, simply cannot be decided yet, largely for lack of evidence. For example, evidence on the relative independence of retrieval and procedural components (see the discussion of modularity below, based on evidence from brain-damaged individuals) might argue against such interference across components. Interference might, however, be a natural prediction within, say, a neural network approach (e.g., Churchland, 1990; McCloskey & Lindemann, 1992). It of course remains to be seen whether an integration of the models, such as that sketched here, will yield new insights into these issues.

Two larger issues

Rules of arithmetic

This review, like the bulk of the research, has focused on the simple, basic facts of arithmetic, especially addition and multiplication. Yet clearly, our understanding of arithmetic and mathematical cognition is impoverished if we neglect more complex computations, and the rules of arithmetic by which the computations are performed. That is, for problems like $147 + 259$ or 90×94, there is a sequencing of steps or components, including retrieval of course, but also operations such as carrying or borrowing, keeping track of place in the sequence, maintaining intermediate answers, and so forth. Consider the following lines of research, presented only briefly here, as examples of this larger scope of arithmetic investigations.

Work by Widaman et al. (1989; also Geary & Widaman, 1987) has examined adults' processing in complex arithmetic situations, in particular evaluating RT to verification problems in terms of a componential model of performance. The model assesses the relative contributions of several elementary information processes, including simple fact retrieval, the encoding of multiple digits, and carrying.[6] For a problem like $48 + 16 = 64$, for instance, total RT is composed of fact retrieval time for the $8 + 6 = 14$ step, an increment of time for comparing the 4 in the ones column to the value stated in the problem, time for carrying the 1 to the tens column, and so forth. The analyses also included a self-terminating decision component, to reflect subjects' shorter RTs to incorrect problems in which the error appears in the 1s column versus the 10s column of the answer (e.g., faster RT to $46 \times 8 = 366$ vs. 358).

Part of the strength of this work is that it relates performance on the separate components of RT to more traditional measures of ability and performance, for example, assessments of working memory capacity or resources, perceptual or spatial abilities, and the like (e.g., Geary & Widaman, 1987). Such demonstrations confirm the utility of RT tasks and models for discovering the underlying cognitive bases of arithmetic skills, and also provide an exploitable avenue of investigation for studies of mathematics disabilities; for example, the Geary et al. (1991) study of math-disabled children.

A second approach to this broader class of "rules of arithmetic" is the now classic work by Brown and Burton (1978; Brown & VanLehn, 1980; VanLehn, 1990) on children's subtraction. This research, and the cognitive modeling based on it, relies on extensive analysis of the patterns of errors that children commit when working complex subtraction problems in a pencil-and-paper task. In particular, the approach diagnoses *bugs* in children's procedural or rule knowledge – (mis)understandings of one or another component of the complete set of subtraction rules.

Consider the problem 42–3, to which a child might respond with the answer 41. Brown and Burton's analysis claims that the given answer reflects a bug in the child's understanding of the borrowing rule: "subtract the smaller from the larger value", i.e., subtract 2 from 3. A more subtle bug is apparent in the problem

[6]The table-retrieval network model proposed for fact retrieval by Widaman et al. was not considered in the Current models section because it fails to specify a number of important mechanisms. For example, a problem is retrieved by a spread of activation into the area of the network defined by its operators; for 7×6, the entire area beginning at the origin (0, 0), bounded by the (7) column and (6) row and the intersection at (7, 6). The explication of the model does not indicate, however, whether all the intersections contained in the activated area are themselves activated, and thus whether any problem within that area would be facilitated as a target of the activation spread. If they are activated, then the model predicts much wider priming effects than are found; if they are not activated, then no apparent mechanism predicts the error and relatedness effects reviewed above.

801 − 158 = 553. Here, the bug avoids borrowing from zero, by taking both borrows from the leftmost column, the 8 in 801.

The important point here is that children will sometimes err not in their simple performance on facts, but instead on the procedural aspects of the operation, that is, on the rules for borrowing. Such demonstrations are rich in empirical and educational implications, of course. For present purposes, though, note that the theoretical ramifications are especially interesting. The Brown and VanLehn research, along with the work by Widaman et al., adds significantly to the evidence for two classes of arithmetic knowledge, facts and procedures, each with its own characteristics for accuracy, each acting in at least partial autonomy from the other. In ways just beginning to be explored, this evidence speaks to the overall organization and architecture of the cognitive arithmetic system.

Architecture

Reconsider for a moment the simple rules for adding or multiplying with an operand of 0 or 1. The evidence from RT studies was somewhat equivocal on the nature of processing these problems. Some evidence indicated that these problems were retrieved in the same way as larger facts, and some indicated they were processed via special rules. Stazyk et al. (1982; also Parkman, 1972) found that verification performance to $N \times 0$ problems was especially slow and error prone. They concluded that "zero problems" are routinely performed by means of a rule, rather than straightforward retrieval. On the other hand, Miller et al. (1984) found that these problems were not performed unusually slowly in a production task. Adults' normative judgments indicate that these are low-difficulty problems. On the other hand, these problems appear with substantially lower frequency in arithmetic textbooks than problems with non-zero small operands (Hamann & Ashcraft, 1986).

Research by McCloskey and his colleagues (e.g., McCloskey, this issue; McCloskey et al., 1985; Sokol et al., 1991) has found that performance to the zero problems in multiplication can be very informative about the architecture of the mental arithmetic system. In their case studies of patients with brain damage, intriguing patterns of dissociations between fact retrieval and rule-based performance suggest that these two components may in fact be autonomous modules in the arithmetic system (see also Luria, 1980; Warrington, 1982, presents evidence of a similar dissociation).

The case studies reported in Sokol et al. (1991; also McCloskey, Aliminosa, & Sokol, 1991) are particularly informative. One of these patients (PS), with left temporal lobe damage, erred fairly frequently in basic fact retrieval – for non-zero problems, approximately 15% errors. Her performance on complex multiplication showed frequent errors in basic fact retrieval but hardly any errors on the procedures of multiplying (e.g., carrying, shifting columns). Remarkably, she was almost always inaccurate on $N \times 0$ problems in early test sessions (97.5% error rate), and yet was 100% *correct* on the zero problems embedded in multi-column multiplication problems, for example, 90×94.

The Sokol et al. analysis of this pattern involved two steps. First, the similarity of PS's errors in simple and complex multiplication suggested that a basic fact retrieval component was common to both problem types, and was fairly impaired in PS's case. Her accuracy on the multiplication algorithm, however, suggested that a rule-based component was relatively unimpaired, with one exception. This exception appeared to be the special case rule $N \times 0 = 0$. In some fashion, PS seemed to have "lost" this rule, accounting for the 97.5% error rate. In multi-column multiplication, however, she followed a different method, the common "bring down the zero" rule, and thus avoided her incorrect $N \times 0$ rule. Sokol et al.'s second case study also demonstrated uniform inaccuracy on $N \times 0$ facts (both patients advanced N as the answer to these problems), yet very accurate computational performance when the multi-digit problem involved a zero (93%).

As described elsewhere (e.g., McCloskey, this issue: see also Deloche & Seron, 1987), these two patients are not by any means the only source of evidence for such dissociations in arithmetic performance. For the present review, however, this evidence is sufficient to make the following point. Theories and models in mental arithmetic have long speculated that fact retrieval, procedures such as carrying and borrowing, and special rules like $N \times 0 = 0$, reflect separate components in the overall processing system. Yet, standard methodologies gave only indirect and somewhat inconclusive evidence about this division of labor. The reports of dissociations among neurologically impaired individuals, however, indicate quite strongly that separate, autonomous systems are responsible for retrieval and computation-based performance.

McCloskey et al. (1985), in fact, offer a strong interpretation of these data, claiming they are evidence for a clearcut modularity in number and arithmetic processing (see also McCloskey, this issue; McCloskey, Sokol, & Goodman, 1986: Sokol, Goodman-Schulman, & McCloskey, 1989; Sokol et al., 1991). One of the modules in the normal system, termed Arithmetic Facts, accomplishes those functions attributed to the long-term memory representation discussed earlier. Note that McCloskey et al. (1991) expressed serious reservations about the three current models discussed earlier, yet did not make a theoretical commitment to any alternative set of representation and processing assumptions for their Arithmetic Fact component (but see McCloskey & Lindemann, 1992). A second autonomous module, Calculation Procedures, involves the various carrying, borrowing, sequencing, and other procedural rules discussed earlier. Sokol et al.'s two patients, according to this interpretation, experienced neurological damage that affected one of the modules but not, for the most part, the other.

Suggestions of modular cognitive architecture, based on the selective pattern of impairments shown by brain-damaged patients, and case histories on such impairments (see Deloche & Seron, 1987, for a collection of reviews) are not new, of course, although such effects in arithmetic are certainly less widely known than those on language disruptions. This by no means indicates that a modular approach to arithmetic processing is universally accepted, however. In particular, Campbell and Clark (1988; Clark & Campbell, 1991) have presented arguments against the McCloskey et al. modular theory, and especially against the proposal that a single, abstract semantic representation underlies all processing. Such a proposal, Clark and Campbell (1991) argue, ignores data that show format-specific effects; for example, different error patterns to multiplication problems presented in digit versus word format imply the opposite of a format-free, abstract representation.

Regardless of the specific strengths or weaknesses of either approach, a few more general comments may be abstracted from this debate. First, it is probably premature to make strong claims about modularity of processing, or the lack thereof, given our current understanding of number representation and calculation, and of their neurological realization. After all, even existing data pose some difficulties for a strictly modular explanation.[7] Thus, modularity might be viewed as an appropriate and useful working hypothesis, a framework to be explored and exploited, and to be discarded (Sokol et al., 1989) if and when evidence accumulates that the assumptions are no longer useful.

Concluding remarks

In my earliest paper on cognitive arithmetic (Ashcraft & Battaglia, 1978), I challenged the Groen and Parkman *min* model on intuitive grounds; I asked in essence, "Why would adults continue to count in order to solve the basic addition facts, even after years of experience?" The answer, "They don't," led to a productive set of new questions, "Well, what *do* they do, and how do children eventually get to that level?" This paper has described the history of that set of questions, and the answers that can now be offered.

A related question is still a puzzle: why is there a problem size/difficulty effect? Why, after so many years of experience, are larger problems difficult enough that even adults continue to show this effect in their performance?

I suspect that larger problems are at a somewhat permanent disadvantage.

[7]Midway through the testing sessions, PS apparently noticed her inconsistency on zero problems, and thereafter was 94.8% correct on these problems. Sokol et al. speculated that she may have then reformulated the previously incorrect $N \times 0$ rule. While altogether plausible, the explanation also raises the question of how communication and feedback take place among and within the presumably autonomous modules.

They clearly occur less frequently, and are acquired somewhat later than their smaller counterparts, in the early years of formal schooling (Hamann & Ashcraft, 1986). It may be that this relative imbalance continues, as we experience problems naturally across the years, thus explaining the effect. Such an argument advances the analogy between problem difficulty and the well-known effects of word frequency. Just as frequently occurring words enjoy a privilege in lexical and semantic access, frequent problems enjoy an advantage in the memory system for arithmetic.

Two objections can be raised to this line of reasoning. The first is that we are merely guessing about a "continued imbalance." That is, perhaps it is just as likely that later schooling and adult experience will counteract, rather than perpetuate, the early imbalance. In other words, a bias in first-grade texts "does not imply that such a bias is present in the full set of problems that college-age adults have encountered in their experience with arithmetic facts" (McCloskey et al., 1991, p. 382). Of course, if later experience roughly equates the problems on frequency, then the current models are robbed of their explanation of the problem size/difficulty effect, unless some other appeal, say to age of acquisition, is made.

This counterargument is possibly true. And yet, there is at least suggestive evidence that the assumption of imbalance and the analogy to word frequency are both tenable. Dehaene and Mehler (1992; see also Benford, 1932), for example, tabulated the frequency of occurrence of numbers across many naturally occurring samples (and across several different languages). They found that frequency is generally a decreasing function of magnitude. Further, we have recently tabulated the frequency of basic addition and multiplication facts in textbooks from grades 3 through 6 (Ashcraft & Christy, 1991), thus extending the earlier tabulation for grades kindergarten through 3 (Hamann & Ashcraft, 1986). The same patterns were found; smaller facts, both in isolation and as components to larger problems, occurred more frequently than larger ones, and facts with zeros occurred the least frequently of all. Thus, numbers appear naturally as a decreasing function of magnitude, as do arithmetic problems, at least as they are represented in texts through sixth grade.

The second objection involves the apparent tautology that more difficult problems are processed more slowly. That is, when subjects rate problems as more difficult, is this a legitimate predictor of RT, or are the ratings instead based on subjects' subjective awareness of how long it took to solve the problem? (See Washburne & Vogel, 1928, for an earlier illustration of truly circular reasoning; some problems were judged "inherently more difficult" because children made more errors to them.) I argue here that the textbook data provide at least a partial escape from the circular logic. That is, the correlation between strength in memory, as indexed by difficulty ratings, and RT cannot disentangle the cause–effect relationship we need. But consider a completely plausible causal pathway, and the supporting analogy to word frequency. Differential frequency of exposure, especially in elementary school, and possibly in adult experience as well, causes differential memory strength. Memory strength, the inverse equivalent of problem difficulty, in turn influences the duration of retrieval. If this causal pathway is supported with additional normative data, and if direct manipulations of frequency show a lawful frequency–difficulty–performance relationship, then the hypothesis will hold.

I suspect that this will be the case. The literature claims repeatedly that arithmetic processing is similar to language processing, but has the interesting characteristic that its learning and development are more easily investigated because they are deliberately, formally taught. If this claim is true, and if the word frequency–problem difficulty analogy holds, then the area of cognitive arithmetic could suggest answers to any number of questions about language, and about the overall human cognitive system. In particular, via manipulation of problem frequency, we may not only discover important information about knowledge of number and arithmetic, we may also be able to investigate other effects, confounded or correlated in normal experience, that shape the contents and procedures of human memory and cognition.

References

Allen, P.A., McNeal, M., & Kvak, D. (in press). Perhaps the lexicon is coded as a function of word frequency. *Journal of Memory and Language.*

Ashcraft, M.H. (1976). Priming and property dominance effects in semantic memory. *Memory & Cognition, 4,* 490–500.

Ashcraft, M.H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2,* 213–236.

Ashcraft, M.H. (May, 1985). *Children's knowledge of simple arithmetic: A developmental model and simulation.* Paper presented at the conference on "Formal Models in Developmental Psychology," University of Alberta, Edmonton.

Ashcraft, M.H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C.J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology: Progress in cognitive development research* (pp. 302–338). New York: Springer-Verlag.

Ashcraft, M.H. (1990). Strategic processing in children's mental arithmetic: A review and proposal. In D.F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 185–212). Hillsdale, NJ: Erlbaum.

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., & Christy, K. (1991) *Tabulation of problem frequency in elementary texts: Grades 3–6.* Unpublished manuscript.

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., Fierman, B.A., & Bartolotta, R. (1984). The production and verification tasks in mental addition: An empirical comparison. *Developmental Review, 4,* 157–170.

Ashcraft, M.H., & Stazyk, E.H. (1981). Mental addition: A test of three verification models. *Memory and Cognition*, *9*, 185–196.

Benford, F. (1932). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, *78*, 342–348.

Bisanz, J., & LeFevre, J.A. (1990).; Strategic and nonstrategic processing in the development of mathematical cognition. In D.F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 213–244). Hillsdale, NJ: Erlbaum.

Brown, J.S., & Burton, R.R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, *2*, 155–192.

Brown, J.S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, *4*, 379–426.

Brownell, W.A. (1928). *The development of children's number ideas in the primary grades*. Chicago: University of Chicago Press.

Campbell, J.I.D. (1985). *Associative interference in mental computation*. Unpublished doctoral dissertation, University of Waterloo.

Campbell, J.I.D. (1987a). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 109–123.

Campbell, J.I.D. (1987b). Production, verification, and priming of multiplication facts. *Memory and Cognition*, *15*, 349–364.

Campbell, J.I.D. (1990). Retrieval inhibition and interference in cognitive arithmetic. *Canadian Journal of Psychology*, *44*, 445–464.

Campbell, J.I.D. (1991). Conditions of error priming in number-fact retrieval. *Memory and Cognition*, *19*, 197–209.

Campbell, J.I.D., & Clark, J.M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, and Goodman (1986). *Journal of Experimental Psychology: General*, *117*, 204–214.

Campbell, J.I.D., & Clark, J.M. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 920–929.

Campbell, J.I.D., & Graham, D.J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology*, *39*, 338–366.

Campbell, J.I.D., & Oliphant, M. (1992). Number-fact retrieval: A model and computer simulation. In J.I.D. Campbell (Ed.), The nature and origins of mathematical skills (pp. 331–364). Amsterdam: Elsevier.

Churchland, P.M. (1990). Cognitive activity in artificial neural networks. In. D.N. Osherson & E.E. Smith (Eds.), *Thinking: An invitation to cognitive science* (Vol. 3, pp. 199–227). Cambridge, MA: MIT Press.

Clapp, F.L. (1924). *The number combinations: Their relative difficulty and frequency of their appearance in textbooks*. Research Bulletin No. 1, Madison, WI: Bureau of Educational Research.

Clark, J.M., & Campbell, J.I.D. (1991). Integrated versus modular theories of number skills and acalculia. *Brain and Cognition*, *17*, 204–239.

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Cooney, J.B., Swanson, H.L., & Ladd, S.F. (1988). Acquistion of mental multiplication skill: Evidence for the transition between counting and retrieval strategies. *Cognition and Instruction*, *5*, 323–345.

Dehaene, S. & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29.

Deloche, G. & Seron, X. (Eds.) (1987). *Mathematical disabilities: A cognitive neuropsychological perspective*. Hillsdale, NJ: Erlbaum.

Geary, D.C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, *49*, 363–383.

Geary, D.C., Brown, S.C., & Samaranayake, V.A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology*, *27*, 787–797.

Geary, D.C., Fan, L., & Bow-Thomas, C.C. (in press). Numerical cognition: Loci of ability differences comparing children from China and the United States. *Psychological Science*.

Geary, D.C., & Widaman, K.F. (1987). Individual differences in cognitive arithmetic. *Journal of Experimental Psychology: General*, *116*, 154–171.

Geary, D.C., Widaman, K.F., & Little, T.D. (1986). Cognitive addition and multiplication: Evidence for a single memory network. *Memory & Cognition*, *14*, 478–487.

Geary, D.C., & Wiley, J.G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young and elderly adults. *Psychology and Aging*, *6*, 474–483.

Gelman, R., & Gallistel, C.R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Groen, G.J., & Parkman, J.M. (1972). A chronometric analysis of simple addition. *Psychological Review*, *79*, 329–343.

Hamann, M.S., & Ashcraft, M.H. (1985). Simple and complex mental addition across development. *Journal of Experimental Child Psychology*, *40*, 49–72.

Hamann, M.S., & Ashcraft, M.H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction*, *3*, 173–192.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Koshmider, J.,W., III, & Ashcraft, M.H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology*, *51*, 53–89.

Lave, J. (1988). *Cognition in practice*. Cambridge, UK: Cambridge University Press.

LeFevre, J.A., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory & Cognition*, *16*, 45–53.

Loftus, G.R., & Loftus, E.F. (1974). The influence of one memory retrieval on a subsequent memory retrieval. *Memory & Cognition*, *2*, 467–471.

Luria, A.R. (1980). *Higher cortical functions in man* (2nd edn.). New York: Basic Books.

McCloskey, M, Aliminosa, D., & Sokol, S.M. (1991). Facts, rules, and procedures in normal calculation: Evidence from multiple single-patient studies of impaired arithmetic fact retrieval. *Brain and Cognition*, *17*, 154–203.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*, 171–196.

McCloskey, M., Harley, W., & Sokol, S.M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *17*, 377–397.

McCloskey, M., & Lindemann, A.M. (1992). MATHNET: Preliminary results from a distributed model of arithmetic fact retrieval. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 365–410). Amsterdam: Elsevier.

McCloskey, M., Sokol, S.M., & Goodman, R.A. (1986). Cognitive processes in verbal number production: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: General*, *115*, 307–330.

Miller, K.F., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 46–60.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Pachella, R.G. (1974). The interpretation of reaction time in information processing research. In B.H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum.

Parkman, J.M. (1972). Temporal aspects of simple multiplication and comparison. *Journal of Experimental Psychology*, *95*, 437–444.

Parkman, J.M., & Groen, G.J. (1971). Temporal aspects of simple addition and comparison. *Journal of Experimental Psychology*, *89*, 335–342.

Reder, L.M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, *89*, 250–280.

Resnick, L.B., & Ford, W.W. (1981). *The psychology of mathematics for instruction*. Hillsdale, NJ: Erlbaum.

Selfridge, O.G. (1959). Pandemonium: A paradigm for learning. In *The mechanisation of thought processes*. London: H.M. Stationery Office.

Siegler, R.S. (1986). Unities in strategy choices across domains. In M. Perlmutter (Ed.), *Minnesota symposium on child development* (Vol. 19, pp. 1–48). Hillsdale, NJ: Erlbaum.

Siegler, R.S. (1987a). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250–264.

Siegler, R.S. (1987b). Strategy choices in subtraction. In J. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics*. Oxford; Oxford University Press.

Siegler, R.S. (1988a). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*, 833–851.

Siegler, R.S. (1988b). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General, 117*, 258–275.

Siegler, R.S., & Jenkins, E.A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

Siegler, R.S., & Robinson, M. (1982). The development of numerical understandings. In H. Reese & L.P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 16, pp. 241–312). New York: Academic Press.

Siegler, R.S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Simpson, G.B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316–340.

Sokol, S.M., Goodman-Schulman, R., & McCloskey, M. (1989). In defense of a modular architecture for the number-processing system: Reply to Campbell and Clark. *Journal of Experimental Psychology: General, 118*, 105–110.

Sokol, S.M., McCloskey, M., Cohen, N.J., & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 355–376.

Stazyk, E.H., Ashcraft, M.H., & Hamann, M.S. (1982). A network approach to simple multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 320–335.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W.G. Koster (Ed.), *Attention and performance II. Acta Psychologica, 30*, 276–315.

Thorndike, E.L. (1922). *The psychology of arithmetic*. New York: Macmillian.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions*. Cambridge, MA: MIT Press.

Warrington, E.K. (1982). The fractionation of arithmetical skills: A single case study. *Quarterly Journal of Experimental Psychology, 34A*, 31–51.

Washburne, C., & Vogel, M. (1928). Are any number combinations inherently difficult? *Journal of Educational Research, 17*, 235–255.

Wheeler, L.R. (1939). A comparative study of the difficulty of the 100 addition combinations. *Journal of Genetic Psychology, 54*, 295–312.

Widaman, K.F., Geary, D.C., Cormier, P., & Little, T.D. (1989). A componential model for mental addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 898–919.

Winkelman, J.H., & Schmidt, J. (1974). Associative confusions in mental arithmetic. *Journal of Experimental Psychology, 102*, 734–736.

Woods, S.S., Resnick, L.B., & Groen, G.J. (1975). An experimental test of five process models for subtraction. *Journal of Educational Psychology, 67*, 17–21.

Zbrodoff, N.J., & Logan, G.D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115*, 118–130.

Zbrodoff, N.J., & Logan, G.D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 83–97.