An Interactive Activation Model of Arithmetic Fact Retrieval

Timothy C. Rickard, Michael C. Mozer, and Lyle E. Bourne Jr.

Institute of Cognitive Science
University of Colorado
Boulder, CO 80309

Technical Report 92-15

An Interactive Activation Model of Arithmetic Fact Retrieval

Timothy C. Rickard, Michael C. Mozer, and Lyle E. Bourne Jr.

Institute of Cognitive Science
University of Colorado
Boulder, CO 80309

Technical Report 92-15

An Interactive Activation Model of Arithmetic Fact Retrieval

Timothy C. Rickard, Michael C. Mozer, & Lyle E. Bourne Jr.

University of Colorado

Basic arithmetic is a relatively self-contained and well-defined knowledge domain which provides rich opportunities for the detailed study of a naturally occurring cognitive skill. Over the last two decades, psychologists have conducted extensive research into arithmetic performance of both children and adults, and several models have emerged from this work which should prove to have both practical value (e.g., in educational settings), and broad theoretical implications for topics such as skill acquisition and theories of associative memory. In this paper, we contribute to this line of work by presenting preliminary results from a new model of skilled (i.e., adult) arithmetic which, among other things, synthesizes selected components of existing models (Ashcraft, 1987; Campbell & Graham, 1985; Siegler, 1988).

Although our work is intended as a general model of skilled arithmetic, we focus on multiplication in the present paper. There is in fact little empirical data which directly addresses arithmetic operations other than multiplication and addition (but see Campbell, 1985; Rickard, 1992). Further, the majority of the performance phenomena have been reported either solely for multiplication, or for both multiplication and addition. Thus, a model which can account for multiplication performance accounts for most of the data on arithmetic currently available, and is a strong candidate as a more general theory of skilled arithmetic performance.

The paper is organized into four sections. In the first section, evidence is presented in support of a basic underlying assumption of the proposed model-- that skilled performance on simple arithmetic fundamentally reflects retrieval of answers from an associative network of arithmetic facts (although specialized rules and procedures may be also be used under specifiable circumstances). Second, the major empirical phenomena will be reviewed, and existing models of multiplication fact retrieval will be described and evaluated in light of these phenomena. Third, the current model will be described, and the model's accounts for the various phenomena will be discussed. Finally, we will discuss some remaining theoretical issues, outline some possible approaches to exploring these issues, and propose some directions for future research.

Evidence for Fact Retrieval Processes in Skilled Mental Arithmetic

There are two theoretical perspectives from which to consider the processes involved in skilled mental calculation; the procedural perspective and the associative network perspective. The procedural perspective holds that mental arithmetic involves the execution of general algorithms or rules. For multiplication, one of the most basic of these is a repeated adding algorithm, such that, for example, 3 x 7 is performed by adding 7 + 7 + 7. There is direct behavioral evidence from Siegler (1988) that third-grade children use this algorithm frequently. However, in this same study, the children showed no behavioral signs of employing any algorithm on 68% of the problems, and Siegler (1988) concluded that in these instances the problems were likely solved by direct retrieval of answers from associative memory. Thus, even as early as third grade, the basic repeated adding algorithm appears to be by-passed in a majority of cases in favor of an approach which is consistent with fact retrieval processes. Consistent with Siegler's findings, a recent study by Bourne and Rickard (1991) provides evidence that consciously mediated algorithms are not usually involved in adult performance: They found that, for simple problems ranging from 2 x 5 to 8 x 9, adults report that the answer simply "pops to mind" on around 85% of problems.

A perhaps more plausible version of the procedural perspective holds that skilled performers use specialized procedures, or rules, which apply to restricted sub-groups of problems, and which may with experience become "automatized" to the point that they are not (or only occasionally) available to conscious experience. For example, Baroody (1985) suggested that for the class of problems 0 x N (where N is any number), a simple rule is learned which specifies that, if the problem contains zero as a multiplicand, then the answer is zero. Baroody (1985) proposed that similar rules may exist from most or all other arithmetic problems, circumventing the need to commit individual facts to memory. Whereas recent research does suggest that rules are sometimes used for 0 x N problems (Sokol, McCloskey, Cohen & Aliminosa, 1991), and perhaps also for 1 x N problems, it is unclear whether other equally efficient and precise rules exist (or are discovered) for other problems (although there are less precise rules that may be useful in constraining the set of candidate answers, e.g., the odd/even rule proposed by Krueger, 1988). Unless such rules can be identified, this "automatized procedures" perspective can not be considered as a serious model (see Ashcraft, 1985, for further discussion of this issue).

In contrast to the difficulties encountered by a procedural perspective, there is rapidly accumulating evidence which provides direct support for the associative network perspective. First, interference and priming effects, touchstones of associative retrieval processes, have been documented by many researchers across varying experimental contexts (e.g., Campbell, 1987a, b, &

c; Campbell, 1991; Campbell & Clark, 1989; Campbell and Graham, 1985; Koshdimer & Ashcraft, 1991; Miller & Paredes, 1990; Stazyk, Ashcraft, & Hamann, 1982). For example, several researchers (e.g., Campbell and Graham, 1985; Campbell, 1987a) have shown that around 80% of errors that adults make are multiples of one or both multiplicands (e.g., 32 is a multiple of 8, and is a common error to 3 x 8), suggesting interference among problems that share multiplicands. This and related interference and priming effects, to be discussed in more detail later, and will be the focus of many of the simulation results which will be presented in this paper. Second, in recent studies of brain damaged patients, Sokol et al. (1991) found deficits in addition and multiplication performance which appear to support rule-based processing only for 0 x N problems, and retrieval processes for all other problems. One of their patients was consistently wrong on all 0 x N problems (apparently using an alternative, incorrect rule, 0 x N = N), and another subject was consistently wrong on these problems early in practice, and then made a sudden transition to being consistently correct later during practice. Performance of both subjects, then, supports the notion that a single rule does underlie performance on these problems. For the remaining (M x N) problems, however, there was an apparently unsystematic variation in performance from problem to problem such that error percentages ranged from a low of 0% on some problems to 80% or higher on others. Further, in contrast to the 0 x N problems, performance on these problems improved gradually across practice, as would be expected under a network retrieval perspective. Also, Sokol et. al (1991) found that the errors made on M x N problems followed the same patterns as errors made by normal subjects (e.g., the error pattern discussed above). As the authors pointed out, these results are consistent with the idea that, for M x N problems, skilled performance reflects retrieval from an associative network, rather than the application of rules or procedures (i.e., if performance on M x N problems involved extensive application of rules, one would expect systematic patterns in performance deficits for all problems reflecting the disruption of rule-based knowledge, and a step function for any improvement with practice, reflecting reinstatement of the rule).

In sum, the evidence suggests that adult performance on basic multiplication problems primarily reflects retrieval processes, with rule-based processing taking place either in addition to, or instead of, retrieval for special classes of problems (e.g., 0 x N problems). Further, the evidence for priming and interference indicates that retrieval is not simply a process of accessing isolated problem-to-answer associations, but rather reflects associative network processes in which multiple problem and/or answer representations are activated during every retrieval attempt. Several theorists have developed models of this retrieval network (Ashcraft, 1987; Campbell & Graham, 1985; Siegler, 1988). Each of the models has some degree of consistency with the various empirical phenomena, and each incorporates insightful mechanisms which may prove essential in accounting for arithmetic learning and performance. In a recent review of this literature, however, McCloskey , Harley, and Sokol (1991) point out that each of these models falls short of accounting for all of the empirical phenomena, and in some cases they are not explicated in enough detail to allow strong predictions. They suggest that the focus should shift from the demonstration of the basic merit of the associative memory framework to development of a more detailed, fully elaborated model. In this paper, we present an interactive activation model of arithmetic fact retrieval which represents an effort in that direction. Before discussing the model in detail, we review the major empirical phenomena, and discuss existing models designed to account for these phenomena.

Empirical Phenomena and a Review of Existing Models

Basic Tasks and Empirical Phenomena

Three tasks have dominated experimental work on mental multiplication. In the production task, subjects are simply presented the problem (e.g., 3 x 7), and are asked to produce the answer. In a variation of this task, the primed production task, subjects are presented a candidate answer, which may be correct or incorrect, for a brief period (on the order of 200 ms), and are then presented the problem and asked to produce the answer. In the verification task, subjects are presented the problem along with a candidate answer (e.g., 9 x 4 = 32), and are asked to indicate whether the equation is true or false. In each of these tasks, subjects are typically informed that both accuracy and reaction time (RT) are important.

The problem size effect

The problem size effect refers to the robust RT advantage for small problems (e.g., 3 x 4) compared to large problems (e.g., 7 x 8), which has been reported in essentially all experimental investigation of mental multiplication Problem size has been indexed by several correlated variables including the magnitude of the largest operand, and the magnitude of the product. The effect has been found for children and college students (e.g., Campbell & Graham, 1985) and also for college students who have received extensive practice (Fendrich, Healy, & Bourne, in press), although the effect shrinks over practice sessions. Additionally, there is evidence that some special classes of problems, such a squares (e.g., 4 x 4) and 5 x N problems (e.g., 7 x 5) exhibit a problem size effect which is less pronounced than that for other problems (e.g., Siegler, 1988). An analogous problem size effect has also been reported for the verification task (e.g., Zbordoff & Logan, 1990).

The relatedness effect

Several researchers (e.g., Campbell & Graham, 1985; Sokol et al., 1991) have shown that, in both the production and primed production tasks, the frequency with which a given incorrect answer will be stated on a particular problem can be predicted by the relatedness category that answer falls into with respect to the problem. An example of each category of incorrect answers along this relatedness dimension is shown in Table 1 for the example problem 4 x 8.

The most frequent errors (typically 70 to 90% of all errors) involve table-related answers; that is, most errors are the correct answer to another problem in the same multiplication table . Thus, for example, 36 is table-related to 4 x 8, since 4 x 9 = 36 belongs to the same table as 4 x 8 = 32. A special case of table-related answers, which are double table-related, may be especially frequent. In contrast, few errors (typically 10-20%) are table-unrelated (belong to a different table than either of the multiplicands of the problem), and even fewer (less than 10%) are miscellaneous, i.e., not the correct answer to any multiplication problem.

A parallel to the basic relatedness effect in production has been found in both the primed production task and the verification task. Campbell (1987b; 1991) showed that, in the primed production task, RT to respond correctly is slowest if the prime is table-related to the problem, faster if the prime is table-unrelated, and fastest if the prime is correct. Also, the prime is stated as the answer (in error) much more frequently if it is table-related than if it is table-unrelated. Koshdimer and Ashcraft (1991) found analogous results using a verification task. When the candidate answer is table-related, RT to respond "false" is slower than when the candidate answer is table-unrelated, and both of these "false" responses are slower than "true" responses when the candidate answer is correct.

These relatedness effects have been shown to interact with problem size. In the primed production task, Campbell (1991) found that the effect of a table-related prime, relative to a table-unrelated prime, was greater for large problems than for small problems (as indexed by both RT and error frequency). Koshdimer and Ashcraft (1991) found the same problem size by relatedness interaction using the verification task.

The Operation Confusion Effect

Winkelman and Schmidt (1974, see also Miller & Paredes, 1990) reported that errors in both multiplication and addition are often the correct answer to the corresponding problem in the other operation (e.g., 4 + 7 = 28; 4 x 7 = 11). This effect has been found for both children and adults, in both mixed block experimental conditions (i.e., conditions in which multiplication and addition problems are randomly mixed and seen in sequence), and pure block conditions (i.e., conditions in which many problems of only one operation are solved in sequence). The fact that the effect occurs in pure block conditions suggests that it is not solely due to misreading of the symbol (x or +), or to subject hypotheses which anticipate the operation of the next problem, but rather reflects an overlap of addition and multiplication knowledge at the cognitive level. Results from the verification task (Zbordoff & Logan, 1987) converge on this interpretation: RT is slower and the error rate is higher for problems in which the candidate answer given is correct for the corresponding problem in the other operation (3 + 6 = 18), than when it is not (e.g., 3 + 6 = 20). Miller and Paredes (1990) reported that the operation confusion effect is asymmetrical, that is, subjects are significantly more likely to make an a multiplication response on an addition problem than they are to make an addition response on a multiplication problem.

The Split Effect

An additional variable which predicts the degree to which an incorrect answer influences performance is the split, or magnitude differential, between the incorrect answer and the correct answer. In the production task, several researchers (e.g., Miller et al., 1984; Campbell & Graham, 1985; Sokol et al., 1991) found that most answers which occur as errors to a given problem are close in magnitude to the correct answer (e.g., 63 is a more frequent error than 21 given the problem 8 x 9). In the verification task, RT is slower when the split is small than when it is large (e.g., Zbordoff & Logan, 1990).

Correlation Between RT and Error Rate

Several researchers (e.g. Miller et al., 1984; Campbell & Graham, 1985) have shown that RT correlates positively with error rate. Thus, the more time subjects take on average to answer a given problem correctly, the more likely they to err on that problem.

Transfer Effects

Extended practice on one subset of multiplication problems does not transfer positively to other problems (Campbell. 1987a). On the contrary, Campbell's (1987a) results indicate slight negative transfer of learning: Practice results in a slight increase in RT for the subset of problems on which the subjects are not practiced.

Fendrich et al. (in press, see also Rickard, 1992) investigated transfer across a reversal of operand order. Subjects were practiced on one operand order (e.g., 4 x 7), and then tested on these same problems, problems with reversed order (7 x 4), and new problems not seen during practice (e.g., 5 x 8). Results showed performance in the reversed condition to be substantially better than performance on new problems, indicating positive transfer, but also slightly poorer than performance on non-reversed problems.

General Models of Production and Verification

Both production and verification tasks have been used to study arithmetic under the assumption that each task has unique strengths in revealing the underlying cognitive process. The production task has the advantage that it is exactly the task typically faced in naturalistic settings. The verification task has the advantage of being methodologically simpler, because only two responses, true and false, are involved (see Campbell, 1987b, for a detailed discussion of the relative merits of the two tasks). Verification was originally assumed to involve two sequential cognitive processing stages: (1) production of an answer just as in the production task, and (2) comparison of the produced answer with the candidate answer provided. According to this production-comparision perspective, the characteristics of the candidate answer relative to the problem (e.g., the relatedness level) do not influence the production stage, although they can influence processing during the comparison stage (Ashcraft, 1987). More recently, however, evidence has been introduced supporting two alternative models. First, Campbell (1987b) hypothesized that the candidate answer does influence the production stage of processing, facilitating production when it is correct, and interfering with and thus slowing production when it is incorrect, with the degree of of interference determined by the degree of similarity (e.g., the relatedness level and the split) between the problems being solved and the candidate answer. To investigate this hypothesis, Campbell (1987b) compared performance in the primed production and verification tasks. He found that the prime in the primed production task does influence processing in the manner described above, and that performance in the verification task follows the same patterns. Thus, he concluded that the prime in the primed production task and the candidate answer in the verification task both influence the production process, and in analogous ways. Consistent with these findings, we will refer to Campbell's model as the primed production-comparision model of verification.

Zbordoff and Logan (1990) proposed a model of verification which is a more radical departure from the original production-comparison model. They suggested that subjects produce no answer on at least some trials, but rather evaluate the extent to which the presented equation resonates with memory, where resonance roughly reflects the overall level of activation in memory in response to the equation. We refer to this as the resonance account of verification performance. Zbordoff and Logan (1990) assume that subjects have a criterion against which they evaluate resonance: If the equation resonates more strongly than the criterion, it is accepted as true, if its resonance is below the criterion, it is rejected as false. Further, they hypothesize that RT is inversely proportional to distance from the criterion. True equations (e.g., 3 x 4 = 12), will resonate, and thus will be responded to relatively quickly and accurately. Conversely, false equations in which the candidate answer is not associated with the problem (e.g., 8 x 9 = 17) well resonate little and will thus be rejected quickly. On false equations in which the candidate answer is in some way partially associated with the problem (e.g., for 4 x 8 = 24, 24 is table-related by way of both multiplicands), however, the answer will be at least partially consistent with the arguments, and thus resonance will fall somewhere between the levels for the two cases described above, resulting in slower RT, and more frequent errors, as has been observed. Zbordoff and Logan (1990) present evidence that a resonance process is necessary to account for performance at least some of the time, across problems and subjects, but they also acknowledge that it is unclear from their data how frequently this approach is used relative to other possible approaches (e.g., production-comparison).

The evidence for resonance in verification indicates that the various tasks used to explore mental multiplication may make different cognitive demands on subjects. Zbordoff and Logan (1990) propose a framework for conceptualizing the similarities and differences among processes involved in performing these tasks. Their framework specifies representations, microprocesses and macroprocesses. The representation and microprocesses (in a network model, these correspond to the nodes in the model, the links between nodes, and the rules by which nodes send activation to one other) are assumed to be the same in all tasks. But the macroprocess, the process by which an answer is actually selected, is assumed to be different in the production/primed production and verification tasks. In the production and primed production tasks, a selection macroprocess selects an answer (e.g., the most activated answer) after the representation and microprocess have operated on the input. Under the resonance hypothesis of verification, the resonance macroprocess computes the overall resonance in the system after the representation and microprocess have operated, and then compares this to a criterion value as described above.

Models of Skilled Mental Multiplication

The first network models of skilled mental multiplication were the table-search models (e.g., Stazyk et al. 1982; Miller et al., 1984; Geary, Widaman, & Little, 1986). A central assumption of these models is that activation from each of the multiplicands spreads through an associative network organized similarly to multiplication tables. Activation spreads more quickly to small problems than to large problems, because small problems are closer to the origin of the table. This in turn yields the problem size effect in adult performance. Although these initial models clearly inspired many of the constructs included in later models, there are currently strong empirical arguments against the table-search mechanism (Siegler, 1988; Sokol, et al., 1991), and thus these models will not be reviewed in detail. We are most interested in more recent models, e.g., the associative network model (Ashcraft, 1987), the network-interference model (Campbell & Graham, 1985), and the distribution of associations model (Siegler, 1988). An additional review of these models as well as the table-search models can be found in McCloskey et al. (1991).

We will discuss each model within the framework of microprocesses, representations, and macroprocesses outlined by Zbordoff and Logan (1990). While none of the theorists use all of these terms in describing their models, it will be clear from the discussion that the various components of the models can be categorized quite naturally within this framework.

The Network Retrieval Model

Ashcraft's network retrieval model (1987) holds that associations directly linking operands and answers are formed during initial learning and are strengthened with practice (see Figure 1). In the model, there is a separate set of operand nodes for each of the two operands of a problem. Thus, representations of the operands are token rather than type (i.e., there is a token for each operand). Each of the pairs of nodes, one taken from each set, is connected to a unique answer representation. Thus, representations at the answer level are also token rather than type (for example, there is a separate 24 node for 4 x 6 and 3 x 8 , as shown in Figure 1). The microprocess involves the spreading of activation from each of the operands of a problem to all answers connected with the operands. The correct answer receives activation from both of the multiplicands, and thus typically reaches a higher activation level than the other answer nodes. Ashcraft further postulates that, after activation has propagated to the answers, activation will spread from the most activated answer node to other answer nodes in the same row or column of a multiplication table (i.e., table-related answers). More activation spreads to answers that are nearby in magnitude than to answers distant in magnitude. Ashcraft makes three assumptions about the selection macroprocess in the production task. First, an answer is only a candidate for selection if its activation exceeds some threshold. Second, the most highly activated node is selected. Third, the higher the winning node's activation above threshold, the more quickly it is selected. Two additional assumptions are made in modeling the verification task. First, the candidate answer is assumed to act as a prime, sending activation to its answer node. Second, a decision stage is assumed to operate which compares the selected answer to the candidate answer given in the problem. If the two are the same, decision time is a constant. If the two are different, then the decision time is inversely related to the difference in activation between the selected answer and the candidate answer.

The network retrieval model covers both the production and verification tasks, and can account (either as demonstrated directly through simulations or by way of strong logical argument) for several of the essential RT phenomena, including the problem-size effect in production, and the relatedness effect, the split effect and the operation confusion effect in verification (see McCloskey et al., 1991; Ashcraft, 1987). In contrast, the model cannot account directly for errors, because the correct answer is always the most active node. Although the relative activation patterns of incorrect answers do approximately parallel the empirical patterns of errors, there are likely to be some difficulties with extending the model to take advantage of the activation of incorrect answers to generate errors (see McCloskey et. al., 1991, for further discussion).

There are additional effects for which the model as currently formulated is unable to account. First, the model does not predict the relatedness and relatedness by problem size interactions in the primed production task (Campbell, 1991). Recall that interference from incorrect nodes occurs in the model at the decision stage, when the retrieved answer is compared to the candidate answer. Because there is no decision stage in the primed production task, no relatedness effects on RT are predicted. Second, the model cannot account for the correlation between RT and error rate, because it simply does not incorporate a mechanism for generating errors. Third, the resonance macroprocess in verification is outside the scope of the model. Fourth, the negative transfer effects (Campbell, 1987a) are not predicted by the model, and the model predicts no transfer of learning across operand order, a prediction which is clearly inconsistent with available data (Fendrich et al., in press, Rickard, 1992). On another point related to operand order, it is not clear whether different operand orders access the same or different answer representations. Because answer nodes are assumed to have token representations in the case where two problems have the same answer (e.g., 3 x 8 and 4 x 6 = 24), it is most consistent to allow for different answer representations for each operand order. While this assumption gives the model internal consistency, it also introduces an unparsimonious representational scheme in which up to four different representations exist for each answer (e.g., 3 x 8 = 24, 8 x 3 = 24, 4 x 6 = 24 and 6 x 4 = 24). It is not clear why a cognitive system would develop multiple, completely independent representations for one answer response. This problem is greatly increased if the model is extended to addition, because there are many problems corresponding to each addition answer (e.g., 7 is the answer to 2 + 5, 3 + 4, 4 + 3, 6 + 1, etc.).

<u>The Distribution of Associations Model</u>

Siegler (1988) proposed the distribution of associations model to account for (1) aspects of strategy use during initial learning, (2) direct retrieval by skilled performers, and (3) the transition from strategy use into retrieval. The retrieval component of the model specifies a unique problem-based representation for each problem (see Figure 2). On each attempt to solve a problem (by way of either direct retrieval or some non-retrieval strategy), an associative link forms or is strengthened between the problem representation and the representation for the answer stated, regardless of whether that answer is correct or incorrect. Thus, a distribution of associations forms for each problem, and the strength of the link to the correct answer relative to other answers is directly proportional to the number of times the correct answer has been produced relative to other answers. This simple scheme, considered in the context of children's strategy use during initial learning, provides an elegant mechanism for accounting for many of the empirical effects outlined above. For example, Siegler (1988) noted that, during initial stages of learning, children are often taught (or spontaneously engage in) an adding strategy (e.g., 4 x 7 is seven added to itself four times). The primary error he observed in children making use of this strategy was to add by too many or too few operand magnitudes (e.g, adding 7 to itself only three times to get 21 as an answer to 4 x 7). Siegler (1988) also observed that the errors generated by this strategy are usually close in magnitude to the correct answer. Thus, the incorrect answers that become a part of the distribution of associations for a given problem during initial learning are dominated by table-related answers and answers that are close in magnitude to the correct answer. When children begin retrieving answers directly from memory, these false answer associations compete with the correct answer for retrieval. Occasionally, one of the false answers wins the competition, leading to relatedness and split effects in error patterns.

Siegler (1988) developed a simulation in which he implements this retrieval theory. The simulation involves first establishing a distribution of associations for each problem by having a program execute algorithms based on children's strategies. To simulate children's errors, the program is designed to mis-calculate a specified percentage of the time. After the system is trained to a desired level of skill, its performance is tested in an attempt to reproduce children's performance in experimental contexts. Siegler (1988) makes several assumption about the macroprocesses that operate on the network in the test phase. First, the probability that a given answer will be retrieved on a given attempt is directly related to its associative strength relative to the total associative strength of all answers. Second, on each retrieval attempt, the child (or adult) sets a criterion for the associative strength a retrieved answer must have in order to be selected as a response. If the answer retrieved on a given attempt has an associative strength lower than the criterion set for that attempt, no answer is stated. Third, before the first retrieval is attempted for a given problem, a search length is set which limits the number of retrieval attempts that will take place before a back-up strategy (e.g., repeated adding, or a "sophisticated guessing" approach in which the criterion is set to zero before attempting retrieval) is employed.

As discussed above, the model provides a good account of the relatedness and split effects in production task error patterns, and also accounts for the problem size effect, and the correlation between RT and error rate in the production task (see Siegler, 1988). There are many additional effects, however, which the model accounts for only indirectly, or not at all. First consider the remaining effects in the production task. Siegler (1988) suggests that the model can account for operation confusion errors by assuming that the answer from the corresponding problem in addition is a part of the distribution of associations. However, he fails to specify the mechanism by which the addition answer becomes a part of the distribution. This stands in contrast to his explanations of the relatedness and magnitude effects, where he provides an elegant description of systematic strategy failures which bias the distribution of associations for each problem to include incorrect answer associations which give rise to these effects. Additional effects in the production task for which the model as currently formulated offers no account include inter-trial priming effects, negative transfer effects, and operand order transfer effects. With respect to operand order transfer, Siegler (1988) does not specify whether there is a separate problem node for each order, or one node which represents both orders. Clearly, the model's predictions concerning operand order transfer depend on this specification.

The model also has difficulty directly accounting for the various effects reflecting the influence of the primed answer in the primed production task and the candidate answer in the verification task. The model might be modified to account for these effects, but it is unclear whether such modifications could be made without adversely affecting the model's ability to account for other phenomena.

The Network-Interference Model

Over several papers (Campbell and Graham, 1985; Campbell, 1987 a, b, c; Graham, 1987), Campbell and Graham have postulated many types of representations and associative links, and it is thus difficult specify the exact architecture which they prefer. The best existing description of a network-interference model is, in our opinion, provided by McCloskey et al. (1991), and we will work from this description (see Figure 3). The model specifies element-to-answer associations as central to retrieval. Thus, each multiplicand activates its multiples. The authors also postulate problem-to-answer associations, which are necessary to circumvent two problems with the element-to-problem associations. First, because the answer representations are type representations, the element-to-answer component of the model is not able to distinguish correct answers from answers that are table-related by way of both multiplicands. For example, when presented with 4 x 8, the element-to-answer associations send two doses of activation to both the correct answer, 32, and the double-table-related answer, 24. Second, the element-to-problem associations give the same amount of partial activation to all table-related answers, regardless of how close these answers are to the correct answer. The empirical data, however, show a split effect, such that incorrect answers close in magnitude influence performance more than incorrect answers distant in magnitude. This effect was accounted for under the network retrieval model by assuming spreading activation among answer nodes, and under the distribution of associations model by assumptions about the characteristics of the distribution of associations. In the network interference model, the authors assume that problem representations access a magnitude representation that gives extra activation to answer nodes that are close in magnitude to the correct answer (see Figure 3). Finally, answer-to-answer associations are postulated, such that answers with overlapping digits have associative links (not shown in Figure 3).

This model is potentially able to account for most of the empirical phenomena which we have discussed. This potential derives, however, from the multiple classes of representations and associations that are assumed, in combination with the occasionally vague or inconsistent descriptions of important aspects of the model, such as the microprocesses that operate on the representations, and the macroprocesses by which answers are selected. These factors appear to render the model, as currently formulated, difficult or impossible to test definitively, and thus not a serious theoretical alternative (see McCloskey et. al, 1991 for a more detailed discussion). Nevertheless, at a broad qualitative level, the model contributes important ideas. For example, it is the only existing mode which assumes that interference caused by competition among similar representations directly influences RT (see Campbell and Graham, 1985; Campbell, 1987a). A strongly related idea is also central to our model.

An Interactive-Activation Model of Skilled Mental Arithmetic

General Theoretical Framework

As a basic theoretical approach to number processing and mental calculation we adopt the framework outlined by McCloskey, Caramazza and Basili (1985). They propose three functionally autonomous processing systems as shown in Figure 4. The number comprehension system is composed of format specific subsystems which transform stimulus items into abstract internal representations, providing input to the calculational system. The calculational system is in turn composed of two independent subsystems; a fact retrieval subsystem and an algorithmic, or procedural subsystem. Finally, the number production system translates internal representations of numbers (i.e., the answers) into an appropriate output code. The authors and several other researchers have generated support for this model from studies of brain damaged patients (e.g., Sokol et al., 1991). Broadly speaking, they have shown that brain damage can cause deficiencies in one of these processing systems, while leaving the others intact. For example, in support of the claim that the calculational system operates on abstract representation, Sokol et al. (1991) presented one of their subjects with multiplication problems in a variety of input and output formats and showed that the patterns of errors (i.e., the frequency of errors overall and at each relatedness level) were nearly identical for each input/output combination, suggesting that the same, abstract type of representation is being accessed regardless of the input or output format.

Our intent in this paper is to model the details of the retrieval sub-system (the system containing arithmetic facts) as proposed in outline form by McCloskey et al. (1985). We will therefore adopt the working assumption that all representations in the retrieval network are abstract.

An Interactive Activation Model

We assume the retrieval sub-system can be best described as consisting of representations, microprocesses and macroprocesses (Zbordoff & Logan, 1990). We will refer to the representations and microprocesses in combination as constituting the retrieval network. The macroprocesses are assumed to operate on this network according to the demands of each experimental task (production, primed production and verification).

Representations

A diagram of the retrieval network is shown in Figure 5. We start with the assumption that abstract representations of the elements of a problem serve as input to the retrieval network. For example, for the problem 4 x 7, the elements of the problem, 4, 7, and x, are represented abstractly as input to the network. Both operands are represented within a single pool of element nodes as shown. Note that just as the the operand nodes (e.g., 4 and 7) are assumed to represent numerical values abstractly, the x node abstractly represents the operation of multiplication. Thus, this node would be activated regardless of the symbol actually used in the problem (e.g., x or *), as long as it is known to the subject that the symbol specifies multiplication. The x node is included in the current model description and simulations for completeness. Because this node will always be active for multiplication, however, it plays a computationally negligible role in the reported simulations. As we will discuss later, though, representations for the various operations will be critical in future attempts to extend the model to account for multiple operations.

There are both theoretical and empirical motivations for specifying the elements of a problem as the input level of the network. First, representations of individual numbers predate the acquisition of multiplication skill, and they are central to the application of algorithms during initial learning. It is thus parsimonious to assume that these representations are also involved in retrieval processes at higher levels of skill. Second, as we have discussed, there is empirical support for the position that the representations in the retrieval network are abstract, that is, independent of stimulus modality and format. This position allows for optimal generality of the retrieval skill: When subjects are faced with a novel input modality or format, they first translate the problem into abstract representations and then proceed with retrieving the answer. Converting the novel problem context into an abstract representation almost surely involves parsing the problem into its elemental components, and then using the abstract representations of these components to retrieve the answer. Indeed, it is difficult to conceive of a processing system in which a problem presented in a novel stimulus modality/format could directly activate an abstract problem representation without first processing the problem at the level of the individual elements. Finally, there are several empirical phenomena in the arithmetic literature (e.g., operation confusion effects, relatedness effects) which are most easily explained by a model which includes element-based representations, and which assumes that activation spreads from each of these representations to associated problem and/or answer representations.

Each of the nodes at the element level has feed forward connections directly to problem nodes appropriate for that element, e.g., 4 is connected to the problem node (4, 7, x), but not to (5, 7, x). The strengths of these connections are shown in Figure 5. When a problem is encountered, all multiplication problems receive activation form the x node, table-related problem nodes receive activation form the x node and one of the two operands, and the correct problem representation receives activation from all three elements. Note that the complements of a problems (e.g., 4 x 7 and 7 x 4) feed into the same representations at both the element and problem levels. Thus, operand order information is not preserved at any level of the model.

The problem level is motivated primarily by the important role it plays in the model's overall ability to account for the various empirical phenomena, to be described later. In addition, there have been several successful models of associative memory processes in other domains (e.g., Laberge & Samuels, 1974; McClelland & Rumelhart, 1986) in which hierarchical part-to-whole representations have been specified. Thus, the current model, which has a hierarchical element-to-problem representational scheme, incorporates a relatively standard architecture for memory representation.

At the problem level there is complete mutual inhibition, that is, each problem node has an inhibitory, or negative, connection to each of the other problem nodes. As we will discuss in more detail later, retrieval is assumed to involve a process in which each node cycles through successive states of activation. On the first cycle, input from the element level reaches the problem level, and there is substantial competition among the problem nodes. As the system continues to iterate, however, the most activated problem node sends relatively more inhibition than other nodes, and this serves to strengthen that nodes activation advantage. This process results in the correct problem ultimately receiving the most activation under most conditions.

As explicated thus far, however, the system encounters computational difficulties with squares problems (e.g., 3 x 3). We will demonstrate this by way of example. First, consider the case in which the system is to solve the problem 3 x 4. Here, the 3 and 4 nodes, and the x node, are all activated by external input, in accordance with the microprocesses discussed in more detail below. Activation then propagates from these nodes to the problem level. After the first cycle, the (3, 4, x) node has received three doses of activation (one from the 3, one from the 4, and one from the x), and each of the other problem nodes connected to 3 or 4 has received two doses of activation (one from the x node, and one from either the 3 or the 4 node). As will be demonstrated in simulations, the (3, 4, x) node can quickly suppress activation to these other nodes, and in turn send activation to the correct answer node, 12. Consider an example, however, in which a square problem, 3 x 3, is presented to the system. In this case, two doses of activation (one dose from the 3 node, and one from the x node) reach the correct problem node after the first cycle. Two doses of activation, however, also reach other nodes to which the 3's are connected and thus there is a stalemate in the competition within the problem level. This problem can be addressed by including a separate "squares" node, as shown in Figure 3. When a squares problem is encountered, the squares node is assumed to become active, and in turn send activation to all squares problems. The correct squares node will then get three doses of activation, and will thus have the advantage over both the table-related problem nodes (each of which still receive two doses of activation), and the other squares problem nodes (which will also receive two doses, one from the squares node, and one from the "x" node). Thus, the squares nodes provide a mechanism by which to overcome the stalemate of activation at the problem level for squares problems.

Recently, there has been substantial accumulating evidence pointing toward such specialized representations for squares problems. It has long been known to arithmetic researchers that squares problems are solved more quickly and accurately than other problems, by both children and adults. Historically, an influential account of this has been that squares problems have been encountered more frequently than other problems, and thus are retrieved more efficiently. At least two recent empirical findings, however, are inconsistent with this account. First, Graham and Campbell (in press, as reported in Campbell & Oliphant, in press) taught subjects an artificial arithmetic task, which they termed "alphaplication", and found that squares problems were learned faster than other problems, even with relative frequency of exposure to the two types of problems equated. Similarly, Reder and Ritter (1992) trained college students on more complex multiplication problems (e.g., 17 * 34), and found that, with frequency of exposure equated, squares were committed to memory faster than non-squares. These findings are clearly inconsistent with a frequency account of the squares advantage, but rather point toward a more intrinsic uniqueness in the representations/processes (such as that provided by the squares node in this model) that underlie performance on these problems.

The connections between problem nodes and the answer nodes implement a distribution of associations for each problem analogous to that specified by Siegler (1988). For example, the problem node for (3, 4, x) in Figure 5 has a correct association to 12 (indicated by the solid line), and several incorrect associations to other answer (indicated by the dashed line). Recall that under the distribution of associations model (Siegler, 1988), the incorrect problem-to-answer associations reflect error associations formed by incorrect execution of algorithms during initial learning. Thus, the particular incorrect problem-to-answer associations that exist, and the strength of these associations, is assumed to be unique for each individual. Also, like Siegler (1988), we assume that most incorrect problem-answer associations are table-related, and close in magnitude to the correct answer. The connections between nodes at the problem and answer level are bi-directional: Activity flows from the problem nodes to the answer nodes, and also from the answer nodes to the problem nodes. The forward connection from the problem to the answer, however, is twice the strength of the connection from the answer back to the problem. This implements an assumption that associations form asymmetrically in the direction in which facts are typically retrieved, an assumption for which the paired associate learning literature provides some support (e.g., Battig & Koppenall, 1965; Schild & Battig, 1966). As with the problem level, there is complete mutual inhibition among nodes at the answer level.

Retrieval Microprocess

Retrieval follows an iterative cycling process similar to that of other interactive activation models (McClelland & Rumelhart, 1986). As already discussed, when element nodes are activated by external input, they in turn send activation to nodes at the problem level. Nodes at the problem level then send activation on to the answer level. Nodes at both the problem and answer levels then compete until, in most cases, a clear winner at each level emerges. Thus, for example, when 4 x 7 is encountered, the problem node for 4 x 7 and the answer node for 28 typically emerge as the winners. Activation ranges from a minimum of 0 to a maximum of 1.0. Input to each node on a given cycle is simply the sum of the excitatory and inhibitory input being received from other nodes, plus an excitatory self-activation, plus any external input applied to that node. The activation level of each node after each cycle is this summed input, subject to the restriction that activation remains within the range of 0 to 1.0. There are no biases for any nodes (i.e., all biases are set to zero). The output of each node on a given cycle is simply the activation level from that cycle. The cycling is synchronous, such that, on each cycle, the activation levels of all nodes are updated simultaneously.

Retrieval Macroprocesses

Following the framework of Zbordoff and Logan (1990), the representation and microprocess involved in retrieval are assumed to be the same for the production, primed production and verification tasks. Thus, the network described above and depicted in Figure 5 provides the basic memory structure which is tapped in each of the tasks. The external influences on the network, and the macroprocess by which a response is selected are, however, assumed to vary across these tasks. To simulate the production task, the "x" node, and the appropriate multiplicand nodes representing the input elements of the problem to be solved, are given external input of .05 throughout the entire cycling process. The macroprocess for selecting an answer is a mechanism which continually monitors the activation of the answer nodes. When one answer node reaches an activation criterion (in these simulations, an activation level of .8), and all other nodes are suppressed to below a second activation criterion (in these simulations, an activation level of .2), the macroprocess selects the most activated answer node as the response. The primed production task is modeled by giving external activation to the primed node and letting the system cycle 6 times (to simulate brief exposure to the prime). The external input for the prime is then removed, and the input nodes for "x" and for the appropriate operands are given external input as in the production task. The answer is selected by the same macroprocess as for the production task.

To model the primed production-comparison account of verification, we assume that, during the production stage, attention is focused on the problem statement (e.g., "4 x 8"), and the candidate answer is relatively unattended. This assumption is reasonable because it is the problem statement that provides the information needed for production of the answer. Despite this selected attentional allocation, however, we also assume that the candidate answer causes some activation of its corresponding answer representation throughout the cycling process, although this activation is significantly less than that for the elements of the problem statement. In the current simulations, we implemented these assumptions by providing external activation of .05 to the element nodes (the same as in the production task) and an external activation of .01 to the node corresponding to the candidate answer. The answer is then selected by the production macroprocess.

We also present simulations demonstrating that the model can perform verification based on a resonance macroprocess, similar to that described by Zbordoff and Logan (1990). The resonance macroprocess implies that attention is allocated relatively equally across both the problem statement and the candidate answer. Thus, in these simulations, the nodes for both the elements of the problem and the candidate answer receive external activation of .05 throughout the cycling process. We adopt the Zbordoff and Logan (1990) notion that resonance is calculated continuously throughout cycling, and that the resonance value on each cycle is compared to a criterion resonance. In other ways, our conceptualization of resonance differs importantly from that of Zbordoff and Logan (1991), who suggest that resonance reflects the overall level of activation in memory. Under their conceptualization, activation of multiple, mutually inconsistent nodes (e.g., multiple answer nodes), results in large positive resonance. Intuitively, however, this represents a state of confusion in the subject, i.e., a state of "non-resonance", which should be reflected in the resonance calculation. Thus, we equate resonance with harmony (Smolensky, 1986), which measures the degree of self-consistency of active nodes. Harmony will be high when only nodes which have positive connections to one another are active, and will be low when many nodes that have negative connections to one another (i.e., multiple problem or answer nodes) are active. Thus, the greater the harmony in the network, the greater the self-consistency of the network. The formula, adapted to the current model, is:

$$\text{Harmony} = \text{Sum}\,[\, a(i)\, a(j)\, w(ij) \,] \; + \; \text{Sum}\,[\, ext(i)\, a(i) \,],$$

Where the summation is taken across all i and j.

Here, a(i) and a(j) is the activation of nodes i and j, respectively, w(ij) is the strength of the weight connecting nodes i and j, and ext(i) is the external input to node i.

For all tasks, RT is modeled directly in these simulations by assuming that each cycle of processing corresponds to a fixed interval of processing time in the fact retrieval system. Thus, the total number of cycles that take place between the initial activation at the element level and the selection of a response by the macroprocess is assumed to be directly proportional to the network retrieval component of RT in experimental data. This characteristic of the model sets it apart from existing models, all of which assume RT is more or less determined by a macroprocess which selects an answer after the network has processed the input and generated a set of activated answers.

Errors are not modeled directly in these simulations. Following McClelland & Rumelhart (1986), however, the degree to which a given answer node competes for activation during cycling can be taken as a rough correlate of the probability that that error will occur. It should be possible to take advantage of these relative activation levels in future work to model errors explicitly; that is, it should be possible to extend the model such that it generates incorrect responses on some simulation runs. The manner in which this might be achieved will be considered in the Discussion.

Accounts of Empirical Phenomena and Simulations

In the preliminary simulations discussed below, we will compare performance of an idealized model with other versions of the model in which individual, theoretically motivated alterations have been made to simulate various phenomena one at a time. This idealized model is a special case of the system described above in which there are no incorrect problem-answer associations, and in which activation levels for all nodes are set to zero before a retrieval attempt is simulated. The strengths for each class of connection in the idealized model are shown in Figure 5.

This simulation approach has the advantage of demonstrating clearly the contribution of various mechanisms in the model to overall performance. It has the disadvantage that possible interactions among parameters are not elucidated. Thus, although the simulations reported below clearly demonstrate the potential of the model to account for various detailed effects, a future research priority will be to develop versions which more directly simulate a complete human retrieval network. We will consider this issue in more detail in the Discussion.

In these preliminary simulations, we will limit our discussion to non-squares problems. Although the special squares node discussed previously should allow the system to solve squares problems as well, this node has not yet been implemented.

Simulations of the Production Task

The Problem-size Effect. A primary factor assumed by many researchers (e.g., Ashcraft, 1987; Campbell & Graham, 1985) to underlie the problem-size effect is frequency: The more frequently a problem has been solved, the stronger the relevant associations become, and thus the more quickly the correct answer can be selected. This frequency mechanism can be incorporated quite naturally into our model. To demonstrate this with a simulation, we first arbitrarily defined all problems with correct products of 30 or smaller to be small problems, and all problems with products of 32 or larger to be large problems. To simulate lesser experience with large problems, we decremented the strengths of the each of the appropriate connections (the element-to-problem and problem-to-answer excitatory connections, as well as the all inhibitory connections emanating from the appropriate problem and answer nodes) by 10% compared to the values specified under the idealized model. After this manipulation, the number of cycles required to reach criterion (the cycle time) for small problems was 42 and for large problems was 51. Before this frequency manipulation, the cycle time was 43 for all problems.

A second account for the problem-size effect proposed by Siegler (1988) is that incorrect problem-to-answer associations in the distribution of associations for each problem interfere with and slow retrieval. The greater the number and strength of the incorrect associations, the greater the interference. In the next section, we will show that incorrect problem-to-answer associations influence cycle time required to reach criterion in this model.

The Relatedness Effect. To investigate the effect of element-to-problem spreading activation on the retrieval process, we presented the problem 4 x 8 as input to the system and observed activation as a function of cycle time at both the problem and answer levels. (We will use 4 x 8 for all discussions and demonstrations of simulations of the relatedness effect. All other problems produce essentially the same results across each of the manipulations which we will discuss.) Miscellaneous and table-unrelated problem nodes never reach a state of positive activation at any time during cycling. Also, while table-related and double table-related nodes do compete early during cycling, they never reach an activation level of more than .03, and they are soon suppressed by the correct problem node. At the answer level, no answer nodes other than the correct node ever reach an activation level above .01. The correct problem and answer nodes, however, soon reach an activation level of 1.0. In sum, then, element-to-problem spreading activation, by itself, plays a negligible role in the performance of the idealized model in the simulated production task. Inhibition at the problem level acts as a filter, or damping mechanism, which powerfully constrains the impact that spreading activation from elements to problem nodes alone can have on performance. As will be demonstrated in the following simulations, however, element-to-problem spreading activation often has a substantial influence by way of interactions with other factors. Thus, this component of the model can be seen as providing a greater degree of "interference potential" among table-related problems/answers. This limited role for element-to-problem spreading activation may be more psychologically plausible than a role in which spreading activation from elements directly to the answer is the primary mechanism leading to interference, as is the case in the network-interference model. In that model, and also the network retrieval model, there is always substantial activation of table-related answer nodes, even for an "idealized subject" who is highly practiced and has no incorrect associations. In contrast, the current model suggests a cognitive system which is more robustly resistant to these part-to-whole spreading activation effects. As we present additional simulations, it will become clear that it is this computational robustness allows the system to function reasonably well (and thus model human performance) even when factors are manipulated (e.g., incorrect problem-to-answer associations, priming) which yield greatly increased potential for interference.

Interestingly, the system has no difficulty suppressing answers that are table-related by way of both operands (i.e, double table-related answers), because each operand is connected to the double table-related answer by way of a unique problem node. (In the example problem, 4 x 8 is connected to the double table-related answer, 24, by way of the problem nodes 3, 8, x, and 4, 6, x.) Thus, neither of the double table-related problem nodes receives as much activation as the correct problem node (4, 8, x), allowing the correct problem node to suppresses these nodes. As we will demonstrate in later simulations, however, given that other factors are constant, double table-related problem/answer nodes do compete slightly more successfully than other table-related nodes, a phenomenon which also seems to be reflected in human performance (see Campbell and Graham, 1985).

To investigate the effect of incorrect problem-to-answer associations on production, we introduced one incorrect connection between problem and answer nodes at each relatedness level (miscellaneous, table-unrelated, table-related, and double table-related) for the problem 4 x 8 (the specific answers chosen for incorrect problem-answer associations at each relatedness level are the same as those shown in Table 1). Each incorrect association was given a connection strength of .025 from the problem to the answer (and a connection strength of .0125 from the answer back to the problem). Thus, the incorrect problem-answer connections were approximately 40% the strength of the correct problem-answer connections specified in the idealized model. Activation of answer nodes as a function of cycle under these conditions is shown in Figure 6. Note first that there is some competition exhibited by answers at each relatedness level. Also, the degree of competition increases across increasing levels of relatedness. This reflects an interaction between element-to-problem spreading activation, and the distribution of associations assumed for each problem, which is a central and unique characteristic of the model.

Figure 6 also shows that cycle time increases to 48 cycles given these incorrect problem-answer connections. Further simulations showed that cycle time increases monotonically when incorrect connections are introduced one at a time, regardless of the order in which they are introduced. For example, with no incorrect problem-answer connections, cycle time is 43. Cycle time increases to 45, 47, and 48 when table-unrelated, table-related, and double table-related incorrect connections are added, respectively. Adopting the assumption that the frequency and strength of incorrect problem-answer associations increases with problem size (see Siegler, 1988), this simulation demonstrates that one mechanism which can potentially contribute to the problem size effect under this model is the distribution of incorrect problem-answer associations associations.

Accounting for errors implicitly. Our account of errors is currently based on the assumption that the probability of a given error occurring for a given problem is proportional to the activation of that during the cycling process. From the simulations described above, it is clear that, all factors being equal, table-related and double table-related answers compete more effectively during cycling than do table-unrelated answers. Thus, by the criterion described above, the model correctly predicts more frequent table-related errors than table-unrelated errors. Table-unrelated answers and miscellaneous answers reach the same levels of activation during cycling, however, in contrast to the empirical findings that table-unrelated errors are more frequent than miscellaneous errors. If incorrect associations to miscellaneous answers are assumed to be less frequent than incorrect associations to table-unrelated answers, then the model could still be made to predict more frequent table-unrelated errors. Other researchers (McCloskey et al., 1991), however, have pointed out that Siegler's (1988) assumptions about the mechanisms underlying the formation of the distribution of associations actually predicts a reverse state of affairs -- more incorrect associations to miscellaneous than to table-unrelated problems. Other changes to the model that might give rise to more frequent table-unrelated errors include allowing the "x" node to send activation not only to all multiplication problem nodes, but also to all multiplication answer nodes, or stipulating that all correct multiplication answers are be primed simply by the subject's knowledge that he/she is performing a multiplication task. Both of these changes would result in more activation being sent to table-unrelated than to miscellaneous answer nodes, other factors being equal.

The model is also potentially consistent with the finding that most errors are answers close in magnitude to the correct answer. Under the distribution of associations model (Siegler, 1988), incorrect associations that are close to the correct answer in magnitude are stronger and more frequent than incorrect associations distant from the correct answer. Because we have demonstrated that the distribution of associations causes retrieval interference, and because retrieval interference is assumed to be correlated with error rate, this mechanism can provide a potential account of the split effect on errors in the production task. Thus, there is no clear need for an analog magnitude representation of the type proposed in the network interference model (Campbell & Graham, 1985) or for the answer-to-answer spreading activation as proposed in the network retrieval model (Ashcraft, 1987). These mechanisms may indeed contribute to the split effect, but at present they do not appear to be necessary to account for the data.

Transfer Effects. Campbell (1987) has presented empirical evidence of a slight negative transfer effect to unpracticed problems in multiplication. We simulated this effect by incrementing the connection strengths of half the problems (randomly selected) by 10% to simulate practice, and then comparing number of cycles required to reach criterion for the unpracticed (unincremented) problems before and after this manipulation. Before this manipulation, 43 cycles were required to reach criterion for all non-squares problems (i.e., the cycle time for the idealized model). After the manipulation, 38 cycles were required for the practiced problems, and 44 cycles were required for the unpracticed problems. Thus, there was slight (1 cycle) negative transfer to the unpracticed problems.

The model can also account for positive transfer across operand order found by Fendrich et al. (in press) and Rickard (1992). Quite simply, either operand order of a problem (e.g., 4 x 9 and 9 x 4) would activate the exact same nodes at each level of the model. Thus, any practice one one order would transfer completely to the other order. One potential problem with this account is that transfer across operand order, while substantial, is not complete: There is a slightly longer RT for the unpracticed order in the studies reported by Fendrich et al. (in press), and Rickard (1992). Recall, however, that the McCloskey et al. (1985) framework, within which we have motivated our model, assumed autonomous, sequential systems for comprehension, fact retrieval, and answer production. It is possible that the longer RT for the unpracticed order reflects longer comprehension time for that order rather than longer fact retrieval time. A recent experiment in our laboratory confirms this. Subjects were practiced extensively on one operand order of a given problem in one format (e.g., "six times nine equals"), and then tested on both operand orders in the practice format (e.g., "six times nine equals" and "nine times six equals"), as well as on both operand orders in a different format (e.g., "6 x 9 =" and "9 x 6 ="). There was substantial positive transfer of learning to problems in each of these conditions. Also, whereas there was clearly an advantage for the practiced operand order in the practiced format (replicating the earlier findings), there was no difference in performance on the two operand orders when presented in the new format. Changing to the new (unpracticed) format appears to have factored out the comprehension advantage for the practiced operand order, and, with this factored out, there was complete transfer across order, as predicted by our model.

Simulations of the Primed Production Task

Figure 7 shows simulated primed production for the problem 4 x 8, for cases where the prime is correct, table-unrelated, table-related, and double table-related. In these simulations, there are no incorrect problem-answer associations. First, note that when the correct answer is the prime, the selection criterion is reached in 38 cycles, compared to 43 cycles for the idealized model in the standard production task. Thus, priming the correct answer facilitates RT, as has been found with human subjects (e.g., Campbell, 1991). Also, the prime is more competitive across cycles when it is table-related than when it is not: The number of cycles needed to reach criterion for selecting an answer increases across the levels of relatedness (Figure 7 a, b, and c), again directly modeling the RT patterns in the experimental data (Campbell, 1991).

Figure 8 shows the same simulations for the case in which there is an association between the correct problem and the incorrect primed answer which is 50% of the strength of the association between the correct problem and the correct answer. Again, there is an increase in competition from the primed node across increasing levels of relatedness. Also, the time required to reach the selection criterion is greater at each relatedness level than in the preceding simulations in which there were no incorrect problem-answer connections. Indeed, for the two table-related prime conditions, selection is never reached. In these cases, element-to-problem spreading activation, combined with the incorrect problem-answer associations, produces sufficient interference to prevent the system from discriminating between the correct answer and the prime. (Note that, in the analogous production task simulation, Figure 6, the system was able to completely suppress these table-related answers in favor of the correct answer.) We assume for now that, when the selection criterion is not reached for any answer, subjects either relax the selection criterion to the point where a selection can be made, or they resort to a non-retrieval back-up strategy (see Siegler, 1988).

Figure 9 shows cycle time required to reach selection criterion in cases where either there is no incorrect connection between the correct problem node and the primed answer (based on the same simulations depicted in Figure 7), or there is an incorrect connection between these nodes which is 25% the strength of the correct problem-answer connection. The purpose here is to demonstrate the potential of the model to generate relatedness by problem size interactions (Campbell, 1991). In Figure 9, it is clear that the effect of relatedness of the prime on cycle time is greater when there is an incorrect problem-answer association to the primed answer. Recall that these incorrect associations are assumed to be more frequent and stronger for large problems. Hence, these simulations suggest that the model can account for the relatedness by problem-size interaction in the primed production task.

Simulations of the Verification Task

The production plus comparison strategy. Figure 10 shows simulated verification for the problem 4 x 8, for cases where the candidate answer is correct, table-unrelated, table-related, and double table-related. In these simulations, there are no incorrect problem-answer associations. The results mirror those for primed production. When the candidate answer is correct, cycle time to reach criterion is fastest, and cycle time increases across relatedness levels because of increasing interference. These results correspond to human performance patterns reported by Koshdimer and Ashcraft (1991). Figure 11 shows the same simulations for the case in which there is an incorrect association between the correct problem and the candidate answer which is 50% the strength of the connection between the correct problem and the correct answer. Again, the results mirror those for primed production. There is an increase in competition from the primed node across increasing levels of relatedness, and the time required to reach the selection criterion is greater at each relatedness level than in the verification simulations in which there are no incorrect problem-answer associations (see Figure 10). Figure 12 is analogous to Figure 9, suggesting that the model can produce the relatedness by problem size interaction in the verification task (Koshdimer & Ashcraft, 1991).

The Resonance Strategy. Table 2 shows resonance for the problem 4 x 8 at 0, 20, and 40 cycles, in cases where the candidate answer is correct, table-unrelated, or table-related, and where the incorrect problem-answer connection is either 0% and 40% the strength of the correct problem-answer connection. First, note that, in all cases, resonance increases across cycles. Second, at each level of cycles, the resonance is highest when the candidate answer is correct, and lowest when it is table-unrelated with 0% incorrect association strength. Resonance also steadily increases across both relatedness level (from table-unrelated to table-related), and strength of incorrect problem-answer connection. Thus, under the assumption that there is a criterion resonance for responding "true," the probability of responding "true" incorrectly based on resonance in memory is greater when the candidate answer is table-related and/or has an incorrect association to the problem representation. Likewise, correct "false" responses will be slowed when the candidate answer is table-related and/or has an association with the problem, because in this case resonance will be relatively close to the criterion for responding "true."

This simulation, while encouraging, should be viewed primarily as a concrete illustration of how the resonance process defined qualitatively by Zbordoff and Logan (1990) can be implemented quantitatively. Many details of how the process leads to the specific RT and error patterns which have been observed in human performance on the verification task are not yet well-specified. Clearly, further work is needed both to establish the validity of resonance as a process underlying human verification performance, and to identify in detail its functional characteristics.

Additional Phenomena Consistent with the Model

Although the preliminary simulations discussed above show that the model has the potential to account most of the primary empirical findings in the mental multiplication literature, there are several remaining effects which were not addressed. In this section, we will take each of these effects in turn, and discuss how our model can be made to account for them, in most cases with only slight elaboration.

The operation confusion effect. It should be straightforward to extend the current model to account for multiplication/addition confusion effects. First, note that the basic architecture of the model is general enough to be applicable to arithmetic operations other than multiplication. If the current model were expanded to include addition, then nodes representing the symbol for addition (e.g., +) would be needed to direct processing toward the appropriate subset of the system (i.e., either the addition network or the multiplication network). For example, when 4 x 7 is presented to the system, 4 and 7 would attempt to activate all addition and multiplication problem nodes of which they are elements. The x node, however, would provide an excitatory bias to all multiplication problems, and perhaps also an inhibitory bias to all addition problems. The system would then be able to settle on the correct answer, but would experience significant competition from the 11 node (the answer to 4 + 7).

The finding that children's performance on addition problems gets worse temporarily after they learn multiplication (Miller & Paredes, 1990) could also be accounted for, in the following way. As children develop a retrieval network in a given arithmetic domain, they presumably are building appropriate excitatory links among representations within the domain and also learning to suppress (by way of inhibitory connections) answers which correspond to other domains with which they are already familiar. When children learn multiplication, then, they are both creating the multiplication network, and learning to suppress competition from the addition network. When they learned addition, however, there was no multiplication network to suppress. Thus, immediately after learning multiplication, children are relatively unable to suppress multiplication from interfering with addition (i.e., there is relatively little inhibition at the problem or answer levels from addition representations to multiplication representations), and this leads to a temporary increase in RT for addition problems. The argument that the current model would show such an increase in RT (i.e., cycle time) is evident from the negative transfer effects within multiplication which are simulated above: The impact that newly learned multiplication would have on addition performance is analogous to the demonstrated impact that practiced multiplication problems have on unpracticed multiplication problems. With practice, the addition network could adjust to inhibit appropriately the multiplication network, thus improving performance on addition problems. In sum, the reorganization of addition knowledge which researchers (e.g., Miller & Paredes, 1990) have suggested is necessary after learning multiplication may reflect simply the development of inhibitory links from addition to multiplication which slowly takes place after learning multiplication.

This type of cross-operation interference would not be limited under the model to initial learning by children. Given the findings that performance of college students improves considerably with practice (Campbell, 1987a, Rickard, 1992, Fendrich, et al., in press), then practice on one operation (say, multiplication) should at least temporarily worsen performance on other operations (e.g., addition). This possibility would lend itself straight-forwardly to empirical investigation.

Relations Between RT and Error Rate. As we have shown, increases in interference (e.g., placement of incorrect answer-to-problem associations) result in an increase in cycle time. We assume that the degree of interference provided by an incorrect answer during cycling is directly related to the probability of that answer occurring as an error. Thus, the model clearly has the potential to account for the strong positive correlation between RT and error rate.

The model also provides a natural account of speed-accuracy tradeoffs which are very likely to be characteristic of arithmetic performance (although we are not aware of studies which explore this). Recall that incorrect problem/answer nodes compete more strongly early during cycling than late during cycling. A basic assumption that we can make about the selection macroprocess is that the high and low activation thresholds required for response can vary, reflecting task demands. Thus, if the importance of speed is manipulated as a variable, we assume that the subject uses a relatively strict criterion for low speed conditions and a relaxed criterion for high speed conditions. When the criterion is relaxed for the high speed condition, the RT decreases, but the error rate should increase (due to the greater level of competition among answers during earlier cycling). It is not clear that any of the other existing models have this potential to account for speed-accuracy tradeoffs. As an aside, note that the existence of speed-accuracy tradeoffs of this sort are not inconsistent with positive correlations between RT and error rate: At any given point on the speed-accuracy trade-off continuum, an RT-error rate correlation across problems could exist.

Modeling Errors Explicitly

In the current simulations, errors in the production and primed production task are modeled only implicitly; that is, the probability of a given error occurring for a given problem is assumed to be proportional to the level of activation achieved by that answer during the retrieval process. In future simulations, however, it should be possible to model errors explicitly; that is, it should be possible to get the model to actually generate incorrect responses occasionally. Each human error can be intuitively classified into one of three unique cases, and each of these cases corresponds to a unique type of error that the simulation can be made to generate. First, subjects may simply have learned the wrong answer to one or more problems. When this type of error occurs, the subject literally believes he/she has answered correctly. Although this type of error is likely to be infrequent for educated adults solving single digit operand problems, it may not be uncommon for children who are just beginning to rely on network retrieval. In the current system, this type of error could be modeled simply by assuming the connection strength to an incorrect answer is significantly stronger than the connection strength to the correct answer, such that the system reaches criterion for that incorrect answer. The second type of subject error results when network retrieval fails and an error is made in the use of a back-up strategy. Siegler (1988) has shown that children are more likely to make errors when they use a back-up strategy (e.g., repeated addition) than when they use retrieval. The results of Bourne and Rickard (1991) suggests that adults also use non-retrieval strategies on some trials, and errors appear to be much more likely on these trials than on retrieval trials. Thus, failed retrievals may underlie a significant proportion of errors. In the current model, failed retrieval corresponds to any retrieval attempt in which the system fails to reach criterion for selecting an answer. When retrieval fails, we assume subjects employ more error prone back-up strategies. The third type of error, which likely accounts for many adult errors in experimental contexts, is retrieval of an incorrect answer even when the subject "knows" the correct answer. This type of error could occur as a consequence of several factors operating simultaneously. First, note that in most mental multiplication experiments, subjects are given instructions for speed. In order to meet this demand, they may lower their criterion considerably. Changes in response criterion (which may also vary from trial to trial and from subject to subject), combined with interference producing factors such as element-to-problem spreading activation, incorrect problem-to-answer associations, priming in the primed production task, and random noise in the system would potentially result in the network occasionally having a high enough activation of an incorrect answer relative to all other answers for the incorrect answer to be selected.

Discussion

While the arguments set forth in the preceding section show that the current model is competitive with existing models in terms of its ability to account for much of the existing data, they do not speak directly to issues of necessity of the various mechanisms assumed in the model. In this section, we will briefly review several fundamental assumptions, explore possible theoretical alternatives, and outline some experimental approaches which may differentiate among these possibilities.

Representations

By adopting the general framework of number processing and arithmetic proposed by McCloskey et al. (1985), we have limited our model to abstract number representations which are independent of the perceptual characteristics of the problem. There are alternative views (e.g., Campbell & Clark, 1988; Gonzalez & Kolers, 1982) in which more specific (e.g., perceptually-based) representations are involved directly in the retrieval process (cf. Sokol, Goodman-Schulman & McCloskey, 1988). These alternative views may be correct under at least some circumstances. For example, with extensive practice on a particular problem format, a perceptually-based representation of that format may develop and in turn form direct connections to the answer which bypass the abstract network machinery. Such a process may reflect a transition toward automaticity similar to that proposed by LaBerge and Samuels (1974) for reading. There is already evidence that format specific perceptually-based representations do form with practice (e.g., Fendrich et al., in press), and it is not improbable that these representations also form associations directly with answer representations. It remains a central issue to explore these possibilities.

In the current model, we have postulated a network in which element representations are linked to problem representations which are in turn linked to answer representations, and we have presented both empirical and theoretical arguments in support of this architecture. The architecture, however, represents an increase in complexity over models which assume only element-based representations (the network retrieval model) or problem-based representations (the distribution of associations model) at the input of the network. Further evidence uniquely supporting our slightly more elaborate architecture is needed. A unique prediction of our architecture, which should be testable, is that the severity of the relatedness effect is a function of both the degree of element-to-problem spreading activation, and the extent to which the distribution of associations includes incorrect associations to table-related answers (see Siegler, 1988). Experimental work is needed to provide direct evidence for each of these mechanisms independently, and to provide support for their combined role in producing interference.

Microprocesses

A fundamental difference between this model and other models is the microprocess assumed. In other models, there is simply a one time propagation of activation from the input representations (element or problem representations) to the output representations (the answers). In contrast, according to the current model, highly parallel activation of alternative "hypotheses" compete in an iterative manner until, in most cases, a coherent state of activation is reached in which only one pair of problem and answer representations is highly active, with all other representations suppressed. This microprocess has the advantage of accounting for RT, and speed-accuracy tradeoffs, directly as a function of the retrieval microprocess, rather than stipulating them as a function of the macroprocess. The approach used in the other models, however, has the advantage of greater simplicity. These alternative microprocesses represent a central topic in need of further investigation.

Macroprocesses

The selection macroprocess specified in the model for the production and primed production tasks is a discrimination process in which an answer is chosen only if that answer has relatively high activation and all other answers have relatively low activation. It might prove difficult to provide direct experimental support for this macroprocess relative to others that have been outlined for these tasks. Rather, assumptions about macroprocess are most likely to be supported indirectly by a model's demonstrated ability to account for various performance phenomena. Because the current model provides arguably the most complete account of known phenomena, we argue that the macroprocess assumed here is most viable at present.

The production-comparison and resonance hypotheses for verification may, on the other hand, be more open to experimental comparison. Verbal report methodologies (e.g., Ericsson & Simon, 1984) should be of value here. Specifically, a production-comparison strategy by definition involves two sequential stages (production and comparison), and the subject employing this approach at some level surely is aware of performing these stages. In contrast, a resonance approach would involve only a single evaluation of resonance. Thus, subjects using a production-comparison approach should report some consciously mediated processes which reflect the two stages, whereas subjects using resonance should report few if any thoughts, but rather should report simply knowing the answer (true or false).

Summary

We have described a model of skilled arithmetic performance which constitutes an improvement over existing models in that it is more explicitly defined, more complete and self-consistent theoretically, and better able to account for various empirical phenomena. Perhaps the most important contribution of the model is that it begins to synthesize theoretical perspectives and issues which have previously been considered independently (i.e., general models of number processing and arithmetic, the various network models of arithmetic fact retrieval, the resonance hypothesis for verification, interactive activation models). It is our hope that this effort will provide a theoretical foundation which will motivate research leading to a fully integrated and precise theory. Toward that end, we have outlined several important remaining theoretical issues and proposed future directions for extending the current model.

Acknowledgements

This research was supported by the the United States Army Research Institute Contract Number MDA903-90-K-0066 to the Institute of Cognitive Science at the University of Colorado. We gratefully acknowledge Alice Healy for her helpful comments on an earlier version of this manuscript.

References

Ashcraft, M. H. (1985). Is it farfetched that some of us remember our arithmetic facts? Journal for Research in Mathematics Education, 16, 99-105.

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

Baroody, A. J. (1985). Mastery of basic number combinations: Internalization of relationships or facts? Journal for Research in Mathematics Education, 16, 83-98.

Battig, W. F. & Koppenall, R. J. (1965). Associative asymmetry in S-R recall of double-function lists. Psychological Reports, 16, 287 - 293.

Bourne, L. E., Jr. & Rickard T. C. (1991). Mental calculation: The development of a cognitive skill. Inter-American Congress of Psychology. San Jose, Costa Rica, July, 1991.

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

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 & Cognition, 15, 349-364.

Campbell, J. I. D. (1987c). The role of associative interference in learning and retrieving arithmetic facts. In J. Sloboda & D. Rogers (Eds.), Cognitive processes in mathematics (pp. 107-122). London: Oxford University Press.

Campbell, J. I. D. (1991). Conditions of error priming in number fact retrieval. Memory & 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. (in press). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In Campbell, J. I. D. (editor), The nature and origins of mathematical skills (pp. #-#). North Holland: Elsevier.

Ericsson, K. A., & Simon H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT press.

Fendrich, D. W., Healy, A. F., & Bourne, L. E. Jr. (in press). Mental arithmetic: Training and retention of multiplication skill. In Izawa (Ed). Applied cognitive psychology: Applications of cognitive theories and concepts. New York: Erlbaum.

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.

Gonzalez, E. G., Kolers, P. A. (1982). Mental manipulation of arithmetic symbols. Journal of Experimental Psychology: Learning, Memory, and cognition, 8, 308-319.

Graham, D. J. (1987). An associative retrieval model of arithmetic memory: How children learn to multiply. In J. Sloboda & D. Rogers (Eds.), Cognitive processes in mathematics (pp.123-141). London: Oxford University Press.

Graham, D. J., & Campbell, J. I. D. (in press). Network-interference and number-fact retrieval: evidence from children's alphaplication. Canadian Journal of Psychology.

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

Krueger, L. E., (1986). Why 2 X 2 = 5 looks so wrong: On the odd-even rule in product verification. Memory & Cognition, 14, 141-149.

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

McClelland, J. L., &Rumelhart D. E. (1986). Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1). Cambridge, MA: MIT press.

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 brain-damaged subjects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 377-397.

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

Miller, K. F., & Paredes, D. R. (1990). Starting to add worse: Effects of learning to multiply on children's addition. Cognition, 37, 213-242.

Reder, Lynne M., & Ritter, Frank E. (1992). What determines feeling of knowing? Familiarity with question terms, not with the answer. Journal of Experimental Psychology: Learning, Memory and Cognition, 18, 435 - 451.

Rickard, T. C. (1992). Acquisition and transfer of mental arithmetic skill. Unpublished Masters thesis. University of Colorado, Boulder, Colorado.

Rickard, T. C., & Bourne L. E. Jr. (1990). Network interference processes in mental multiplication and division skills. Paper presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, LS.

Schild, Mary E., & Battig, William F. (1966). Directionality in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 5, 42 - 49.

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

Smolensky, Paul (1986). Information processing in dynamical systems: Foundations of harmony theory. In J. McClelland & D. Rumelhart (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition. (Vol 1.). Cambridge, MA: MIT press.

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: Evidence from the performance of brain-damaged patients. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 355-376.

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

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

Zbordoff, 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.

Table 1

An Example of Each Relatedness Level for the Problem, 4 x 8

Problem	Incorrect Problem/Answer		Relatedness Level
4 x 8	–	37	Miscellaneous
4 x 8	5 x 7	35	Table-Unrelated
4 x 8	4 x 9	36	Table-Related
4 x 8	3 x 8 4 x 6	24 24	Double Table-Related

Table 2

Resonance as a function of relatedness level, strength of incorrect problem answer association, and number of cycles

Relatedness Level	Strength of incorrect problem-answer association: .0 (0%) — No. of Cycles 0	20	40	.025 (40%) — No. of Cycles 0	20	40
Correct	.00	.26	.40	--	--	--
Table-Unrelated	.00	.17	.28	.00	.20	.34
Table-Related	.00	.19	.32	.00	.22	.36

Figure Captions

Figure 1. Part of the network postulated by Ashcraft's (1987) network retrieval model.

Figure 2. Examples of the distribution of associations postulated by Siegler's (1988) distribution of associations model.

Figure 3. Part of the network postulated by Campbell and Graham's (1985) network interference model.

Figure 4. A depiction of the McCloskey, Caramazza, and Basili (1985) model of number processing specialized to demonstrate the model's account of arithmetic.

Figure 5. Part of the interactive activation model of arithmetic fact retrieval.

Figure 6. Activation of nodes at the answer level as a function of cycle time for the problem 4 x 8 (production task). Relatedness level is a parameter.

Figure 7. Activation of the correct and primed nodes at the answer level as a function of cycle time for the problem 4 x 8 (primed production task), for the case of no incorrect problem-to-primed answer associations. Figure 7 A, correct prime; Figure 7 B, table-unrelated prime (35); Figure 7 C, table-related prime (28); Figure 7 D, double table-related prime (24). The dashed lines show cycle time required for the correct answer to reach the activation level (.8) required for selection by the macroprocess.

Figure 8. Activation of the correct and primed nodes at the answer level as a function of cycle time for the problem 4 x 8 (primed production, for the case of incorrect problem-to-primed answer associations of 40% the strength of the correct problem-to-answer association. Figure 7 A, correct prime; Figure 7 B, table-unrelated prime; Figure 7 C, table-related prime; Figure 7 D, double table-related prime (24). Dashed lines show cycle time required for the correct answer to reach the activation level (.8) required for selection by the macroprocess (in cases where this occurs).

Figure 9. Cycle time for the correct answer to reach criterion for selection by the macroprocesses for the problem 4 x 8 (primed production task), as a function of strength of the incorrect problem-to-primed answer association (either 0% or 20% the strength of the correct problem-to-answer association). Relatedness level is a parameter.

Figure 10. Activation of the correct and candidate answer nodes as a function of cycle time for the problem 4 x 8 (verification task), for the case of no incorrect problem-to-candidate answer associations. Figure 7 A, correct candidate; Figure 7 B, table-unrelated candidate (35); Figure 7 C, table-related candidate (28); Figure 7 D, double table-related candidate (24). The dashed lines show cycle time required for the correct answer to reach the activation level (.8) required for selection by the macroprocess.

Figure 11. Activation of the correct and candidate answer nodes as a function of cycle time for the problem 4 x 8 (verification task), for the case of incorrect problem-to-candidate answer associations

of 40% the strength of the correct problem-to-answer association. Figure 7 A, correct candidate; Figure 7 B, table-unrelated candidate (35); Figure 7 C, table-related candidate (28); Figure 7 D, double table-related candidate (24). Dashed lines show cycle time required for the correct answer to reach the activation level (.8) required for selection by the macroprocess (in cases where this occurs).

Figure 12. Cycle time for the correct answer to reach criterion for selection by the macroprocess for the problem 4 x 8 (verification task), as a function of strength of the incorrect problem-to-candidate answer association (either 0% or 20% the strength of the correct problem-to-answer association). Relatedness level is a parameter.

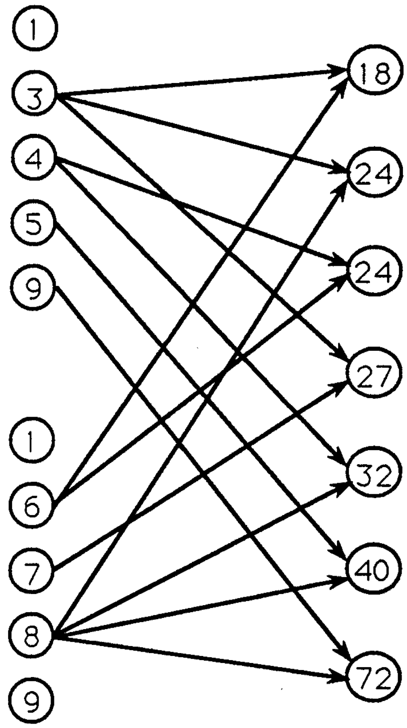


Figure 1

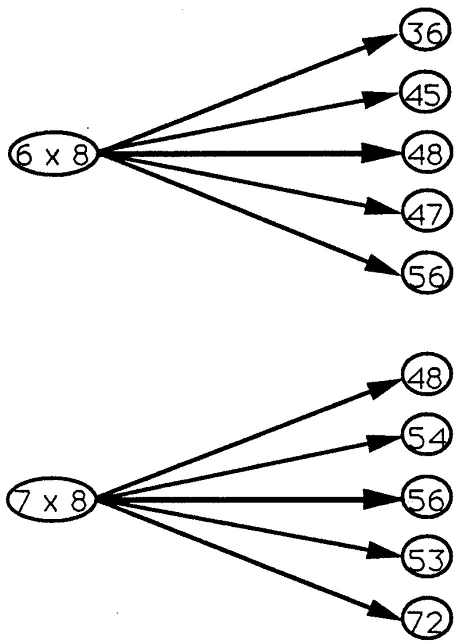


Figure 2

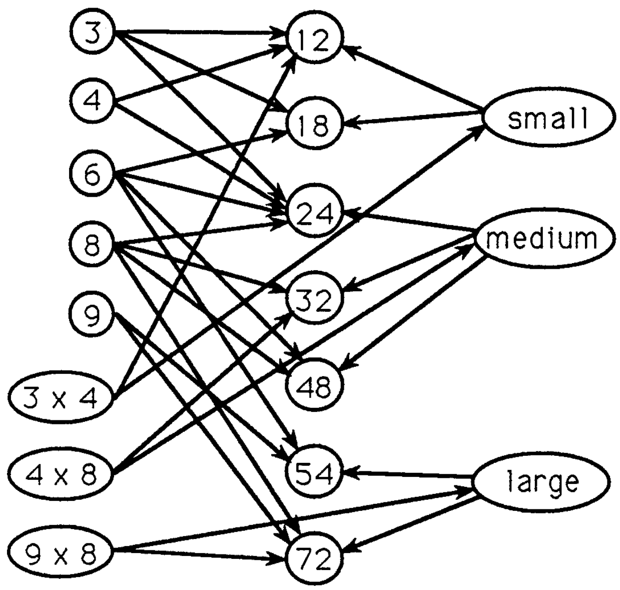


Figure 3

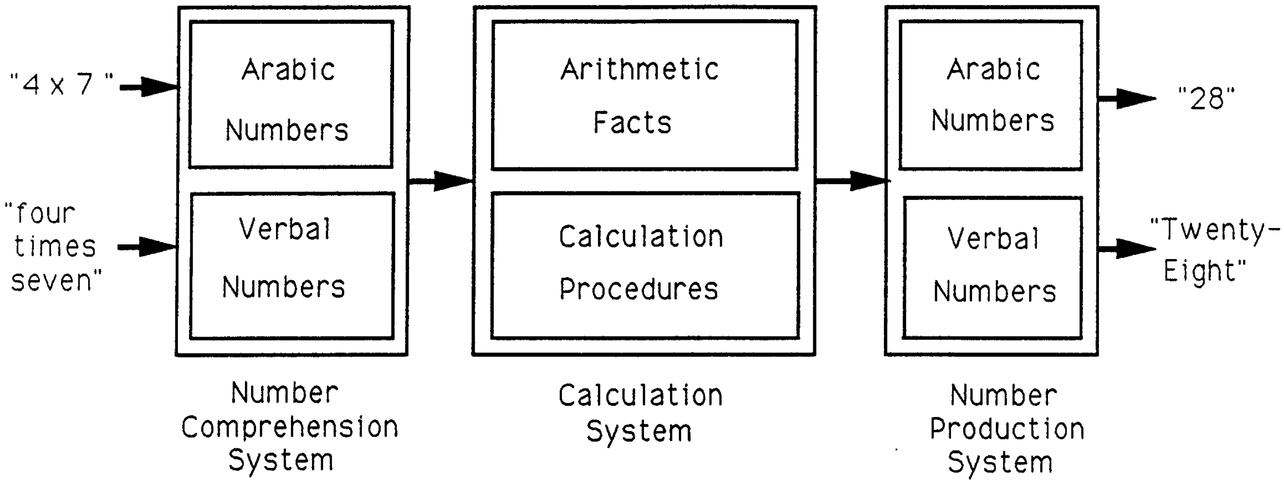

"4 x 7"
"four times seven"
Arabic Numbers
Verbal Numbers
Number Comprehension System
Arithmetic Facts
Calculation Procedures
Calculation System
Arabic Numbers
Verbal Numbers
Number Production System
"28"
"Twenty-Eight"

Figure 4

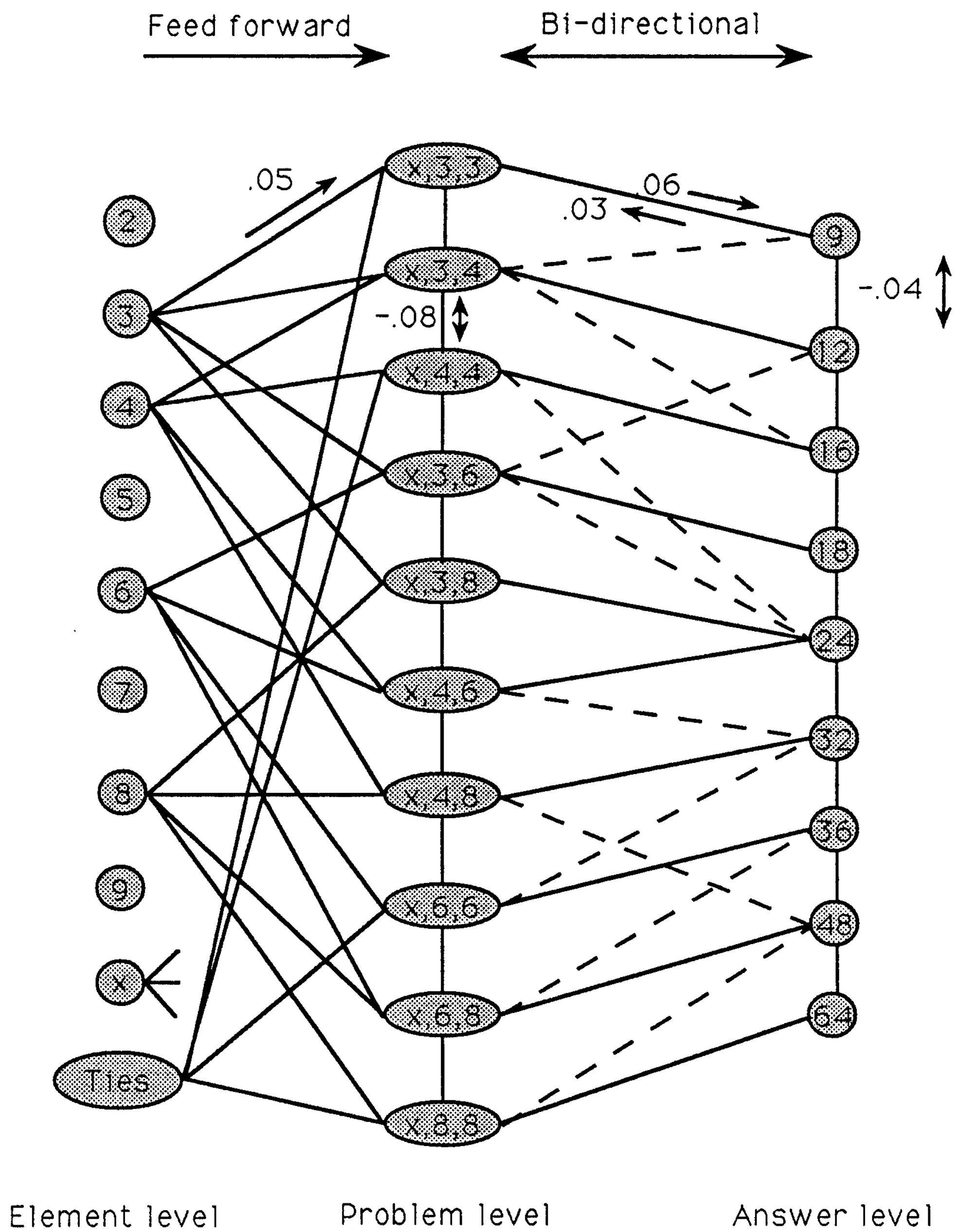
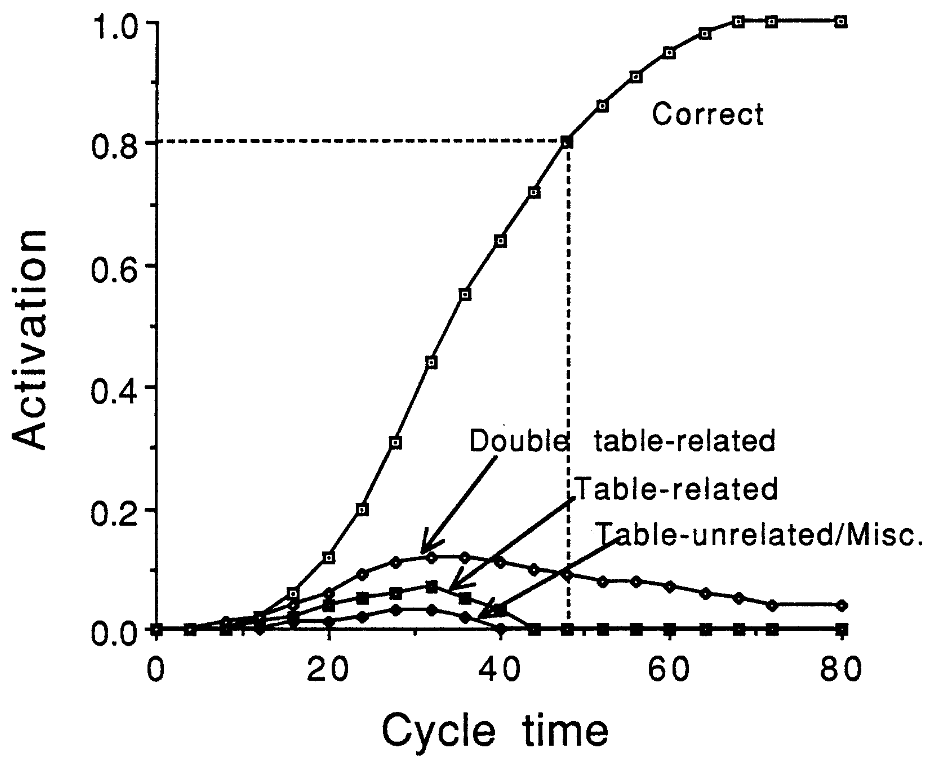
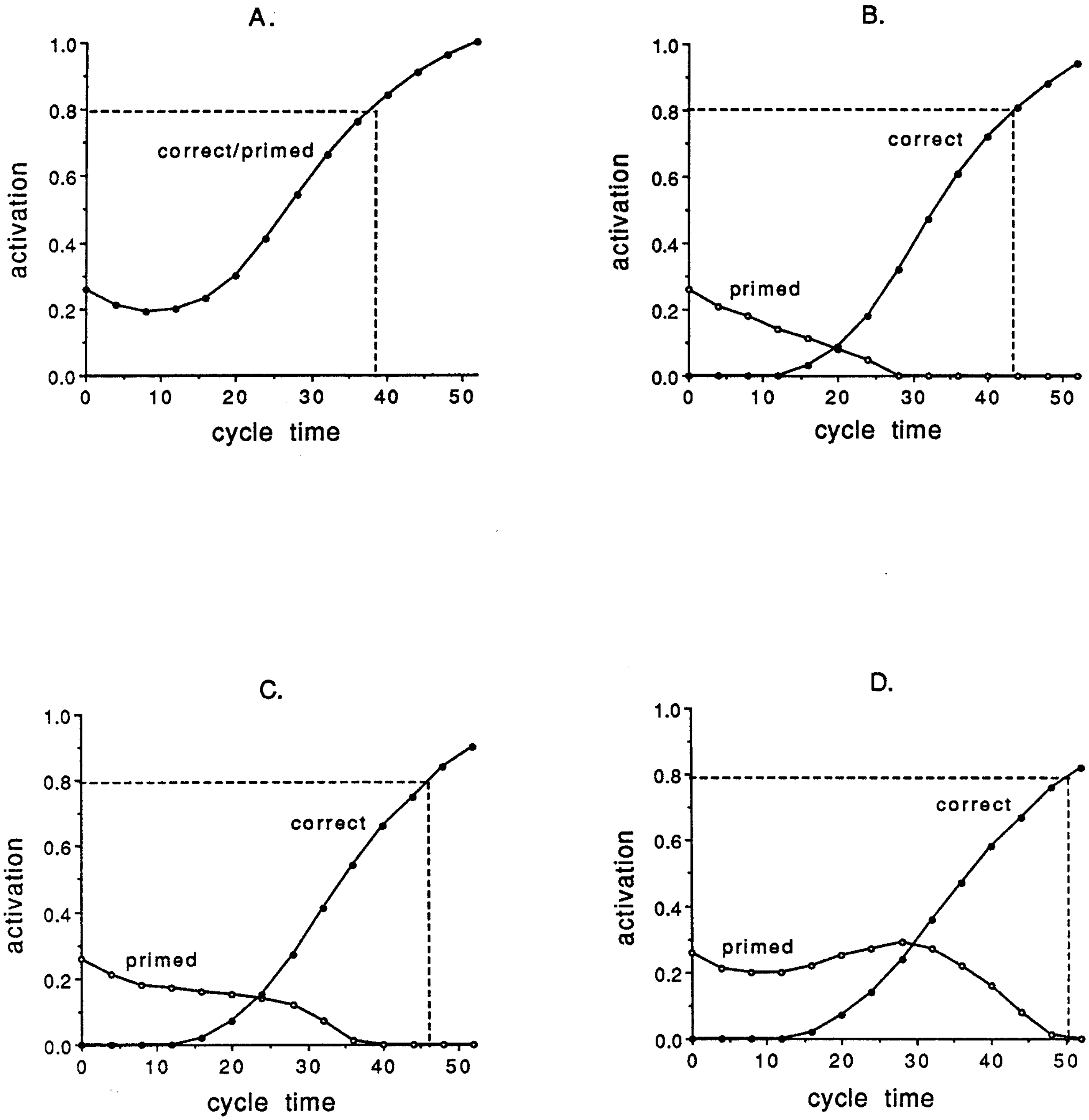
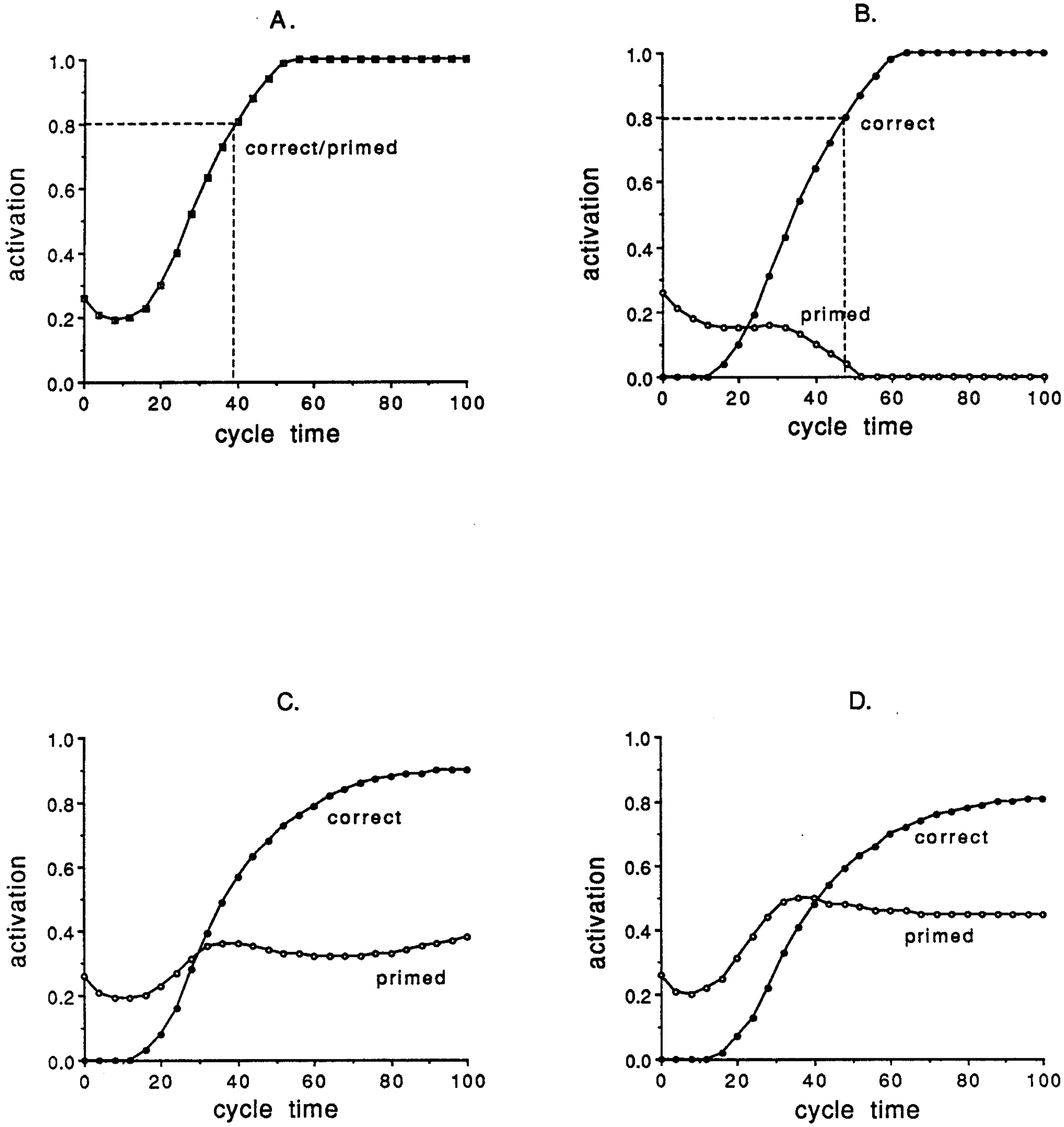
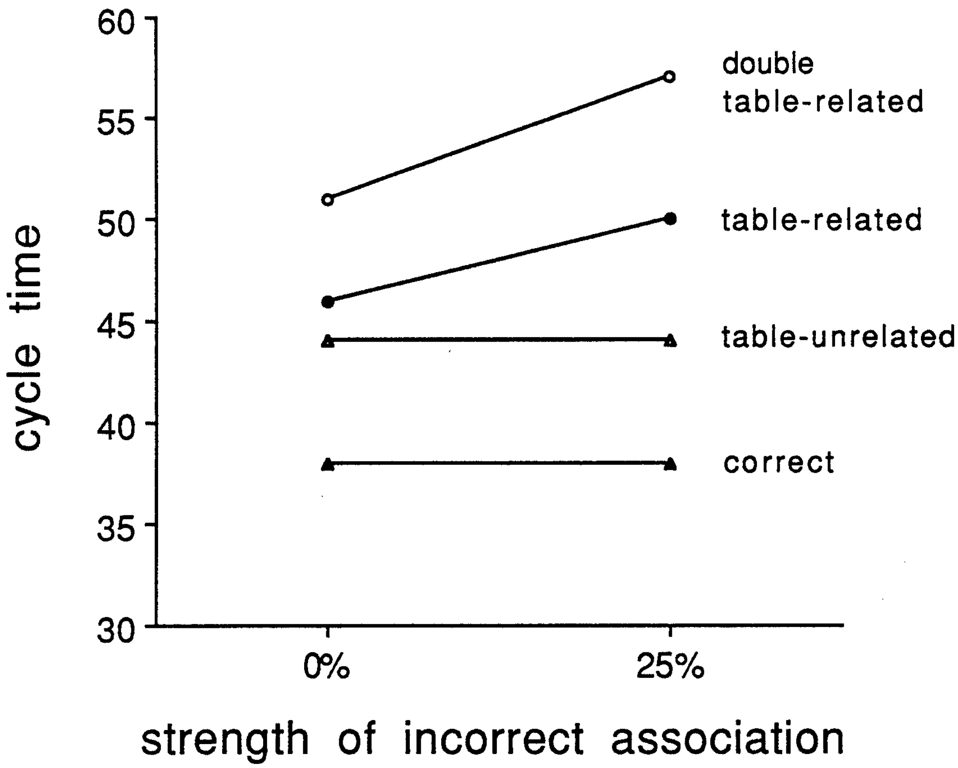

Feed forward
Bi-directional
.05
.06
.03
-.08
-.04
2
3
4
5
6
7
8
9
x
Ties
x,3,3
x,3,4
x,4,4
x,3,6
x,3,8
x,4,6
x,4,8
x,6,6
x,6,8
x,8,8
9
12
16
18
24
32
36
48
64
Element level
Problem level
Answer level

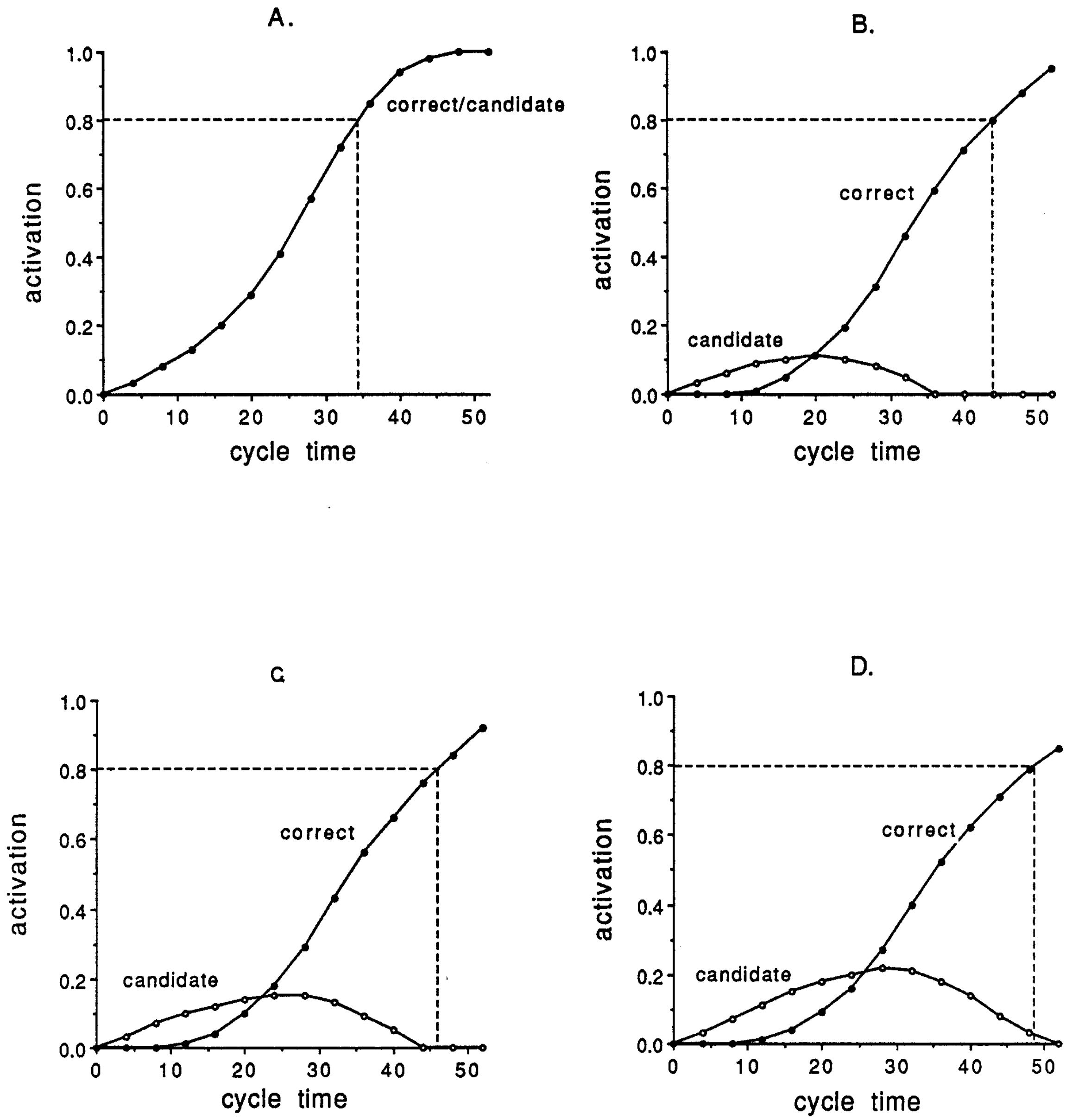
Figure 5

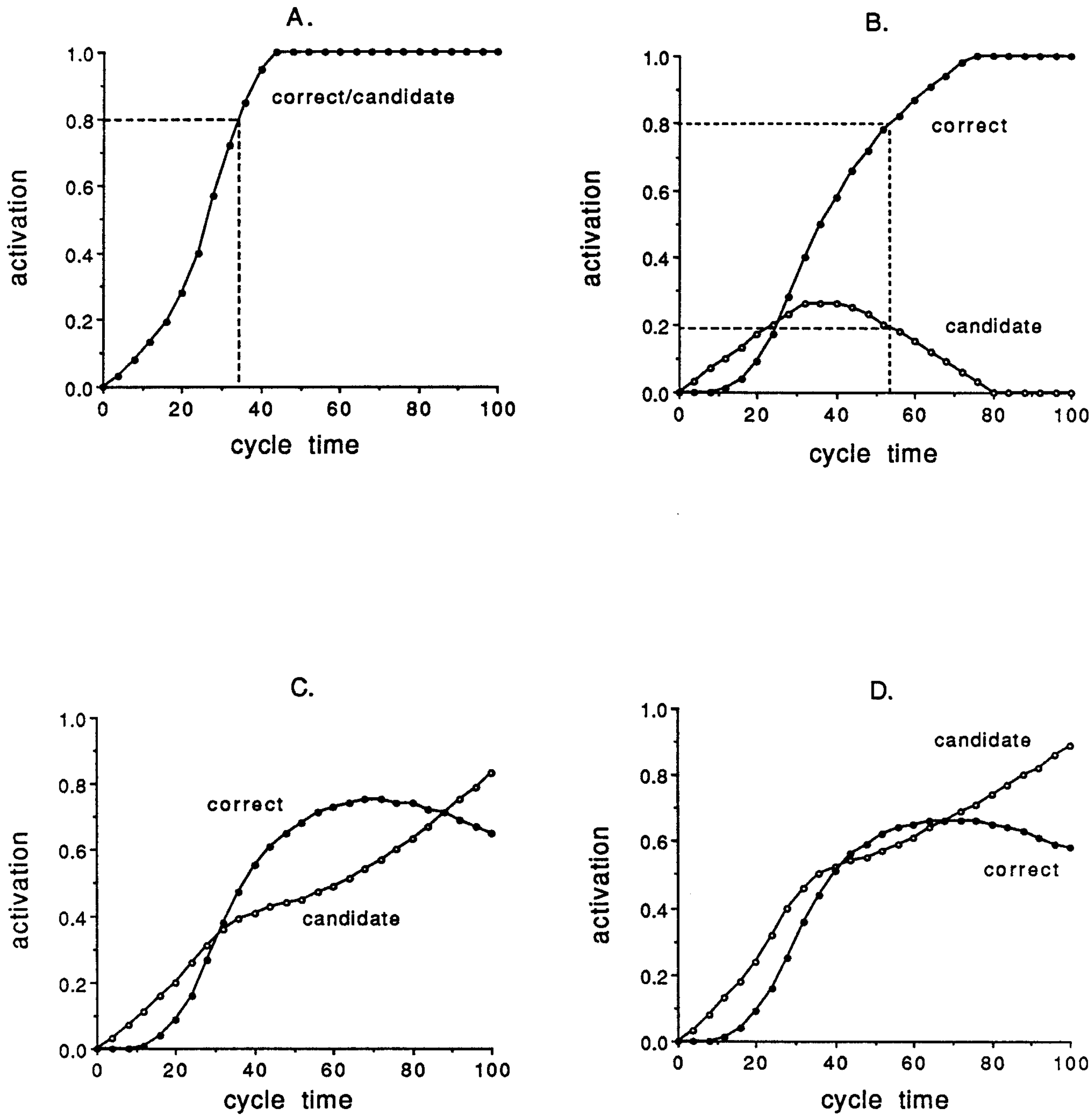

1.0
0.8
0.6
0.4
0.2
0.0
Activation
Correct
Double table-related
Table-related
Table-unrelated/Misc.
0
20
40
60
80
Cycle time

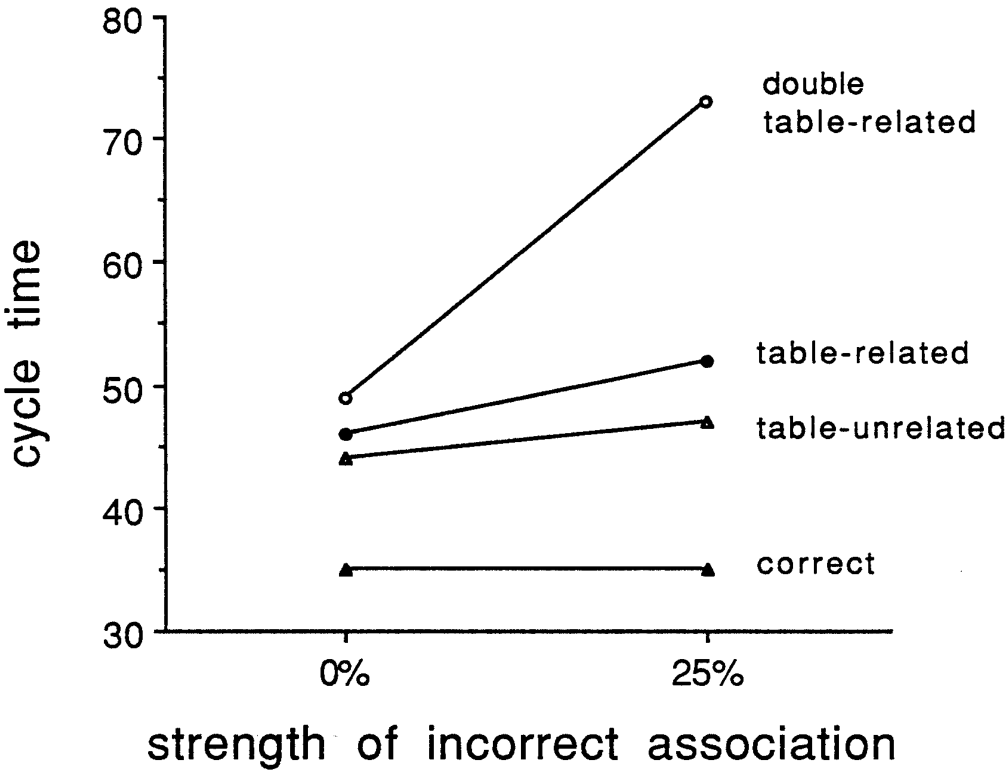

A.
B.
C.
D.
activation
cycle time
correct/primed
correct
primed
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50


A.
B.
C.
D.
activation
cycle time
correct/primed
correct
primed
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100


cycle time
60
55
50
45
40
35
30
double
table-related
table-related
table-unrelated
correct
0%
25%
strength of incorrect association


A.
B.
C.
D.
activation
cycle time
correct/candidate
correct
candidate
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50


A.
correct/candidate
activation
cycle time
B.
correct
candidate
activation
cycle time
C.
correct
candidate
activation
cycle time
D.
candidate
correct
activation
cycle time


80
70
60
50
40
30
cycle time
double
table-related
table-related
table-unrelated
correct
0%
25%
strength of incorrect association