On the Whats and Hows of Retrieval in the Acquisition of a Simple Skill

Michael J. Wenger
University of Notre Dame

Two general views on the role of memory in cognitive skills—an instance-based theory and an associative perspective—were compared with respect to their general assumptions about the information involved and the processes that operate on that information. Characteristics of memory information were examined in terms of predictions for transfer to various stimulus forms as a function of 2 types of learning conditions. Characteristics of memory processes were examined using a set of general process models. Results of 4 experiments indicate that (a) neither theoretical perspective was capable of accounting for all the observed transfer effects, indicating needed refinements to informational assumptions, and that (b) 1 class of process assumptions was consistently supported, whereas other classes were consistently contradicted, indicating a general set of process characteristics that can be used in further model development.

The assumption that memory—retained information and the processes that operate on it—plays an important role in the acquisition and expression of cognitive skills (such as mathematics, chess, music composition, etc.) is a common one. Yet, despite the agreement that memory is critical, the nature of the memory information and the processes that operate on it have tended to be considered separately (e.g., see the general discussions in Massaro, 1998; Thomas, 1996) and in model-specific rather than contrastive terms (e.g., Anderson, 1987; Anderson, Fincham, & Douglass, 1997; Logan, 1988b; Nosofsky & Palmeri, 1997; Richman, Staszewski, & Simon, 1995; Rickard, 1997; Rickard, Healy, & Bourne, 1994). The work presented here assessed predictions about both information and process. A critical conclusion of this effort is that, by considering assumptions about memory information and processes at a reasonably general level (i.e., above the level of specific models), it is possible to gain insights that can help distinguish among competing theoretical alternatives and guide further model development.

The starting point for this work is Logan's (Lassaline & Logan, 1993; Logan, 1988a, 1988b, 1992, 1998; Logan & Etherton, 1994; Logan & Klapp, 1991; Logan, Taylor, & Etherton, 1996, in press) instance theory. Although there is an impressive body of evidence supporting instance theory (e.g., Logan, 1988b, 1992, 1998; Logan & Etherton, 1994; Logan et al., 1996), there is substantial evidence favoring alternatives, including those founded on associative assumptions (e.g., Anderson, 1992; Anderson et al., 1997; Ashcraft, 1987; Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981; Blessing & Anderson, 1996; Rickard, 1997). Unfortunately, this evidence (among others) has been compiled using reasonably distinct paradigms, and the number of cross-model comparisons has been small (however, see Palmeri, 1999; Rickard, 1997, 1999). Furthermore, there does not seem to have been a systematic exploration of either (a) the possible implications of associative information on the acquisition and performance of the tasks that have been addressed by instance theory (however, see Campbell, 1987, 1991; Zbrodoff, 1995) or (b) the general process characteristics implied by either instance theory or associative alternatives. Two of the goals of the work presented here were to take the initial steps in exploring these questions.

Specifying the Hypotheses: Memory Information

According to instance theory, memory information consists of a set of instance representations pertinent to the task, with each embodying the current goal, the interpretation of the stimulus with respect to that goal, and the specific response made to the stimulus.1 Continuing efforts by Logan

---

1 It is intriguing to note that this definition implies that encoding is a process that either cannot start or cannot be completed until after a response is made. Exploration of the implications of this definition, although important, is beyond the scope of this article.
and colleagues have refined the conception of an instance (e.g., Logan, 1998; Logan & Etherton, 1994; Logan et al., in press). For example, analysis of encoding and retrieval conditions has suggested that aspects of the processing episode stored in an instance are a function of attentional selection, with not all attributes being explicitly represented. Recent work (e.g., Logan & Etherton, 1994) further suggests that the co-occurrence of properties is a critical component. This is generally consistent with traditional associative accounts of learning (e.g., Bower & Hilgard, 1981; Hull, 1943) and contemporary propositional accounts (e.g., Kintsch, 1988), which has led to the notion that instances might be thought of as propositions (Logan, 1998; Logan & Etherton, 1994).

These refinements suggest that transfer should be supported to the degree to which the transfer stimuli preserve the abstract "wholeness" of the originally encoded information (e.g., Lassaline & Logan, 1993; Logan, 1990; Palmeri, 1997). Violation of this wholeness, or any stimulus configuration that goes beyond the limits of the propositional representation, should result in highly delimited or negligible transfer.

These refinements also suggest logical challenges to instance theory. For example, the possibility of a range of training contexts and tasks pertinent to any given task suggests a need for some mechanism for selecting instances for retrieval or for modifying their retrieval speed as a function of the match to the transfer task, such as the process outlined by Zbrodoff and Logan (1990) or the mechanism proposed and developed by Nosofsky and Palmeri (see, e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999). The data, such as those on category-level influences on automatic responding (e.g., Egeth, Atkinson, Gilmore, & Marcus, 1973; Fisk & Schneider, 1983; Jonides & Gleitman, 1972; Schneider & Fisk, 1984; Schneider & Shiffrin, 1977; also see Palmeri, 1997) or those pertaining to interitem (associative) relations (e.g., Fisk, Oransky, & Skedsvold, 1988), reinforce the need to consider such mechanisms.

To draw out the implications of these ideas, consider the skilled performance of a simple cognitive task: alphabet arithmetic (e.g., Klapp, Boches, Trabert, & Logan, 1991; Logan, 1988b; Logan & Klapp, 1991). In this task, the alphabet is treated like a number line. Statements involving addition and subtraction (e.g., \( B + 2 = D, D - 2 = B \)) can be constructed using this "number line" and presented to participants for verifications (e.g., Fisk, Oransky, & Skedsvold, 1988), which has led to the notion that instances might be thought of as propositions (Logan, 1998; Logan & Etherton, 1994).

These refinements suggest that transfer should be supported to the degree to which the transfer stimuli preserve the abstract "wholeness" of the originally encoded information (e.g., Lassaline & Logan, 1993; Logan, 1990; Palmeri, 1997). Violation of this wholeness, or any stimulus configuration that goes beyond the limits of the propositional representation, should result in highly delimited or negligible transfer.

These refinements also suggest logical challenges to instance theory. For example, the possibility of a range of training contexts and tasks pertinent to any given task suggests a need for some mechanism for selecting instances for retrieval or for modifying their retrieval speed as a function of the match to the transfer task, such as the process outlined by Zbrodoff and Logan (1990) or the mechanism proposed and developed by Nosofsky and Palmeri (see, e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999). The data, such as those on category-level influences on automatic responding (e.g., Egeth, Atkinson, Gilmore, & Marcus, 1973; Fisk & Schneider, 1983; Jonides & Gleitman, 1972; Schneider & Fisk, 1984; Schneider & Shiffrin, 1977; also see Palmeri, 1997) or those pertaining to interitem (associative) relations (e.g., Fisk, Oransky, & Skedsvold, 1988), reinforce the need to consider such mechanisms.

To draw out the implications of these ideas, consider the skilled performance of a simple cognitive task: alphabet arithmetic (e.g., Klapp, Boches, Trabert, & Logan, 1991; Logan, 1988b; Logan & Klapp, 1991). In this task, the alphabet is treated like a number line. Statements involving addition and subtraction (e.g., \( B + 2 = D, D - 2 = B \)) can be constructed using this "number line" and presented to participants for verifications (e.g., \( B + 2 = D \rightarrow \text{TRUE} \)).

Now consider two possible ways of acquiring the memory information that would be useful in this task (i.e., two training conditions). In the first, a set of alphabet arithmetic facts is memorized (as in Logan & Klapp, 1991, Experiment 2). It is reasonable to assume that the interpretation of each stimulus (e.g., the final element being \( n \) steps to the right of the initial element), the goal (e.g., to learn a set of true alphabet arithmetic facts), and an implicit response (e.g., acknowledging each presented fact as true) can all be assumed to be encoded as a result of memorization.

In the second training condition, associations between letters and a set of numbers (representing the ordinal positions of those letters, such as \( B = 2, D = 4 \)) are memorized. Although participants in this second condition are not encoding complete alphabet arithmetic statements, they can be assumed to be encoding instances embodying propositional knowledge pertinent to the relationship between the letters and numbers. If these participants are later asked to verify alphabet arithmetic statements, they can retrieve the number information, translate the alphabet statement to a numeric statement, and then use the translated statement as a retrieval cue for previously encoded information regarding its truth value (G. D. Logan, personal communication, June 10, 1998).

Various types of alphabet arithmetic statements can be constructed to assess the ability of the information acquired in each condition to support transfer. These statements vary in the degree to which they preserve the physical and conceptual match to training stimuli (see Table 1). Specifically, reversed addition statements preserve the semantics of the normal statements while varying surface form. Inverted subtraction statements preserve the alphabetic "distance" while varying the surface form, and noninverted subtraction statements vary both alphabetic distance and surface form. The statements involving new first elements use alphabetic distances similar to those in the normal statements and are composed from letters in a well-learned region of the alphabet but involve different surface forms. Finally, the symbol statements allow for a test of the ability to learn and perform this task in the absence of relevant preexisting knowledge.

Table 2 shows the transfer predictions derived for instance

### Normal and Transfer Statements Used in Experiments 1–4

<table>
<thead>
<tr>
<th>Normal statements</th>
<th>Example transfer statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1</td>
<td>Set 2</td>
</tr>
<tr>
<td>( B + [2, 3, 4] )</td>
<td>( H + [2, 3, 4] )</td>
</tr>
<tr>
<td>( N + [2, 3, 4] )</td>
<td>( T + [2, 3, 4] )</td>
</tr>
<tr>
<td>( 2 + B = D (Set 1) )</td>
<td>( 3 + H = K (Set 2) )</td>
</tr>
</tbody>
</table>

**Experiment 1**

\( B + [1, 2, 3, 4] \) \( H + [1, 2, 3, 4] \) \( D - 2, E - 3 = B (Set 1, inverted) \)

\( N + [1, 2, 3, 4] \) \( T + [1, 2, 3, 4] \) \( F - [2, 3, 4] (Set 1, noninverted) \)

\( L - [2, 3, 4] (Set 2, noninverted) \)

**Experiment 3**

\( B + [1, 2, 3, 4, 5] \) \( H + [1, 2, 3, 4] \) \( C + [2, 3, 4] (Set 1) \)

\( N + [1, 2, 3, 4, 5] \) \( T + [1, 2, 3, 4] \) \( I + [2, 3, 4] (Set 2) \)

**Experiment 4**

Sequence: \( \Psi \) \( \not\Delta \epsilon \cap \not\Theta \rightarrow Y \)

\( \Psi + [2, 3, 4, 5] \) \( \not\Theta + [2, 3, 4, 5] \)

\( \not\Psi + [2, 3, 4, 5] \) \( \not\Theta + [2, 3, 4, 5] \)

Note. For participants who studied Set 1, Set 2 acted as the unstudied set, and vice versa. The symbols shown here are for example only; the actual symbols were drawn from the DOS ASCII character graphics set.
theory and an associative perspective\(^2\) for each of the training conditions (referred to as the "facts" and "pairs" conditions). Consider first the predictions of instance theory for the facts condition. Consistent with the character of instances (e.g., Logan, 1988b; Logan & Etherton, 1994; Logan et al., 1996, in press) and the data (e.g., Logan, 1988b; Logan & Klapp, 1991), transfer to the normal addition statements (alphabet and symbol) should be robust. Transfer to the reversed addition statements should also be robust, but it may be diminished from that observed for the normal statements because previous work has shown slightly and regularly diminished transfer when surface characteristics of the transfer stimuli differ from those of the training stimuli (e.g., see Figures 3, 4, and 5 in Lassaline & Logan, 1993). This possibility has also been suggested by the fact that verification can be considered an indirect test of memory (e.g., Richardson-Klavehn & Bjork, 1988), and indirect tests are generally sensitive to manipulations of stimulus surface characteristics (e.g., Blaxton, 1989; Wenger & Payne, 1997).

According to instance theory, transfer to the inverted subtraction statements may be questionable. Although both the normal addition and the inverted subtraction statements represent the same alphabetic distance, the direction of the distance may be important. Data consistent with this possibility include observations of the increased difficulty of subtraction relative to addition (e.g., Barrouillet & Fayol, 1998; Fuson, 1984), differences in response times (RTs) for different symbols representing the same quantity (e.g., Gonzalez & Kolers, 1982; Sciana, Semenza, & Butterworth, 1999), and asymmetries in rule application (e.g., Anderson & Fincham, 1994). In addition, Lassaline and Logan (1993) showed that a 180° change in orientation (i.e., flipping the stimulus) produced an impairment in performance. Finally, the facts training should not allow for any transfer to the noninverted subtraction statements or the addition statements with new first elements because these statements share little specific propositional content or surface features with the encoded instances.

Instance theory predicts positive transfer to the normal and symbol statements after the pairs training. However, transfers to these statements should be less than those observed for the same statements after facts training because the additional retrieval operation and the fact that the context and goal of retrieving the number associated with each letter is verification differ from the context and goal of the original learning. Furthermore, because the number of retrievals should be the same for all the transfer statements, instance theory predicts equivalent levels of transfer across all of the other types of statements in the pairs condition.

According to the associative view, the facts training should produce high levels of transfer to the normal, reversed, and symbol statements while also predicting robust but reduced transfer to the inverted subtraction statements. This is because the semantic similarity between the training and transfer statements should be higher in the former cases than in the latter. There should be no transfer to the noninverted subtraction statements or addition statements with the new first elements after the facts training because there is little if any semantic relationship between the training and transfer statements in these cases.

Considering the pairs training, the associative view (like instance theory) requires retrievals of the numeric information and then the use of that information as retrieval cues for verification. Although the predictions of the associative perspective are similar to those of instance theory for this training condition, there is a potentially important difference. Specifically, according to the associative perspective, training in this condition involves a systematic pairing of the letter and the number that may be sensitive to the context (e.g., Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981) but insensitive to the goal of the task in which the retrieval is embedded. Training should produce rapid and reliable retrieval of the initial number association, speeding the entire task in spite of the need for the second retrieval operation. This logic also suggests that the amount of transfer obtained with the reversed addition statements should be equal to that observed for the normal statements in both training conditions.

\(^2\) The predictions for the associative perspective were assessed using an extension of the search of associative memory model developed by Shaffrin and colleagues, most notably, Raaijmakers and Shiffrin (1981) and Gillund and Shiffrin (1984). Details of the simulations and a summary of the results of those simulations are available on request from the author.
Examination of the predictions in Table 2 reveals that, whereas there are points on which the two theoretical perspectives differ, there is substantial overlap. This suggests that it may be difficult to construct strong tests of the competing accounts on the basis of assumptions about memory information, a conclusion that might be anticipated given the difficulties in distinguishing competing theoretical perspectives that are based on differing representational assumptions in other domains (e.g., Ashby & Maddox, 1993; Barsalou, 1990; Busey, in press). Perhaps additional inferential leverage could be gained by considering general assumptions about memory processing.

Specifying the Hypotheses: Process

Instance theory specifies that all the stored instances participate as independent “runners” in a race that itself competes with a task-specific algorithm. By assuming that (among other things) the distributions of retrieval and algorithm execution times conform to a generalized Weibull distribution (Colonius, 1995; Logan, 1995; Weibull, 1951), instance theory predicts that the mean and the standard deviation of observed RTs will reduce as a power function of the number of stored instances (Logan, 1992), predictions that have been supported by several studies (e.g., Logan, 1992; Palmeri, 1999; Rickard, 1999).

As noted earlier, retrieval needs to be supported by some process that allows the selection of instances that participate in the race (e.g., Logan, 1990; Logan & Etherton, 1994; Logan et al., 1996; Palmeri, 1997). Details of this process were not specified in earlier versions of instance theory, but at a minimum one must assume that this process operates in a sequential arrangement with retrieval (see, e.g., Logan, 1988b, p. 512). Similar processing assumptions can be found in associative models (e.g., Ashcraft, 1982, 1987; Ashcraft & Stazzyk, 1981). Alternatively, the selection process may be “embedded” in the process of accumulating evidence in favor of one of a set of responses, as in recent extensions of instance theory by Nosofsky and Palmeri (Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999).

Rickard (1997, 1999) recently challenged the processing assumptions of instance theory, suggesting that observers choose either the algorithm or retrieval, modifying that strategy as required (e.g., if retrieval is selected but later fails). A critical characteristic of Rickard’s alternative is the proposal that the algorithm and retrieval cannot be executed simultaneously across the course of a trial. Yet, as noted by Colonius (1995), the consistent pattern of data obtained in tests of instance theory seems to support the race model (see also recent work by Palmeri, 1997, 1999).

To derive predictions on the basis of these alternative assumptions, a general approach to developing process models, known as precedence networks (Schweickert, Fisher, & Goldstein, 1992), was used. Briefly, precedence networks are based on an analysis of a task in terms of its component processes. Specification of the temporal and logical relations among these component processes allows for predictions about total completion times (for related work, see Ehrenstein, Schweickert, Choi, & Proctor, 1997; Liu, 1996; Miller, 1993). The specific type of precedence network used here was the order-of-processing (OP) network (e.g., Fisher & Goldstein, 1983; Goldstein & Fisher, 1991; Schweickert et al., 1992), a less restrictive variant of program evaluation and review technique (PERT) networks (cf. Schweickert, 1978, 1983).

Seven different OP network models were developed to instantiate the processing assumptions of instance theory, a variety of associative approaches, and a simplified version of Rickard’s (1997, 1999) model; the models are presented schematically in Figure 1. Details for each of the models and derivations of the predictions for the various factorial manipulations used in the experiments are provided in the Appendix.

Consider first the models developed for the facts training (Models 1 and 2). In Model 1, the outputs of either an algorithmic (a) or a memory retrieval process (m) are operated on by a comparison process (c) that compares the output to the test stimulus in order to select a response. This retrieval-with-comparison captures the basic processing assumptions of earlier versions of instance theory, Ashcraft’s model for numeric arithmetic, and responding on the basis of item familiarity in various global memory models. Model 2 differs from Model 1 in that memory retrieval can directly support a response without the need for an intervening

3 Although Logan (e.g., 1988b, 1992) argued that this ability to predict the power-function decrease in the mean and standard deviation of response times (RTs), on the basis of a small set of assumptions, is unique to instance theory, Colonius (1995) demonstrated that it is the power function that actually “predicts” the assumptions of instance theory. That is, the finding of a power-function decrease in RT is a necessary and sufficient condition for instance theory’s assumption of the underlying Weibull distribution (for an alternative interpretation, see Logan, 1995). Consequently, the ability to predict the power function becomes less critical for evaluating alternative accounts. Colonius further noted that the consistent findings of power-function reductions in RTs and good fits of the Weibull distribution to RT distributions fail to contradict instance theory’s race model, something that more recently has been challenged, but in a far less general way, by Rickard (1997). Complicating the issue further, there are long-standing problems associated with artifacts of averaging across participants generally (e.g., Ashby, Maddox, & Lee, 1994; Estes, 1956) and specific to power functions (e.g., Myung, Kim, & Pitt, 1998).

4 Specification of this process would also need to include some way of defining if and how an error could be made because this was unspecified (and perhaps impossible) in earlier versions of instance theory. Note that this was not a problem for Shiffrin and Schneider’s models (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), which were able to make predictions about both response latency and probability.

5 One of the more important simplifications of Rickard’s (1997, 1999) model is the assumption that the continuous activation flow and feedback that characterize that model can be analyzed in terms of stages of processing. For general considerations of some of the issues related to this simplification, see Ashby and Townsend (1980) and Townsend and Fikes (1995). For a demonstration of the consistency of a continuous flow approach with the network analyses used here, see Wenger and Townsend (1996, in press-b). As an additional simplification, it was assumed that retrieval failure would be negligible.
RETRIEVAL PROCESSES IN SKILL

Model 1

\[ a \rightarrow c \]

\[ m \]

Model 3

\[ a \rightarrow c \]

\[ x \rightarrow d \]

\[ y \]

Model 4

\[ a \rightarrow c \]

\[ x \rightarrow y \rightarrow d \]

Model 2

\[ a \rightarrow c \]

\[ m \]

Model 5

\[ a \rightarrow c \]

\[ x \rightarrow d \]

\[ y \]

Model 6

\[ a \rightarrow c \]

\[ x \rightarrow y \rightarrow d \]

Model 7

\[ a \]

\[ a' \]

\[ m \]

\[ m' \]

\[ c \]

AND gate

OR gate

Figure 1. Schematic representations of the process models developed for the experiments (see the Appendix for details concerning the models and the predictions for the mean interaction contrasts). Models 1 and 2 were developed for the facts training condition, whereas Models 3, 4, 5, and 6 were developed for the associative training condition. Models 1, 3, 4, and 7 require that the products of retrieval be acted on by a comparison process before responding, whereas Models 2, 5, and 6 allow retrieval to directly support responding. Model 7 is a representation of Rickard's (1997) model. For the component processes of Models 1-6, \( a \) is the algorithmic process, \( c \) is the comparison process, \( m \) is a general memory retrieval process, \( x, y \) are the two associative retrievals, and \( d \) is a translation process. For the component processes of Model 7, \( a \) is the initial step of the algorithm, \( a' \) is the remaining steps of the algorithm, \( m \) is the initial portion of retrieval, \( m' \) is the remaining portion of retrieval, and \( c \) is a comparison process.

Comparison processes. This would correspond to the mechanisms in later refinements of instance theory (specifically those of Nosofsky & Palmeri, 1997; Palmeri, 1997), single-production responding in adaptive control of thought-revised (ACT-R; e.g., Anderson, 1993; Anderson & Fincham, 1994), and the sampling and recovery process of a global associative memory model such as the search of associative memory (SAM) model.

In the models developed for the pairs training (Models 3-6), an algorithmic process outputs in all cases to a comparison process. These models assumed two retrievals (\( x, y \)) for the two letter-number associations rather than a single memory retrieval. In Models 3 and 4, the outputs of these retrievals are combined and used as a cue in a subsequent retrieval (\( d \)), which in turn leads to the comparison process. In Model 3, the initial retrievals take place in parallel, with both being required to complete before the second retrieval and comparison processes can begin. In Model 4, the initial retrievals take place sequentially. Models 5 and 6 differ from Models 3 and 4 in only one respect: For these models, completion of the retrieval processes can lead directly to a response.

In Model 7 (the simplified representation of Rickard's 1997, 1999, model), a strategy choice is made on the basis of a competition between the initial step of the algorithm (\( a \)) and the initial steps of retrieval (\( m \)). The outcome of this competition determines which process (algorithm \( a' \) or retrieval \( m' \)) will be used to complete processing, with each providing inputs to a final comparison process. As such, the algorithm and retrieval do not operate concurrently for the duration of the trial.

To develop empirical tests for the models, a metatheoretical approach known as systems factorial technology (e.g., Schweickert & Townsend, 1989; Townsend & Nozawa, 1995; Townsend & Schweickert, 1989) was used. This approach is designed to investigate and reveal the structure of processing systems by assessing the impact of factorial manipulations of experimental variables, each of which can be assumed to have selective influence on one of the processes of interest (see also Townsend, 1984; Townsend & Thomas, 1994).^6

^6 The approach has been applied extensively to the study of processing architectures, being particularly informative with respect to the arrangement of processes (e.g., parallel vs. serial); it
One of the diagnostics of systems factorial technology is the mean interaction contrast. Assume for the moment that there are two experimental factors, each with two levels. Assume that in moving from the first to the second level for either factor, the duration of the process affected by that factor decreases. Let $\text{RT}_{i,j}$ be the mean RT when the first factor is at level $i$ and the second factor is at level $j$ with $i,j = 1$ denoting the factor at its slowest level and $i,j = 2$ denoting the factor at its fastest level. The mean interaction contrast can then be defined as

$$\Delta^2(\text{RT}) = (\text{RT}_{1,1} - \text{RT}_{1,2}) - (\text{RT}_{2,1} - \text{RT}_{2,2}).$$  

(1)

This is the familiar “difference of differences” used to assess the presence and sign of an interaction in a $2 \times 2$ factorial experiment and is the same form used by Sternberg (1969) in his additive factors method. Three outcomes are possible for this contrast: The contrast may be zero (additive), the contrast may be positive (over- or superadditive), or the contrast may be negative (under- or subadditive). The pattern of the signs of the interaction contrasts across all possible pairs of component processes provides the basis for distinguishing both simple (e.g., Townsend & Nozawa, 1995; Wenger & Townsend, in press-b) and complex (e.g., Schweickert et al., 1992) alternatives.

In the experiments that follow, manipulation of the algorithmic process corresponds to manipulation of the value of the addend in the statements: Increasing the addend requires additional counting. Manipulation of the comparison process corresponds to manipulation of the truth value of the statement: Given that all statements in all of the experiments were false by $\pm 1$ letter, it was assumed that verification of false statements would take longer than true statements (e.g., Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981; Krueger, 1986; Krueger & Hallford, 1984; Logan & Klapp, 1991). Finally, manipulation of the memory or retrieval processes corresponds to manipulation of the training status (trained, untrained) under the assumption that training should lead to a reduction in RT.

The predicted signs of the interaction contrasts for the effects of pairwise manipulation of the component processes in each of the OP network models are presented in Table 3. Note that these predictions are reasonably general, in that they were derived without making any assumptions about the form of the distributions of the component process durations (i.e., they are zero-parameter predictions). As can be seen in Table 3, although the set of models is not completely identifiable, it is possible to obtain distinct predictions for different sets of the alternative models. In particular, it is possible to distinguish (at the level of the means) models that do require a comparison process interposed between retrieval and responding from those that do not. It is also possible to distinguish models that allow full concurrency of the algorithm and retrieval from the model that does not. Thus, it appears that, by considering assumptions about processing along with assumptions about information, one can gain some inferential leverage.

Experiment 1

The four experiments used the basic procedure outlined by Logan and Klapp (1991, Experiment 2), with two modifications. First, whereas Logan and Klapp used a single type of training (i.e., memorization of alphabet arithmetic statements), the present work used two types of training (i.e., the facts and pairs conditions described earlier). Second, performance in the transfer phase was assessed using (a) trained and untrained alphabet addition statements (referred to here as “normal addition statements”) and (b) alphabet addition statements involving the same elements as the normal statements but with the initial two elements reversed (e.g., $B + 2 = D \Rightarrow D + 2 = B$).

\[\text{Note that the developments presented by Schweickert, Fisher, and Goldstein (1992) allow for predictions to be derived for all moments of the response time distribution, such as contrasts on the cumulative distribution and survivor functions } F(t) \text{ and } S(t).\]

\[\text{Note that the slowing for the false statements relative to the true ones is specific to the small amount of error in the false statements. Previous work, particularly that of Ashcraft (Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981), systematically varied the degree of incorrectness in the false statements, such that the false statements could often be rejected faster than the true statements could be verified. Such an approach could be used to test the models presented here. However, if this were done, the true statements would correspond to the slow operation of the comparison process and the false statements would represent the fast operation.}\]
Method

Participants. A total of 48 participants was recruited from introductory psychology courses and participated voluntarily in exchange for credit toward completion of a course requirement for laboratory or library research. All participants reported normal or corrected-to-normal vision.

Design. The experiment was conducted as a 2 (training condition: facts, pairs) \times 2 (transfer condition: trained, untrained) \times 3 (digit addend: 2, 3, 4) \times 2 (statement type: normal addition, reverse addition) \times 2 (truth value: true, false) mixed-factorial design, with digit addend, statement type, and truth value manipulated within subject.

Materials. The stimuli were similar to those used by Logan and Klap (1991, Experiment 2). Two sets of eight letters and three digits were combined to form (in each set) six true alphabet arithmetic statements (the normal addition statements; see Table 1). Each statement was a single-operator (two-addend) addition statement (e.g., \( B + 2 = D \)). Set 1 was formed by combining the digits 2–4 with the letters B and N; Set 2 was formed by combining the digits 2–4 with the letters H and T. For each set, a total of 12 false statements was constructed by adding and subtracting one from the correct answer for each of the six statements in the sets. Six true reverse addition statements were formed for each set by exchanging the first and second addends (e.g., \( B + 2 = D \Rightarrow 2 + B = D \)). Two false statements were constructed for each of these reverse additions statements, with these false statements being incorrect by \( \pm 1 \) letter.

Apparatus. All stimuli were displayed and all responses were collected and timed (to \( \pm 1 \) ms) using a PC-compatible microcomputer and a 33-cm amber monochrome monitor. All stimuli were produced with the DOS character set on an 80-column line and were presented amber on black.

Procedure. Participants were tested individually in sessions lasting approximately 1 hr. Sessions were divided into two phases. The first consisted of training with the appropriate materials and practice using the appropriate response keys. The objective was to give participants robust knowledge of the information (either complete alphabet arithmetic statements or letter-number associations) they would be relying on during the transfer phase. This was accomplished by memorization followed by criterion-based tests for memory. The second (transfer) phase consisted of the true–false verification task involving the normal and reverse addition statements.

At the beginning of each session, participants were presented with the full set of training stimuli as a list of items to be memorized. In the facts condition, participants were shown the list of alphabet arithmetic facts, one at a time, centered on the screen, in ascending order with respect to the value of the digit addend. In the pairs condition, participants were shown the association between the first letter in the fact and its numeric value (its position in the alphabet), followed by the association between the final letter in the fact and its numeric value. For example, the presentation of the fact \( B + 2 = D \) took the form \( B = 2 \), followed by \( D = 4 \). Participants in both conditions had control over the pacing of list presentation and could repeat the list any number of times. However, they were required to repeat the list a minimum of three times and were instructed to repeat the list until they were confident that they had completely memorized it. For the participants in the pairs condition, no mention was made of the alphabet arithmetic task; they were told only that their memory for the letter-number associations would be tested.

After memorization, participants were tested for their memory for the list. In the facts condition, participants were presented with the first portion of each fact (e.g., \( B + 2 = ?? \)) and were required to verbally respond with the correct answer (e.g., \( D \)). In the pairs condition, participants were presented with a letter (e.g., \( B = ?? \)) and were required to verbally respond with the paired number (e.g., \( B = 2 \)). The items in the set were presented one at a time, in a random order, centered on the screen. Responses were collected and timed by having the experimenter press a key on the computer keyboard corresponding to the correct response.

After each response in the memory test, participants were given auditory feedback (a 1,320- or a 200-Hz tone presented for 250 ms for correct and incorrect responses, respectively). At the end of the set, participants were given feedback about their level of accuracy and median RT. Participants repeated the set until they were able to make three consecutive passes through the set with no errors and had a median RT for correct responses of 2,000 ms or less for the facts condition and 1,500 ms or less for the pairs condition. These RT values were chosen (on the basis of pilot work) to equate the amount of training experienced in each of the conditions.

After meeting the performance criteria, participants were given practice making the true–false responses required in the verification task. A total of 20 stimuli (10 instances each of the words TRUE and FALSE, in uppercase letters) was presented in the center of the screen. Participants were instructed to press the appropriate key (\( z \) for true and \( / \) for false) as quickly and accurately as possible, without looking at the keyboard, and were given auditory feedback on their responses. After each response, the screen was cleared for 500 ms. After this, participants were tested two additional times for their memory of the list using the procedure and criteria described previously. In each of these two additional memory tests, participants were required to make three successive passes through the list with no errors using the appropriate RT criterion.

After the second memory test, participants performed the verification task with either the trained or untrained set of normal addition statements plus the set of reverse addition statements. This transfer phase began with response instructions describing the types of statements that would be seen and instructing participants in a response strategy appropriate to the training condition. Participants in the facts condition who were presented with the trained set of stimuli were instructed to rely on their memory for the statements learned during training to respond to the normal arithmetic statements (the normal addition statements; see Table 1). RT values were chosen (on the basis of pilot work) to equate the amount of training experienced in each of the conditions.

Note that this procedure required that the first element in the facts receive multiple presentations; this was done to equate the number of presentations of the first element with the number of times the first element was presented in the other conditions.

To ensure the accuracy, speed, and consistency of experimenter responding, two steps were taken. First, a keyboard template was constructed mapping the set of correct responses onto a set of adjacent keys on the A row of the QWERTY keyboard. Second, all experimenters were given 1–2 hr of training and practice collecting responses before testing participants. These latencies were of interest only for documenting sufficient training and were not used in any of the analyses reported for this or any of the remaining experiments.

---

9 Eighty-one participants were tested originally. However, 9 participants (distributed approximately equally across the four conditions) had to be replaced because of highly variable response times (i.e., standard deviations were equal to or greater than means for each participant) and extremely low levels of accuracy (i.e., less than 80% correct responding). Twenty-four of the remaining 72 participants were in two comparison groups unrelated to the present effort; data from these conditions are not discussed.

10 Note that this procedure required that the first element in the facts receive multiple presentations; this was done to equate the number of presentations of the first element with the number of times the first element was presented in the other conditions.

11 To ensure the accuracy, speed, and consistency of experimenter responding, two steps were taken. First, a keyboard template was constructed mapping the set of correct responses onto a set of adjacent keys on the A row of the QWERTY keyboard. Second, all experimenters were given 1–2 hr of training and practice collecting responses before testing participants. These latencies were of interest only for documenting sufficient training and were not used in any of the analyses reported for this or any of the remaining experiments.
addition statements and to use the method of counting through the alphabet to respond to the reverse addition statements. Participants in the pairs condition who were presented with statements constructed from the trained set of stimuli were instructed to rely on their knowledge of the associations between the letters and numbers to translate the normal and reverse addition statements into numeric addition statements. Participants in both conditions who were presented with untrained stimuli were instructed to use the counting algorithm for all statements. For those participants instructed to use the counting algorithm on any or all statements, an example of its use was presented by the experimenter, who then gave them a single probe statement ($A + 3 = D$ for all participants) to assess their understanding of the algorithm. The explanation and probe were repeated if participants were at all unsure of the use of the counting algorithm.

The true normal addition statements were each presented eight times, and the two types of false normal addition statements were each presented four times. The true reverse addition statements were each presented four times, and the false reverse addition statements were each presented twice. Thus, a total of 144 statements was presented.

Each trial began with presentation of the orienting message "Get ready..." centered on the screen for 670 ms. The screen was then blank for 670 ms, after which the statement to be verified was presented centered on the screen. The statement remained on the screen until participants made their response, which was followed by auditory feedback and either the word "correct" or "incorrect" presented centered on the screen for 1,250 ms. Participants were also given feedback about their RT on each trial. For RTs less than 3,000 ms, the message "Time was good" was printed centered horizontally on the screen on the line directly below the accuracy feedback. For RTs between 3,000 and 4,000 ms, the feedback phrase was "Time was OK but should be faster," and for RTs greater than 4,600 ms the feedback phrase was "Time was too slow." The screen was then blank for 1,750 ms, after which the orienting message for the next stimulus was presented. These timings, RT ranges, and orienting and feedback statements were based on those used by Logan and Klapp (1991, Experiment 2). The order of statement presentation was randomly determined for each participant. At the halfway point in the transfer phase (i.e., after 72 statements had been presented), participants were given a short break.

**Results**

**Training.** The central question with respect to the training data was whether there would be systematic differences as a function of training condition that could confound the interpretation of the remainder of the data. Two aspects of the training data were examined: the mean study time per item on the initial three repetitions of the list during memorization and the number of memory verification blocks required to reach criterion. An alpha level of .05 was used throughout the experiments.

Mean study times per item for the first three list repetitions were analyzed with a 2 (training condition: facts, pairs) × 3 (repetition: 1, 2, 3) mixed-factorial analysis of variance (ANOVA). The ANOVA showed a main effect for training condition, $F(1, 46) = 20.06, MSE = 16.29$, with participants in the facts condition requiring more study time per item (6.89 s) than participants in the pairs condition (3.88 s). This is not surprising given that the items to be memorized in the facts condition were entire statements, whereas the items to be memorized in the pairs condition were simple letter-number associations. There was also a main effect for repetition, $F(2, 92) = 26.16, MSE = 8.41$, indicating that study times decreased across the three list repetitions (7.84, 4.44, and 3.88 s for Repetitions 1, 2, and 3, respectively). The Condition × Repetition interaction was marginally significant, $F(2, 92) = 2.62, MSE = 8.41, p = .08$, reflecting the fact that the decrease in study times was slightly more pronounced for participants in the facts condition (10.12, 5.60, and 4.94 s, respectively) than it was for participants in the pairs condition (5.55, 3.27, and 2.82 s, respectively).

Participants in the facts and pairs conditions required virtually identical levels of exposure (10.4 vs. 10.5 blocks, ns) to reach criteria. This indicated that the training criteria allowed the amount of verification testing required in each of the conditions to be equated, providing additional evidence that participants in the two conditions had reached approximately equal levels of memory for the target information by the end of the training phase.

**Transfer: Informational hypotheses.** Although the main data of concern were the RTs, the accuracy data were examined for the presence of any evidence of speed-accuracy trade-offs (e.g., Pachella, 1974). None was found, for either the normal or reverse addition statements, and accuracy for the trained (.94) and the untrained (.91) statements in both training conditions was uniformly high.

To control for outlier values in the RT data, each observation for each participant was compared with the median RT for that participant at the appropriate level of the design before analysis. If an individual observation was more than 3 SDs away from the participant’s median RT at that level, that RT was replaced with the median RT for that participant at that level of the design. This procedure resulted in 15 observations (0.2% of the total) being replaced in the data from Experiment 1. For the analyses of the RT data, the median RT for each participant’s correct responses at each level of the design was calculated and analyses were conducted on the means of these medians.

Correct response RTs were analyzed with a 2 (training condition: facts, pairs) × 2 (training status: trained, untrained) × 2 (statement type: normal, reverse) × 2 (truth value: true, false) × 3 (addend: 2, 3, and 4) mixed-factorial ANOVA. Only the initial four presentations of the normal statements were included in this analysis to equate number of presentations across statement types. The main questions to be addressed in this analysis concerned the predictions for transfer to the normal and reverse addition statements as a function of each of the training conditions (see Table 2).

The analysis revealed that the two types of training speeded responding, reduced the influence of the value of the digit addend on RT, and led to equivalent levels of positive transfer. Trained participants responded faster over-
all (1,493 ms) than did untrained participants (1,835 ms), $F(1, 44) = 10.15$, $MSE = 1,664,586.39$. Although increasing the value of the digit addend did lead to reliable changes in RTs (1,448, 1,780, and 1,764 ms for statements with addends of 2, 3, and 4, respectively), $F(2, 88) = 51.59$, $MSE = 130,678.68$, the effect was much more pronounced for untrained statements (1,573, 1,906, and 2,027 ms for statements with addends of 2, 3, and 4, respectively) than it was for trained statements (1,323, 1,654, and 1,501 ms for statements with addends of 2, 3, and 4, respectively).14 $F(2, 88) = 9.24$, $MSE = 130,678.68$. Critically, neither the main effect of training condition nor any of the interactions involving training condition, training status, or digit addend were reliable ($F < 1.00$). In addition, neither the main effect of statement type nor any of the interactions involving statement type, training condition, training status, and digit addend were reliable ($F < 1.50$, $p > .20$). Essentially, both training conditions were effective and resulted in equivalent positive transfer to both types of statements. Considering the predictions for transfer in Table 2, the observed effects were more consistent with the predictions of the associative view than with those of instance theory.

The last effect of interest concerned the truth value of statements. Overall, true statements were responded to faster (1,599 ms) than were false statements (1,729 ms), $F(1, 44) = 9.24$, $MSE = 47,165.05$. This finding parallels similar effects found in numeric addition verification tasks (e.g., Ashcraft & Battaglia, 1978; Ashcraft & Stazyk, 1981; Krueger, 1986; Krueger & Hallford, 1984).

Transfer: Process hypotheses. The predictions of the OP network models (see Tables 3 and 4) are presented graphically in Figure 2, along with a summary of the results for all four of the experiments. Table 5 shows the mean RTs for each of the conditions in Experiment 1, organized according to the manipulations of memory and the algorithm.15

All of the model predictions are based on the critical assumption of monotonic effects of factor manipulations on RT (Schweickert et al., 1992). The presence of the endpoint effects (as a function of digit addend) meant that this assumption was violated in some cases. Consequently, it was necessary to focus on that subset of the data in which the effects of the factor manipulations were monotonic.

All of the models propose that memory operates concurrently with the algorithm for some portion of the trial and predict that the sign of the mean interaction contrast for the manipulation of memory and the algorithm will be positive. Overall, 14 of the 16 possible interaction contrasts were significantly greater than zero (see Table 5). However, 7 of these contrasts involved nonmonotonic changes in RT across factor levels. Of the 9 contrasts involving monotonic changes, 7 were positive. Consequently, the predictions of all of the models were supported for this set of interaction contrasts.

All of the models propose that the algorithmic and comparison processes are arranged sequentially. Models 1, 3, and 4 hold that the algorithm and memory transition to the comparison process by way of an OR gate (see Figure 1), although Model 7 holds that a comparison process will follow the process selected on the basis of the initial competition between the algorithm and retrieval. In contrast, Models 2, 5, and 6 hold that only the algorithm communicates with the comparison process, with both being concurrent to the memory process. Consequently, Models 1, 3, 4, and 7 differ from Models 2, 5, and 6 in terms of the predicted sign of the mean interaction contrast for the manipulation of the algorithmic and comparison processes: Models 1, 3, 4, and 7 predict an additive interaction, whereas Models 2, 5, and 6 predict a subadditive interaction.16 Of the 16 possible contrasts, 13 were negative; however, 4 of these contrasts

Table 4

<table>
<thead>
<tr>
<th>Model</th>
<th>$g_h$</th>
<th>$R_h$</th>
<th>$\Delta^2$ (RT)</th>
<th>$a, c$</th>
<th>$a, m$</th>
<th>$m, c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$t_a + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$t_a + t_m + t_c$</td>
<td>$t_a &lt; t_m$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td></td>
</tr>
</tbody>
</table>

Note. $g_h$ and $R_h$ refer to the components of Equation A4. For Models 1–6, $a, m$, and $c$ refer to the algorithmic, memory, and comparison processes, respectively. For Model 7, $a$ refers to the initial steps of the algorithm, $a'$ refers to the remaining steps of the algorithm, $m$ refers to the initial part of memory retrieval, $m'$ refers to the remaining part of memory retrieval, and $c$ refers to the final comparison process. RT = response time; OP = order of processing.

14 This "bow" in the response time function was, to foreshadow, something observed to varying degrees in all four experiments. It might best be understood as an instance of the general finding of endpoint effects in tasks involving memory for a series (e.g., Hintzman, Block, & Summers, 1973; Zimmerman & Underwood, 1968) or discrimination along a dimension (e.g., Braida et al., 1984; Macmillan, Braida, & Goldberg, 1987; Macmillan, Goldberg, & Braida, 1988; Pollack, 1952, 1953; Shiffrin & Nosofsky, 1993), with the dimension here involving the symbolic distances represented by the alphabet as number line (see also Neath, 1998, chap. 14). Similar effects have been observed in other studies using similar tasks (e.g., Logan & Klapp, 1991; Palmeri, 1997).

15 To conserve space, only one set of interaction contrasts is presented for each of the experiments. The remaining interaction contrasts (i.e., those for the manipulation of the algorithm and the comparison process, and memory and the comparison process) can be recovered by reorganizing the presented data.

16 These predictions are consistent with the results for sequential processes in a network with OR gates, in which the arrangement of the component processes does not involve a Wheatstone bridge (see Schweickert & Wang, 1993, p. 28).
involved nonmonotonic changes in RT across factor levels. Of the 12 remaining contrasts, 9 were negative. Consequently, these data support the models that allow memory to directly lead to responding, with the comparison process operating only on the output of the algorithm and with memory and the algorithm operating in parallel across the course of the trial.

The interaction contrasts for the manipulation of memory and the comparison process provide an additional basis for discriminating among the sets of models. The models that place a comparison process between retrieval and responding predict that the interaction contrasts for memory and comparison will be additive. In contrast, the models that allow retrieval to directly support responding predict that this interaction will be superadditive. Of the 12 possible contrasts, 8 were positive, 3 were zero, and 1 was negative. One of these contrasts involved nonmonotonic changes in mean RTs across factor levels, and of the 11 remaining interaction contrasts, 7 were positive, 3 were zero, and 1 was negative. Consequently, these results are consistent with the results from the preceding interaction contrasts in that they support Models 2, 5, and 6 and contradict Models 1, 3, 4, and 7.

Discussion

In terms of information, the results of Experiment 1 are more consistent with the predictions of the associative perspective than with those of instance theory. Both training conditions were effective and resulted in equivalent transfer to the normal and reverse addition statements. Although this is the pattern of transfer predicted by the associative model, instance theory predicted positive but reduced transfer for the associative training condition relative to the fact training condition. Critically, the equivalent levels of transfer to the normal addition statements were obtained in the first four item presentations, meaning that the equivalence was not an artifact of averaging over eight item presentations (double the number used by Logan & Klapp, 1991, Experiment 2).

In terms of process, the results clearly favor the models that allow retrieval to directly support responding (Models 2, 5, and 6) and that allow retrieval and the algorithm to operate concurrently across the course of the trial. These outcomes support hypotheses for processes such as the evidence accumulation process in recent extensions of instance theory (Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999), the single-production operations in ACT-R, and the

Figure 2. Graphical summary of the predictions for the order-of-processing network models for each pair of process manipulations and summary of the results obtained in each of the experiments individually and overall. Numbers within each of the bars indicate the number of interaction contrasts involving monotonic changes in response time across factor levels that were positive (clear bars), 0 (gray bars), or negative (black bars). In each panel, the left set of bars represents the outcomes for the manipulation of the algorithm and memory (aim), the middle set of bars represents the outcomes for the manipulation of the algorithm and comparison (alc), and the right set of bars represents the outcomes for the manipulation of memory and comparison (m/c). Expt. = Experiment.
**Table 5**

*Mean Interaction Contrasts (Equation 1) for the Manipulation of Memory (Trained, Untrained) and the Algorithm (Digit Operand) in Experiment 1*

<table>
<thead>
<tr>
<th>Manipulation of memory</th>
<th>True statements addend</th>
<th>False statements addend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Facts condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,133</td>
<td>1,387</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,302</td>
<td>1,650</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>94</td>
<td>401</td>
</tr>
<tr>
<td>Asociative condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,197</td>
<td>1,428</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,317</td>
<td>1,644</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>96</td>
<td>243</td>
</tr>
<tr>
<td>Facts condition, transfer statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,248</td>
<td>1,623</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,421</td>
<td>1,799</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>3</td>
<td>193</td>
</tr>
<tr>
<td>Asociative condition, transfer statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,280</td>
<td>1,384</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,499</td>
<td>1,706</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>103</td>
<td>187</td>
</tr>
</tbody>
</table>

*Note.* All of the other interaction contrasts can be recovered by reorganizing the data in this table. *Interaction contrast not significantly different from zero.

Experiment 2

Experiment 2 was intended to allow another test of the process predictions while examining performance with transfer stimuli that require different predictions for the effects of the two training conditions (see Table 2). Specifically, Experiment 2 investigated the ability of the knowledge gained in the two training conditions to support transfer to two types of subtraction statements, along with the normal addition statements.

**Method**

**Participants.** A total of 48 participants was recruited for Experiment 2. All participants reported normal or corrected-to-normal vision, and none had participated in Experiment 1.

**Design.** Experiment 2 was conducted as a 2 (training condition: facts, pairs) × 2 (training status: trained, untrained) × 4 (digit operand: 1, 2, 3, and 4) × 3 (statement type: normal addition, inverted subtraction, and noninverted subtraction) × 2 (truth value: true, false) mixed-factorial design, with digit operand, statement type, and truth value manipulated within subject.

**Materials.** The normal addition statements for Experiment 2 were identical to those in Experiment 1, except that each initial letter was paired with the Digit Addend 1 as well as the addends 2–4. Because the Operand 1 statements were not used in the transfer phase (see below), no corresponding false statements were constructed for these statements. Two types of subtraction statements were constructed (see Table 1). First, six true inverted subtraction statements were formed by exchanging the first and final elements (e.g., $B + 4 = F → F - 4 = B$) in the set of six normal addition statements involving the digit addends 2–4. Second, an additional four true noninverted subtraction statements were formed by using the last letter in the range of the trained set as the initial element in the subtraction statement and subtracting the digits 2 and 3 (e.g., $F - 2 = D$ and $F - 3 = C$). It was not necessary to construct six noninverted subtraction statements because the subtraction statements in this group involving the digit 4 were also inverted subtraction statements (e.g., $F - 4 = B$ is both an inverted and noninverted transfer statement). Two types of false statements (with answers being off by ±1) were constructed for each of the transfer statements.

**Apparatus and procedure.** The apparatus and procedure used in Experiment 2 were identical to those used in Experiment 1 with the following exceptions: First, participants repeated the memory verification test a total of four times using the same accuracy and RT criteria used in Experiment 1. Pilot work and prior research (e.g., Logan & Klapp, 1991) have suggested that with longer lists (i.e., more facts), additional study and verification time would be required for participants to become accurate and rapid. Second, the normal addition statements in the transfer phase of the experiment were discarded using the criteria noted for Experiment 1.
were those involving only addends of 2, 3, and 4. Pilot work indicated that Operand 1 statements could be verified so quickly that the resulting RT functions would be distorted at Operand 1. Normal addition and subtraction statements of each type (true and false) were presented the same number of times as in Experiment 1; however, because there were two types of subtraction statements (inverted and noninverted), the total statements presented in Experiment 2 was 176.

Results

Training. As in Experiment 1, two aspects of the training data were examined: the mean study time per item on the initial three repetitions of the list and the number of memory verification blocks required to reach criteria. Mean study time per item for the first three list repetitions was analyzed with a 2 (training condition: facts, pairs) × 3 (repetition: 1, 2, and 3) mixed-factorial ANOVA. This analysis revealed a main effect for repetition, F(2, 92) = 26.50, MSE = 7.45, with study times decreasing across the three list repetitions (6.42, 3.11, and 2.74 s, respectively). However, there was no main effect for training condition and no interaction between condition and repetition (Fs < 1.00).

As in Experiment 1, there was no significant difference in the number of verification blocks required by participants in the two training conditions to reach the training criteria (9.6 vs. 11.5 blocks, respectively, ns).

Transfer: Informational hypotheses. Accuracy on the normal addition statements was uniformly high for both conditions for both the trained (.94) and the untrained (.89) statements, and there was no evidence suggesting speed-accuracy trade-offs. Before analysis, each participant’s RT data (for the normal and the subtraction statements) were checked for outliers using the criteria and procedures used in Experiment 1. This resulted in a total of 23 observations (0.3%) being replaced.

Correct response RTs were analyzed with a 2 (training condition: facts, pairs) × 2 (training status: trained, untrained) × 3 (statement type: normal, inverted subtraction, and noninverted subtraction) × 2 (truth value: true, false) × 3 (addend: 2, 3, and 4) mixed-factorial ANOVA. As in Experiment 1, only the initial four presentations of the normal statements were involved to equate the number of presentations across statement types. Predictions for transfer for each of the training conditions are presented in Table 2.

As before, trained participants responded faster (1,877 ms) than did untrained participants (2,161 ms), F(1, 44) = 4.74, MSE = 4,372,782.83. Increasing the value of the digit operand led to overall increases in RTs (1,884, 2,122, and 2,065 ms for statements with operands of 2, 3, and 4, respectively), F(2, 88) = 15.90, MSE = 308,077.04. This effect, however, was less pronounced for the trained participants (1,789, 2,013, and 1,804 ms for statements with operands of 2, 3, and 4, respectively) than it was for the untrained participants (1,979, 2,232, and 2,326 ms for statements with operands of 2, 3, and 4, respectively). As in Experiment 1, there was a pronounced bow in the RT function for trained participants. Finally, although there were differences among the statement types (1,859, 2,013, and 2,183 ms for normal, inverted, and noninverted subtraction statements, respectively), F(2, 88) = 47.51, MSE = 162,440.09, statement type interacted with both training condition and training status, F(2, 88) = 3.56, MSE = 162,440.09.

As can be seen in Figure 3, although both training conditions resulted in reductions in RTs for all three statement types, the difference between trained and untrained participants was most pronounced for the pairs training. For the facts training, trained participants were faster than untrained participants for the normal addition and both types of subtraction statements, but these latter statements produced RTs that were longer than those obtained with the normal addition statements. For the pairs condition, trained participants were faster than untrained participants on all three types of transfer statements. In addition, performance on the inverted subtraction statements following the pairs training was indistinguishable from performance on the normal statements. Performance on the noninverted subtraction statements was reliably slower than that observed for either the normal addition or the inverted subtraction statements in this condition. Thus, although both instance theory and the associative perspective correctly predicted the transfer effects for the facts training, only the associative view predicted the patterns of transfer for the pairs training.

There was also a three-way interaction among statement type, training status, and digit, F(4, 176) = 3.22, MSE = 88,113.21. The form of the interaction is shown in Figure 4, which shows the bow in the RT function across digit

![Figure 3](image-url)
RETRIEVAL PROCESSES IN SKILL

Figure 4. Mean response times (RTs) for the normal, inverted subtraction, and noninverted subtraction statements as a function of digit operand and training status in Experiment 2.

operands for trained participants for all three statement types.

True statements were responded to faster (1,968 ms) than were false statements (2,169 ms), $F(1, 44) = 15.90$, $MSE = 209,767.78$, replicating a difference observed in Experiment 1. In addition, the difference between true and false statements varied across the three types of transfer statements, with the effect being most pronounced for the normal addition (1,768 and 1,950 ms for true and false statements, respectively) and inverted subtraction statements (1,941 and 2,068 ms for true and false statements) and absent for the noninverted subtraction statements (2,186 and 2,180 ms for true and false statements).

Transfer: Process hypotheses. Table 6 shows the mean RTs for each of the conditions for the manipulation of memory and the algorithm. Tables 3 and 4 and Figure 2 show the predictions of the OP network process models for all of the manipulations.

For the manipulation of the algorithm and memory, of the 16 total possible interaction contrasts, 12 were positive, 4 were 0, and none were negative. Excluding the 5 contrasts that involved nonmonotonic changes in RT across factor levels resulted in 11 possible contrasts, 8 of which were positive. Positive interaction contrasts were predicted by all of the models under consideration for this pair of processes.

For the manipulations of the algorithm and comparison processes, of the 16 total possible interaction contrasts, 12 were negative, 3 were 0, and 1 was positive. Of these 16, 5 involved nonmonotonic changes in RT across factor levels, resulting in 11 possible interaction contrasts, of which 7 were negative, 1 was positive, and 3 were 0. These results favor the models that allow retrieval to lead directly to responding (Models 2, 5, and 6) and contradict the predictions of the models that interpose a comparison process between retrieval and responding (Models 1, 3, 4, and 7) or that hold that the algorithm and retrieval cannot be executed in parallel across the course of the trial (Model 7).

For the manipulation of the memory and comparison processes, of the 12 total possible interaction contrasts, 6 were positive, 3 were 0, and 3 were negative. However, one of these contrasts was obtained in the context of nonmonotonic changes in RTs across factor levels. Of the 11 remaining interaction contrasts, 5 were positive, 3 were 0, and 3 were negative. As with the preceding interaction contrasts, these outcomes support those models in which retrieval leads directly to responding (Models 2, 5, and 6) and contradict the models that require the outputs of retrieval to be acted on by a comparison process before responding (Models 1, 3, 4, and 7). They also contradict the model that disallows full concurrency of the algorithm and retrieval (Model 7).

Discussion

The results of Experiment 2 challenge the predictions of both instance theory and the associative view and support the set of process models that was supported in Experiment 1. Both theoretical views failed to predict the positive transfer that was obtained for the noninverted subtraction statements following the facts training. Furthermore, instance theory underpredicted the positive transfer obtained for the inverted subtraction statements after the pairs training. The mean interaction contrasts supported the models (2, 5, and 6) that were supported in Experiment 1. Thus, in spite of a variation in the statement types, a consistent set of conclusions about processing was supported. In particular, the data from Experiment 2 provide additional substantial support for a processing architecture in which retrieval can directly support responding without requiring an intervening comparison process. They also contradict (on a second data set) the hypothesis that algorithmic and retrieval processing cannot occur in parallel across the course of the trial (Rickard, 1997). Instead, models such as the recent extensions of instance theory (e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999), ACT-R, or retrieval in a global memory model such as SAM are supported for a second time.
Table 6
Mean Interaction Contrasts (Equation 1) for the Manipulation of Memory (Trained, Untrained) and the Algorithm (Digit Operand) in Experiment 2

<table>
<thead>
<tr>
<th>Manipulation of memory</th>
<th>True statements addend</th>
<th>False statements addend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Facts condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,352</td>
<td>1,541</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,418</td>
<td>1,731</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>124</td>
<td>495</td>
</tr>
<tr>
<td>Associative condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,323</td>
<td>1,637</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,409</td>
<td>1,870</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>147</td>
<td>326</td>
</tr>
<tr>
<td>Facts condition, noninverted subtraction statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>2,042</td>
<td>2,145</td>
</tr>
<tr>
<td>Untrained</td>
<td>2,094</td>
<td>2,367</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>170</td>
<td>-</td>
</tr>
<tr>
<td>Associative condition, noninverted subtraction statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,875</td>
<td>1,808</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,999</td>
<td>2,117</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>185</td>
<td>-</td>
</tr>
<tr>
<td>Facts condition, inverted subtraction statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,706</td>
<td>1,987</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,802</td>
<td>2,420</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>337</td>
<td>-</td>
</tr>
<tr>
<td>Associative condition, inverted subtraction statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,573</td>
<td>1,593</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,544</td>
<td>2,041</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>477</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. All of the other interaction contrasts can be recovered by reorganizing the data in this table.
*Interaction contrast not significantly different from zero.

Experiment 3

Experiment 3 was intended to extend the findings of Experiments 1 and 2 by determining whether knowledge about the ordinal relationships among stimuli could be used to support transfer. For example, if participants acquire knowledge about the relative numeric values associated with the letters B–G, then it might be possible to obtain positive transfer to addition statements using a letter within this range that was never used as an initial letter during training. Use of this relative knowledge would have implications for the specificity of instance representations and the types of associative relations that may need to be encoded in instances.

The addition statements used to address this question (see Table 1) were constructed by replacing the initial element of the normal addition statements with the second letter from the range. This letter was trained in the set of normal statements as a final element (e.g., $B + 1 = C$) but did not appear as an initial element until the transfer phase. Both views predict (see Table 2) that there will be little if any positive transfer to the new first element statements after the facts training. Also, both predict positive transfer after the pairs training but below the level obtained for the normal statements.

Method

Participants. Forty-eight participants were recruited for Experiment 3. All participants reported normal or corrected-to-normal vision, and none had participated in the previous experiments.\(^\text{18}\)

Design. Experiment 3 was conducted as a 2 (training condition: facts, pairs) $\times$ 2 (transfer condition: trained, untrained) $\times$ 5 (digit operand: 1, 2, 3, 4, and 5) $\times$ 2 (statement type: normal addition, new addition) $\times$ 2 (truth value: true, false) mixed-factorial design, with digit operand, statement type, and truth value manipulated within subjects.

Materials. The normal addition statements in Experiment 3 were identical to those used in Experiment 2, with the addition of an Addend 5 statement. The Addend 1 statements were used only in training to provide participants with exposure to the second letter element of the stimulus range, which would appear in the transfer phase as an initial element. Two types of false statements (with

\(^\text{18}\) Fifty-one participants were tested in total, with the data for 3 participants (all in the facts training condition) discarded using the criteria used in the previous experiments.
answers being off by \( \pm 1 \) were constructed for each Addend 5 statement in each set. Six true new addition statements were constructed using the second letter in the range in each set and adding the digits 2, 3, and 4. For example, if the trained set consisted of the letter \( B \) plus the digits 1–5, then the transfer set consisted of the letter \( C \) plus the digits 2, 3, and 4 (see Table 1). Two false transfer statements were constructed for each true new addition statement by adding and subtracting one from the correct answer.

Apparatus and procedure. The apparatus and procedure used in Experiment 3 were identical to those used in the preceding experiments with the following exceptions. First, participants repeated the memory test a total of four times using the same accuracy and RT criteria used in Experiment 1. Second, the normal addition statements in the transfer phase of the experiment were those involving only addends of 2, 3, 4, and 5. Normal and new addition statements of each type (true, false) were presented the same number of times as in Experiments 1 and 2, resulting in a total of 176 item presentations in the transfer phase.

Results

Training. The mean study times per item for the initial three list repetitions were analyzed with a 2 (training condition: facts, pairs) \( \times 3 \) (repetition: 1, 2, and 3) mixed-factorial ANOVA. Results show a main effect for repetition, \( F(2, 92) = 47.19, MSE = 2.33 \), with study times decreasing across the three list repetitions (5.99, 3.56, and 3.20 s, respectively). There was a main effect for training condition, \( F(1, 46) = 4.14, MSE = 13.96 \), with participants in the facts condition requiring more study time per item (4.88 s) than participants in the pairs condition (3.62 s). As noted earlier, this most likely was the result of the items in the facts condition being more complex than items in the pairs condition. Finally, there was no interaction between repetition and condition, suggesting that the magnitude of the decrease in study time was equivalent for the two training conditions. Participants in the facts condition required 13.3 blocks of verification to reach the training criteria and participants in the pairs condition required 14.2 blocks, a difference that was nonsignificant.

Transfer: Informational hypotheses. There was no evidence of speed-accuracy trade-offs in the data, with the accuracy for normal addition statements being uniformly high for both conditions for both trained (.93) and untrained (.88) statements. As before, each participant’s RT data (for correct responses to each type of statement) were checked for outliers, resulting in a total of 32 observations (0.4% of the total) being replaced.

RTs for correct responses for the initial four item presentations were analyzed with a 2 (training condition: facts, pairs) \( \times 2 \) (training status: trained, untrained) \( \times 3 \) (statement type: normal, new first element) \( \times 2 \) (truth value: true, false) \( \times 3 \) (addend: 2, 3, and 4) mixed-factorial ANOVA. Statements with addends of 5 were excluded from this analysis to allow for comparison of transfer across statement types.

As was the case in Experiments 1 and 2, trained participants responded faster (1,845 ms) than did untrained participants (2,520 ms), \( F(1, 44) = 28.06, MSE = 3,008.630.02 \). Increasing the value of the digit addend had the overall effect of increasing RTs (1,841, 2,200, and 2,464 ms for statements with addends of 2, 3, and 4, respectively), \( F(3, 88) = 110.97, MSE = 144,153.39 \). However, this effect was attenuated in the performance of trained participants (1,626, 1,942, and 2,094 ms for statements with addends of 2, 3, and 4, respectively) relative to the performance of untrained participants (2,055, 2,457, and 2,833 ms for statements with addends of 2, 3, and 4, respectively), \( F(3, 88) = 19.58, MSE = 144,153.39 \).

Overall, participants responded faster to normal addition statements (2,104 ms) than they did to new first element statements (2,288 ms), \( F(1, 44) = 62.77, MSE = 211,949.46 \). However, this main effect needs to be interpreted in the context of a set of interactions. First, there was an interaction between statement type and training status, \( F(1, 44) = 25.48, MSE = 211,949.46 \). Trained participants were slower to respond to the new first element statements (2,054 ms) than they were to the normal statements (1,688 ms). However, their RTs to both of these statements were faster than those of the untrained participants, with there being no difference between the two statement types (2,519 and 2,522 ms) for the untrained participants. Second, there were interactions among statement type, training condition, and addend, \( F(3, 88) = 3.88, MSE = 107,840.98 \), statement type, training status, and addend, \( F(3, 88) = 11.27, MSE = 107,840.98 \), and among statement type, training condition, training status, and addend, \( F(3, 88) = 3.58, MSE = 107,840.98 \).

The form of this four-way interaction is presented in Figure 5. The typical pattern associated with trained versus untrained participants can be seen for the facts training condition with the normal statements and for the pairs training with the normal and new first element statements. Although the facts training did reduce the overall RTs for new first element statements, it did not attenuate the effect of the addend. The pairs training also reduced RTs on the new first element statements and attenuated the effect of addend on RTs. Still, RTs for these statements were slower than those for the normal statements. These interactions suggest that although there was some positive transfer to the new first element statements in both training conditions, more transfer was obtained for the pairs training than for the facts training. Although both theoretical views (see Table 2) predicted positive transfer after the pairs training, neither predicted any positive transfer after the facts training. The final effect of interest in this analysis was one obtained in the preceding experiments. Specifically, true statements were responded to more quickly (1,951 ms) than false statements (2,104 ms), \( F(1, 44) = 8.48, MSE = 141,354.26 \).

Transfer: Process hypotheses. Table 7 shows the means for the manipulation of memory and the algorithm. Tables 3 and 4 and Figure 2 show the predictions for all of the manipulations.

For the manipulation of the algorithm and memory, of the 20 total possible mean interaction contrasts, 19 were positive and 1 was negative. Excluding the 4 interaction contrasts that involved nonmonotonic changes across factor levels left 16 possible interaction contrasts, of which 11 were positive and 1 was negative. As noted earlier, all seven models predict overadditive means interaction contrasts.
For the manipulation of the algorithmic and comparison processes, of the 20 total possible interaction contrasts, 18 were negative, 2 were 0, and none were positive. Two of these interaction contrasts involved nonmonotonic changes in RTs across factor levels; of the remaining 18, 16 were negative and 2 were 0. These outcomes are consistent with the predictions of Models 2, 5, and 6 but contradict the predictions of Models 1, 3, 4, and 7, repeating the patterns found for this pair of processes in Experiments 1 and 2.

Finally, for the manipulation of memory and the comparison process, of the 14 possible interaction contrasts, 10 were positive, 4 were 0, and none were negative. This pattern of outcomes replicates those observed for the same manipulations in Experiments 1 and 2, being consistent with the predictions of Models 2, 5, and 6 and contradicting the predictions of Models 1, 3, 4, and 7.

Discussion

The results of Experiment 3 provide another set of challenges to the informational predictions of both theoretical views. Although both views correctly predicted reduced positive transfer for the new first element statements (relative to the normal statements) after the pairs training, neither correctly predicted the low but significant level of transfer obtained for these statements after the facts training. This would seem to indicate that some sort of relative knowledge was available after memorization of alphabet arithmetic facts.

The mean interaction contrasts in Experiment 3 are consistent with the patterns observed in the first two experiments. Specifically, and in the context of a new set of transfer stimuli, they support the predictions of a process architecture in which retrieval directly supports responding, contradict the predictions of models that require a comparison operation to be interposed between retrieval and responding, and contradict the predictions of the model that disallow simultaneous execution of the algorithm and retrieval across the complete course of a trial.

Experiment 4

A factor that could potentially compromise the interpretation of the results of Experiments 1–3 is the lack of difference between the two training conditions on the normal statements. This equivalence may be due in part to the fact that participants may not have been naive with respect to letter–number relations, leading to more positive transfer for the pairs training than would exist otherwise. That is, participants (college students) have a great deal of experience with the letters of the alphabet (at least the initial six) and the ordinal relations among them. It is possible that participants in the pairs condition were able to make use of their preexisting knowledge to aid them in performing the task and that this preexisting knowledge was important in supporting positive transfer. Experiment 4 tested this possibility by using materials that carried no preexisting associative content.

Method

Participants. A total of 48 participants contributed data to the analyses of Experiment 4. All participants reported normal or corrected-to-normal vision, and none had participated in the previous experiments.19

19 Seventy-seven participants were tested in total, with the data for 29 participants (distributed approximately equally across the two conditions) discarded according to the criteria used in the preceding experiments. This dramatic increase in the rate of participant discards may indicate that the stimuli used in Experiment 4 were much more difficult to learn than the stimuli used in
**Table 7**

*Mean Interaction Contrasts (Equation 1) for the Manipulation of Memory (Trained, Untrained) and the Algorithm (Digit Operand) in Experiment 3*

<table>
<thead>
<tr>
<th>Manipulation of memory</th>
<th>True statements addend</th>
<th>False statements addend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Facts condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,186</td>
<td>1,504</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,660</td>
<td>2,174</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>196</td>
<td>358</td>
</tr>
<tr>
<td>Associative condition, normal statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,376</td>
<td>1,698</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,580</td>
<td>2,146</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>244</td>
<td>615</td>
</tr>
<tr>
<td>Facts condition, transfer statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,567</td>
<td>2,010</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,943</td>
<td>2,464</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>78</td>
<td>-110</td>
</tr>
<tr>
<td>Associative condition, transfer statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trained</td>
<td>1,788</td>
<td>2,010</td>
</tr>
<tr>
<td>Untrained</td>
<td>1,927</td>
<td>2,310</td>
</tr>
<tr>
<td>$\Delta^2$</td>
<td>161</td>
<td>485</td>
</tr>
</tbody>
</table>

*Note.* All of the other interaction contrasts can be recovered by reorganizing the data in this table.

**Design.** Experiment 4 was conducted as a 2 (training condition: facts, pairs) × 2 (transfer condition: trained, untrained) × 4 (digit operand: 2, 3, 4, 5) × 2 (truth value: true, false) mixed-factorial design, with digit operand and truth value manipulated within subject. Experiment 4 did not require transfer statements other than the normal addition statements, as the question of interest was specific to the effects of the two training conditions on performance with the normal addition statements.

**Materials.** Table 1 shows examples of the types of addition statements that were used in Experiment 4. An arbitrary set of unfamiliar symbols was drawn from the DOS character graphics symbol set (ASCII Codes 1-254) and placed in an arbitrary order. Symbols were selected to have minimal associative relations or content; a similar use of these types of materials can be found in Fisk et al. (1988) and Blessing and Anderson (1996).

Two sets of eight true addition statements were formed, with the first set using the initial symbol and four symbols at distances of 2, 3, 4, and 5 and the second set using the second symbol and four symbols at distances of 2, 3, 4, and 5. Two false addition statements were constructed for each true statement by replacing the symbol in the final position with a symbol that was incorrect by ±1 position.

**Apparatus and procedure.** The procedure used in Experiment 4 was identical to that used in Experiment 3, with the following modifications. First, stimuli were presented on a color monitor and elements were presented using the 40-column character size. This was done to make the symbols clearly visible. All stimuli were presented in white on a black background. The number of item presentations was also reduced relative to in Experiments 1-3. If the presence of preexisting associative knowledge in the alphabet task led to the lack of differences between training conditions, then removing the preexisting associative component should allow any differences between the conditions to be apparent within the first four presentations. Thus, each true statement was presented four times during the transfer phase, whereas each of the two types of false statements was also presented four times, resulting in a total of 64 item presentations in the transfer phase.

**Results**

**Training.** The mean study times per item for the initial three list repetitions were analyzed with a 2 (training condition: facts, pairs) × 3 (repetition: 1, 2, and 3) mixed-factorial ANOVA. As in Experiments 1–3, there was a main effect for repetition, $F(2, 92) = 73.03, MSE = 8,977,070$, with study time decreasing across the first two list repetitions (11.5, 5.0, and 5.2 s, respectively). Both the main effect for training condition and the interaction between condition and repetition failed to reach standard levels of significance ($Fs < 1.00$). There was no difference between the two training conditions in the number of verification blocks required to meet the RT and accuracy criteria, with participants in the facts condition requiring 12.5 blocks and participants in the pairs condition requiring 11.0 blocks. These means are consistent with the mean number of blocks required in the previous experiments.

**Transfer: Informational hypotheses.** Accuracy for the symbol addition statements was high for both training conditions for both trained (.93) and untrained (.93) statements. As was the case for the alphabet arithmetic statements used in Experiments 1–3, there was no evidence suggesting any speed–accuracy trade-offs. Each participant’s RT data (for correct responses) were checked for experiments 1–3. Although it is possible that a selection bias might have been operating, the amount of training required by retained participants was equivalent to that required by participants in the preceding experiments, suggesting that the general implications of these data may not be compromised.
outliers, with a total of 34 observations (0.6% of the total) being replaced.

The RT data for correct responses were analyzed with a 2 (training condition: facts, pairs) × 2 (training status: trained, untrained) × 2 (truth value: true, false) × 4 (addend: 2, 3, 4, and 5) mixed-factorial ANOVA. The analysis revealed no main effect for training condition (F < 1.00). There was a main effect for training status, F(1, 44) = 106.25, MSE = 408,601.15, and no interaction between training condition and training status (F < 1.00). Trained participants responded faster (1,660 ms) than did untrained participants (2,341 ms), with the advantage being independent of the type of training.

As in the alphabet tasks, there was a main effect for digit addend, F(3, 132) = 35.80, MSE = 38,406.07, and an interaction between digit addend and training status, F(3, 132) = 41.96, MSE = 38,406.07. Overall, increasing the digit addend produced an increase in RT, but this increase was attenuated in the trained conditions (1,639, 1,681, 1,728, and 1,593 ms for statements with addends of 2, 3, 4, and 5, respectively) relative to the untrained conditions (2,054, 2,244, 2,442, and 2,622 ms for statements with addends of 2, 3, 4, and 5, respectively). In summary, both training conditions resulted in positive transfer to the symbol statements, and they did so to equivalent degrees. This is consistent with the predictions of both theoretical perspectives.

The final significant result was the main effect for true-false status, F(1, 44) = 48.27, MSE = 32,997.19, with true statements producing shorter RTs (1,943 ms) than false statements (2,173 ms).

Transfer: Process hypotheses. Table 8 shows the mean RTs in each condition of Experiment 4, organized according to the manipulation of memory and the algorithm. The predictions for all of the manipulations are presented in Tables 3 and 4 and Figure 2.

For the manipulation of memory and the algorithm, all of the 12 possible interaction contrasts were positive. Of the 9 interaction contrasts that involved monotonic changes in RT across factor levels, 5 were positive and three were 0. These outcomes were consistent with the predictions of the models that allow retrieval to directly support responding (Models 2, 5, and 6), contradict the predictions of the models that require the outputs of retrieval to be passed to the comparison process before responding (Models 1, 3, 4, and 7), and contradict the predictions of the model that disallow concurrent execution of the algorithm and retrieval across the course of the trial (Model 7).

Consistent with this conclusion were the data for the manipulation of the memory and comparison processes. Of the eight possible interaction contrasts (all of which involved monotonic changes in RT across factor levels), five were positive and three were 0. These outcomes were consistent with the predictions of the models that allow retrieval to directly support responding (Models 2, 5, and 6), contradict the predictions of the models that require retrieval to be sequential with a comparison process (Models 1, 3, 4, and 7), and contradict the predictions of the model in which retrieval and the algorithm are concurrent for only the initial portion of the trial (Model 7).

Discussion

The results of Experiment 4 provide strong evidence against the notion that preexisting associative knowledge was the source of the consistent lack of difference between the facts and pairs training conditions in Experiments 1–3. In fact, the lack of a difference in Experiment 4, an experiment in which there was almost no chance for preexisting associative knowledge to be at work, was observable given only four item presentations. Overall, the results of Experiment 4 were similar to those observed in the previous experiments, which suggests that the evidence supporting the predictions of the associative view was not conditional on the presence of preexisting knowledge.

As with the reversed addition statements in Experiment 2, the symbol statements provided a point of difference for the

<table>
<thead>
<tr>
<th>Table 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Interaction Contrasts (Equation 1) for the Manipulation of Memory (Trained, Untrained) and the Algorithm (Digit Operand) in Experiment 4</td>
</tr>
<tr>
<td>Manipulation of memory</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Facts condition, normal statements</td>
</tr>
<tr>
<td>Trained</td>
</tr>
<tr>
<td>Untrained</td>
</tr>
<tr>
<td>Δ²</td>
</tr>
<tr>
<td>Associative condition, normal</td>
</tr>
<tr>
<td>statements</td>
</tr>
<tr>
<td>Trained</td>
</tr>
<tr>
<td>Untrained</td>
</tr>
<tr>
<td>Δ²</td>
</tr>
</tbody>
</table>

Note. All of the other interaction contrasts can be recovered by reorganizing the data in this table.
predictions of the two theoretical views. Both instance theory and the associative perspective predicted positive transfer for both training conditions but differed in terms in terms of the magnitude of positive transfer that should be observed. Instance theory predicted less positive transfer for the pairs than for the facts training, whereas the associative view predicted equivalent positive transfer for the two conditions. In actuality, both training conditions produced substantial and equivalent positive transfer.

The pattern of interaction contrasts obtained in Experiment 4 parallels the patterns observed in Experiments 1–3 and supports the conclusions regarding processing obtained in those earlier experiments, even with a reasonably dramatic change in stimuli. That is, the data supported those models that allow retrieval to directly support responding, contradicted those models that require a comparison operation on the outputs of retrieval, and contradicted the model that disallows concurrent execution of the algorithm and memory retrieval across the course of a trial.

General Discussion

The results of the four experiments suggest that both instance theory and the associative perspective fall short of being able to account for the range and ordering of transfer effects. These results are particularly informative with respect to the types of information that may be available in skilled performance. In addition, the consistent pattern of outcomes obtained for the interaction contrasts across the four experiments lends strong support to one class of process models, a class that comprises the general process assumptions of recent extensions of instance theory (e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999), a set of associative alternatives (e.g., Anderson, 1993; Anderson et al., 1997; Raaijmakers & Shiffrin, 1981), and more general conceptions of memory retrieval (e.g., Ratcliff, 1978).

Consider first the results pertinent to the informational issues. With respect to the effects of the facts training, instance theory successfully predicted the positive transfer for the normal and symbol addition statements and the robust but reduced (relative to the normal statements) positive transfer for the inverted subtraction statements. However, although instance theory predicted no positive transfer to the noninverted subtraction and new first element addition statements, reliable positive transfer was obtained for these statements but at a level below that observed for the normal statements. With respect to the pairs training, instance theory accurately predicted positive transfer in each of the cases in which it was observed. However, it did not predict the obtained ordering of positive transfer across statement types. Specifically, instance theory predicted equivalent levels of positive transfer across all statement types, as the retrieval operations should have been the same for each of these stimuli. Instead, the pairs training led to high levels of positive transfer for the normal, symbol, and reversed addition statements, and the inverted subtraction statements, and robust but reduced transfer for the noninverted subtraction and new first element addition statements.

The associative approach fared better in its informational predictions, yet it failed to predict all of the observed effects. With respect to the effects of the facts training, the associative view failed to predict the positive transfer obtained for the noninverted subtraction and new first element addition statements. With respect to the pairs training, the associative view accurately predicted the levels and ordering of positive transfer across the four experiments.

Both views also failed to predict the reliable effects of the end points observed in all four experiments. Instance theory cannot account for this effect because the end point items were presented the same number of times as the other items. Some additional information seems to be implicated as being available, information that would need to be specified in the propositional representation of the relevant instances and that would be operative independent of the training condition. The associative view cannot account for end point effects without assuming an additional model-specific memory mechanism (e.g., the rehearsal buffer in SAM and associated strengthening operations as a function of number of items in the buffer). These end point effects have additional implications relative to the process models and are discussed further below.

The results relative to the process models were regular (see the summary of outcomes in Figure 2) and consistently supported one class of models. The models that allow for parallel operation of the algorithm and memory retrieval across the course of the trial, with the algorithm operating sequentially with a comparison operation and memory retrieval being able to directly support responding, were consistently supported. The models that required a comparison operation to follow memory retrieval and the model that disallowed the concurrent operation of the algorithm and memory retrieval across the entire trial were consistently contradicted.

Recall that earlier versions of instance theory required a selection or comparison process of some kind, in sequence with retrieval, to delimit the set of instances that would be involved in the retrieval race. That is, the theory required some type of mechanism to restrict the race to those instances relevant to the particular task; otherwise, all instances in memory could participate. Thus, the basic processing assumptions of earlier versions of instance theory seem inadequate, as does the associative model for arithmetic skill proposed by Ashcraft (e.g., 1982, 1987).

In contrast, these findings support mechanisms such as the process structure described by Zbrodoff and Logan (1990), the direct retrieval mechanism (the exemplar-based random walk) proposed by Nosofsky and Palmeri (1997; Palmeri, 1997, 1999), and the sampling and recovery processes within a global associative memory model such as SAM (e.g., Raaijmakers & Shiffrin, 1981; see also Busey, in press). Specifically, the random-walk mechanism of the exemplar-based random walk encapsulates the comparison operation within the retrieval process (T. J. Palmeri, personal communication, September 7, 1998), and the sampling and recovery mechanisms in SAM do not require the type of threshold comparison mechanism required in familiarity-based retrieval; instead, sampling and recovery can directly support responding without a general comparison process.
These findings are also congruent with the ACT-R notion that practice results in the compilation of knowledge, with the general assumptions of Schneider's model for skilled performance (e.g., Schneider & Detweiler, 1988) and one of the two types of retrieval processes proposed by Campbell and colleagues (e.g., Campbell & Tarling, 1996; Meagher & Campbell, 1995).

The interaction-contrasts also consistently contradicted the predictions of the simplified version of Rickard's (1997) alternative to instance theory. It should be emphasized, however, that the predictions were derived for a simplified version of that model, one that captured the essential assumption of an initial strategy selection. It remains to be seen whether either the feedback mechanisms or the continuous flow of activation in the full model would lead to fundamentally different predictions for the process manipulations used here.

The results of the interaction contrasts do, however, indicate that the present set of process models may be in need of critical embellishments. Specifically, the regular finding of end point effects suggests memory information or processes that are not accounted for by the current set of process models and thus represents a significant challenge for model refinement.

Whereas this challenge, and the others represented by the findings reported here, can and should be addressed within the context of a specific model (e.g., Logan, 1988b; Palmeri, 1997; Rickard, 1997), the present research suggests the utility of examining the questions in a more general and less model dependent manner. That is, it is possible to examine competing hypotheses (and models) in terms of their general characteristics and to construct tests that address these general characteristics and place constraints on all of the possible candidates (see, e.g., Ratcliff, Shu, & Gronlund, 1992). For example, any instantiation of a process architecture (e.g., one based on counters, random walks, continuous flow networks, etc.) that corresponds to the assumptions realized in Models 2, 5, and 6 would be supported by the current results. Conversely, any instantiation of the process architectures represented by Models 1, 3, 4, and 7 should be contradicted by the current results.

One of the most obvious challenges to be faced comes from the fact that although the models and manipulations used in the present research address the processing architectures within which retrieval operates, they were not analytic with respect to the retrieval process itself. For example, it is impossible to point to evidence in these experiments that might differentially support either sequential or concurrent retrieval processing of the component associations in the pairs condition and, as a consequence, further restrict the class of candidate process models. However, note that the present approach can be easily extended to address this question.

Another challenge comes from the fact that, in general, the levels of practice used in these experiments are far below what are normally used in studies of skilled performance, simple or otherwise (e.g., Chase & Ericsson, 1981; Wenger & Payne, 1995). Thus, it will be important to show that the conclusions drawn here, particularly with respect to process architectures, hold under conditions involving much greater practice (e.g., Logan, 1988b, 1992). This would also enable use of the more elegant tools of contrasts at the level of RT distributions (e.g., Townsend, 1990b; Townsend & Nozawa, 1995; Wenger & Townsend, in press-b) and allow consideration of the process contrasts used here within the data of a single participant, allowing for potentially much stronger inferences and conclusions.

The aspects of processing considered here really represent only a subset of the important issues that can and should be addressed with respect to the retrieval processes involved in skilled performance. Of particular importance would be the capacity of the retrieval system (see Townsend & Ashby, 1978; Townsend & Nozawa, 1995; Wenger & Townsend, in press-a). Instance theory, for example, assumes that retrieval is either an unlimited or supercapacity process at the level of the individual instances. Furthermore, the knowledge compilation process within ACT-R implies either unlimited capacity or supercapacity processing, although specification of the particular levels and types of representation important at different levels of practice would be necessary to derive a more precise characterization. Such general questions can be posed in experimental paradigms that allow for strong inferences and that are compatible with the types of tasks used in these and preceding studies (see, in particular, Townsend & Nozawa, 1995; Wenger & Townsend, in press-b). Answers to questions such as these can only help refine conceptualization of the role of memory information and processes in cognitive skill.

References


Appendix

Mean Processing Times for the Order-of-Processing Network Models

The expected processing times for each of the models shown schematically in Figure 1 were developed according to the methods described by Schweickert, Fisher, and Goldstein (1992). Then, under the assumption of selective influence (e.g., Schweickert & Townsend, 1989; Townsend & Schweickert, 1989), predictions for the effects of pairwise factorial manipulations of a subset of the component processes were obtained (Equation 1).

For the models developed for the facts condition (Models 1 and 2), three component processes were required: an algorithmic process \( a \) with (random) duration \( T_a \); a comparison process \( c \) with duration \( T_c \); and a memory retrieval process \( m \), with duration \( T_m \). For the models developed for the pairs condition (Models 3–6), five component processes were required: an algorithmic process \( a \) with duration \( T_a \); a comparison process \( c \) with duration \( T_c \), two retrieval processes \( x, y \) with durations \( T_x, T_y \); and a translation (i.e., an additional retrieval) process \( d \), with duration \( T_d \). For Model 7, the simplification of Rickard’s (1997) model, the initial steps of the algorithm were represented by process \( a \) with duration \( T_a \); the initial steps of memory retrieval were represented by process \( m \), with duration \( T_m \); the remaining steps of the algorithm and memory retrieval were represented by processes \( a' \) and \( m' \), with durations \( T_{a'} \) and \( T_{m'} \); respectively; and the comparison operation was represented by process \( c \), with duration \( T_c \). Using these component processes, the schematic models in Figure 1 were represented as order-of-processing (OP) network models using procedures described in detail elsewhere (e.g., Fisher & Goldstein, 1983; Goldstein & Fisher, 1991, 1992; Schweickert et al., 1992).

For each of the resulting representations, a complete set of paths (from the presentation of the stimulus to the selection of the response) was specified. From these specifications, it was possible to (a) derive the expression for the mean response time (RT) and (b) derive the predicted sign for the mean interaction contrast (Equation 1). An outline of the approach follows, with the description based on that developed by Fisher and Goldstein and presented in Schweickert et al. (1992). Let \( H \) be the complete set of paths for a particular model and \(|H| \) be the total number of different paths in that model. Then, letting \( P(P = h) \) be the probability that path \( h \) was taken in that model, the expected value of the processing time, \( E(RT) \), for the model can be expressed as

\[
E(RT) = \sum_{j=1}^{|H|} E(RT|P = h)P(P = h). \tag{A1}
\]

Let \( X \) be a vector of process durations for a particular model (e.g., the durations of that model’s component processes) and let \( f_x(x_1, \ldots, x_1^{|x|}) \) be the joint density for those durations. For present purposes, both within- and between-states independence of the component processes were assumed, without any assumptions about the specific form of those component distributions, other than assuming a lower bound for each of zero. Now let \( h \) index a particular path in a model and \( j \) indicate the \( j \)th state on any path. Let \( x_h \) be the duration of the process that completes first in any state \( x \) (see Fisher & Goldstein, 1983, and Goldstein & Fisher, 1991, 1992, for details on state representations in OP network models). Let \( T_{hi} \) be the duration of the \( j \)th state on path \( h \). As described by...
Schweickert et al. (1992), this is equal to the duration of the process that completes first in that state minus the duration of each of the preceding states in which it was active. Then (see Schweickert et al., 1992, Equation 10),

\[
T_{hl} = \begin{cases} 
  x_{hl} & j = 1 \\
  x_{hl} - \sum_{i=1}^{j-1} a_{ij} T_{hi} & j > 1,
\end{cases}
\]  

(A2)

where \( a_{ij} = 1 \) if \( x_{hl} \) is current in state \( S_h \); otherwise it is zero. Finally, let

\[
g_h = t_{h1} + t_{h2} + \cdots + t_{hh'},
\]

(A3)

where \( h' \) indexes the last state on path \( h \) prior to the finish state. Then (Schweickert et al., 1992, Equation 11), the expected value for the total task time \( \langle E[RT]\rangle \) can be written as

\[
E(RT) = \sum_{k=1}^{N} \left| \int \cdots \int g_{kl}(x_1, \ldots, x_N) dx_1 \cdots dx_N \right|.
\]

(A4)

where \( R_k \) is the region in which the durations of all of the states on path \( h \) are positive. The mean RT for each of the models was then expressed as a linear combination of terms containing the densities and cumulative distribution functions (CDFs) of the various component processes.

The analysis of the models with respect to the means and mean interaction contrast (Equation 1) was grounded on a small set of assumptions about the ordering of the process distribution functions. First, ordering at the level of the CDF, \( F(t) \), was assumed, such that for any process \( x \), with Levels 1 and 2, where the process at Level 1 was assumed to be slower than at Level 2, or \( F_x(t) < F_x(t) \) for all \( t > 0 \). An ordering at this level ensures an ordering at the level of the means; the reverse, however, is not guaranteed (see Townsend, 1990a). Second, it was assumed that the densities for any process at the two levels will cross exactly once and in a nontrivial manner (see Townsend & Nozawa, 1995). That is, there exists some time \( t^* \), such that for all \( t < t^* \), \( f_2(t) > f_1(t) \), with \( f_2(t) < f_1(t) \) for all \( t > t^* \). As noted elsewhere (Townsend, 1990a; Townsend & Ashby, 1978; Townsend & Nozawa, 1995), this property implies an ordering in the survivor functions, \( S_2(t) > S_1(t) \), which in turn implies \( F_2(t) < F_1(t) \) for all \( t > 0 \).

A small set of simplifying assumptions was possible specific to the experiments in the present research. First, although Models 3–6 distinguished between concurrent and sequential retrievals \((x, y)\), the experiments did not involve factorial manipulation of the component associations. Consequently, it was possible to let \( z = \max(x, y) \) in Models 3 and 5 and \( z^* = x + y \) in Models 4 and 6. Second, although these models specified a translation process \( d \) (i.e., a second retrieval process), this process was not manipulated in any of the experiments. Consequently, an additional simplification was allowed. Letting \( q = z + d \) and \( q^* = z^* + d \), an ordering at the levels of \( F_q\) and \( F_{q^*} \) was assumed. Finally, to derive predictions for Model 7, it was assumed that manipulation of the algorithm selectively affected process \( a' \), implying that the duration of process \( a' \) (the initial step of the algorithm) was invariant across levels of \( a' \) (the remaining steps of the algorithm). It was also assumed that manipulation of memory retrieval affected both \( m \) and \( m^* \).

Table 8 shows the resulting OP network and models in terms of components of the mean RT (see Equation A4) and the predictions for the mean interaction contrast (Equation 1). As can be seen in the table, given the preceding simplifications, Models 1, 3, 4, and 7 make essentially identical predictions, as do Models 2, 5, and 6; however, the two groups of models make distinctly different predictions from each other.

Received March 16, 1998
Revision received October 15, 1998
Accepted March 30, 1999