

# Strategy transitions during cognitive skill learning in younger and older adults: Effects of interitem confusability

ANDREA S. WHITE

*Kenyon College, Gambier, Ohio*

AND

JOHN CERELLA AND WILLIAM J. HOYER

*Syracuse University, Syracuse, New York*

Groups of young and old adults were trained for four sessions on a set of 24 alphabet-arithmetic problems. Problem sets were either highly confusable or highly distinct. Power-function and mixture-model fits to the means and standard deviations of the acquisition data, resolved at the participant problem level, were compared. “Shallow” power functions signaled that a problem was computed throughout training; “humped” mixture functions signaled a shift from slow computed solutions to fast retrieved solutions. Not surprisingly, shifts to retrieval occurred later for confusable problems, but there were also fewer shifts in that condition. Failures to shift, even after extended practice, suggest that retrieving problem solutions is an elective strategy, and not an automatic concomitant of skill training. Participants can be viewed as choosing between strategies that trade off benefits in speed against costs in accuracy. Older adults showed few retrieval solutions in either condition, perhaps because of their emphasis on accuracy.

Fluency in carrying out cognitive skills is often based on a shift from the application of a rule or algorithm to direct retrieval of the solutions to familiar problems. Three influential theories, the instance theory of Logan (1988), the component power laws (CMPL) theory of Rickard (1997), and the source activation confusion theory (SAC) of Schunn, Reder, Nhouyvanisvong, Richards, and Stroffolino (1997), portray this process as primarily a form of incidental learning, with answers being recorded in associative memory as an automatic concomitant of repetitions of problems and their solutions during rule execution. Once the memory representation of a repeated problem is sufficiently strong, a retrieval solution will occur in lieu of computation.

These are essentially “bottom-up” models of learning, and have proven to be both parsimonious and powerful; however, there is evidence to suggest that “top-down” factors influence skill acquisition. Such evidence challenges the completeness of the bottom-up models. For example, not all participants demonstrate a shift from computed solutions to retrieval solutions with training. Ackerman and Woltz (1994, Experiment 1) found that one third of their participants (“nonretrievers”) continued to look up answers in a table of noun–noun pairs after training. The split outcome suggested that “To retrieve or to compute?” was a strategic question—a question that individuals exposed to the same training regimen might answer differently.

In support of this view, Ackerman and Woltz (1994, Experiment 3) trained another group on the same list, and administered an explicit memory test five times over the course of training. The percentage of nonretrievers dropped to 20% in this condition. Ackerman and Woltz argued that memory testing influenced strategy choice by heightening awareness of the retrieval option. Participants who became aware of the retrieval option chose it because of the savings in effort it represented, compared with table lookups. Although memory-trace formation may be an inevitable consequence of repetition training, the shift to retrieval solutions was an elective strategy controlled by the participant.

Results of Experiment 2 in Ackerman and Woltz (1994) pointed to a similar conclusion. The noun–noun pairs used in Experiment 1 were confusable (i.e., all cue words were drawn from one semantic category and all response words from another). In Experiment 2, the words were distinct (i.e., drawn from different categories). Trained on the easy different-categories list, only 3% of the participants remained nonretrievers, compared with 33% on the hard same-categories list. This result can be accounted for by assuming that retrieval of confusable associates was error prone. When participants weighed the benefits in speed and the costs in accuracy, most continued to look up pairs from the hard list instead of shifting to retrieval. Memory based responding was a strategy choice.

---

W. J. Hoyer, [wjhoyer@syr.edu](mailto:wjhoyer@syr.edu)

---

The present study examines the “elective” character of retrieval solutions. We revisit Ackerman and Woltz’s (1994) finding that responses to confusable word pairs were less likely to be memory based than were responses to distinct word pairs. It is possible that this finding reflects a strategy choice not to use memory retrieval for hard items. However, the result can be more simply attributed to the difficulty that hard items imposed on the processing resources of the associative memory system. Memory-trace formation and/or utilization may have been automatic or obligatory in both lists, but was simply less successful in the hard list. Zbrodoff (1995) interpreted performance differences in easy-and-hard alphabet arithmetic problems in just this way; Zbrodoff’s (1995) work is discussed below (see Discussion).

In the experiment reported here, we attempted to sort through these alternatives. Participants were repeatedly required to solve a fixed set of alphabet-arithmetic problems over many blocks of training. This task has been used extensively in laboratory studies of skill formation (Logan, 1988; Rickard, 1997). Problems take the form of equations in artificial arithmetic, such as “ $S + 4 = W$ .” Participants must count letters to determine whether an equation is true or false (see Method); with continued training, participants may remember the answer to particular equations and respond without counting. We looked for performance differences on two sets of problems. One experimental group was trained on a distinct-equation set in which each leading letter was unique, and another group was trained on a confusable-equation set, in which several equations shared a leading letter.

From the performance data we extracted two measures: (1) Whether a given item or participant exhibited a shift from computation to retrieval over the course of training or remained compute-only throughout; and (2) the timing of the compute-to-retrieve transition for those items and/or participants that did exhibit a shift. Thus (1) measures the frequency of retrieval strategies at the item-participant level, and (2) measures the ease of retrieval. These two measures were confounded in the Ackerman and Woltz (1994) studies. Participants would be misclassified as “computing only” if they shifted to retrieving late in training—too late to bring down their mean response time (RT), averaged over five sessions of training, below the 1-sec cutoff used by Ackerman and Woltz to define “retrievers.”

It seems to us that the critical desideratum regarding the strategic character of the memory strategy is the frequency with which it occurs in hard and easy conditions, not its early or late emergence. Our analysis separated the two outcomes, counting occurrences of the retrieval strategy independently of their precocity. If retrievals are elective, fewer participants and/or items will exhibit the strategy on a hard list; conversely, if retrieval is an intrinsic concomitant of repetition training, its frequency will not change with list difficulty (although the shift to retrieval may occur later in training). Participants in our study were trained for a total of 96 blocks, scheduled over four sessions to allow ample opportunity for items to be learned. It turned out that all retrieval shifts occurred in the first two

of the four sessions, validating our pilot observations that learning occurs relatively early in training on alphabet-arithmetic problems.

Both younger and older adults were tested because of the known negative effects of aging on the frequency of retrieval solutions over the course of skill training (e.g., Hoyer, Cerella, & Onyper, 2003; Rogers & Gilbert, 1997; Rogers, Hertzog, & Fisk, 2000; Tournon & Hertzog, 2004a, 2004b; Tournon, Hoyer, & Cerella, 2004). Having data from both younger and older adults provided an opportunity to examine the effects of list difficulty on participants who differed in their baseline levels of retrieval.

## METHOD

### Participants

Sixteen younger adults (8 men and 8 women) and 16 older adults (8 men and 8 women) participated. Younger adults (mean age, 19.96; age range, 18–26) were recruited from the human subjects pool at Syracuse University. Older adults (mean age, 69.16 years; age range, 61–75) were recruited from community organizations in central New York State. Participants who rated their health as good or excellent and who had 20/30 visual acuity or better (corrected) were eligible to participate. Several brief tests of memory, arithmetic skill, and perceptual speed were administered for the purpose of sample description; scores for these measures and for other characteristics of the samples are reported in Table 1.

### Stimuli and Procedure

The task involved the verification of alphabet-arithmetic problems such as “ $S + 4 = W$ .” The example gives a true equation, which can be ascertained by counting through four letters of the alphabet, starting with S: S–T–U–V–W. The final letter, W, agrees with the answer given in the problem; therefore, the correct response to this problem is “true.” “ $U + 3 = Y$ ” is an example of a false problem; the correct answer would be “X.” There were 24 problems altogether, 12 true and 12 false. Two lists of 24 problems were constructed; one list was composed of distinct problems, in which each true equation

**Table 1**  
Means (and Standard Deviations) for Measures of the Characteristics of the Research Participants

Measure	Younger Adults		Older Adults	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age (years)*	19.96	2.0	69.16	3.6
Education*	12.97	0.9	15.1	2.5
Health	1.6	0.5	1.7	0.7
Limitations	1.2	0.6	1.2	0.7
Digit span	17.63	3.5	16.43	3.4
Digit symbol*	70.70	10.7	53.13	8.7
Number comparisons*	54.6	8.9	45.5	9.0
Number calculation*	26.46	8.4	40.3	8.9

Note—Sixteen younger adults and 16 older adults participated. Education, self-reported number of years of formal education; older adults had more years of education than younger adults [ $F(1,31) = 12.04$ ,  $MS_e = 3.661$ ,  $p = .0012$ ]. Health, self-reported, using a scale from 1 (*excellent*) to 5 (*poor*). Limitations, self-reported number of health-related limitations. Digit span combined the forward span and backward span scores for the WAIS-R digit span (Wechsler, 1981). Digit symbol, score on WAIS-R digit symbol substitution subtest (Wechsler, 1981). Number comparisons, number of correct minus incorrect comparisons completed on both parts of the number comparisons test from the *Kit of Factor-Referenced Cognitive Tests* (Ekstrom, French, Harman, & Dermen, 1976). Number calculation, score on the number subtest of the *Primary Abilities Test* (Thurstone & Thurstone, 1949). \* $p < .01$ .

began with a different initial letter. The other list was composed of confusable problems, in which any true equation had the same initial letter as another true equation. For both lists, false equations were created from true equations by changing the correct answer to a letter one position further in the alphabet.<sup>1</sup> In each list, there were four true and four false equations at each of three addends (3, 4, and 5).

Eight participants in each age group were assigned to each list condition. Participants were tested individually in four sessions on consecutive days. The 24 problems of a list were presented one by one in random order in a block of 24 trials. Each session consisted of 24 blocks (576 trials), for a total of 96 blocks (2,304 trials) over the four sessions. Each trial consisted of the presentation of a fixation cross in the center of the screen for 500 msec, followed by the presentation of an alphabet-arithmetic equation. A true/false keypress response terminated the display, and was followed by a 1,000-msec intertrial interval. Incorrect responses elicited a brief low tone. On-screen and oral instructions emphasized both speed and accuracy of responding.

### RESULTS

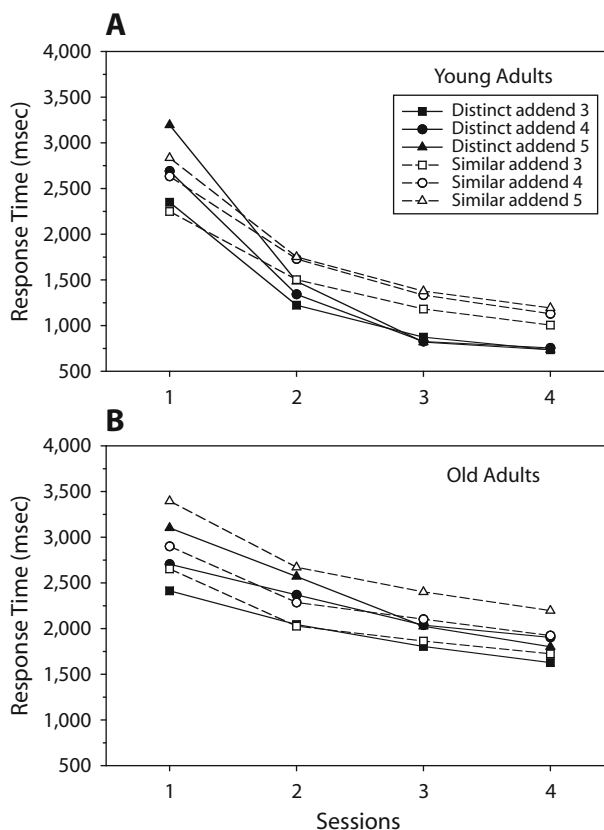
RTs longer than 15,000 msec were considered outliers and were deleted from the data set, resulting in the removal of fewer than .01% of responses. Error trials were also removed. The percentage of error trials by age and list condition are shown by sessions in Table 2. Overall, the error rate was low: 3.2%. An age (young, old) × list type (distinct, confusable) × sessions ANOVA (Greenhouse-Geisser adjustments needed) revealed a significant effect of sessions [ $F(2.072, 54.729) = 54.729, MS_e = 4.582, p = .004$ ] (accuracy improved with training), and a significant effect of age [ $F(1,27) = 5.379, MS_e = 16.815, p = .028$ ]. Older adults were more accurate (2.3%) than young adults (4.1%), perhaps due to their continued reliance on computed solutions as opposed to memorized solutions, as described below. There was no effect of list type (distinct condition, 2.84%; similar condition, 3.58%), and there were no interactions.

#### RTs As a Function of Session

RTs were analyzed at three levels of aggregation. For the first and coarsest analysis, means were calculated by session (1–4), broken down by list type (distinct, confusable), addend (3, 4, 5), and age (young, old). These data are illustrated in Figure 1 (A, young; B, old). In an ANOVA on these factors, the following effects were significant at the  $p < .05$  level or better: A main effect was found for session [ $F(3,87) = 120.6, MS_e = 71,005,328$ ].

**Table 2**  
Percent Errors, With Standard Deviations, by Age and List Condition

	Sessions							
	1		2		3		4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<b>Younger Adults</b>								
Distinct items	5.87	4.6	4.06	2.9	2.75	1.5	2.94	1.8
Confusable items	4.25	2.0	4.50	4.3	3.38	1.8	4.50	3.6
<b>Older Adults</b>								
Distinct items	2.50	1.7	1.36	0.6	1.21	0.8	1.43	0.8
Confusable items	4.25	3.0	2.94	2.1	2.25	1.7	2.63	2.6



**Figure 1.** Mean response times by session, broken down by addend (Addend 3, squares; Addend 4, circles; Addend 5, triangles) and by list type (distinct items, filled symbols; similar items, open symbols). (A) Data from 16 young adults. (B) Data from 16 older adults. Effects due to session, addend, list type, and age were statistically significant. Similar items showed less improvement with training than did distinct items; older adults showed less improvement with training than did young adults. Both interactions reflect continued reliance on slow computations, and reduced propensity for fast retrievals.

Performance improved with training, this being the most basic outcome of any training study, the cause of which we seek to explain. A main effect was found for addend [ $F(2,58) = 47.81, MS_e = 9,741,650$ ]. Larger addends took longer, in keeping with our understanding of their computational complexity. The addend effect interacted with sessions [ $F(6,174) = 20.70, MS_e = 883,104$ ], reducing, with training, the spread across addends. This convergence of the addend curves is in line with our understanding of skill development, in which fast retrieval responses come to replace slow computational responses with continued training. The addend × session effect in turn interacted with list type [ $F(6,174) = 3.94, MS_e = 167,966$ ]. The curves for confusable items show less convergence than those for distinct items. This suggests that fewer retrievals replaced computes in the confusable list.

There was also a main effect due to age [ $F(1,29) = 40.14, MS_e = 102,684,027$ ]. The old were slower overall. Two interactions involving age were significant: age × session [ $F(3,87) = 10.75, MS_e = 6,327,292$ ], indicating

that the old improved less with training; and age  $\times$  addend [ $F(2,58) = 3.53$ ,  $MS_e = 719,599$ ], indicating that the addend spread was expanded in the old. Both of these interactions can be understood as a consequence of continued reliance on computed solutions in the old, and a reduced level of compute-to-retrieve replacements.

### RTs As a Function of Addend

The addend effect looms large in these interpretations. A large addend effect is taken as an indication of computed solutions; a small one, as an indication of retrieved solutions. Such interpretations pervade the skill learning literature. The classical formulation is that of the “load” function, which captures the impact of the computational load at any point of training, in the form of a linear relationship between RT and a load variable such as addend or numerosity. Logan (1988), Palmeri (1997), and Rickard (1997) have all seized on the slope of the load function as an index of task automatization or retrieval use. On this derived measure, large values (steep slopes) have been shown to reflect computational solutions, and near-zero values (flat slopes) have been shown to reflect retrieval solutions. Intermediate values are interpreted as reflecting a mixture of computational solutions and retrieval solutions in the aggregated data. A cornerstone observation is

the finding that slopes drop from large to small over the course of repetition training.

For our second analysis, slope values were extracted from the data. The analysis was at an intermediate level of aggregation, with values calculated for each of 12 “epochs,” defined by averaging every eight list repetitions. For each epoch, RTs were averaged by age, list, problem type, and addend. In each subcondition, the effect of addend was approximately linear, and was summarized by a linear slope. In Figure 2, slope values are plotted as a function of epoch, separately for each age group, list condition, and problem type. The progression of values is of interest. In the first epoch, all values are close to 400 msec/addend, a value typical of the alphabet enumeration process that underlies computational solutions (Logan & Klapp, 1991). As training progresses, the eight curves in Figure 2 diverge; three patterns can be discerned. Young adults on the easy list show the classical result of a rapid drop to zero slopes, for both true and false problems. The pattern suggests a shift from all computation to all retrieval, over the course of training. Four of the remaining curves (young adults, hard list; old adults, easy list) drop from 400 msec/addend to an intermediate value of about 100 msec/addend. We interpret the latter value as reflecting a mixture of computes and retrieves over the last training epochs. The two

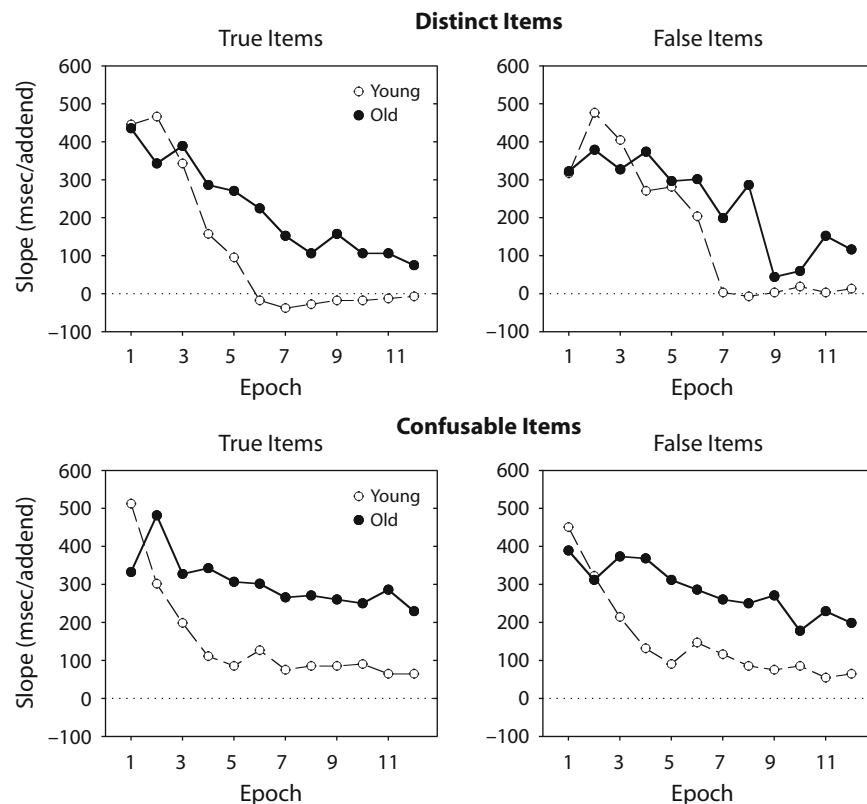


Figure 2. Addend slopes by epoch, broken down by age (young or old), item similarity (distinct or confusable) and item type (true or false). Each epoch represents eight list repetitions. Addend slope is calculated from the linear trend relating addend (with values of 3, 4, and 5) and mean response time. Slope values of 400–500 msec/addend reflect computational solutions; values close to zero reflect noncomputational retrieval solutions. Both list difficulty and adult age limited the shift to retrieval solutions.

remaining curves (old adults, hard list) show much less of a drop, to a final value of 200 msec/addend. Perhaps there are fewer retrieves in this mix, or perhaps this small improvement is due to a speedup in the enumeration rate of the computes.

Reinforcing the results of the initial analysis, the several patterns suggest that with training, both list difficulty and adult age limited the shift to retrieval solutions. But neither analysis takes into account interindividual differences. Intermediate addend effects or slopes may be due not to within-subjects mixing of computes and retrieves at the end of training, but to all computes from some participants, and all retrieves from others, or even to a simple speedup in the computes alone. Sorting through these possibilities requires analysis at a higher resolution than the group means.

### RTs As a Function of Items

At the most detailed level of aggregation, data were averaged over 24 “blocks” of four list repetitions, separately for each participant. For each block, addend, and trial type, there were 16 RT observations (minus errors); the mean and the standard deviation (*SD*) of these values were computed. This resulted in 12 “acquisition functions” or “traces” per participant,  $RT(N)$  and  $SD(N)$  for each addend and trial type, defined over 24 blocks. Illustrative data from three participants are shown in Figure 3. Two nonlinear models, described in the sections that follow, were fitted to these data.

Logan’s (1988) instance theory was the earliest model of cognitive skill learning detailed enough to derive the form of acquisition functions of the sort illustrated in Figure 3. These traces, both  $RT(N)$  and  $SD(N)$ , were predicted to be power functions, at least for tasks in which computation times were not substantially longer than retrieval times. In a critique of instance theory, Rickard (1997) pointed out that in tasks such as alphabet arithmetic, computations are likely to take considerably longer than retrievals, at least for larger addend values. In such cases Rickard (1997) argued that  $RT(N)$  and  $SD(N)$  will be given by the formulas for the mean and *SD* of a statistical mixture—a mixture, for any block  $N$ , of compute trials and retrieve trials. These formulas depend on five quantities, on the mean and *SD* of the computes and of the retrieves, and on the proportion of retrieves in the mix (or the complementary proportion of computes).

In Rickard’s (1997) view, an adequate theory of skill learning must specify how compute times change with practice (both the mean and the *SD*), how retrieve times change with practice, and how the proportion of retrieves changes with practice. These are the five “components” of Rickard’s (1997) CMPL theory. By means of postresponse strategy probes, Rickard (1997) examined compute trials and retrieve trials separately, and observed that the four latency components were all power functions of  $N$ , and that indeed, the compute function lay above the retrieve function (computes took longer than retrieves throughout training). Independently, Delaney, Reder, Staszewski, and Ritter (1998) came to the same conclusion. In a reanalysis of studies using preretrieve strategy probes, they too

found that computes and retrieves were best described by separate power functions.

An interesting consequence of the CMPL model is that when power-function components are passed back through the mixture formulas, the result is no longer a power function. In log–log coordinates, the composite *SD* exhibits a distinct hump above the best-fitting power function, and the composite RT displays a rising, then falling deviation from a power function. Rickard (1997) went on to show exactly these nonclassical distortions in the composite data from his experiments, and retrospectively, in Logan’s (1988) own data.

More recently, both Haider and Frensch (2002) and Rickard (2004) capitalized on the potential of power-function distortions to act as markers that indicate a change in item processing. At a participant or item level of analysis, “pure” power-function acquisition can be taken as a signal that a single solution process was deployed throughout training (computation). Conversely, a “humped” acquisition trace signals a shift at some point in training from one process or component to a faster process (retrieval).

This is the approach we took to analyze the acquisition functions collected here. Traces were coupled into  $RT(N)$ – $SD(N)$  pairs. Each participant yielded 6 pairs (true and false addends of 3, 4, and 5); each group, 48 pairs. A power function and a mixture function were fitted to each pair, and the pair was classified as either single process (PF) or dual process (MIX), according to the better fit (adjusted for the degrees of freedom). An inventory was thereby assembled of single-process and dual-process outcomes in each of the four cells of our design (list  $\times$  age). The frequency and nature of those outcomes were then examined at the parameter level.

Applied to the six trace pairs of any participant, a simple power-function model requires 24 parameters, a coefficient and exponent for each trace. A simple mixture model requires more than twice that number of parameters, 60 in all: 24 to specify the compute functions, 24 to specify the retrieve functions, and 12 to specify the mixture proportions. Models as large as these cannot be fitted to data with any degree of confidence. Parameter interdependencies make the fitted values somewhat arbitrary. Worse still, noise in the data will almost certainly distort the true value of many parameters, as the fitting algorithm exploits chance perturbations. In consideration of these issues, we reduced the number of free parameters and imposed constraints among the remaining parameters, following the lead of Rickard (1997). After examining the fit of many model variations to a selected subset of participants, we finally settled on a 20-parameter power-function model (PF), and a 24-parameter mixture model (MIX) (see the Appendix).<sup>2</sup> These versions, it seemed to us, were most successful in both following the meaningful trends in the sample data and avoiding chance aberrations.

The final models have much of the flavor of Rickard’s, with one exception. Rickard used a negative-exponential function to specify the growth in the retrieval proportion with  $N$ . This function failed to capture the “humps” in some of our traces that occurred late in training. A deferred

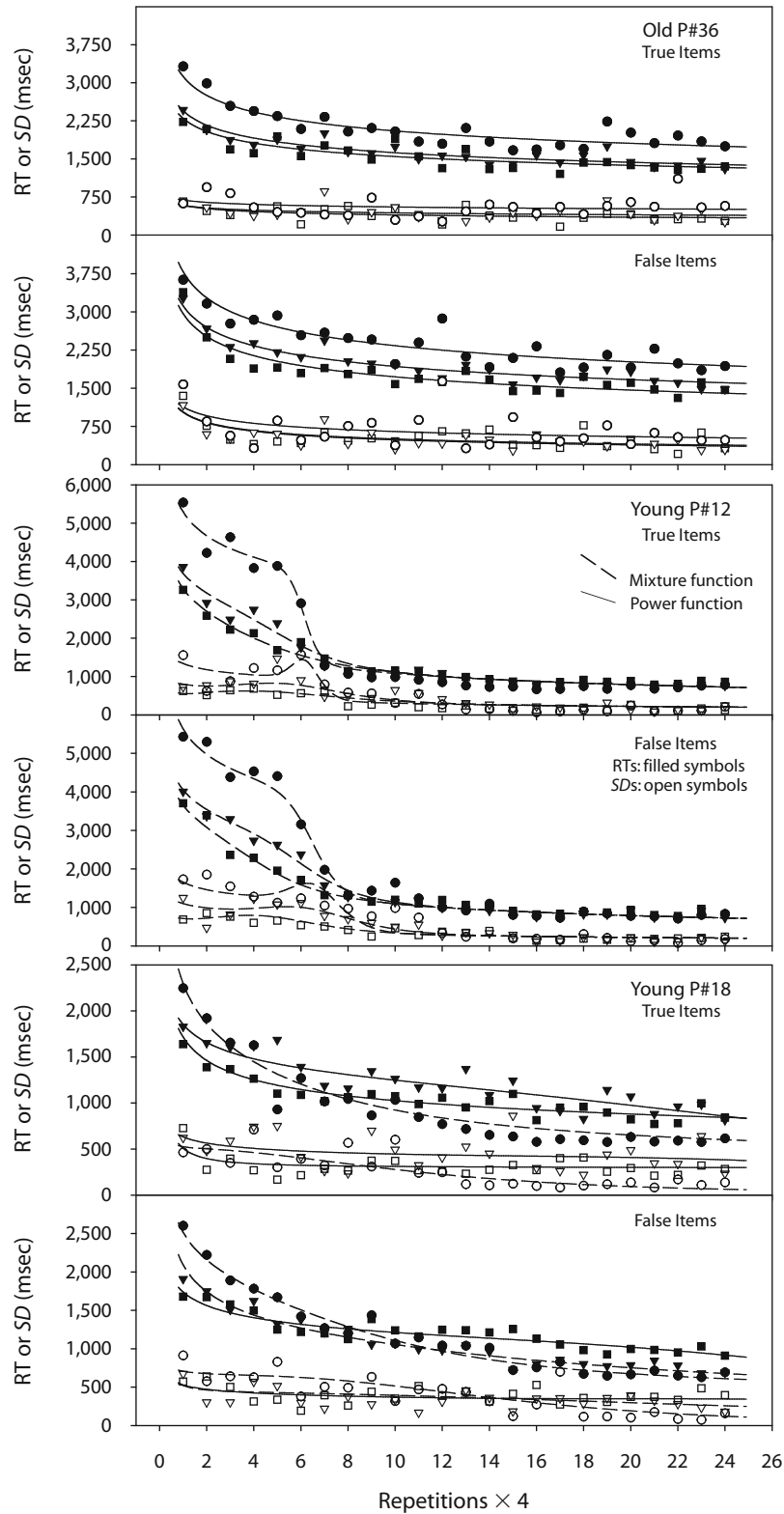


Figure 3. Acquisition data and fits for three representative participants (Ps). Data for true and false items are shown in separate panels. The 6 data sets in each panel show mean response times (RTs; filled symbols) and standard deviations (SDs; open symbols) for Addend 3 (squares), Addend 4 (triangles), and Addend 5 (circles). Each of the six smooth curves in the panel shows the fit of either a power function (solid lines) or a mixture function (dotted lines) to a coupled RT-SD data set. All the data sets for Participant 36 (confusable list) were best fit by power functions. All the data sets for Participant 12 (confusable list) were best fit by mixture functions. Participant 18 (distinct list) exhibits a combination of power and mixture functions across data sets.

exponential would work in all cases, but instead we used a logistic function with two readily interpretable parameters: a shift parameter marking the 50% point in the compute-to-retrieve transition, and a rate parameter marking the slope at the midpoint—that is, the transition rate.

### Comparisons Between Acquisition Models

The model-fitting exercise delivered two accounts of each participant's data. The fit of both models was quite good. Across the 16 young adults, the median variance accounted for by PF was .868 (range, .779–.949), and by MIX, .941 (range, .864–.974). Across the 16 older adults, the median variance accounted for by PF was .930 (.882–.977), and by MIX, .933 (.841–.976). Because MIX was a superset of PF with four additional parameters, it would be expected to account for more variance on purely mathematical grounds. An incremental  $F$  test was needed to compare the models while controlling for the number of parameters in each. This test indicated that PF was the better fit for 19 participants, and MIX was the better fit for 13 participants. But inspection of individual traces showed that model reckoning at the participant level was misleading. Many participants, for whom some traces were clearly PF and others clearly MIX, displayed a combination of PF and MIX traces. Accordingly, the model comparison was redone at the trace level (that is, for each of the six RT–SD trace pairs per participant), using a conservative  $p < .01$  alpha criterion for accepting a MIX fit over a PF fit. At the trace level, 118 traces were classified as PF (62%), and 74 traces were classified as MIX (38%).

With traces classified in this manner, we returned to the participant level. About half of the participants (15 of 32) showed composite PF–MIX fits, a quarter of the remainder (8 of 32) showed pure PF fits, and the remainder (9 of 32) showed pure MIX fits. Figure 3 shows the complete data from one individual of each type. Participant 36 (top two panels: older group, confusable items) was pure PF; Participant 12 (middle panels: young group, confusable items) was pure MIX; Participant 18 (bottom panels: young group, distinct items) was composite. Three traces were PF and three MIX.

Figure 3 bears study. The upper three traces in each panel give the mean RTs and the lower three traces give the SDs. Focus on the RT traces. Note that the power functions of Participant 36 show only a small amount of speedup over 24 blocks; in contrast, the mixture functions of Participant 12 drop precipitously early in training. As each of Participant 12's RT traces drops, a telltale “hump” surfaces in the corresponding SD trace below. Participant 18 shows both of these patterns: shallow PF traces and more highly modulated MIX traces. These patterns provide strong intuitive support for the conceptualization advanced above: Power-function acquisition reflects modest speedup in computational solutions, whereas mixture-function acquisition signals a switch from long computes to short retrieves at some point in training.

Table 3 shows the frequency of MIX outcomes (and the complementary frequency of PF outcomes) for the four cells of our design. These are our key findings. The MIX counts for young and old were quite different. For the

**Table 3**  
Numbers of RT–SD Trace Pairs Classified As Power Function or Mixture Function by Age Group and List Condition

Model	Distinct Items		Confusable Items	
	Young	Old	Young	Old
Power function	6	41	28	43
Mixture function	42	7	20	5

young, list type had a substantial effect; MIX frequencies dropped from 87.5% for distinct items to 42% for confusable items. Given our framework, this is evidence that a retrieval strategy was adopted by younger participants less often in the condition in which it would have been more difficult to implement. For the old there were few MIX traces in either list: 15% for distinct items, 10% for confusable items. In our framework, it seems that older participants found even an “easy” list of 24 alphabet-arithmetic items so problematic that they rarely opted to retrieve items, and instead continued to compute. These effects were supported by a log-linear analysis of the Table 3 contingency data. The effect of list was significant [ $\chi^2(1) = 12.85, p < .01$ ], as was the effect of age [ $\chi^2(1) = 64.53, p < .01$ ], together with the list  $\times$  age interaction [ $\chi^2(1) = 5.29, p < .05$ ].

A further analysis was done on the 62 MIX traces from young participants, since there were too few MIX traces from the old to analyze. Table 3 shows that these 62 traces broke down into 42 from the easy list and 20 from the hard list. For each trace, one of the MIX parameters ( $b$  in the *retrievePROB* equation in the Appendix) specifies the midway point of the compute-to-retrieve transition, in units of blocks of training, and another parameter (the inverse of  $d$  in the same equation) specifies the slope at the midway point, in units of the change in retrieval probability per change in block. For the two lists, the average shift points were 7.4 blocks for distinct items and 9.7 blocks for similar items. The average transition rates were 0.43 probability/block for distinct items and 0.21 probability/block for similar items. The effect of list was significant in both cases [ $t(60) = 2.0, p < .05$ , for the shifts; and  $t(60) = 3.5, p < .001$ , for the rates]. This is evidence that over and above its effect on adoption of the memory strategy, list difficulty affected the success of the strategy.

## DISCUSSION

Much evidence has accumulated that the development of skill in many cognitive tasks involves a shift from computation-based solutions to memory-based solutions, when the same problems are encountered repeatedly. Three formal models of skill formation have portrayed the emergence of memory representations as a primarily automatic concomitant of repetition training (the instance theory of Logan, 1988; the CMPL theory of Rickard, 1997; and the SAC theory of Schunn et al., 1997). But some researchers have raised questions about the autonomy of the compute-to-retrieve shift, and have sought to demonstrate the influence of strategic factors in skill attainment. For example, Ackerman and Woltz (1994) reported reduced levels of

retrieval in a list of confusable items (word pairs drawn from the same semantic category) compared with a list of distinct items (word pairs from different categories), and conjectured that participants more often adopted a retrieval strategy for the easier list; retrieval solutions were, therefore, strategic, not automatic.

A simpler interpretation of Ackerman and Woltz's (1994) results is that memory trace development and utilization were automatic for both easy and hard items, and were simply less successful in the hard list. Palmeri (1997) developed an account along just these lines, for easy list/hard list acquisition differences. Palmeri's (1997) participants reported the numerosity of repeatedly presented random-dot patterns; patterns were either similar or dissimilar. Palmeri's (1997) account was an elaboration of instance theory; as such, it assumed that memory traces formed and were activated automatically throughout training. List-difficulty effects were explained by adding to instance theory a more realistic trace retrieval mechanism, a random walk process that would be derailed by confusable traces.

Zbrodoff (1995) anticipated Palmeri (1997) in her interpretation of a skill-learning experiment involving alphabet-arithmetic problems much like ours. In a between-groups design, four problem sets were created by crossing two manipulations: distinct items versus confusable items, and equally frequent items versus a mixture of rare and common items. Zbrodoff's (1995) results looked much like those of our young adults in Figure 2: With training, problem-size slopes dropped to zero in the easier conditions (distinct, equal-frequency items), but leveled out well above zero in the hardest condition (confusable, mixed-frequency items). Like Palmeri (1997), Zbrodoff (1995) argued that item learning and retrieval proceeded apace, autonomously, in all conditions, but that the retrieval times of high-interference items were elevated. This interpretation works in the context of Logan's (1988) race model; computes would continue to beat slow retrievers throughout training. But what of the Ackerman and Woltz (1994) alternative? Perhaps persistent problem-size effects reflect continued computation on the part of participants who were reluctant to retrieve rare-and-confusable items.

How can "bottom-up" and "top-down" accounts of list difficulty be distinguished, if both predict fewer successful retrievals for difficult lists over the course of training? Our approach to this question was to examine the form of the acquisition traces for each item (or average of four related items), defined jointly by the RT and *SD* measurements.<sup>3</sup> Uninflected power-function traces (PF) were taken as the signature of a single, computational, solution process employed throughout training. Humped mixture-function traces (MIX) were taken as the signature of a shift from a slow computational process to a fast retrieval process at some point in training. This framework was first proposed by Rickard (1997) and exploited systematically by Haider and Frensch (2002) and by Rickard (2004). It derives strong intuitive support from the examples given in Figure 3—the fitted PF traces are shallow, showing little speedup with training, and the MIX traces drop precipitously from high values to low values.

This analysis separates two issues: whether or not a retrieval strategy is deployed; and, if deployed, whether the shift to retrieval occurs early or late. It seemed to us that a reduced frequency of retrieval would point to a top-down strategy aversion, whereas delayed retrieval could be referred more parsimoniously to bottom-up memory limits. For young participants our results were clear. In the easy list, 42 out of 48 traces signaled a shift to retrieval, but in the hard list only 20 of 48 traces showed this signal. This finding gives a firmer basis for Ackerman and Woltz's (1994) claim that participants' strategies can vary, and, in particular, that memory-based responses will be chosen less often if the list characteristics do not favor them.

Zbrodoff (1999), in a follow-up to her 1995 study, also came to the conclusion that participants' strategies can vary. The 1999 experiment again involved training on a set of alphabet-arithmetic problems, repeated over the course of 13 sessions. Participants were asked about their use of counting and retrieval at the end of each session. By the end of training, 8 of 11 participants reported using retrieval on most but not all trials (between 90% and 100%), whereas 3 participants reported retrievals on fewer than 40% of trials. Zbrodoff (1999) concluded that "it could be that 13 days of practice were insufficient for a complete transfer [from counting to retrieving]. However, it may be that a complete transition would never occur." Returning to her earlier 1995 study, it seems to us that differences in the frequency of retrieval, as opposed to differences in the speed of retrieval, provide an equally plausible explanation of the problem-size effects that persisted in the most confusable problem set.

In our study, we went on to show that when items were retrieved from the hard list, the retrievals appeared later in training: The average midway point of the compute-to-retrieve transition shifted from Block 7.4 for the easy list to Block 9.7 for the hard list. The rate of transition was also more gradual for hard items (a gain of .21 probability increments per block for young adults) than for easy items (a gain of .43 probability increments per block). Both effects would be expected on the basis of nonstrategic associative-memory limitations—weaker traces or heightened interference from competing traces. It is worth noting that, even for the hard list, when memory-based responding did occur it emerged in the first half of the four-day training regimen (that is, in Block 9.7 of 24 blocks). This argues against the possibility that every trace would eventually show a shift if training were continued long enough. Indeed, this may be so, but it could well be due to a participant's reevaluation of his or her original strategy after Day 4, and the formulation of a new strategy. (In an interesting aside, Logan, 1988, Experiment 4, speculated that some of the participants in his alphabet-arithmetic procedure might have done just that, switching from a less efficient strategy to a more efficient strategy between session 4 and session 5.)

### Composite Skill Models

The conclusion that item retrieval represents a strategy elected by a participant after evaluating its feasibility has implications for skill models. Bottom-up learning models



are in need of extension or further elaboration to allow for top-down influences. What form might this take? Rickard's (1997) CMPL model, for example, is specified as a neural network. It has a dual-pathway architecture: Either a compute pathway or a retrieve pathway is activated after the stimulus is registered, depending in part on the strength of the leading node in each pathway. Strategy effects such as those observed in the present data could perhaps be accommodated within CMPL by allowing an "executive" access to the lead nodes, as Rickard (2004) himself has noted. Strategy decisions would be translated into a static bias in the baseline activation of one or the other lead node, so as to favor either the compute or retrieve pathway. Except for this bias and its top-down origin, skill learning would proceed as before, in a bottom-up fashion.

It is also possible to extend instance theory to accommodate strategy differences, but a fix might require two steps. First, the model must be altered to produce inflected (humped) acquisition functions. This may turn out to be easy. In an important aside mentioned without follow-up, Palmeri (1999) described an instance model simulation that generated mixture-like humps in the acquisition function by the simple expedient of specifying a long compute process coupled with a short retrieval process, as opposed to the equal- (or nearly equal) duration processes used in Logan's (1988) simulations. Along these lines, Cerella and Hoyer (2001) reported a successful application of the "humped" instance model to data. Second, the model must be expanded to provide purchase for a strategy variable. Again, this may be easy. Palmeri's version of instance theory (the exemplar-based random walk model, or EBRW) appends a random-walk retrieval mechanism to Logan's instance memory, as mentioned above. Integral to the random walk is a strength-of-evidence criterion. By giving an "executive" access to this criterion, retrieval solutions could be either fostered or discouraged, to prevent, for example, the acceptance of misleading memory traces from confusable items. Except for this criterion shift and its top-down origin, skill learning would proceed as before, in a bottom-up fashion.

The third quantitative acquisition model, SAC theory (Schunn et al., 1997), comes closest to accommodating strategy effects in its current form. As with the other two models, memory representations in the SAC model develop autonomously as training proceeds. A given representation will actually be used as the basis for a problem solution only after its strength exceeds a fixed threshold. What is notable is that Schunn et al. found that threshold differences were the key to the successful prediction of individual performances. The six learning parameters of the SAC model could be held constant across individuals, provided that this one decision parameter was allowed to vary. The threshold parameter provides a natural locus for top-down control. As yet, however, the SAC model has only been used to predict strategy choice—only one of the components needed to fully specify skill performance.

In these three cases, we have elaborated "composite" models that combine top-down and bottom-up influences. The models are exciting, because the play of strategy is so localized in them. The three models all retain Logan's

(1988) original notion that encoding is obligatory, and tie strategy to the decision to retrieve the accumulated problem information, or, if retrieval is also obligatory, whether to base a solution on the retrieved information.<sup>4</sup>

As revised, these models would seem to have the capability of reproducing our basic result: a reduction in the deployment of a retrieval strategy above and beyond reductions in associative learning, in response to a mnemonically challenging list. It is less clear, however, whether they will be able to reproduce other aspects of strategy use exposed in our data. One question is whether strategy decisions are made at the list level or at the item level. Haider, Frensch, and Joram (2005, Experiments 1A, 1B, 1C) demonstrated list-level strategy election in an alphabet verification task. Participants who switched from less efficient to more efficient verification routines did so across the board, largely for all strings within a single block of trials at some point in the eight-block training sequence. About half of the 32 participants in our study can be characterized in a similar fashion. They pursued a single strategy over the entire course of training; of these, 8 computed only, and 9 shifted to retrieval for every item. In either case, the strategy decision was unanimous, embracing every item. But the other 15 participants demonstrated split strategies, computing some items and retrieving others. Models capable of splitting strategies are likely to require an additional level of control logic, tied, perhaps, to the characteristics of individual items. Such considerations have been a focus in the development of models of mental arithmetic (Lemaire, Arnaud, & Lecacheur, 2004; Siegler & Lemaire, 1997), and in a broader context, the ACT-R model of Lovett and Anderson (1996).

### Age Effects

In contrast to the performance of the young adults, older participants rarely adopted the memory strategy, with only 7 traces (out of 48) signaling a switch to retrieval in the easy list, and 5 in the hard list. What is known about skill learning in the elderly? In a recent study, Touron et al. (2004) presented a set of 10 artificial-arithmetic problems over 60 repetitions, and polled participants about their solution method after a problem was given. At the group level, age differences in these data were open to a simple description. The probability of retrieval solutions (as opposed to computed solutions) rose on a negative-exponential function of repetitions for both age groups, but the exponential rate parameter was smaller for old adults than for young adults. This account mirrors that of Rickard (1997) for young adults. Indeed, it is a classical portrait of bottom-up learning, as developed in the first mathematical models of paired-associate learning (Atkinson, Bower, & Crothers, 1965). These models credit each stimulus presentation with a small probability that the stimulus-response association will lodge in long-term memory. Learning is encapsulated in this single probability parameter, defined by the exponential rate value, and the age deficit is fully accounted for by a reduction in this parameter. (The parallel between age effects in paired-associate learning and skill learning is explored by Cerella, Onyper, & Hoyer, 2006).

What does this one-parameter model predict for the old in the least controversial, easy-list, condition? The prediction is that the old should not differ from the young in the frequency of retrieval traces, but that the compute-to-retrieve shift should be delayed. This prediction, of course, fails spectacularly. One additional factor that may contribute to age-related differences in strategy deployment is the participant's degree of confidence in retrieval, as suggested by Touron and Hertzog (2004a, 2004b). Like us, Touron and Hertzog found reduced levels of retrieval in elderly adults' skill learning of noun-noun pairs. At the same time they were able to demonstrate considerable knowledge of the pairs, either before or after training. Touron and Hertzog concluded that the old lacked the confidence to rely on their knowledge base, as long as they were given the option to look up the answers in a table. They appeared to be reluctant to trust their memories, even for items whose answers they knew.

The implication here is that older adults are burdened by a dual deficit: an associative-memory deficit, revealed by lower rates of acquisition in paired-associate training (Cerella et al., 2006; Naveh-Benjamin, 2000) and in short skill-problem sets (Touron et al., 2004); but over and above the associative deficit, a second "deficit," a reluctance to utilize accumulated item information in situations that tax confidence (Touron & Hertzog, 2004a, 2004b). It is this second deficit, unanticipated by us, that may lie behind the low level of memory usage by older adults in our study. A list of 24 alphabet-arithmetic problems, even distinct problems, may have been enough to compromise the confidence of older participants and dissuade them from using item knowledge.

The differences in performance due to age are extreme, but they can be accommodated by the top-down/bottom-up composite models as readily as were the differences due to item confusability. In either case, retrieval reluctance can be accounted for by increases in the threshold level of a parameter which governs the solution pathway selected or activated following presentation of a skill item. In both cases, the parameter setting expresses a strategic decision on the part of the participant—an evaluation based on the intersection of perceived ability and the difficulty of the problem set.

#### AUTHOR NOTE

This research was supported by Research Grant AG11451 to W.J.H. from the National Institute on Aging. We thank Hilde Haider, Alice Healy, and Tim Rickard for valuable comments on an earlier draft, and also Michael Kane, Gordon Logan, and an anonymous reviewer for help in sharpening our arguments. Correspondence concerning this article should be addressed to W. J. Hoyer, Department of Psychology, Syracuse University, Syracuse, NY 13244-2340 (e-mail: wjhoyer@sy.edu).

#### REFERENCES

- ACKERMAN, P. L., & WOLTZ, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task constraints, individual differences, volition, and motivation. *Journal of Educational Psychology, 86*, 487-515.
- ANDERSON, R. B. (2001). The power law as an emergent property. *Memory & Cognition, 29*, 1061-1068.
- ATKINSON, R. C., BOWER, G. H., & CROTHERS, E. J. (1965). *Introduction to mathematical learning theory*. Oxford: Wiley.
- CERELLA, J., & HOYER, W. J. (2001, November). *Cognitive skill learning: Races, mixtures, and humps in the power law*. Paper presented at the Annual Meeting of the Psychonomic Society, Orlando, FL.
- CERELLA, J., ONYPER, S. V., & HOYER, W. J. (2006). The associative-memory basis of cognitive skill learning: Adult age differences. *Psychology & Aging, 21*, 483-498.
- DELANEY, P. F., REDER, L. M., STASZEWSKI, J. J., & RITTER, F. E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science, 9*, 1-7.
- EKSTROM, R. B., FRENCH, J. W., HARMAN, H. H., & DERMEN, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- HAIDER, H., & FRENCH, P. A. (2002). Why aggregated learning follows the power law of practice when individual learning does not. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 392-406.
- HAIDER, H., FRENCH, P. A., & JORAM, D. (2005). Are strategy shifts caused by data-driven processes or by voluntary processes? *Consciousness & Cognition, 14*, 495-519.
- HEATHCOTE, A., BROWN, S., & MEWHORT, D. J. K. (2000). The power law revealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*, 185-207.
- HOYER, W. J., CERELLA, J., & ONYPER, S. V. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition, 31*, 1260-1270.
- KIRK, E. P., & ASHCRAFT, M. H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 157-175.
- LEMAIRE, P., ARNAUD, L., & LECACHEUR, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology & Aging, 19*, 467-481.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.
- LOGAN, G. D., & KLAPP, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 179-195.
- LOVETT, M. C., & ANDERSON, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31*, 168-217.
- MYUNG, I. J., KIM, C., & PITT, M. A. (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition, 28*, 832-840.
- NAVEH-BENJAMIN, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 1170-1187.
- PALMERI, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 324-354.
- PALMERI, T. J. (1999). Theories of automaticity and the power law of practice. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 543-551.
- RICKARD, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126*, 288-311.
- RICKARD, T. C. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 65-82.
- ROGERS, W. A., & GILBERT, D. K. (1997). Do performance strategies mediate age-related differences in associative learning? *Psychology & Aging, 12*, 620-633.
- ROGERS, W. A., HERTZOG, C., & FISK, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 359-394.
- SCHUNN, C. D., REDER, L. M., NHOUYVANISVONG, A., RICHARDS, D. R., & STROFFOLINO, P. J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy

- selection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 3-29.
- SIEGLER, R. S., & LEMAIRE, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, **126**, 71-92.
- SMITH-CHANT, B. L., & LEFEVRE, J. (2003). Doing as they are told and telling it like it is: Self-reports in mental arithmetic. *Memory & Cognition*, **31**, 516-528.
- THURSTONE, L. L., & THURSTONE, T. G. (1949). *Examiner manual for the SRA Primary Mental Abilities Test* (Form 10-14). Chicago: Science Research Associates.
- TOURON, D. R., & HERTZOG, C. (2004a). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology & Aging*, **19**, 452-466.
- TOURON, D. R., & HERTZOG, C. (2004b). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition*, **32**, 298-310.
- TOURON, D. R., HOYER, W. J., & CERELLA, J. (2004). Cognitive skill learning: Age-related differences in strategy shifts and speed of component operations. *Psychology & Aging*, **19**, 565-580.
- WECHSLER, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio: Psychological Corporation.
- ZBRODOFF, N. J. (1995). Why is  $9 + 7$  harder than  $2 + 3$ ? Strength and interference as explanations of the problem-size effect. *Memory & Cognition*, **23**, 689-700.
- ZBRODOFF, N. J. (1999). Effects of counting in alphabet arithmetic: Opportunistic stopping and priming of intermediate steps. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 299-317.

## NOTES

1. In the confusable list, not all letter pairs were unique: Some true and false items shared the same letters, and differed only in the digit. This would make letter-based mnemonics harder to implement by participants (although Logan, 1988, Experiment 4, reports the use of mnemonics in a list of similar composition). This would add to the difficulty of the confusable items, and thus strengthen our manipulation. We thank Hilde Haider for this observation.
2. Heathcote, Brown, and Mewhort (2000) demonstrated that component functions are closer to negative exponentials than to power functions in their mathematical form, at the level of individual items. At the same time, power functions better describe data averaged over items. Myung, Kim, and Pitt (2000) were able to show analytically that negative exponentials converge to a power function very rapidly as traces are added to the average. This result, confirmed by simulations (Anderson, 2001), underlies our choice of power-function components to describe the four-trace aggregates of our analysis.
3. Postresponse strategy probes ("Did you compute your answer or remember it?") were not used because of the possibility that the probes themselves might prompt participants to adopt a retrieval strategy, in much the same way as interpolated memory tests enhanced retrieval use in Ackerman and Woltz's (1994) experiments. Strategy suggestions of this sort have been found to influence strategy use in mental arithmetic tasks (Kirk & Ashcraft, 2001; see also Smith-Chant & LeFevre, 2003).
4. Not acknowledged within these models is the possibility that strategy may influence encoding. This is the distinction between incidental (nonstrategic) and intentional (strategic) learning. An intentional regimen may foster stronger or more elaborate memory traces than otherwise. This possibility remains to be explored.

## APPENDIX

Mixture Model (24 Parameters)		Power-Function Model (20 Parameters)	
Add3true	$computeRT = a3t N^{-cc}$ $computeSD = a3tsd N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b3t)/(.55d3t)))$	Add3true	$RT = a3t N^{-c3t}$ $SD = a3tsd N^{-c3tsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			
Add3false	$computeRT = (a3t + af) N^{-cc}$ $computeSD = (a3tsd + afsd) N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b3f)/(.55d3f)))$	Add3false	$RT = (a3t + af) N^{-c3f}$ $SD = (a3tsd + afsd) N^{-c3fsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			
Add4true	$computeRT = a4t N^{-cc}$ $computeSD = a4tsd N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b4t)/(.55d4t)))$	Add4true	$RT = a4t N^{-c4t}$ $SD = a4tsd N^{-c4tsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			
Add4false	$computeRT = (a4t + af) N^{-cc}$ $computeSD = (a4tsd + afsd) N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b4f)/(.55d4f)))$	Add4false	$RT = (a4t + af) N^{-c4f}$ $SD = (a4tsd + afsd) N^{-c4fsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			
Add5true	$computeRT = a5t N^{-cc}$ $computeSD = a5tsd N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b5t)/(.55d5t)))$	Add5true	$RT = a5t N^{-c5t}$ $SD = a5tsd N^{-c5tsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			
Add5false	$computeRT = (a5t + af) N^{-cc}$ $computeSD = (a5tsd + afsd) N^{-cc}$ $retrieveRT = a N^{-cr}$ $retrieveSD = asd N^{-cr}$ $retrievePROB = 1/(1 + \exp(-(N - b5f)/(.55d5f)))$	Add5false	$RT = (a5t + af) N^{-c5f}$ $SD = (a5tsd + afsd) N^{-c5fsd}$
$mixtureRT = (1 - retrievePROB) computeRT + retrievePROB retrieveRT$ $mixtureSD = \sqrt{(1 - retrievePROB) computeSD^2 + retrievePROB retrieveSD^2 + (1 - retrievePROB)(retrievePROB)(computeRT - retrieveRT)^2}$			