Journal of Experimental Psychology:
Learning, Memory, and Cognition
1991, Vol. 17, No. 2, 196–209

0278-7393/91/$3.00

Automatizing Alphabet Arithmetic: II. Are There Practice Effects After Automaticity Is Achieved?

Stuart T. Klapp, Catherine A. Boches, and Marilyn L. Trabert
California State University, Hayward

Gordon D. Logan
University of Illinois

Subjects made true/false judgments regarding alphabet-arithmetic (AA) statements like C + 2 = E (a true response because E is two steps beyond C in the alphabet). Logan and Klapp (1991) studied reaction time (RT) for novice and automatized performance on this task. The present experiments extended this analysis by testing interference between concurrent speech and AA verification at three levels of training: novice, automatized, and overtrained. At the novice level, RT depended on the digit addend and concurrent speech was disruptive. At the automatized level, RT was independent of addend and repetitive speech did not interfere. However, nonrepetitive speech (reciting the months of the year) interfered. At the overtrained level, RT was shorter, and interference from nonrepetitive speech was reduced, showing that there were practice effects after automaticity was achieved. These findings support an interpretation of automaticity in which memory retrieval is speeded during training. When retrieval is faster than algorithmic processing (e.g., alphabet counting), performance becomes automatized. With further training, memory retrieval becomes even faster, thereby reducing further the interference from concurrent tasks. We discuss the extent to which this account represents a general theory of automaticity.

Alphabet-arithmetic (AA) statements indicate "addition" of a number to a letter to obtain another letter. Addition can be accomplished by advancing through the alphabet from the first letter by the number of steps indicated by the digit "addend." For example, B + 2 = D because D is obtained by advancing two steps from B. Logan and Klapp (1991) concluded that novice subjects verify AA statements as true or false by performing this explicit counting operation because: (a) they reported doing so, and (b) reaction time (RT) increased substantially as a linear function of the digit addend. The present experiments included a test of an additional prediction of alphabet counting.

Logan and Klapp (1991) studied the transition from this novice algorithmic mode of performance to automatic performance in the AA verification task, basing their analysis on RT data. The present experiments extended this analysis in two ways. First, they examined the possibility that there may be continued practice effects with additional training beyond that needed to achieve initial automaticity. Second, they examined interference between concurrent tasks in addition to RT.

The notion of automaticity is not crisply or consistently defined in the literature. Definition is often in terms of a list of properties (differing in detail among authors) that are assumed to co-occur when automaticity is achieved or in terms of freedom from demands on vaguely specified "resources" (see Logan, 1988; Logan & Klapp, 1991). Our approach is not within either of these traditions. Rather, we take the view that automaticity may be usefully defined as the situation in which performance is supported by direct access to memory, which can replace the slower algorithmic modes of processing that support nonautomatic performance. This approach is consistent with several modern theories that propose memory retrieval as the mechanism underlying automaticity (e.g., Logan, 1988; Schneider & Detweiler, 1987; Strayer & Kramer, 1990). Applying this view to AA verification, Logan and Klapp (1991) assumed that automaticity is achieved when the algorithmic processing (counting through the alphabet) is replaced by single-step, direct-access retrieval of solutions from memory. Thus, a criterion for automaticity in this task is that the RT for verification should become short and independent of the addend, indicating that counting through the alphabet is no longer involved.

Our present extension of this analysis to handle the effects of additional training beyond that needed to achieve automaticity is based largely on interference of concurrent tasks with AA verification. The study of dual-task interference is important because elimination of interference is an often-discussed property of automaticity in the traditional analysis.

We acknowledge support from California State University, Hayward, through the School of Science to Stuart T. Klapp and from National Science Foundation Grant BNS 87-10436 to Gordon D. Logan. We thank Jeffrey Bisanz, Tom Carr, and Keith Rayner for insightful comments and suggestions on an earlier version of this article. We acknowledge the very substantive contributions of Julie Stein to this research. Other students who have contributed include Joy Cottril, Ruth Canady, Carolyn Harrison, Nancy Shimomoto, and Marilyn Yee; and also Christine Acacio, James Barnak, Marzieh Janatpour, Barbara Macauliffe, Pamela Means, Jean Pendleton, Anthony Pope, and Patti Woodside. Some aspects of this research resulted from discussions during meetings of a United States Air Force working group sponsored by the Air Force Office of Scientific Research.

Correspondence concerning this article should be addressed to Stuart T. Klapp, Department of Psychology, California State University, Hayward, California 94542.

However, our analysis, unlike traditional approaches, does not require that dual-task interference must vanish with automaticity. Rather, our approach makes articulated predictions concerning which tasks will and will not interfere with AA verification at various stages of practice.

Effects of Training

The memory theory of automaticity implies that automatic processing is possible if the AA facts (e.g., B + 2 = D) become available in memory. Training to achieve this memory representation can be based on extensive practice with the verification task itself (as might be expected in some traditional accounts of automaticity), but it need not be. Training on the facts outside of the verification-task context should be equally effective as Logan and Klapp (1991) demonstrated. Rapid acquisition of automaticity was achieved with only 15 min of rote-memory training on true statements before the subjects had been exposed to the verification task.

It is important that automaticity can be achieved by rote memorization without training on the criterion-verification task, because training on the task itself is a central idea in most traditional views of automaticity. It is also important that extensive practice is not necessary to achieve automaticity. Traditional views often consider extensive practice to be the fundamental requirement for development of automaticity.

Training Beyond Achievement of Automaticity

The memory theory proposes a specific mechanism for the transition from controlled to automatic processing: replacement of the counting algorithm by memory access. Because this theory holds that automaticity is attributable to development of memory, we expect that phenomena associated with learning should apply also to the acquisition of automaticity. A classic finding is the overlearning effect in which performance improves as learning continues beyond the level needed to achieve a criterion of correct performance (see McGeoch & Irion, 1952, pp. 377–384). Modern analyses also demonstrate continued improvement with extended practice (e.g., Logan, 1988).

We now report experiments that test for the continued improvement in AA verification after automaticity has been achieved. This is expected because automaticity is assumed to occur when retrieval from memory becomes faster than the counting algorithm for each of the facts. Further increases in retrieval speed beyond that criterion should be possible.

Interference of Concurrent Tasks With AA Verification

How are we to assess improvement beyond the level of performance needed for memory retrieval to replace algorithmic processing as assessed by elimination of the slope of RT as a function of addend? Once this slope becomes zero, no further improvement is possible in this criterion. Of course, the overall RT may become shorter even after the slope becomes zero, but this is not the only way that additional improvement might be observable. We predict that dual-task interference can be reduced with further training.

Specifically, we investigated interference from concurrent tasks, and how the patterns of interference change during acquisition of automaticity and during continued training beyond initial automaticity. We used two different concurrent tasks, both of which involved overt speech. One of these—repetitive month saying—required the subject to pronounce the same word (one month of the year, e.g., January) repetitively. The other—sequential month saying—required pronunciation of the months in succession (i.e., January, February, March, and so on).

Logan and Klapp (1991) concluded that, before the facts are memorized, AA verification is accomplished by counting through the alphabet for the required number of steps. This may be a process corresponding to internal speech. According to this interpretation, either month-saying task should interfere with novice AA verification because alphabet counting and month saying share a common speech-like process. By contrast, after the facts have been memorized (and thus AA becomes automatized), AA verification is assumed to be supported by a single memory retrieval rather than by alphabet counting. Because continuous internal speech is not required for automatized AA verification, repetitive month saying should not produce interference. Consider, however, what should happen for sequential as opposed to repetitive month saying. Whereas repetitive month saying may not interfere with automatized AA verification, sequential month saying should interfere according to the assumption that concurrent retrievals from memory can interfere with each other. Sequential month saying requires successive retrievals of the months in sequence, and automatized AA verification requires a single retrieval. Experiments 1 and 2 confirmed these predictions and thereby provided converging evidence for the proposed contrast between alphabet counting in novice performance and memory retrieval in automatized performance.

Experiments 3, 4, and 5 were designed to study the possibility of improvement with overlearning training beyond the level needed to achieve initial automaticity. As retrieval of AA facts becomes faster, there is less opportunity for a concurrent task that also requires retrieval to interfere. Thus, a concurrent task (sequential month saying) that interferes with initially automatized AA verification should show greatly reduced interference after additional training beyond that needed to achieve automaticity.

Experiment 1

We begin with examination of interference between AA verification and sequential month saying (January, February, March, and so on) contrasting novice performance (pretest) with performance after learning the AA facts, thereby achieving automaticity (posttest). Determination that automaticity has been achieved in the posttest was based on the RT data, following the approach of Logan and Klapp (1991). For novice performance (pretest), RT for AA verification should increase as a function of digit addend. This relation should vanish if the task has become automatized by the posttest.

The automaticity-as-memory theory predicts that sequential month saying should interfere with AA verification in both pretest and posttest but for different reasons. The pretest interference is expected because overt speech (month saying) and inner speech (alphabet counting) interfere. The posttest interference is expected because both tasks require retrieval of information from memory (AA facts and month sequence). By contrast, many traditional approaches tend to suppose that, when automaticity is achieved, all properties of automaticity should emerge together. In particular, interference from concurrent tasks including sequential month saying should vanish at automaticity.

Speech articulation and manual action can interfere with each other because of conflicting response timing (Klapp, 1981). Because we were concerned with central interference rather than with response incompatibility in the present research, we included a control condition that would be subject to interference between articulation and the motor actions used in AA verification. We predicted that month saying will produce more interference with AA verification than with this control task.

Method

Experiments 1 to 4 involved similar methodology, which is described in detail only for Experiment 1.

Subjects. The 16 subjects were students in Introductory Psychology at California State University, Hayward, who participated as one option of a course requirement. All claimed to be native speakers of English, and all were able to recite the months of the year and the letters of the alphabet in sequence.

Alphabet–arithmetic task. The AA verification task involved true versus false judgments on AA statements, for example, B + 2 = D. This fact is true because, starting after B and counting up, two more letters (C, D) lead to D. Half of the displays were true (as in the example), and the remaining displays were false because the indicated letter was advanced either one less than the correct answer (e.g., B + 2 = C) or one more than the correct answer (e.g., B + 2 = E). Subjects made yes (true) or no (false) responses using F or J keys of the computer keyboard.

Two sets of six AA facts appear in Table 1. Half of the subjects in each condition were tested with each set of facts. The untested set was used for initial training on the general procedure of AA verification. The facts in Table 1 would, of course, elicit true as the correct response. Stimuli for which the answer is "false" were generated by using the letter before or after the correct letter.

Consistency of mapping of stimuli to responses is a major theme in automaticity theory (Schneider & Shiffrin, 1977). Note that our mapping of stimuli to true and false answers was consistent with respect to the overall equation (e.g., B + 2 = D always corresponded to a true response). However, no element of the equation taken alone was consistently mapped to a particular response. For example, the initial letter *B* led to false as well as to true responses. The only exception was that some answer letters (e.g., C and G) appeared only with false statements.

All alphanumerical stimuli subtended a visual angular height of 0.6 degrees as viewed on the monitor of an Apple II microcomputer. Each trial began with the display "get ready," which appeared for approximately 1.3 s, followed by a blank delay of 1.3 s. Then the problem appeared (e.g., M + 2 = N). This remained until the subject responded with the F or J key. Assignment of F and J to true and

Table 1
Sets of Alphabet–Arithmetic Facts

Fact set	Initial letter	Addends	Responses
1	B	2, 3, 4	D, E, F
	N	2, 3, 4	P, Q, R
2	H	2, 3, 4	J, K, L
	T	2, 3, 4	V, W, X

Note. For example, the first fact in Set 1 was B + 2 = D.

false, respectively, was balanced across subjects in each condition. Subjects kept their hands on these keys during the block of 24 trials. RT was measured from the onset of the problem until the key press occurred. Then the median RT was determined for each subject at each digit addend (and also separately for true and false displays). The descriptive statistics reported are the means of these medians. The rate of errors was determined for each condition, and the RT for error trials was excluded from the RT analysis.

Immediately after the subject responded, the display "correct" or "error" appeared as appropriate. If the RT was less than 3 s, the display "time was good" appeared below. For RT between approximately 3 s and 4.6 s, the display was "time was ok, but should be faster." Slower responses received the display "time was too slow." These displays remained for about 2.5 s, and were followed by an intertrial interval (ITI) of about 3.5 s.

Each block of 24 trials was composed of 12 true problems, using two instances of each of the six facts, and 12 false problems. These were six instances of one letter beyond the correct version of each fact, and six instances of one letter before the correct version of each fact. These problems were presented in a new random order for each block tested.

Control task. Performance on a yes–no task in which subjects press one of two keys in response to a display of yes or no was used as a control for assessing interference between motor responses of key pressing and articulation for the month-saying task (sequential months for Experiment 1).

Month saying. For the dual-task condition, the subjects recited the months of the year aloud sequentially, always starting with January (i.e., January, February, March, and so on). In this dual-task condition, tones sounded at the rate of two tones per second to pace the speaking. The command "recite months" replaced the "get ready" command of the single-task condition. This command appeared approximately 2.6 s before the letter–number problem, ensuring that month saying would be in progress before the problem appeared. The command "stop months" appeared after the subject pressed the response key for the letter–number problem. There was a marked tendency to try to skip months to permit analysis of letter–number problems. Therefore, the experimenter sat next to the subject throughout testing and insisted on consistent month recitation. The experimenter also sat next to the subject in the single-task condition.

Training of the letter–number facts. Subjects learned the facts by computer-controlled paired-associate learning. Rote recitation of the facts was used initially, and then responding in anticipation of the answers was required. The rate of input was rapid to discourage elaboration and mediation. Throughout the 20-min training program, the experimenter sat next to the subject to maintain motivation. Each cycle in which the six facts were presented represented an independent randomization of the order of occurrence of the facts. The stages of training were as follows:

1. Each fact was presented four times in succession, and the subject was to pronounce each. The fact equation remained visible for 1,200 ms, and the ITI was 600 ms.

2. Each fact was presented once for pronunciation and twice for anticipation. In the anticipation case, the stem (A + 2 = ?) appeared for 2,000 ms, and the subject was to respond orally before this time had expired. Then the complete equation appeared for 750 ms.
3. Each fact was presented once for pronunciation and once for anticipation for 5 cycles through the six facts.
4. Each fact was presented twice for anticipation for 12 cycles through the facts.

In this training, each AA fact had been presented 41 times. Logan and Klapp (1991) showed that 36 to 72 presentations per fact was sufficient to achieve automaticity. At this point in the procedure, each subject's level of learning was tested by the anticipation procedure in which each fact appeared once. The criterion was three successive passes through the six facts without error. For subjects who failed to reach criterion on the sixth cycle, Step 4 of the training was repeated. Then another criterion test was attempted. Subjects who failed to reach criterion on the second attempt were dismissed. In Experiment 1, it was necessary to replace 4 subjects for failure to meet this criterion and 1 because of equipment failure.

Design. Two independent groups of 8 subjects each were tested in the single-task (AA verification only) and dual-task (AA verification with month saying) conditions; alternate subjects were assigned to the conditions as they qualified for the experiment. All subjects were tested on both control (yes–no) and AA verification tasks before training (pretest) and after training (posttest). The order in which the control and AA verification tasks were presented was balanced across subjects in each condition using a pattern in which the last task tested before training was the first task tested after training. Half of the subjects in each condition were assigned to each of the sets of facts (Table 1) for testing.

Subjects were first trained on the months task (if in the dual-task condition), with emphasis on not skipping months. Then they received 24 trials on the yes–no task to learn the assignment of keys to responses. They then practiced the AA verification task for 12 trials using the set of facts that would not appear in the scored trials. Next, they received the pretest of 24 scored trials on each of the yes–no and AA verification tasks; the dual-task group recited the months concurrently. Next, the training procedure for the facts was performed, followed by the posttest for both yes–no and AA verification tasks, again accompanied with reciting months for the dual-task condition.

Results

Reaction time and error rate as a function of addend. The mean of median RTs for correct AA verifications in the single-task (no month saying) condition appear in Figure 1. There was a marked interaction of digit addend (N) by pretest versus posttest, $F(2, 14) = 12.0$, $p < .001$, $MS_e = 42{,}777.66$. Reaction time increased linearly as a function of N for the pretest (slope of 360 ms, $r^2 = .97$), but not for the posttest, $F(2, 14) < 1$, $MS_e = 33{,}911.38$. Also note that RT was shorter for the posttest than for the pretest, $F(1, 7) = 17.7$, $p < .01$, $MS_e = 422{,}416.57$. These findings demonstrate that automaticity had been achieved by the posttest.[1]

The trend in error rates corresponded to the significant effects for RT, as can be seen in Table 2. Although the rate of errors tended to increase as a function of addend in the pretest, this trend was nonsignificant, $F(2, 14) = 2.57$, $p = .11$, $MS_e = 0.696$, and the trend toward interaction, which had the sense that error rates appear to increase with addend

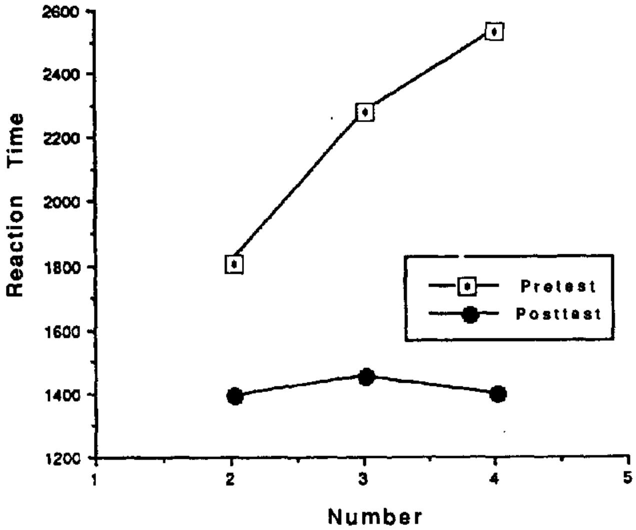

Figure 1. Reaction time (RT) (ms) as a function of digit addend, Experiment 1.

more for pretest than for posttest, also was nonsignificant, $F(2, 14) = 1.94$, $p = .18$, $MS_e = 0.559$.

Because we are using RT to assess whether automaticity has been achieved, the significant effects for single-task RT makes our point. However, a similar pattern appeared in the dual-task RT trends. The apparent interaction of pretest versus posttest by digit addend was not quite significant, $F(2, 14) = 3.4$, $p = .06$, $MS_e = 107{,}986.88$, but the trend corresponded to the single-task data. The RT increased as a function of N for the pretest, slope = 355 ms, $r^2 = .99$, $F(2, 14) = 4.3$, $p < .05$, $MS_e = 234{,}395.34$, more than for the posttest, slope = 52.1 ms, $r^2 = .87$, $F(2, 14) = 1.1$, $MS_e = 21{,}679.31$. Overall RT was shorter in the posttest than in the pretest, $F(1, 7) = 12.3$, $p < .01$, $MS_e = 345{,}504.15$. Reaction times for pretest as a function of increasing N were 1,678, 2,061, 2,388 ms, and for posttest, 1,383, 1,469, and 1,487 ms. In all of the experiments we report, the dual-task data reflect the general pattern of the single-task data but with more statistical "noise." To conserve space, we do not report dual-task RTs for the remaining experiments.

Interference of month saying with AA. Considering only the pretest, there was a nonsignificant trend for the mean RT for AA verification in the month-saying condition (2,042 ms) to be shorter than the mean for the single-task condition (2,202 ms). This trend was reversed for the posttest (1,415 ms for single vs. 1,447 ms for dual). Overall, the presence of month saying did not appear to have a consistent effect on RT.

Month saying had a dramatic effect on the rate of errors in AA verification (Table 2), with a higher rate of errors for the

[1] We used median RT because of the limited number of observations. When means and medians from a similar experiment were compared, the pattern of results was found to correspond between the two measures (Logan & Klapp, 1991, Experiment 2).

Table 2
Error Rate for Alphabet–Arithmetic Task of Experiment 1

Task	Addend (N) 2	3	4	M
Single				
Pretest	3.1%	12.5%	14.1%	9.9%
Posttest	4.7%	4.7%	3.1%	4.2%
Dual				
Pretest	25%	32.8%	43.8%	33.9%
Posttest	17.2%	21.9%	9.4%	16.2%

dual-task condition than the single-task condition, $F(1, 14) = 22.7$, $p < .001$, $MS_e = 6.57$. This dual-task decrement (interference) was more pronounced in the pretest (decrement of 24.0%) than in the posttest (12.0%) as confirmed by a Pretest Versus Posttest × Single Versus Dual interaction, $F(1, 14) = 4.9$, $p < .05$, $MS_e = 3.41$. However, the dual-task decrement remained significant even in the posttest, $F(1, 14) = 13.3$, $p < .01$, $MS_e = 2.49$. Also the overall rate of errors was higher for the pretest than for the posttest, $F(1, 14) = 22.7$, $p < .001$, $MS_e = 6.57$.

The RTs and error rates for AA verification in single-task versus dual-task (month saying) conditions present some difficulty for interpretation. Interference appeared in error rates but not in RT. We consider this to indicate that subjects have great difficulty in the dual-task situation, causing them to guess on AA verification and thereby producing somewhat shorter RTs and substantially higher error rates. This interpretation agrees with the subjects' reports and with our impressions when we attempted to combine the tasks. If month saying is maintained to keep up with the pacing signal, alphabet counting seems to be virtually impossible.

One way that RT and error data may be combined is to determine the average number of correct AA responses per second in each of the conditions (Table 3). Note that interference is evident in this measure for both pretest and posttest, in agreement with the conclusions that would be reached if only error data were considered. Therefore, we take error rate as the primary measure of interference in the reports of the other experiments.

Yes–no control task. Data from 1 subject in the single task pretest were lost as a result of a computer malfunction. The mean of median RTs and the error rates for the yes–no task appear in Table 4. This control task shared physical features with AA verification, but did not require solution of AA problems. Thus, it is not surprising that RT was shorter than for AA verification. Reaction times improved from the pretest to posttest, $F(1, 14) = 5.0$, $p < .05$, $MS_e = 4{,}776.28$, but this improvement was limited to the dual-task condition, yielding a significant Single–Dual × Pretest–Posttest interaction, $F(1, 14) = 4.6$, $p < .05$, $MS_e = 4{,}776.28$. The overall trend toward a dual-task decrement in RT was nonsignificant, $F(1, 14) = 3.7$, $.1 > p > .05$, $MS_e = 26{,}043.50$.

Table 3
Mean Number of Correct Alphabet–Arithmetic Responses per Second of Reaction Time in Experiment 1

Task	Pretest	Posttest
Single	0.41	0.68
Dual (month saying)	0.32	0.58

Table 4
Mean of Median RT and Error Rates for The Yes–No Control Task of Experiment 1

Task	RT (ms)	Error rate
Single		
Pretest	519	1.8%
Posttest	517	1.0%
Dual		
Pretest	681	5.7%
Posttest	574	0.5%

Note. RT = reaction time.

Although the error rates for the pretest dual-task condition tended to be higher than the others, the trend toward a dual-task decrement in the pretest data (higher errors for dual task compared with single task) was nonsignificant, $F(1, 14) = 3.2$, $p = .09$, $MS_e = 1.27$. An important feature of these control data is the lack of even a trend toward dual-task decrement in the posttest error rate. This contrasts markedly with the results of the AA verification task for which the dual-task decrement is strong in both pretest and posttest error data.

Discussion

Rote memorization of the AA facts eliminated the effect of digit addend on RT for AA verification, indicating that automaticity had been achieved at the posttest. The concurrent task of pronouncing months sequentially (January, February, March, and so on) led to a marked increase in rate of errors in both pretest and posttest AA verification. This dual-task decrement cannot be attributed to interference between pronouncing the months and pressing the response keys because month saying had only minimal effect on the yes–no control task.

Our interpretation of the interference of month saying and AA verification implies different mechanisms for interference in the pretest and the posttest. Whereas pretest interference was attributed to conflicts of concurrent speech processes, posttest interference was attributed to concurrent retrievals. This implies that repeating the same word would not interfere with posttest AA verification because repetitive speech would not demand successive retrievals from memory. However, repetitive speech should produce interference with pretest AA verification. These predictions were tested in Experiment 2.

Experiment 2

Whereas Experiment 1 investigated the effects of sequential month saying (January, February, March, and so on), Experiment 2 involved repetitive month saying (e.g., January, January, January). Pretest AA verification is assumed to require speech-like alphabet counting, and repetitive month saying obviously involves speech. Because both tasks are assumed to involve speech processes, interference is predicted between repetitive month saying and pretest AA verification. However, our theory implies that alphabet counting is not required in automatized posttest AA verification. Furthermore, unlike sequential month saying, repetitive month saying does not involve retrieval of a sequence of months from memory. Therefore, we predict no interference between repetitive month saying and posttest AA verification.

Method

The method corresponded exactly to that of Experiment 1, except that the month-saying task was saying the same month repetitively rather than naming the months in succession. The month used was rotated among subjects, except that the word *February* was not used because of the difficulty people report in repetitively pronouncing it.

Results

Reaction time and error rate as a function of addend. The means of median RT for AA verification in the single-task condition appear in Figure 2. There was an interaction of digit addend (N) by pretest versus posttest, $F(2, 14) = 6.3$, $p < .05$, $MS_e = 28{,}220.91$. Reaction time increased linearly as a function of N for the pretest (slope of 165 ms, $r^2 = .84$), yielding a significant effect of N, $F(2, 14) = 5.4$, $p < .05$, $MS_e = 48{,}387.81$. By contrast, for the posttest (after the facts were learned), RT decreased nonsignificantly, $F(2, 14) = 1.6$, $p = .24$, $MS_e = 16{,}028.42$, and nonlinearly ($r^2 = .57$) with N. Overall RT was shorter for posttest than for pretest, $F(1, 7) = 27.3$, $p < .001$, $MS_e = 78{,}062.42$. We conclude that automaticity had been achieved in the posttest.

Error rates (Table 5) corresponded to these RT results. The rate of errors increased as a function of N, $F(2, 28) = 11.3$, $p < .001$, $MS_e = 1.041$. Also there was an interaction of N by pretest versus posttest, $F(2, 28) = 9.99$, $p < .001$, $MS_e = 0.529$, which may be described in terms of a larger effect of N for pretest than for posttest. There was no significant three-way interaction of Dual Versus Single × Pretest Versus Posttest × N, $F(2, 28) = 1.49$, $p = .24$, $MS_e = 0.529$.

Comparison of Experiments 1 and 2 suggests that subjects may have operated at different points along a speed–accuracy trade-off

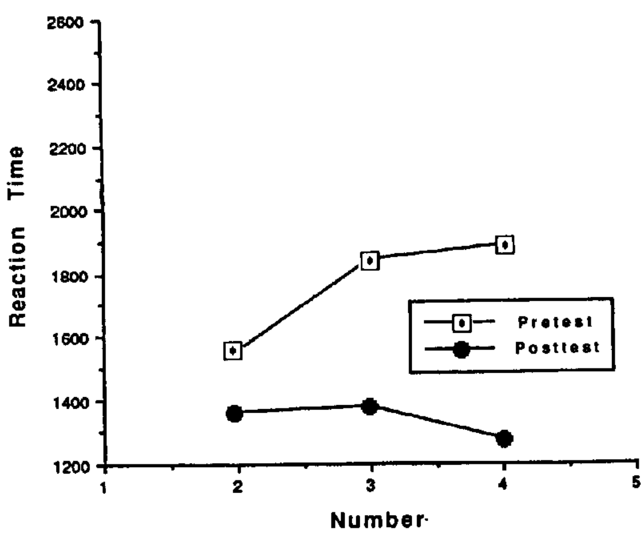

Figure 2. Reaction time (RT) (ms) as a function of digit addend, Experiment 2.

Table 5
Error Rate for Alphabet–Arithmetic Task of Experiment 2

Task	Addend (N) 2	3	4	M
Single				
Pretest	10.9%	20.2%	35.9%	22.3%
Posttest	1.5%	9.4%	10.9%	7.3%
Dual				
Pretest	18.8%	39.0%	43.8%	33.9%
Posttest	6.3%	9.4%	6.3%	7.3%

function in the two experiments. For the single-task conditions, the procedures (and even the software) were nominally identical, and yet the overall RT and error rates differed substantially. In Experiment 1, the mean RT was 1,809 ms and error rate was 7% in comparison to Experiment 2, in which the RT was 1,548 ms and error rate was 14.9%. We do not know what aspect of the situation may have produced this difference. However, because we are interested in comparing the effects of independent variables on RT and error rates rather than in the absolute magnitude of these measures, this difference between experiments is not as much of a problem as it may at first appear to be.

Interference from month saying. In the pretest, RT for the dual task (1,897 ms) tended to be longer than RT for the single task (1,759 ms), $F(1, 14) < 1$, $MS_e = 423{,}050.93$. For the posttest data, there was almost no difference in RT between dual task (1,272 ms) and single task (1,337 ms), $F(1, 14) < 1$, $MS_e = 201{,}162.02$. This lack of consistent effect of month saying on RT for AA verification corresponded to the findings of Experiment 1.

We take error rates (Table 5) as our primary measure of interference. Experiment 1 demonstrated marked interference between sequential month saying and AA verification for both pretest and posttest. By contrast, and as expected, Experiment 2 presented no hint of interference from repetitive month saying in the posttest (identical error rates). Treating the contrast between Experiments 1 and 2 as an independent variable involving type of month saying allows us to text this difference in the effect of the two types of month saying. For the posttest, the interaction of sequential versus repetitive tasks by month saying versus control was significant, $F(1, 28) = 9.3$, $p < .01$, $MS_e = 1.78$, indicating that the disruption was stronger for sequential than for repeating month saying after automaticity had been achieved.

However, we did expect interference in the pretest because both repetitive month saying and alphabet counting involve speech. There was some statistical ambiguity regarding this point because the rate of errors differed only nonsignificantly as a function of the presence versus absence of repetitive month saying in the pretest data, $F(1, 14) = 4.4$, $p = .052$, $MS_e = 2.30$. For the pretest data, the interaction of Type of Month Saying × Single Task Versus Dual Task was nonsignificant, $F(1, 28) = 2.5$, $p = .12$, $MS_e = 7.20$, indicating that, as expected, the amount of disruption did not differ significantly between the two types of month saying in the pretest.

Yes–no control task. Results for the yes–no control task (Table 6) correspond to those of Experiment 1 (Table 4), except that RT tended to be shorter overall in Experiment 2. In both experiments, the highest error rate was associated with the dual task pretest; the other error rates were low and nearly equivalent. In Experiment 2, this pattern of error data led to significant effects for both pretest versus posttest, $F(1, 14) = 12.6$, $p < .01$, $MS_e = 0.357$, and dual task versus single task, $F(1, 14) = 4.5$, $p < .05$, $MS_e = 0.696$, and to a significant interaction, $F(1, 14) = 12.6$, $p < .01$, $MS_e = 0.357$. Taken together, the control data of the two experiments indicate that any

Table 6
Mean of Median Reaction Time (RT) and Error Rates for the Yes–No Control Task of Experiment 2

Task	RT (ms)	Error rate
Single		
Pretest	493	1.5%
Posttest	495	1.5%
Dual		
Pretest	472	7.3%
Posttest	467	1.25%

dual-task decrement in the AA task for the posttest performance cannot be attributed to motor interference between month saying and hand pressing. In the case of Experiment 2, however, it is possible that the nonsignificant trend toward interference between repetitive month saying and AA in the pretest could be due to response interference.

Discussion

As predicted, posttest performance was automatized by the RT criterion, and was not disrupted by concurrent repetitive month saying. We predicted that repetitive month saying would interfere with pretest AA verification because both tasks are assumed to involve a speech-like process. The statistical ambiguity concerning this feature of the data is examined as one aspect of Experiment 3.

Experiment 3

The major purpose of Experiment 3 (and the subsequent experiments) concerns determining whether there are practice effects after automaticity is achieved. The strategy of Experiment 3 was to take advantage of extensive overlearning that our subjects have received in the facts of regular numerical arithmetic. Performance on a number–arithmetic (NA) analog of the procedure used for AA verification was investigated with concurrent sequential month saying, repetitive month saying, and control (no month saying) conditions. The possibility of practice effects after achieving automaticity leads us to predict that, although sequential month saying interferes with automatized AA verification (Experiment 1), this interference may vanish for NA verification. Of course, interference of repetitive month saying with NA should not occur because even initially acquired automaticity eliminated such interference (Experiment 2).

We included novice (pretest) AA verification as a control condition in which both forms of month saying should interfere. This lets us assess our null prediction for interference in NA as an interaction of Month Saying Versus Control × AA Versus NA. Also a secondary purpose of Experiment 3 was to resolve a statistical ambiguity in Experiment 2 concerning whether repetitive month saying interferes with pretest AA verification. In Experiment 2 the trend toward interference was nearly, but not quite, significant ($p = .052$).

Method

Design. Two independent groups of 12 subjects each performed AA verification without training on AA facts or NA verification. Subjects were assigned to either AA or NA verification in the order that they reported for the experiment. (It was necessary to replace 1 subject because of equipment failure.) All subjects were tested in three conditions: sequential month saying, repeating month saying, and control (no month saying). The order of these three conditions was balanced within each group of subjects. Each condition was represented by 24 trials, 8 of which were at each level of addend, N, the second term of the equations "$A + N =$ " or "$3 + N =$."

Procedure. For NA verification, the program that had been used for AA verification was changed by replacing the letters with numbers. In particular, with reference to Table 1, B and N were replaced by 1 and 2, and H and T were replaced by 3 and 4, yielding two sets of six facts each. The timing and other details of the program were unchanged. For all subjects, true responses were made with the right hand using J key, and false responses were assigned to the left hand and the F key. Subjects first practiced responding for 24 trials to *yes* or *no* stimuli. Then subjects practiced both month-saying tasks. Next they practiced for 24 trials on the NA or AA verification task using a different set of facts from those to be tested. No training of the relevant facts was provided. The three scored blocks of 24 trials each followed this preliminary practice.

Results and Discussion

Reaction time. Figure 3 displays RT for AA and NA verification as a function of addend (N) for the single-task control condition. Consistent with previous results for pretest AA verification, RT increased monotonically and linearly as a function of N, with a slope of 259 ms per N and with a close fit to a linear relation ($r^2 = .977$) and a highly significant effect of N, $F(2, 22) = 7.98$, $p < .01$, $MS_e = 103{,}067.86$. By contrast, for NA verification, RT did not increase as a function of N. In fact, there was a weak effect such that RT decreased as N increased, $F(2, 22) = 3.89$, $p < .05$, $MS_e = 4{,}375.48$. The slope of this relation was −36 ms per N, with a reasonable fit to linear regression ($r^2 = .91$).[2] The difference in the relation between N and RT for the two verification tasks was confirmed as a significant $N \times$ Verification Type (AA vs. NA) interaction, $F(2, 22) = 9.79$, $p < .001$, $MS_e = 53{,}721.67$. This pattern of results is consistent with the premise that whereas AA verification was performed by counting, NA verification was based on memory.

Overall RT for verification, averaged across levels of N, appears in Table 7. The most striking finding is that RT was far shorter for NA verification than for AA verification, $F(1, 22) = 51.8$, $p < .001$, $MS_e = 2{,}702{,}039.11$. Consider, for example, the RT for the control, no-month-saying condition of NA verification for which the mean RT, 983 ms, was less than that observed for AA verification in the other experiments even after the facts were learned. The shortest of these comparable values was 1,337 ms (Experiment 2, posttest, single-task condition). This fast performance for NA verification may reflect the years of experience with arithmetic facts compared with the brief training on AA.

Reaction time was shorter in the month-saying conditions than in the control, no-month-saying conditions for both AA

[2] The small but negative slope for NA was unexpected. It may be due to the presence of more paired problems (e.g., 2 + 2, 3 + 3) in the high-addend cases compared with the low-addend cases. Parkman and Groen (1971) reported that paired problems usually produce shorter RTs than others.

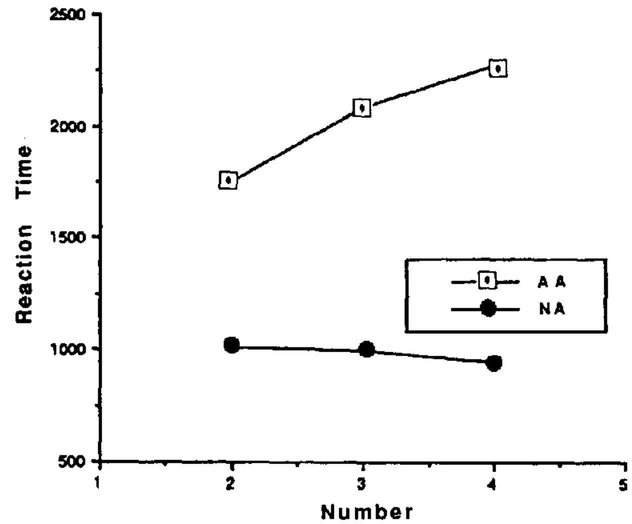

Figure 3. Reaction time (RT) (ms) as a function of digit addend, control (single task) condition of Experiment 3. (AA = alphabet-arithmetic; NA = number-arithmetic.)

and NA verification, with all comparisons yielding at least $F(1, 22) = 4.9$, $p < .05$, $MS_e = 548{,}439.23$. There was no significant interaction of Type of Arithmetic Task × Presence Versus Absence of Month Saying, $F(2, 44) = 1.4$, $p = .25$, $MS_e = 468{,}913.09$.

Errors. As in the other experiments, we take the error rate for verification as the measure of interference from month saying (Table 8). First consider AA verification. The rate of errors was higher when AA verification was accompanied by month saying compared with the control for both sequential month saying, $F(1, 11) = 23.4$, $p < .001$, $MS_e = 9.50$, and for repetitive month saying, $F(1, 11) = 31.5$, $p < .001$, $MS_e = 7.04$. There was no difference at all in the error rate between the two month-saying conditions.

Next consider NA verification. In marked contrast to the interference observed in AA verification, performance in NA verification was not significantly influenced by the month-saying variable, $F(2, 22) = 2.2$, $p > .1$, $MS_e = 2.17$. This difference between NA and AA verification in the effect of month saying was confirmed as a significant interaction of Month Saying × Type of Verification, $F(2, 44) = 10.5$, $p < .001$, $MS_e = 4.75$. However, because the overall rate of errors was much smaller for NA compared with AA verification, $F(1, 22) = 41.0$, $p < .001$, $MS_e = 12.48$, the possibility that the interaction was due to ceiling effects cannot be ruled out.

It is also useful to compare (between experiments) the effects of sequential month saying on error rate for initially automatized AA verification versus NA. Whereas in Experiment 1 the interference from sequential month saying produced an increase in errors of 12.0% in posttest AA verification, the interference was only 4.0% for NA verification in Experiment 3. This between-experiment comparison is consistent with the premise that, even after automaticity is achieved, AA verification is more sensitive to interference from sequential month saying than is NA verification.

Table 7
Overall Reaction Time (ms) for Verification in Experiment 3

Verification task	Type of month saying: Sequential	Repeating	Control
Alphabet-arithmetic	1,791	1,670	2,032
Number-arithmetic	885	839	983

Table 8
Percent Error on Verification Averaged Across N for Experiment 3

Type of verification	Type of month saying: Sequential	Repeating	Control
Alphabet-arithmetic	35%	35%	10%
Number-arithmetic	6%	7%	2%

Note that, although nonsignificant and small, there was a trend suggesting possible interference between month saying and NA verification. Because this trend is not consistent with our proposal that interference should be virtually eliminated for this situation, it needs to be evaluated. The magnitude of this difference in error rates (4.5%) is not markedly different from the corresponding trends indicating that month saying interferes with the yes–no control task of Experiments 1 and 2, for which the difference in pretest performance for month saying versus control ranges from 0.25% to 5.8% (mean: 3.6%). We conclude that the small and nonsignificant trend toward interference between month saying and NA verification can be attributed, at least in part, to interference between the motor responses of articulation and key pressing.

Discussion

In Experiment 1, AA verification was disrupted by sequential month saying, even after performance was automatized. In Experiment 3, we asked whether this disruption would be reduced with additional practice. The finding that neither type of month saying interfered markedly with NA verification suggests that overlearning may reduce or eliminate the interference observed in a newly automatized task; there may be practice effects after automaticity is achieved.

Both forms of month saying interfered equally and significantly with novice AA verification, thereby clarifying the statistical ambiguity in the pretest data of Experiment 2. Pretest AA verification is sensitive to interference from repetitive month saying. (Another study, not reported here, also indicated essentially equal and significant interference of the two types of month saying with novice AA verification.)

Experiment 4

In Experiment 3, NA verification was assumed to correspond to AA verification when based on highly overlearned facts. Because NA was less sensitive to interference from sequential month saying than was newly automatized AA, we concluded that there may be practice effects after automaticity. However, NA differs from AA in many ways in addition to degree of learning, and these differences, rather than overlearning, might have produced the difference in amount of interference. Therefore, in Experiment 4, we attempted to achieve overlearning of AA facts beyond the level of training needed to initially achieve automaticity. One manifestation of practice effects after automaticity would be reduction in interference from sequential month saying with overlearning.

Experiment 4 provided overlearning training on only two AA facts (critical facts) to produce extensive training in a limited time. The critical facts involved $N = 2$ and $N = 4$ to maximize the sensitivity to effects of N on RT. Performance on AA verification, using the six-fact paradigm of the other experiments, was measured at three stages of training on these critical facts: (a) novice, before training (b) after learning the two critical facts to achieve automaticity on those facts, and (c) after overlearning of the critical facts (beyond automaticity).

Method

All 16 subjects were tested on 24 trials of AA verification with concurrent sequential month saying (January, February, March, and so on) and on 24 trials without month saying. The order of these conditions was balanced across subjects. Half of the subjects were tested on each set of AA facts (Table 1). Assignment of sets of two facts to be the critical set was rotated among subjects such that all possible combinations of $N = 2$ and $N = 4$ facts were used equally often, with each subject receiving one critical fact with $N = 2$ and one with $N = 4$.

The experiment was conducted on four sessions on different days. The first session, which was much longer than the others, included five phases of training and testing. First, general training on AA verification and month saying (as in the other experiments) was provided. The second phase was testing on AA verification with and without month saying (Level 1). The third phase was training on critical facts to criterion in which the two critical facts were each displayed with answers 12 times in an intermixed pseudorandom sequence. Subjects pronounced the facts when displayed. Then each fact was displayed three times, once with the answer and twice for anticipation of the answer. This phase concluded with testing for memory of the two facts using answer anticipation. If not correct, three exposures of additional training were allowed. Thus, the total training involved from 15 to 18 exposures to each critical fact. The fourth phase was testing on AA verification with and without month saying (Level 2). In the fifth phase, the first cycle of overlearning training on the two critical facts occurred. The procedure for this 16-min cycle was similar to that used in Experiment 1, except that only two facts were trained.

Sessions 2 and 3 each involved two additional cycles of overlearning training on the two critical facts; each cycle corresponded to the last procedure of the first session. The final session included one more cycle of training. Thus, there were six cycles of overlearning training involving a total of 160 exposures to each fact. Finally, subjects received the final test (Level 3) of AA verification with and without month saying.

Results

Reaction time for critical facts.[3] Reaction time for AA verification in the single task (no month saying) condition appears in Table 9. At Level 1 (before training), RT increased as a function of addend (N), $F(1, 15) = 18.6$, $p < .001$, $MS_e = 146,670.19$. By Level 2 (after initial learning of the facts), overall RT was reduced, $F(1, 15) = 21.2$, $p < .001$, $MS_e = 218,774.72$, and the effect of N on RT became smaller and nonsignificant, $F(1, 15) = 2.0$, $p > .1$, $MS_e = 62,238.75$. This reduction in the effect of N was confirmed as a significant Testing Level × N interaction, $F(1, 15) = 8.3$, $p < .05$, $MS_e = 100,555.44$. These RT results correspond to those reported earlier, and are consistent with the view that, whereas Level 1 performance on the critical facts was based on alphabet counting, Level 2 performance was automatized and based on memory retrieval.

Next consider the contrast between Levels 2 and 3. We assume that this contrast should be similar in many ways to the contrast between newly automatized posttest AA of Experiments 1 and 2 and NA of Experiment 3. Consistent with that expectation, there was a trend toward further reduction in overall RT in Level 3 compared with that achieved in Level 2, but that trend was nonsignificant, $F(1, 15) = 1.5$, $p = .23$, $MS_e = 65,674.55$. Because automaticity had already been achieved by Level 2, there was no further decrease in the slope of RT as a function of N by Level 3 (actually there was a nonsignificant increase in slope), $F(1, 15) < 1$, $MS_e = 43,583.14$. The effect of N on the RT for AA verification remained nonsignificant at Level 3, $F(1, 15) = 2.1$, $p > .1$, $MS_e = 93,351.43$, and, of course, this effect was reliably smaller than the effect of N at Level 1 as confirmed as a Levels (1 vs. 3) × N interaction, $F(1, 15) = 9.4$, $p < .01$, $MS_e = 78,329.04$.

The single-task RT pattern (just described) was reflected in a corresponding pattern of RT for the dual-task condition (Table 9). In particular, RT depended on N at Level 1 of practice, $F(1, 15) = 6.7$, $p < .05$, $MS_e = 258,161.90$, but there was no such effect at Levels 2 and 3. Consistent with most of the previous data, the overall RT in the dual-task condition was slightly shorter than in the single-task condition.

Errors. The error rates for critical facts appear in Table 10. The overall rate of errors declined among the three levels of training, $F(2, 30) = 23.4$, $p < .001$, $MS_e = 0.895$. Also the interference (difference in error rate between month saying and control conditions) declined with training, yielding a significant Training × Month Saying interaction, $F(2, 30) = 3.5$, $p < .05$, $MS_e = 0.810$. The effect of month saying was significant at Level 1, $F(1, 15) = 18.4$, $p < .001$, $MS_e = 1.06$, and nearly significant at Level 2, $F(1, 15) = 4.0$, $p = .06$, $MS_e = 1.53$, and Level 3, $F(1, 15) = 4.44$, $p = .052$, $MS_e = 0.258$.

The comparison between Level 2 (automatized) and Level 3 (overtrained) was of particular interest. The overall error rate decreased between these levels, $F(1, 15) = 8.5$, $p < .01$, $MS_e = 0.595$, but the trend toward interaction of Levels × Month Saying was nonsignificant, $F(1, 15) = 1.25$, $p = .28$, $MS_e = 0.80$.

[3] Data from control, noncritical facts were also examined. Because these data are redundant with, and correspond to, the results of other experiments (Logan & Klapp, 1991), they are not reported in detail. The RT increased with N at all levels, with gradually decreasing slope as the facts were slowly acquired. Errors were higher when AA was accompanied by month saying than for the control condition throughout.

Table 9
Reaction Time (ms) for Alphabet–Arithmetic Verification Involving the Critical Facts in Experiment 4

Training level	N 2	N 4	M	Slope
Single task (no-month-saying condition)				
1: novice	1,715	2,299	2,007	292.0
2: automatic	1,397	1,524	1,461	63.5
3: overtrained	1,306	1,461	1,384	77.5
Dual task (sequential month-saying condition)				
1: novice	1,702	2,134	1,918	216.0
2: automatic	1,421	1,424	1,423	1.5
3: overtrained	1,401	1,394	1,398	−3.5

Discussion

Whereas the RT for critical-fact AA verification depended strongly on addend at Level 1 of practice, this relation was essentially gone by Levels 2 and 3 of practice. This result is consistent with the interpretation that whereas Level 1 AA verification was based on alphabet scanning, AA verification at Levels 2 and 3 was automatized and based on memory retrieval.

Massive interference from sequential month saying occurred at Level 1 (novice) as in the other experiments. The results from Level 2 of training (intended to be just enough to produce automaticity) were somewhat disappointing. The interference from month saying was statistically marginal. However, based on Experiment 1, we are confident that sequential month saying does interfere with AA verification after learning of facts to criterion (but before overlearning beyond criterion).

At Level 3 (overtrained beyond automaticity), the error rate was reduced below that of Level 2, and interference from sequential month saying was reduced to approximately the same level (4.8% more errors when saying months) as was observed for NA in Experiment 3 (4%). This small effect of saying months may be attributed to peripheral interference (see control condition of Experiments 1 and 2). Therefore, we conclude that either overlearning of facts of NA during everyday life (Experiment 3) or specific training on AA facts (Experiment 4) can reduce error rate and interference from sequential month saying in comparison to the results observed for initially automated performance (Experiment 1). There are practice effects after automaticity is achieved.

Table 10
Percent of Errors in Alphabet–Arithmetic Verification Using Critical Facts in Experiment 4

Level of training	Sequential month saying	Control	Interference
1: novice	33.5%	14.0%	19.5%
2: automatic	16.4%	5.4%	11.0%
3: overtrained	6.3%	1.5%	4.8%

Experiment 5

Experiment 5 was a further test of the hypotheses that extensive training would eliminate interference from sequential month saying. Whereas Experiment 4 involved learning by remembering (i.e., training on the AA facts), Experiment 5 involved learning by doing (i.e., training on the AA verification task itself) using the procedure of Logan and Klapp (1991; Experiment 3). Previously we showed that these two types of training had equivalent effects on the reduction of the slope of RT as a function of digit addend. We now seek to extend this conclusion of equivalence to the elimination of dual-task interference. If memory strengthens with repeated presentations of either type so that retrieval becomes progressively faster, there should be reduced opportunity for interference.

Method

Subjects. The subjects were students from the University of Illinois who had served for five sessions in the six-fact group of Logan and Klapp (1991; Experiment 3). They were paid $4 to return for an additional session. Eight of the original 9 subjects were able to return.

Sequential month-saying task. The month-saying task was paced by a metronome program that played a 100-ms 500-Hz tone every 600 ms (i.e., 100 ms of tone; 500 ms of silence). This was slightly slower than the 500-ms rate used in the previous experiments.

Single-task (no month saying) trials were conducted as described in Logan and Klapp (1991; Experiment 3). Stimulus presentation was paced by the experimenter with the next trial beginning 2.5 s after the subject had responded to the previous AA problem. Dual-task trials required a more complicated procedure. They began with the onset of the metronome signals. Subjects were told to synchronize their month saying with the metronome and to begin reciting on the fifth beat. Four beats after they began recitation (when they were saying "April") the experimenter initiated the AA trial. The fixation stimulus was presented for 500 ms, then replaced by the AA equation, which remained on until the subject responded. This cycle was repeated for each trial.

Training procedure. Details of the training procedure during the first five sessions are in Experiment 3 of Logan and Klapp (1991). In that experiment, three groups of 9 subjects were trained on each of 6, 12, and 18 facts for five sessions. The 8 subjects we now test were from the 6-fact group. These subjects had been trained on the 6 facts for two 576-trial sessions involving 48 repetitions of each fact per session. By the end of the second session, the slope of the linear function relating RT to digit addend had reached asymptote, averaging 27 ms per count. Then they were transferred to 18 facts: the 6 on which they were originally trained (old–old), 6 new facts that associated a new digit addend with each of the original 6 letters (new–old), and 6 new facts about different letters (new–new). The 18 facts were trained for three 576-trial sessions, involving 16 repetitions of each fact each session. At the end of the fifth session, the 6 original facts had been presented 144 times, and the 12 new facts introduced on Session 3 had been presented 48 times. The slope for old–old facts was 33 ms per count; the slope for new–old facts was 110 ms per count; and the slope for the new–new facts was 106 ms per count.

The contrast between the 6 original facts and the 12 facts introduced on Session 3 provided a conceptual replication of the design of Experiment 4, comparing dual-task interference at two different levels of practice. In Experiment 4, the same facts were tested at three different times, at three different levels of practice; in Experiment 5, different facts at different levels of practice were tested at the same time. However, whereas Experiment 4 involved learning by remembering the facts, training for Experiment 5 was on AA verification (i.e., learning by doing).

Design. On the final session, each subject performed 576 trials with the 18 facts that they had experienced during training: 288 trials under single-task conditions and 288 trials under dual-task conditions. The order of single-task and dual-task conditions was balanced across subjects.

Each of the 9 original subjects was assigned a different set of six facts on the first training session, and the two additional sets assigned on Session 3 were related according to the original experimental design. Because 1 subject did not return for the session reported here, the nine sets of facts could not be completely balanced across the 8 subjects who did return.

Results and Discussion

Reaction time. Mean Reaction time was 1,135 ms in the single-task condition and 1,324 ms in the concurrent month-saying condition, but this difference was nonsignificant, $F(1, 7) = 2.14$, $p > .10$, $MS_e = 118{,}763.28$. It is worth noting again that RT data have not revealed reliable interference effects in the AA paradigm, and that we rely on error data to demonstrate interference. Table 11 displays single-task RT as a function of digit addend and level of training. Reaction time did not increase systematically as a function of digit addend, indicating that automaticity had been achieved at all levels of training. However, overall RT improved as a function of training, $F(2, 14) = 9.78$, $p < .01$, $MS_e = 54{,}854.41$, showing again that performance can improve after automaticity has been achieved.

Error rates and dual-task interference. The rates of errors in AA verification, averaged across digit addend, appear in Table 12. As in the other experiments, we take the difference in error rates for the dual-task condition (sequential month saying) compared with the single-task control as our measure of interference. The trends suggest that there may be some interference at all levels of training, with somewhat more interference at the lower levels of training. However, neither the overall effect of month saying on errors, $F(1, 7) = 2.94$, $p > .10$, $MS_e = 683.84$, or the interaction of Training × Month Saying, $F(2, 14) < 1$, $MS_e = 147.99$, were significant. The overall rate of errors did not differ significantly as a function of training, $F(2, 14) < 1$, $MS_e = 239.80$.

At the highest level of training, the rate of errors was about 4% higher in the dual-task condition compared with the single-task condition. Because we predicted that interference should be eliminated at this level of training, this small and

Table 11
Single-Task Reaction Time (ms) as a Function of Training and Addend in Experiment 5

Training	Addend 2	Addend 3	Addend 4	*M*	Slope
Old–old (144 trials/fact)	1,075	1,160	971	1,069	−52
New–old (48 trials/fact)	1,156	1,269	1,184	1,203	14
New–new (48 trials/fact)	1,140	1,214	1,053	1,136	−44

Table 12
Percent Error on Alphabet–Arithmetic (AA) Verification in Experiment 5

Training on AA verification	Concurrent task: Sequential months	Concurrent task: Control	Difference
Old–old (144 trials/fact)	11.7%	7.3%	4.4%
New–old (48 trials/fact)	13.7%	6.8%	6.9%
New–new (48 trials/fact)	10.7%	5.3%	5.4%

nonsignificant trend needs to be evaluated. It corresponds to the difference in error rates observed after overlearning in Experiment 4 and for NA (Experiment 3), and was much less than the interference observed in other experiments at initial automaticity. The amount of interference (4%) corresponds to that observed in the control task of Experiments 1 and 2 (3.6%). We conclude that interference from sequential month saying was essentially eliminated by overlearning by either learning by remembering (Experiment 4) or by learning by doing (Experiment 5).

We were somewhat surprised, however, that significant interference was not observed for the facts with less training. Interference of sequential month saying was observed in Experiment 1 after 46 exposures per fact. Why then was there so little interference for the new–old and new–new conditions of Experiment 5 for which each fact had received about the same number of exposures (48)? Perhaps the training of Experiment 5 was more efficient than that of Experiment 1 because analysis of the stimulus and generating appropriate responses was required on every exposure in Experiment 5, but not in Experiment 1 for which reciting the displayed fact was all that was required on many exposures. Also the facts were often presented in blocked rather than random sequence in Experiment 1.

General Discussion

Summary of Findings

With increased training, the RT for AA verification became shorter and less related to digit addend. Also interference from concurrent month saying was reduced. However, improvement on the three measures (RT and interference from two types of month saying) was not correlated in that some aspects of the improvement reached asymptote before others as training progressed. The details of the findings can be summarized in terms of the pattern of results at three levels of training.

Level 1, novice. At this level, RT for AA verification was long, and increased linearly and strongly as a function of digit addend. Both types of month saying, sequential (January, February, March, and so on) and repetitive (January, January, January), interfered markedly (and apparently equally) with AA verification.

Level 2, automatized. At this level, RT for AA verification became shorter and independent of addend. Repetitive month saying produced little interference with AA verification, but sequential month saying continued to interfere.

Level 3, trained beyond what is needed for automaticity. As was true at Level 2, RT was independent of addend at Level 3. However, RT was shorter at Level 3 than at Level 2. As was true at Level 2, there was essentially no interference from repetitive month saying. However, unlike Level 2, interference from sequential month saying was also eliminated.

Analysis of Changing Processes as Training Continues

The contrast between Levels 1 and 2 (Experiments 1 and 2) replicated our previous results for RT (Logan & Klapp, 1991), and provides data on interference that converge with the RT data to support our analysis. At Level 1, repetitive speech interfered with AA verification and RT increased with addend (N), as expected if AA verification was mediated by speech-like counting through N steps of the alphabet. At Level 2, interference from repetitive speech vanished, and RT became independent of N, as expected if the speech-like process was no longer needed. Furthermore, a concurrent task involving memory retrieval (sequential month saying) interfered with AA verification at Level 2, as expected if AA was supported by memory retrieval at that stage of training.

Interference from sequential month saying, which occurred at Level 2, was reduced by Level 3, and overall RT was reduced from Level 2 to Level 3 (Experiments 3, 4, and 5). These results answer the question posed by the title of this article. Yes, there are practice effects after automaticity is achieved. This conclusion fits the expectation that automaticity should manifest the types of findings that have been observed in the study of learning and memory in other contexts. In particular, it corresponds to the classic finding that memory performance can be enhanced by overlearning (i.e., by training beyond the level needed to achieve an initial memory representation).

According to our analysis, the concurrence cost associated with pairing month saying and AA verification is due to different types of conflicts at different levels of training. At the novice level, the interference is assumed to be between concurrent speech-like processes for alphabet counting and overt month saying. However, after automaticity, AA verification is assumed not to be supported by alphabet scanning, thereby eliminating the conflict between two concurrent speech processes. Therefore, repetitive month saying did not produce interference. However, both sequential month saying and automatized AA verification are assumed to involve memory retrieval. One would expect that the month sequence would be retrieved quickly because it is highly overlearned, but that the AA facts would be retrieved slowly at Level 2 because these facts are not well mastered. Because it is unlikely that the fast month retrieval can occur at a time not also occupied by the slow retrieval of AA facts, interference occurs at Level 2. However, we assume that, after the AA facts become overlearned at Level 3, they also can be retrieved quickly. When both retrievals occur quickly, the probability that they will overlap in time is decreased, thereby reducing interference.

Although it seems justifiable to attribute the greater difficulty of sequential as opposed to repetitive month saying to the additional process of retrieval needed to arrive at the next month in the sequence, other possibilities should be considered. Sequential month saying may be regarded as globally more difficult than repetitive month saying; yet these two tasks produced nearly equivalent interference at Level 1 of practice, suggesting the need for a more articulated interpretation such as the one we propose. Another interpretation could be based on the possibility that whereas repetitive month saying may not require that the speech act be reprogrammed for each utterance, sequential month saying clearly requires new programming for articulating each month. However, programming may be a special case of memory retrieval (Klapp, 1976). Retrieval of the month name (as in our preferred interpretation) and retrieval of the articulatory motor program cannot be distinguished by the present data.

Memory Retrieval Theory of Automaticity

In our analysis, automaticity in AA verification is achieved when the slow, controlled algorithm of alphabet counting is replaced by direct access to the information. This fits the essential notion of automaticity: that the correct response is available without the need to "compute" it. The theory is that automaticity develops when the correct AA statements are learned strongly enough so that access to memory replaces the algorithm of counting through the alphabet (Logan & Klapp, 1991). Improvement can continue as learning progresses beyond initial achievement of automaticity because memory retrieval can become more efficient than needed to just barely replace the counting algorithm.

This interpretation of automatizing AA is consistent with theories in which automatic processes are assumed to be supported by memory retrieval (e.g., Logan, 1988; Schneider & Detweiler, 1987; Strayer & Kramer, 1990). This approach specifies a mechanism underlying automaticity (memory retrieval), and can be contrasted with traditional approaches that describe automaticity in terms of its properties (e.g., fast, effortless, and based on extensive practice), but do not specify a mechanism (e.g., Hasher & Zacks, 1979; LaBerge & Samuels, 1974; Logan, 1978; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Logan and Klapp (1991) used the term "property list approaches" to refer to these views, which share the strategy of listing properties, but may disagree concerning which properties to place on the list.

The results we report in this and our previous study (Logan & Klapp, 1991) correspond to predictions of the memory approach. Extensive training is not required for automaticity. Training need not be on the true–false AA verification task itself; it is only necessary that the true AA statements be memorized. Because development of automaticity is due to learning, it should exhibit other properties of memory. In particular, additional training beyond automaticity can further improve performance. These results are not predicted by approaches that imply that extensive training on the criterion task itself is needed to achieve automaticity, and that all of the properties of automaticity emerge simultaneously and fully developed. Although it may be possible to accommodate our results by adaptations and extensions of traditional approaches, these findings are handled much more directly and in a principled way by automaticity as memory theory.

This contrast between memory-retrieval theory and traditional approaches to understanding automaticity may be illustrated further by comparing their treatments of the property of lack of effort in the sense of not being sensitive to interference from concurrent physically compatible tasks. One difficulty with insisting that effortlessness is an all-or-none attribute of automaticity is methodological. This view must predict that all tasks, even those not yet tested, will fail to produce interference when automaticity has been achieved. The present data on interference of month saying with AA verification illustrate the difficulty involved in testing this assertion. At Level 2, AA verification was free of interference from repetitive month saying, a task that interfered with novice AA verification. However, another concurrent task—sequential month saying—did interfere at Level 2. If we had tested only repetitive month saying, we might have concluded (incorrectly) that all dual-task interference vanished by Level 2. By contrast to the confusion that can follow from assuming that lack of interference is a necessary property of automaticity, the memory theory accommodates the different patterns of interference that were observed at different levels of practice, including further reduction in interference after automaticity had been achieved.

Does Our Analysis Generalize Beyond AA Verification?

We claim that AA verification becomes automatized when memory retrieval replaces an algorithmic process (counting through the alphabet). Although our conclusion that automaticity represents replacing an algorithm with fast memory retrieval was developed for AA verification, there is no reason to suppose that this principle is applicable only to that one task. However, other examples that intuitively seem to be automatic appear to require a very different analysis. Consider, for example, a sensory–motor skill such as bicycle riding. Surely performance on this task is not based on retrieval of verbal associates from memory as in our analysis of AA verification.

The distinction between declarative and procedural memory (Squire, 1987; Willingham, Nissen, & Bullemer, 1989) may be relevant to understanding the difference between automated AA verification and automated bicycle riding. Our AA verification task involved declarative rather than procedural memory. We have shown that subjects can be taught the AA statements using a flash-card approach that does not provide practice in the AA verification task itself (Logan & Klapp, 1991, Experiment 2). This type of training would produce declarative memory. Furthermore, subjects report that they know the AA facts after automaticity is acquired (Logan & Klapp, 1991, Experiment 1). By contrast, sensory–motor tasks such as bicycle riding can be performed without a declarative understanding of the memory representation that underlies performance.

Thus, we propose that tasks like AA verification may become automatic as the result of replacing algorithmic procedures (such as alphabet counting) by reliance on declarative knowledge, and that sensory–motor tasks may become automatized when performance is based on retrieval from procedural memory. Reports that procedural skills can develop in the absence of declarative knowledge even in normal people (Cohen, Ivry, & Keele, 1990; Stadler, 1989; Willingham et al., 1989) support the notion that some tasks can rely on procedural memory with no corresponding declarative memory. For both forms of memory, automaticity is assumed to be achieved when fast, direct-access memory retrieval replaces slower mechanisms that can underlie action. Although automatization of these tasks may differ in terms of the form of memory underlying performance, each task become automatized as the result of the development of fast memory retrieval.

An implication of this interpretation is that lack of awareness is not a necessary property of automaticity. Whereas automaticity based on declarative knowledge (such as our AA tasks) is open to awareness of the memory underlying performance, automaticity based on procedural memory does not require awareness of the relevant memory trace. This possibility points out once again the advantage of the analysis of automaticity in terms of processes rather than in terms if rigidly specified lists of properties. In our analysis, lack of awareness is not a necessary property of automaticity because awareness is expected when automaticity is supported by retrieval from declarative memory.

The influential experiments reported by Shiffrin and Schneider (1977) may have included examples of each of these cases of automaticity. Two forms of scanning were simultaneously involved in most of these experiments: (a) scanning a set of items held in memory (Sternberg scanning), and (b) scanning a set of items presented concurrently on a visual display (visual search), for which targets "pop-out" after training. With practice on consistently mapped versions of these tasks, performance on each became independent of the number of items in memory or in the display. However, the mechanisms underlying practiced performance on these two aspects of the task may be different (Flach, 1986). Performance on memory scanning may be supported by declarative knowledge in that subjects may have associated some items with a positive response and some with a negative response, thereby eliminating the need to scan a set of items in memory. By contrast, the visual-search aspects of the task may have involved development of a filtering procedure based on physical features (Flach, 1986; Krueger, 1984).

Conclusion

Automaticity as memory theory reduces the problem of understanding automaticity to the problem of understanding "ordinary" memory input and retrieval. It assumes that memory retrieval can replace algorithmic processing and that retrieval speed increases with practice (thereby leading to practice effects after automaticity is achieved), and it incorporates the distinction between declarative and procedural memory systems. It also assumes that multiple retrieval processes interfere with each other, leading to dual-task interference when retrieval processes overlap temporally. These assumptions are hardly radical, and are not restricted in application to automaticity paradigms. Thus, our analysis encourages the unification of memory and automaticity theory, and offers no encouragement for the development of a theory of automaticity apart from application of the general theory of learning and memory.

References

Cohen, A., Ivry, R. I., & Keele, D. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

Flach, J. M. (1986). Within-set discrimination in a consistent mapping search task. *Perception & Psychophysics, 39,* 397–406.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Klapp, S. T. (1976). Short-term memory as a response preparation state. *Memory & Cognition, 4,* 721–729.

Klapp, S. T. (1981). Temporal compatibility in dual motor tasks. II. Simultaneous articulation and hand movements. *Memory & Cognition, 9,* 398–401.

Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics, 35,* 558–564.

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

Logan, G. D. (1978). Attention in character classification: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107,* 32–63.

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: 2. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 177–193.

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning* (2nd ed.). New York: Longmans, Green.

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

Posner, M. I., & Snyder, C. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 21,* (pp. 53–119). San Diego, CA: Academic Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1–66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

Stadler, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1061–1069.

Strayer, D. L., & Kramer, A. (1990). An analysis of memory-based theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 291–304.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1047–1060.

Received December 27, 1989
Revision received July 30, 1990
Accepted July 31, 1990 ■