Memory & Cognition
1998, 26 (1), 135-142

The time-course of the generation effect

RODERICK W. SMITH and ALICE F. HEALY
University of Colorado, Boulder, Colorado

The *generation effect*, in which items generated by following some rule are remembered better than stimuli that are simply read, has been studied intensely over the past two decades. To date, however, researchers have largely ignored the temporal aspects of this effect. In the present research, we used a variable onset time for the presentation of the to-be-remembered material, thus providing the ability to determine at what point during processing the generation effect originates. The results indicate that some benefit from generation attempts occurs even when subjects have only a few hundred milliseconds in which to process the stimulus, but that more of the benefit occurs later. This finding suggests that the generation effect results from continuous or multiple discrete stages of information accrual or strengthening of memory traces over time, rather than from a single discrete increment upon final generation.

The *generation effect*, first described by Slamecka and Graf (1978), is a robust phenomenon in which recall or recognition of a stimulus list is enhanced if a person must generate the list by using some rule (e.g., rhyming, synonym–antonym relations, multiplication) as opposed to simply reading the list. Theories describing this phenomenon include effort (see, e.g., Griffith, 1976; McFarland, Frey, & Rhodes, 1980) or arousal (Jacoby, 1978); semantic activation (e.g., Graf, 1980; McElroy & Slamecka, 1982); relational processes (e.g., Donaldson & Bass, 1980; Rabinowitz & Craik, 1986); multiple-factor explanations (e.g., Hirshman & Bjork, 1988; McDaniel, Riegler, & Waddill, 1990; McDaniel, Waddill, & Einstein, 1988); and a procedural account (e.g., Crutcher & Healy, 1989; McNamara & Healy, 1995).

To date, however, researchers have largely ignored the time-course of the generation effect. It is not clear whether the benefit from generation is derived from a continual strengthening of memory traces during the generation process as a whole, from a small number of discrete changes during the course of generation, or from a single discrete change in memory representation at the time when generation is completed. Slamecka and Fevreiski (1983) did study the related question of whether recall or recognition would be improved in the event of generation failures. They manipulated the difficulty of an antonym generation and found that subsequent recall levels for stimuli that had been attempted but not fully generated were roughly as high as the levels for completely successful generations, but that recall was much lower for stimuli that had only been read. Recognition tests showed intermediate levels of the generation effect for unsolved antonyms in this research. Slamecka and Fevreiski interpreted these findings as evidence that generation in this task involved at least two processes: generation of a semantic code and subsequent association of this code with a surface representation (an English word, in this case). According to this interpretation, the generation failures were largely failures in the second stage; the first stage was successful, and the generation effect derived largely from this first stage.

Although Slamecka and Fevreiski's (1983) research suggests that completion of generation may not be necessary for the generation effect to occur, it leaves open the question of whether the memorial advantage for generated items results from a single discrete change, a small number of discrete changes, or a continuous change. All of these patterns have been observed in other phenomena. The *speed–accuracy decomposition* (SAD) technique (Meyer, Irwin, Osman, & Kounios, 1988; see also Smith, Kounios, & Osterhout, 1997), for example, measures *partial information*, or the ability of subjects to provide an accurate response before processing has been completed. This measure is related to discreteness or continuity, in that certain types of continuous processes would be expected to produce rising levels of partial information over time, whereas certain types of discrete processes would not. Specifically, processes that continuously transmit information to subsequent processes, to use Miller's (1988) terminology, would likely produce partial information, whereas those that transmit information discretely would not. (This reasoning is described in greater detail in Smith & Kounios, 1996.)

SAD research has provided evidence for processes producing a slow growth or intermediate levels of partial information in subjects performing word recognition (Meyer et al., 1988), study–test recognition memory (Ratcliff, 1988), semantic memory access (Kounios, Osman,

This research was supported in part by Army Research Institute Contract MDA903-93-K-0010 to the University of Colorado (A.H., Principal Investigator). Correspondence should be sent to R. W. Smith, who is now at the Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut St., Suite 400C, Philadelphia, PA 19104-6228 (e-mail: rodsmith@linc.cis.upenn.edu), or to A. Healy, Department of Psychology, Campus Box 345, University of Colorado, Boulder, CO 80309 (e-mail: ahealy@clipr.colorado.edu).

—**Accepted by previous editor, Geoffrey R. Loftus**

& Meyer, 1987), and complex semantic relations (Kounios, Montgomery, & Smith, 1994). In contrast, Smith and Kounios (1996) applied the same SAD technique to anagram problem solving and found little or no partial information, suggesting that a single discrete shift rather than a continuous change was responsible for anagram solution.

The use of metacognitive measures such as *feeling of knowing* and *feeling of warmth* (both here abbreviated as FOK) has produced results similar to those obtained with SAD. In FOK research, subjects provide, at regular intervals, estimates of the likelihood that they will produce an answer during a difficult "tip-of-the-tongue" recall or problem-solving task. Metcalfe (1986a, 1986b) found that subjects' FOK ratings increased gradually over time and predicted success in later recall of trivia questions, but did not predict correct solution of insight problems or anagrams. Rather, FOK ratings for such problems did not increase gradually over time in Metcalfe's (1986b) research when the solution was correct, but incorrect solutions were predicted by a preceding rise in FOK ratings. Algebra problem solving, as studied by Metcalfe and Wiebe (1987), followed a pattern similar to that for trivia questions. As with the SAD results, these studies suggest that some cognitive processes operate in a slow and continuous way, whereas others produce rapid or possibly truly discrete changes.

The findings from SAD and FOK studies lead to the question of how the generation effect emerges over time. In the studies cited above, measures of processing during retrieval or problem solving were used, whereas the generation effect presumably arises from processing during learning. The question is therefore whether the generation effect emerges in a relatively discrete fashion, as do the solutions to anagram and insight problems or in a relatively continuous fashion, as in retrieval in memory-related tasks. Note that *discrete* and *continuous* represent two ends of a continuum, as discussed by Miller (1988), and it is possible for a process to exist in an intermediate state, if some moderate number of steps are used to process or transmit information. In a mathematical sense, such a process would be considered discrete, but it is useful in psychology to distinguish such a process from those which use only one or some smaller number of intermediate states or stages.

Fundamentally, the generation effect represents an increased level of storage (or conceivably of retrievability as a result of storage processes) in certain conditions as opposed to others. Studying the generation effect thus serves as a window into storage processes in the memory system in general. If it is found that the generation effect arises in a relatively continuous fashion, this would represent a sort of symmetry with the previous results found with time-of-test measures; information in memory would then not only be accessed in small chunks to build a response, but might be stored in a similar piecemeal fashion. A finding of discrete effects, however, would indicate a theoretically interesting difference between storage and retrieval, because it would indicate the possibility of a dissociation: Information would be stored in one form but retrieved in another, as when a librarian shelves whole books but recovers them page by page. (See Crowder, 1976, pp. 264–273, for further discussion of incremental vs. all-or-none learning.)

In the present experiments, we investigated the time-course of the genesis of the generation effect. Because the effect is defined in terms of recall or recognition memory after an entire stimulus list has been presented, and because it is presumably produced at the time of study even though it is measured at the time of recall, time-of-test measures such as SAD and metacognitive judgments cannot be applied to it. We therefore utilized a new extension of a typical generation effect design. In a normal generation effect experiment, each to-be-recalled item is presented to subjects immediately, typically with an associated cue; or only the cue is presented, and subjects are required to produce the to-be-recalled item. In the present experiments, we added intermediate conditions, in which the associated cue was presented alone for a brief period, but the to-be-recalled item itself was presented after a delay shorter than subjects took to generate the items in the pure-generate condition. These intermediate conditions gave subjects time to work on the production of the to-be-recalled item, but not so much time that they would generally be able to complete the solution before seeing it. Therefore, if the generation effect is due entirely to a single discrete change in memory traces upon completed generation, subjects would gain no benefit from the intermediate conditions; but if the generation effect is the result of ongoing restructuring or elaboration during processing, the intermediate conditions would provide some, but not all, of the benefit of the full-generate condition.

Because the existing explanations of the generation effect are largely mute with regard to its time-course, the present experiments were not intended to help us distinguish between these theories; rather, they provide information that may be useful as one attempts to elaborate the existing theories and to tie the effect to other memory phenomena, as detailed in the General Discussion.

The task selected for these experiments was multiplication problem solution. This task has been studied in the past (see, e.g., Crutcher & Healy, 1989; Gardiner & Rowley, 1984; McNamara & Healy, 1995) and has produced a robust generation effect.

EXPERIMENT 1

This experiment was designed as a straightforward examination of the time-course of the generation effect, with experimental conditions optimized to provide the greatest chance of detecting any effect in the intermediate conditions.

Method

Subjects and Design. Forty-eight students, who were recruited from introductory psychology classes at the University of Colorado, received course credit for participation. Solution lag time (described below) was treated as a within-subjects variable, and the subjects were placed in equal numbers in each of four counterbalanced conditions for assignment of solution lag time to specific stimuli.

Materials and Apparatus. The multiplication problems were drawn from the pool used by Crutcher and Healy (1989). This pool consists of the problems from the 2-times through the 12-times multiplication tables. There are 40 unique two-digit solutions in this pool. The 10-times problems were eliminated as being potentially too easy for the present purposes, and from the thereby reduced pool, 28 problems, each with a unique solution, were selected. In addition, three practice stimuli were generated from outside the 2-times through 12-times multiplication tables, though these answers (26, 51, and 78) were all two-digit numbers. The smaller of the two multiplicands always appeared first (e.g., "4 × 8" rather than "8 × 4"). A single random stimulus ordering was constructed for all subjects.

As in prior generation effect research, some stimuli were assigned to the read condition and others to the generate condition. In addition, we used two intermediate conditions, in which the solutions appeared 250 msec and 500 msec after the appearance of the multiplication problem. We refer to these four conditions as having different solution presentation lag times. The intermediate lag times were selected as less than typical response times for similar multiplication problems in previous research (e.g., Fendrich, Healy, & Bourne, 1993). Within each of the first three eight-stimulus blocks, and within the final four-stimulus block, equal numbers of stimuli were assigned to the four solution lag times (read, 250 msec, 500 msec, and generate). The assignment of lags to stimuli was rotated to produce four counterbalanced conditions, and, across these conditions, an equal number of assignments of each stimulus was made to each solution lag time.

The stimuli were presented on Amdek monochrome (amber) computer monitors driven by IBM PC and PC/XT computers. All displays were in the standard IBM PC monochrome font (80 × 25 character display mode). During the learning phase, the subjects typed in each response on the numeric keypads of the computers, and during the recall phase, they wrote the multiplication problem solutions on sheets of paper bearing instructions, using pens or pencils.

Procedure. Upon entering the laboratory, all subjects were given brief verbal directions and seated at a computer, which displayed the following instructions:

> In this experiment, you will solve simple multiplication problems in your head. You will see the problem appear on the screen. When you know the answer, type it on the numeric keypad (on the rightmost side of the keyboard), followed by the <Enter> key. On some trials, the computer will give you the correct answer, either immediately or after a short delay. If this happens, simply type the answer the computer gives.

Each trial of the learning phase began with three "+" signs in the center of the screen as a fixation point. These were replaced after 1,000 msec by the multiplication problem itself, which remained on screen through the rest of the trial. In the read condition, the solution appeared at the same time as did the problem (e.g., "5 × 12 = 60" would appear from the outset). In the intermediate conditions, the solution appeared 250 or 500 msec after the problem (e.g., "5 × 12 =" would appear, and then "60" would appear to the right of the "=" in the problem). In the generate condition, the problem appeared, but not the answer. The subjects were unaware of the solution lag status of a trial until the solution did or did not appear. In all cases, the words "Your response:" were displayed two lines below the problem, and the subject's response was echoed after this prompt, as it was typed. The problem and, when applicable, the answer, remained on the screen until the subject typed a response, terminated by the <Enter> key. The computer recorded the time to the first keypress and the time to the subsequent keypresses (the second digit and the <Enter> key), as well as the accuracy of each response. The computer provided response accuracy feedback (correct vs. incorrect) to the subjects.

The subjects initially solved the three practice problems (one each of read, intermediate [500-msec lag], and generate conditions) to familiarize themselves with the procedure and then went on to the 28 experimental trials. The subjects were not informed that they would be tested for their memory of the experimental trial solutions. On completion of this learning phase, the computer automatically proceeded to present a filler task, which required the subjects to produce free-associate responses to 35 common English words. This task was designed to take approximately 2 min, as had been confirmed with pilot testing. After finishing the free-associate task, the subjects were given a sheet of paper along with written instructions and a pen or pencil, and they were asked to recall as many of the multiplication problem answers as they could. The written instructions were as follows:

> This is the final part of the experiment. In the space below, please write down as many of the multiplication problem *answers* as you can remember from the first phase of the experiment. Please try not to write down numbers which did not occur in the experiment; only write down answers which you saw or generated yourself.

If subjects wanted clarification, it was provided verbally. The subjects were given as much time as they required for recall, and they were not timed on this task.

Results

Overall levels of recall are presented in Table 1, which shows a typical generation effect for this type of task: The read condition has a substantially lower level of recall than does the full-generate condition. The two intermediate conditions, with presentation lags of 250 and 500 msec, show intermediate levels of recall in this task.

This description is borne out by statistical analyses. A preliminary analysis of variance (ANOVA) including the factor of counterbalanced condition yielded no statistically significant main effect or interaction involving this factor, so subsequent analyses were conducted without it. An overall ANOVA of presentation lag yielded a statistically significant effect of lag [$F(3,141) = 24.52$, $p < .001$]. Further comparisons were performed with a Newman-Keuls test. This test revealed that the read condition differed significantly from the 250-msec condition and that the 500-msec condition differed from the generate condition (both $ps < .05$) but that the 250- and 500-msec conditions did not differ from each other ($p > .10$).

Response times (RTs) in the study phase are of interest primarily as evidence that the answer presentation lags used (250 and 500 msec) were not so late that subjects would already have generated a response by the time the

Table 1
Recall Levels (Proportion Recalled) for All Stimuli and Only Those With Responses 500 Msec or More After the Solution Presentation (Slow Responses) for the Intermediate Lags, Experiment 1

	All Stimuli		Slow Responses	
Solution Onset	*M*	*SEM*	*M*	*SEM*
Read	.310	.032		
250 msec	.399	.034	.401	.034
500 msec	.402	.032	.400	.032
Generate	.619	.033		

answer was presented. If response generation had been completed on a significant number of 250- and 500-msec lag trials, the recall for these intermediate conditions would effectively have been contaminated by full-generate stimuli. Similarly, response accuracies in the study phase are of interest as evidence that subjects were attending to the task. Hence, Table 2 presents the mean time to the first keypress (RT), the minimum RT across all subjects, the mean of all subjects' minimum RTs, and the accuracy of responses. As can be seen, the mean RTs are well above even the longer of the two presentation lags, and even the fastest responses were longer than the associated lag times, so the possibility of contamination of these intermediate conditions by completed solutions seems small. Similarly, accuracies are high, at 94% overall.

Nonetheless, a further analysis was conducted which eliminates the stimuli in the intermediate lag conditions, for each subject, to which a physical response in the study phase occurred within 500 msec of the presentation of the solution (i.e., faster than 750 msec and 1,000 msec in the 250-msec and 500-msec lag conditions, respectively)—thus eliminating the stimuli that would have been the most likely sources of completed processing contamination in the intermediate conditions. Previous research (e.g., Fendrich et al., 1993) has shown that typed answers to simple multiplication problems have mean RTs of under 1.4 sec, with some conditions producing mean responses of under 1 sec, suggesting that the motor control processes involved likely take no more than roughly half a second. (Problems in this study were somewhat harder, and so would have elicited higher RTs because of higher level cognitive factors, but motor control differences presumably would have been minor.) Simple RTs, of course, are much lower than this value (Ollman & Billington, 1972), but also involve less complex cognitive and motor processing. Thus, selecting 500 msec as the cutoff point should have eliminated most of the items that would have been solved before their answers were presented on the intermediate lags, though of course it is impossible to determine with certainty precisely which items might have been contaminated.

The means for this analysis are presented in the "Slow Responses" column in Table 1 (the read and generate values are unaffected by this analysis). This analysis reduced the number of responses in the 250-msec condition to six for only 1 subject; and in the 500-msec condition, the average number of responses was reduced to 6.69 (range, 4–7). As in the initial analysis, the effects of counterbalanced condition were not statistically significant, so the main analyses do not include this factor. The overall results did not change with this new analysis. An ANOVA of these data revealed a significant overall effect of lag [$F(3,141) = 24.53, p < .001$]. A Newman-Keuls test indicated three clusters: read, 250 and 500 msec, and generate; the 250- and 500-msec conditions did not differ from each other ($p > .10$), but all other comparisons were significant ($p < .05$).

Table 2
Response Times (in Milliseconds) and Accuracies (Proportion Correct) for Study Phase, Experiment 1

	Mean RT		Minimum RT			
Solution Onset	*M*	*SEM*	Min	*M*	*SEM*	Accuracy
Read	1,857	90	813	1,252	36	.958
250 msec	1,810	78	499	1,240	40	.946
500 msec	1,927	74	543	1,314	53	.955
Generate	2,744	189	808	1,295	65	.914

Discussion

These results strongly suggest that the generation effect is caused, at least in part, by a continuous or multistage discrete change during stimulus processing, as opposed to a single discrete change upon completion of processing. Although the magnitude of the effect at the intermediate lags of 250 and 500 msec is small compared with that in the full-generate condition, these values are statistically significantly higher than those in the read condition. Comparison of these recall values with those from the full-generate comparison is difficult, since the only measure of processing time available for the full-generate condition is the time to the first keypress during study, and this measure includes variable motor control and other processing times not included in the presentation lag times. Using first-keypress RTs as a basis for comparison is problematic because of possible differences in subject response strategies—as is discussed in the General Discussion. Nonetheless, the fact that first-keypress RTs are over 800 msec longer in the full-generate condition than in the 500-msec lag condition indicates that the smaller magnitude of the generation effect at the shorter lags is not at all surprising; if an ongoing process produces the generation effect in a relatively continuous fashion, one would expect an effect with an intermediate magnitude at these times in processing.

The difference in recall from the 250- to the 500-msec condition is also quite small. This finding may indicate that the intermediate-lag effect is due to a multistage change in representation. For instance, if an operand retrieval strategy were used (see, e.g., McNamara & Healy, 1995), and if retrieval of the operands occurs in a discrete fashion, intermediate levels of recall might fall at one or more constant "plateaus." Another explanation is that the 250-msec time difference between the two intermediate conditions simply is not great enough to reveal much variation at this point in processing. That is, a continuous process need not necessarily provide a linear output, and it could result in a relatively steep initial period of rising recall (0–250 msec) followed by a relatively shallow rise (250–500 msec). Particularly when combined with normal statistical fluctuations, such an explanation could account for the observed results. In either case, the present results do suggest that some ongoing processing produces the generation effect, be it in multiple discrete steps or truly continuously.

EXPERIMENT 2

In the preceding experiment, we utilized a within-subjects design with randomized order of presentation of stimuli in the different lags and a single recall phase at the end of the experiment. Some have argued, however, that such a design may result in an inflation of estimates of the generation effect due to differential attention to the different stimulus types, displaced rehearsal, or similar mechanisms (e.g., Begg & Snider, 1987; Hirshman & Bjork, 1988; Slamecka & Katsaiti, 1987). We therefore conducted a second experiment with a between-subjects design, in order to control for such effects.

Method

Subjects and Design. Seventy-four subjects were recruited from the same pool as in Experiment 1 and were given course credit for participation. One subject's data were discarded for failure to follow instructions, and the data from the final subject tested were dropped to bring the number in each condition equal at 18. A single random ordering of stimuli was used, and lag condition (read only, intermediate lags of 250 and 500 msec, and full generate) was constant for all stimuli for any given subject. The subjects were assigned a lag condition on a fixed rotation based on the order of their entry to the experiment. A free-associate phase, consisting of generation of 35 free associates, occurred after the multiplication problem phase, and a recall phase occurred after the free-associate phase.

Materials and Apparatus. The same multiplication problem and free-associate filler stimuli were used as in Experiment 1. The stimuli were presented on the same computers as had been used in Experiment 1. The subjects recalled the multiplication answers on sheets of paper with brief written instructions.

Procedure. The procedure was identical to Experiment 1's, except for adjustments to the written and verbal instructions, which were based on the lag condition.

Results

The results are presented in Table 3. For the 500-msec lag, 4 subjects produced one or more responses (i.e., first keypress) in the study phase prior to the presentation of the answer (a total of 10 such responses of 504 total in this condition). Thus, the total recallable items for these subjects were reduced to between 23 and 27 instead of the usual 28. As in Experiment 1, recall increased with increasing presentation lag time. An ANOVA of these results indicated a statistically significant overall effect of lag [$F(3,68) = 9.24$, $p < .001$]. A Newman-Keuls test showed that the 500-msec condition differed from the read condition, and the generate condition differed from the read and both intermediate conditions (all ps $< .05$). Recall in the read condition was not different from that in the 250-msec condition, nor did the 250- and 500-msec conditions differ from each other (ps $> .10$).

Table 4 presents the mean RT and both the minimum RT across all subjects and the mean of subjects' minimum RTs observed in the study task in this experiment, as well as the response accuracies in this task. Although overall RTs are comparable in the two experiments (a mean of 2,085 msec in Experiment 1 vs. 2,013 msec in Experiment 2), Experiment 2 produced faster minimum RTs. These fast RTs argue for an analysis that excludes stimuli that might have been the result of completed problem solving.

Table 3
Recall Levels (Proportion Recalled) for All Stimuli and Only Those With Responses 500 Msec or More After the Solution Presentation (Slow Responses) for the Intermediate Lags, Experiment 2

	All Stimuli		Slow Responses	
Solution Onset	*M*	*SEM*	*M*	*SEM*
Read	.361	.032		
250 msec	.421	.030	.421	.030
500 msec	.488	.038	.473	.038
Generate	.615	.042		

Note—Under the "All Stimuli" column for the 500-msec solution onset condition. 10 stimuli were excluded because the first keypress was produced before 500 msec had elapsed.

Table 4
Response Times (in Milliseconds) and Accuracies (Proportion Correct) for Study Phase, Experiment 2

	Mean RT		Minimum RT			
Solution Onset	*M*	*SEM*	Min	*M*	*SEM*	Accuracy
Read	2,038	208	697	1,135	59	.980
250 msec	1,825	142	292	987	70	.986
500 msec	1,683	100	346	853	81	.962
Generate	2,501	186	743	1,113	51	.974

An additional analysis was performed after items were removed in which subjects' responses occurred within 500 msec of the appearance of the answer for the 250-msec and 500-msec lag conditions (faster than 750 msec and 1,000 msec, respectively). The means for this analysis are presented in the "Slow Responses" column of Table 3. This procedure reduced the mean number of recallable items to 27.8 and 24.4 for the 250- and 500-msec conditions, respectively. The ANOVA again revealed a significant overall effect of lag condition [$F(3,68) = 9.08$, $p < .001$]. A Newman-Keuls test showed that the generate condition differed from the read and intermediate conditions at $p = .05$, but that these three conditions did not differ among themselves at this level. The 500-msec condition was marginally significantly different from the read condition, however ($p < .10$).

Discussion

These results confirm the finding that even a few hundred milliseconds' thought can affect recall in a generation effect paradigm. The initial analysis showed statistically significant effects, including differentiation of the 500-msec condition from both the read and generate conditions. A more conservative test excluding recall of possibly generated items retained overall statistical significance and the effects involving generated stimuli, but reduced the read versus 500-msec comparison to marginal statistical significance. The magnitude of the effect in this experiment is comparable to that in Experiment 1; indeed, it is slightly higher, at a .127 difference in recall from the read to the 500-msec condition in Experiment 2 and a .092 difference in Experiment 1 (.112 and .090, respectively, for the more conservative measures).

GENERAL DISCUSSION

The results of these experiments are noteworthy for two reasons. First is the finding of intermediate levels of recall for the intermediate presentation lags. This result strongly suggests that the generation effect derives from ongoing processing at the time of study rather than from major changes at the time of completed generation. This processing could take the form of a small number of discrete accrual "episodes" or of a more continuous process over the course of several hundred milliseconds. Thus, this finding represents a close parallel to findings of ongoing increases in partial information in memory-related phenomena with SAD (e.g., Kounios et al., 1994; Kounios et al., 1987; Meyer et al., 1988) and to findings of accurate metacognitive judgments in answering trivia questions (Metcalfe, 1986a, 1986b) and algebra problem solving (Metcalfe & Wiebe, 1987). Thus, whatever explanation of the generation effect is invoked (e.g., an effort explanation, a procedural account, etc.), the details of this account must include a relatively continuous, rather than a single discrete, memory change as the generated items are produced, or a series of discrete changes in memory traces, at least when the generation rule involves simple multiplication problems.

As an example, consider the procedural account of the generation effect (e.g., Crutcher & Healy, 1989; Healy et al., 1992; McNamara & Healy, 1995). This explanation holds that the generation effect is produced by the fact that subjects are more likely in a generate condition than in a read condition to reinstate mental processes at the time of test that were used at the time of study. In the case of multiplication problems, these processes would consist of the direct retrieval or working out of the arithmetic calculation. If subjects in such an experiment are led to use an operand retrieval strategy at test, in which they recall operands and perform the multiplication operations as an aid to recalling the answers to the multiplication problems encountered during study, then the mental processes used at test are more likely to match those used at study in the generate condition than in the read condition. The finding of intermediate levels of recall elaborates on this framework in that it specifies that the memory advantage most likely comes from several distinct processes, a process with multiple psychologically distinct stages or traces, or a process that produces a continuous change in memory representation. If the process of solving the math problems were a single discrete process that produced the generation effect, the present experiments would not have yielded intermediate levels of recall.

Multiple or continuous processes or stages could derive from any of a number of arithmetic computation strategies. For instance, even in adults and for simple problems of the type used here, multiplication skills may not be entirely automatic but may be composed of a mixture of direct retrieval and algorithmically based performance (see, e.g., Logan, 1988; see also Newell, 1990). Therefore, on some trials, subjects may utilize an algorithm, such as counting or a mediated retrieval strategy (e.g., $4 \times 9 = 4 \times 10 - 4 = 36$; see, e.g., Healy et al., 1993), to solve the problems, and this procedure would leave more potential retrieval cues in memory than would the read condition. In the intermediate conditions, such strategies would have been initiated but not completed, leaving more retrieval cues than in the read condition, but fewer than in the generate condition.

Another possibility is that multiplication skill for simple problems is based largely or wholly on direct retrieval, but that this retrieval activates, over time, a number of additional memory nodes that could be reinstated at the time of recall. For instance, Campbell and Graham (1985) propose that mental multiplication is carried out, in part, through associations of each operand with the solution. Thus, a problem such as 4×7 brings up associations with 8, 12, 16, and other multiples of 4, as well as 14, 21, 28, and other multiples of 7. According to Campbell and Graham, it is the combination of these associations, in addition to associations to the problem as a whole, which results in retrieval of the solution from memory. In the current context, these individual associations could also be used as cues in the retrieval of solutions from memory. In the intermediate conditions, the memory retrieval would presumably produce some useful cues, as in the partial information findings of Kounios et al. (1994; Kounios et al.,1987) or Meyer et al. (1988). The fact that recall levels for the intermediate conditions in the present experiments were not as high as those for the regular generate conditions indicates that this memory retrieval process can be cut short. This is itself a theoretically interesting inference. It is in line with research indicating an ability to inhibit skilled performance (e.g., Logan, 1982), but on a faster time scale; but it is opposed to at least some theories of memory retrieval (e.g., Soar, as described by Newell, 1990), which postulate a single discrete process which cannot be halted once begun, at least for a single simple memory retrieval.

Slamecka and Fevreiski (1983) found recall levels for nongenerated items as high as for generated items in their Experiments 1 and 2, whereas in their Experiment 2 they found three distinct levels of recognition for read, failed generation, and successful generation items. They interpreted this finding to mean that two stages were at work in their antonym task, the first one involving semantic features (meanings and associations), and the second, surface features (letters). The Slamecka and Fevreiski research indicates the possibility of failed generation producing only semantic features in an antonym task, in which a presumably semantically driven recall task was unaffected by the failure to engage in surface processing. The present research maps imperfectly onto the Slamecka and Fevreiski work because of assorted task differences. It involves a task with different types of semantic features (based on the multiplication table rather than word meanings), yet it also shows evidence for partial processing. In this respect, the two experiments complement each other well, because they provide converging evidence on the conclusion that incomplete processing can produce at least a reduced generation effect.

Given the number of differences between the two experiments' designs, it is difficult to interpret the differences in result patterns, most importantly Slamecka and Fevreiski's (1983) finding of full recall even in the incomplete generate condition compared with our finding of intermediate levels of recall. This difference could result from the task variable; it might be that memory traces are laid down differently in the solving of multiplication problems as opposed to the generation of antonyms. If so, our findings may serve as a warning concerning the generalizability of results from generation effect and other memory research utilizing only one type of task. Another possibility is that the method used to produce "partial" generation is critical; we used a task interruption, whereas Slamecka and Fevreiski used item difficulty (low stimulus information) to induce failed generations. It is possible that our task was simply more sensitive to the ongoing memory process, because it was intended as a time-course measure more than was the Slamecka and Fevreiski task. In this view, intermediate presentation lags might ideally be able to produce recall equivalent to that in the full-generate condition if they were placed late enough; but normal RT variability and uncertainty concerning the time to produce a response after completed cognitive processing make this test a practical impossibility.

The second major point of interest in these findings is the fact that the overall generation effect was large in both within- and between-subjects designs. In the within-subjects Experiment 1, recall increased from .310 to .619 from the read to the generate conditions; whereas in the between-subjects Experiment 2, equivalent means were .361 and .615. Previous research (e.g., Hirshman & Bjork, 1988; Slamecka & Katsaiti, 1987) has often yielded much smaller generation effects in between- than in within-subjects experiments. Soraci et al. (1994) found differences in the magnitude of the generation effect for linguistic stimuli by varying the relatedness of the cue and target words. That is, in some conditions, subjects wrote down or read words related to the cue words, and in others, unrelated words were used. In a free recall task, unrelated generations produced a generation effect, whereas related generations did not. Soraci et al. interpreted their research as supporting a "multiple-cue" account of the generation effect, in which the more the possible retrieval cues that exist, the better will be recall; according to Soraci et al., unrelated generations involved deeper processing, leaving more memory traces than did related generations. This account is compatible with the procedural account (e.g., Crutcher & Healy, 1989; McNamara & Healy, 1995), and it explains the presence of a generation effect even in our between-subjects Experiment 2 by the (presumably) rich memory traces laid down while the math problem was being solved, whereas tasks used in some other experiments (e.g., Hirshman & Bjork, 1988; Slamecka & Katsaiti, 1987) might not produce such complex memory traces and thus would yield little or no generation effect in a between-subjects design.

One unexpected pattern in the present results which deserves comment is the dip in mean and minimum RTs from the read to one or both of the intermediate conditions (Tables 2 and 4; note particularly the minima of minimum RTs). In Experiment 1, this dip was not very pronounced for the mean RTs, and in fact differences among these three conditions' RTs were not statistically significant [$F(2,94) = 2.28, p = .108$]. For the minimum RTs, the pattern appears to have been more pronounced for the minimum of minimums, but not for the mean of subject minimums, and the latter yielded only a marginally significant result [$F(2,94) = 2.42, p = .094$]. For Experiment 2, the dip appears to have been more pronounced even for the mean RTs, but this difference was not statistically significant [$F(2,51) = 1.29, p = .283$]. The means of the minimum RTs, however, did differ among the first three conditions [$F(2,51) = 4.00, p = .024$]. This finding suggests the possibility that subjects might have been using different processing strategies for the different conditions, particularly in Experiment 2, which had a between-subjects design. The most plausible explanation for a dip in minimum RTs in the intermediate lags would be that on some trials subjects were able to use the math problem as a frame to fixate the location at which the answer would appear, thus largely ignoring the problem itself and speeding up the initial stages of encoding the visually presented answer. By contrast, subjects in the full-read condition might have read the entire problem. Note that, if this explanation is correct, this strategy might be expected to have produced worse memory for the solution in the intermediate conditions than in the read condition, thus making the statistically significant results that we did find (and which were presumably produced by trials on which subjects did not use this hypothesized strategy) all the more impressive.

How generalizable are these conclusions? In the present research, we utilized only one generation rule: multiplication. If the existing research on partial information in various tasks (e.g., Kounios et al., 1994; Kounios et al., 1987; Meyer et al., 1988) is used as a guide, we would expect to see a benefit from partial generation in a variety of other memory phenomena. Slamecka and Fevreiski's (1983) study on failed generation attempts for antonyms is consistent with this suggestion. In their recent SAD research, Smith and Kounios (1996) found no evidence of partial information when they used anagram problem solving, however. This finding raises the interesting possibility that if such a paradigm were applied to the generation effect, there might be no benefit from brief and incomplete solution attempts.

REFERENCES

Begg, I., & Snider, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 553-563.

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

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Crutcher, R. J., & Healy, A. F. (1989). Cognitive operations and the generation effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 669-675.

Donaldson, W., & Bass, M. (1980). Relational information and memory for problem solutions. *Journal of Verbal Learning & Verbal Behavior*, **19**, 26-35.

Fendrich, D. W., Healy, A. F., & Bourne, L. E., Jr. (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Cognitive psychology applied* (pp. 111-133). Hillsdale, NJ: Erlbaum.

Gardiner, J. M., & Rowley, J. M. C. (1984). A generation effect with numbers rather than words. *Memory & Cognition*, **12**, 443-445.

Graf, P. (1980). Two consequences of generating: Increased inter- and intraword organization of sentences. *Journal of Verbal Learning & Verbal Behavior*, **19**, 316-327.

Griffith, D. (1976). The attentional demands of mnemonic control processes. *Memory & Cognition*, **4**, 103-108.

Healy, A. F., Clawson, D. M., McNamara, D. S., Marmie, W. R., Schneider, V. I., Rickard, T. C., Crutcher, R. J., King, C. L., Ericsson, K. A., & Bourne, L. E., Jr. (1993). The long-term retention of knowledge and skills. In D. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 135-164). New York: Academic Press.

Healy, A. F., Fendrich, D. W., Crutcher, R. J., Wittman, W. T., Gesi, A. T., Ericsson, K. A., & Bourne, L. E., Jr. (1992). The long-term retention of skills. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 87-118). Hillsdale, NJ: Erlbaum.

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 484-494.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning & Verbal Behavior*, **17**, 649-667.

Kounios, J., Montgomery, E. C., & Smith, R. W. (1994). Semantic memory and the granularity of semantic relations: Evidence from speed–accuracy decomposition. *Memory & Cognition*, **22**, 729-741.

Kounios, J., Osman, A. M., & Meyer, D. E. (1987). Structure and process in semantic memory: New evidence based on speed–accuracy decomposition. *Journal of Experimental Psychology: General*, **116**, 3-25.

Logan, G. D. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 778-792.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.

McDaniel, M. A., Riegler, G. L., & Waddill, P. J. (1990). Generation effects in free recall: Further support for a three-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 789-798.

McDaniel, M. A., Waddill, P. J., & Einstein, G. O. (1988). A contextual account of the generation effect: A three-factor theory. *Journal of Memory & Language*, **27**, 521-536.

McElroy, L. A., & Slamecka, N. J. (1982). Memorial consequences of generating nonwords: Implications for semantic-memory interpretations of the generation effect. *Journal of Verbal Learning & Verbal Behavior*, **21**, 249-259.

McFarland, C. E., Jr., Frey, T. J., & Rhodes, D. D. (1980). Retrieval of internally vs. externally generated words in episodic memory. *Journal of Verbal Learning & Verbal Behavior*, **19**, 210-225.

McNamara, D. S., & Healy, A. F. (1995). A procedural explanation of the generation effect: The use of an operand retrieval strategy for multiplication and addition problems. *Journal of Memory & Language*, **34**, 399-416.

Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 288-294.

Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 623-634.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, **15**, 238-246.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review*, **95**, 183-237.

Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica*, **67**, 191-257.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Ollman, R. T., & Billington, M. J. (1972). The deadline model for simple reaction times. *Cognitive Psychology*, **3**, 311-336.

Rabinowitz, J. C., & Craik, F. I. M. (1986). Specific enhancement effects associated with word generation. *Journal of Memory & Language*, **25**, 226-237.

Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, **95**, 238-255.

Slamecka, N. J., & Fevreiski, J. (1983). The generation effect when generation fails. *Journal of Verbal Learning & Verbal Behavior*, **22**, 153-163.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 592-604.

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory & Language*, **26**, 589-607.

Smith, R. W., & Kounios, J. (1996). Sudden insight: All-or-none processing revealed by speed–accuracy decomposition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1443-1462.

Smith, R. W., Kounios, J., & Osterhout, L. (1997). The robustness and applicability of speed–accuracy decomposition, a technique for measuring partial information. *Psychological Methods*, **2**, 95-120.

Soraci, S. A., Franks, J. J., Bransford, J. D., Chechile, R. A., Belli, R. F., Carr, M., & Carlin, M. (1994). Incongruous item generation effects: A multiple-cue perspective. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 67-78.

(Manuscript received June 19, 1996;
revision accepted for publication December 9, 1996.)