Journal of Experimental Psychology:
Learning, Memory, and Cognition
1988, Vol. 14, No. 2, 248–255

0278-7393/88/$00.75

The Mnemonic Value of Perceptual Identification

James S. Nairne
University of Texas at Arlington

In four experiments, subjects were required to name words presented on a CRT screen. On generate trials, the words were presented quickly, at a point where roughly half could be identified correctly; on read trials, the items were presented for a full second, allowing for rapid and easy naming. A surprise recognition test for the presented items then revealed a substantial retention advantage for the briefly presented items, but no similar advantage was produced in recall. It is argued that under rapid viewing conditions subjects may fail to extract enough visual features to allow for immediate resolution, requiring the initiation of a kind of data-driven generation process. This latter process then produces a generation effect for the briefly presented items compared with the read items, but only on a retention test that shows sensitivity to data-driven processing. These results are discussed from the standpoint of current theoretical views on the generation effect.

People will remember a familiar verbal item better if it has been self-generated during study rather than read (e.g., Jacoby, 1978; McFarland, Frey, & Rhodes, 1980; Russo & Wisher, 1976; Slamecka & Graf, 1978). Although this *generation effect* has proven tractable as an empirical phenomenon, there is, at present, uncertainty about its theoretical base. One popular account, the lexical activation hypothesis, argues that generation produces greater activation of an item's semantic attributes than does reading, which may result in a greater number of functional retrieval routes for generated items (e.g., see McElroy & Slamecka, 1982; Nairne, Pusen, & Widner, 1985; for related accounts, see Gardiner & Hampton, 1985; Payne, Neely, & Burns, 1986). A contrasting view, based on the general principle of transfer-appropriate processing (Morris, Bransford, & Franks, 1977), argues that generation is beneficial to performance because it typically uses cognitive operations that are likely to be matched by those engaged during test (e.g., Glisky & Rabinowitz, 1985; Kolers & Roediger, 1984). Most generation tasks require the subject to produce to-be-remembered information from its general context (e.g., the opposite of cold is h___?) and such conceptually driven processing (that is, processing which is driven by contextual inference) is thought to match up well with the processes driving recall and recognition (see Jacoby, 1983b; Roediger & Blaxton, 1987).

Although these accounts are certainly compatible on many levels, the transfer-appropriate processing view makes the unique prediction that the advantages of generation will not be widespread, but rather will be restricted to particular testing environments. In support of this position, Jacoby (1983b) found that prior experience with reading a verbal item transfers more appropriately to its later perceptual identification than does generation, whereas generation shows the performance advantage in recognition. Compatible results have been obtained by Rabinowitz and Craik (1986), who found that generation enhanced performance only when cues at test helped tap the same information that was used to guide the initial generation process. More directly, Glisky and Rabinowitz (1985) and Nairne and Widner (1987) have shown that the size of the generation effect can be increased whenever subjects are required to generate items again at test, prior to a recognition decision. Collectively, these studies document the importance of the actual procedural operations underlying generation, as well as the overlap that generation potentially provides with the operations performed at test. Generation, it seems, may not produce a better or stronger memory trace than reading; instead, reading and generating may simply produce different memory traces that transfer well to some testing environments but not to others (Jacoby, 1983b).

The fact that the generation effect is typically found on recall and recognition tests, as stated earlier, is thought to result from the conceptually driven processing that most generation tasks require. Subjects are asked to produce to-be-remembered information as an inference from context rather than from a direct visual analysis of letters, and these operations are assumed to be compatible with the operations required by recall and recognition. Yet, in principle, it seems likely that generation could be guided by a variety of means, including some that might not easily be classified as conceptually driven. Consider, for example, the completion of isolated word fragments (Glisky & Rabinowitz, 1985) or the switching of two underlined letters in a presented word (Nairne et al., 1985). Both of these tasks would appear to be primarily data driven, in the sense of occurring in the absence of context, yet each produces significant generation effects in recognition and recall. Such findings lend support to lexical activation accounts, which assert that increased semantic activation is an automatic consequence of generation, regardless of the generation means.

To explore these issues further, the present experiments were designed to examine the mnemonic value of perceptual identification, which is conceptualized as another example of a data-driven generation task. In perceptual identification, the

Thanks go to Bob Widner for valuable discussions about these data and to Cynthia Taylor for help in conducting these experiments.

Correspondence concerning this article should be addressed to James S. Nairne, Department of Psychology, Box 19528, University of Texas, Arlington, Texas 76019.

subject is required to name, usually by saying aloud, verbal items that are rapidly presented on a CRT or through a tachistoscope. The identification process differs from simple reading, presumably, in that the former does not allow for immediate resolution of the presented visual information. One can assume that only fragmentary visual information is actually perceived, from which the subject attempts to generate a response through matching the available visual fragments with permanent memory codes (Morton, 1979), prior episodic experiences (Jacoby, 1983a), or both (Feustel, Shiffrin, & Salasoo, 1983). This notion that perceptual identification is primarily a data-driven task is operationally clear (e.g., no conceptual context is provided), but, in addition, it fits in well with the available data using the task as a retention measure. Jacoby and Dallas (1981) showed, for example, that perceptual identification shows little sensitivity to the depth of processing engaged in a prior episode or the amount of study time. If perceptual identification is also appropriately considered as a generation task, guided by overt "data" rather than conceptual means, then the pertinent empirical question asks whether such a task yields a generation effect on standard measures of retention.

Experiment 1

Experiment 1 used a surprise recognition test to assess the retention of words that had earlier been flashed on a CRT screen for the subject to identify. The details of the perceptual identification task were modeled after some earlier work by Jacoby (Jacoby, 1983a; 1983b; Jacoby & Dallas, 1981), in which perceptual identification was used as a retention measure to assess the effects of prior episodic exposure (see also Witherspoon & Allan, 1985). In the present case, subjects were asked to name singly presented words that were preceded by two short horizontal markers and followed by a row of number signs. The critical manipulation, defining a trial type, centered on the duration of exposure for the to-be-identified items. On *generate* trials, the words were presented quickly, for an interval designed to achieve about 50% correct identification; on *read* trials, the words were presented for a full second, allowing for rapid and easy naming. Following the identification phase, all subjects were given a recognition test in which they were required to circle, from a list containing targets and distractors, items that they felt had occurred before in the experiment.

Method

Subjects and apparatus. The subjects were 28 undergraduates from the University of Texas who participated for course credit. During the identification trials, all stimulus materials were presented in lower case letters on a Televideo 950 CRT screen, which was controlled by a Northstar Horizon microcomputer.

Materials. The stimulus materials were four- and five-letter, single-syllable, medium- to high-frequency nouns (Kučera and Francis, 1967, mean frequency count was 77); the items were rotated through the major conditions of the experiment, appearing as both read and generated items across subjects.

Procedure. The experiment was divided into three parts. The first part was designed to give subjects experience with the two trial types, and to establish a timing interval for each subject that would yield about 50% correct performance on the generate trials. On each of 40 trials, subjects were told to name, by saying aloud, the word that was presented on the CRT screen; guessing was encouraged to promote responding even if the subject was uncertain about what had been presented. Trials began with the command "Push the return key when ready," followed, when initiated, by a pair of horizontal bar markers which were separated by 16 mm. These markers always appeared in the center of the screen for 500 ms, surrounding the location where the to-be-named item would be appearing. Next, the item appeared for either one second (read trial) or for a shorter, predetermined interval (generate trial). The word was then replaced by a row of six number signs, in the same location, to act as a mask. The number signs remained on for 1 s and were followed by a second appearance of the word stimulus for another 2 s. Subjects were instructed merely to look at this second occurrence to obtain feedback about the correctness of their response. Items were presented twice on each trial to ensure that everyone would have some experience with each of the presented words, regardless of the trial type.

This initial phase of 40 trials was separated into five eight-trial blocks, with each block containing four read and four generate trials, randomly mixed. The words in the first four generate trials were presented for an interval that allowed for relatively easy identification (determined by pretesting), whereas the durations in later blocks were either shorter or longer to produce the desired level of performance.[1] By the end of the fifth block, a duration was chosen for each subject for the next part of the experiment. In the second part, subjects were presented with 24 critical items for identification; again, half were read items and half were generate, randomly mixed, and subjects were told to name each item. No mention was made of the following retention test. In addition to the 24 critical items, the first three trials and last three trials in the second phase contained buffer items which were not included in subsequent analyses.

Immediately following the second identification phase, subjects were asked to count backward from 100 by 3s (to remove short-term memory influences) prior to administration of the surprise recognition test. Subjects wrote their counting responses on a blank sheet of paper; the task took from 1 to 2 min to complete. The surprise recognition test contained 36 items of which 24 were items presented during the second identification phase (12 generate and 12 read) and 12 were distractors taken from the same item pool. The items were typed on two sheets, with each sheet containing six examples of each item type randomly mixed. Subjects were instructed to circle only those items that they thought had occurred before in the experiment; each subject was given as much time as needed to complete the test. All subjects received the same recognition test.

Results and Discussion

During the critical identification phase, as expected, subjects correctly named 100% of the items presented for a full second. On generate trials, the mean proportion of words

[1] Stimulus presentation durations were controlled in these experiments by a "for loop" in a Pascal computer program rather than through direct access to a real-time clock. Consequently, although timing "numbers" are available for each subject, they do not represent actual real-time values and therefore are not reported. It is possible, however, to estimate the range of timing intervals used on generate trials (including phosphor decay): the durations varied from approximately 20 to 60 ms. Once again, the particular selected values were determined by each subject's performance during the first identification phase.

correctly identified was .58 (range = .33–.92, SD = .16). When an error occurred, 55% of the time subjects substituted an incorrect items as a response. Although these intrusions were truly self-generated by the subject, they were not tested in later recognition memory because of item-selection concerns.

Of principal interest are the recognition data, which are displayed in Figure 1. Because this study is concerned with the mnemonic effects of perceptual identification as a generation task, it is necessary to draw a sharp distinction in recognition performance between "old" items that were identified correctly on generate trials and those that were not. A scoring criterion of this sort is required because presented items that were not named correctly by the subject cannot be said to have been self-generated; compared with the standard generation experiment, such items are classified as generation "failures" (see Slamecka & Fevreiski, 1983). Consequently, proportions were calculated individually for each subject, keeping separate generation successes (G) and generation failures (GF). Figure 1 displays the overall mean proportion correct in each of these conditions, compared with read items (R). The mean false alarm rate was quite low (.04).

The data were first subjected to an overall analysis of variance (ANOVA) which indicated the presence of significant differences among the three conditions, $F(2, 54) = 12.75$, $MS_e = .03$, $p < .001$. A subsequent Neuman-Keuls analysis (alpha was set at .05) revealed that correctly named items from the generate trials (.82) were recognized significantly better than items from either the read trials (.58) or the generation failures

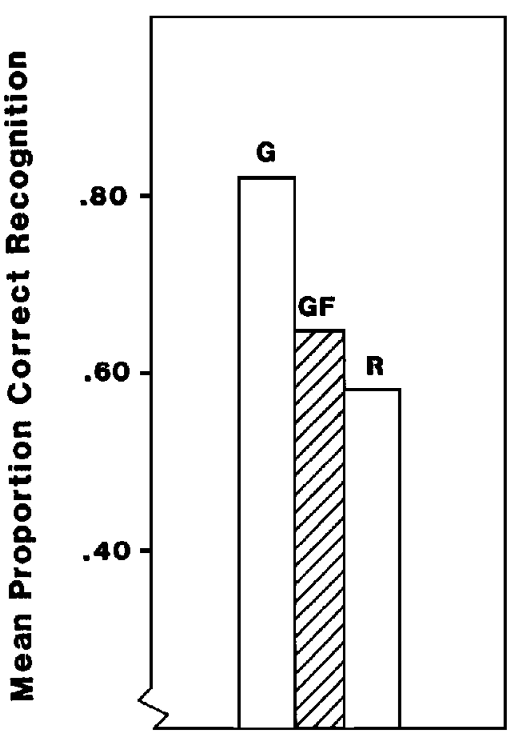

Figure 1. Recognition performance in Experiment 1 for the correctly-named items from the generate trials (G), the generation failures (GF), and the read trials (R).

(.65); these latter two conditions did not differ reliably. For the comparison of major interest, then, one finds a substantial generation effect: Out of 28 subjects, 23 recognized more generated items than read items and there were only three reversals (23+, 3−, 2 ties). Such a finding is consistent with the claim that perceptual identification involves self-generation of material, which presumably serves to enhance later episodic retrieval of the generated items.

One can also more closely examine the comparison between the generation successes and generation failures. Recent research by Slamecka and Fevreiski (1983) has indicated that generation failures, even when assessed through recognition memory, typically produce significant generation effects compared with reading. In the present case, however, generation failures did not differ from read items in recognition performance; specifically, for the 28 subjects, 13 showed better performance for the generation failures, 12 subjects showed the reverse pattern, and there were three ties. According to Slamecka and Fevreiski (1983), who used a generation rule requiring the production of semantic opposites, generation failures can yield mnemonic advantages over reading because the generation process is usually multistaged. Subjects may fail to emit the final surface attributes that define a generation success, but still generate significant semantic attributes that are related to later retention of the item (that is, assuming that surface attributes are at some point made available to the subject). One need only assume that the particular stages that make up the generation process are influenced by the nature of the production rule to explain the present failure to replicate. Because the generation rule was conceptually based in Slamecka and Fevreiski's (1983) study, it seems likely that the generation process involved the generation of semantic attributes; in the present instance, however, the generation task was primarily data driven and thus similar semantic activation seems less likely. Consequently, although no advantage was found for generation failures compared with reading, these results seem consistent with the spirit of Slamecka and Fevreiski's reasoning.

Finally, there is a methodological concern that needs to be mentioned. Because of the high error rate on the generate trials (which, of course, was experimenter induced), one could interpret the reported generation advantage as an item-selection effect. For example, it is possible that the generation failures, which were not included in the G versus R comparison, were simply difficult items to remember. When removed, then, one is left with a set of items that is easier to remember in the generation condition. One can argue against such a claim in several ways: First, an item-selection argument would predict that the generation failures, because they are difficult to remember, should be recognized less well than the read items on the recognition test; this was not the case. Second, it has been shown in a number of studies that subjects are more likely to identify high-frequency words than medium- or low-frequency words in a perceptual identification task (e.g., Jacoby & Dallas, 1981). If this is true, then it would be reasonable to argue that the final group of items in Condition G would tend to be of functionally higher frequency than items in Conditions GF or R. Yet, the word frequency effect in recognition memory shows that higher frequency words are actually less likely to be recognized than low- or medium-frequency words; clearly, this was not the pattern obtained in Experiment 1. Finally, and most directly, one can simply not conditionalize the data; that is, one can compare the read and generate conditions when no distinction is drawn between successful and unsuccessful identification. Under these conditions (G + GF vs. R), once again a significant generation effect emerged. Subjects correctly recognized .74 of the items from generate trials and .58 from the read trials (19+, 2−, 7 ties). Thus, no support is given to an item-selection interpretation of the comparison of interest.

Experiment 2

The major purpose of Experiment 2 was to replicate Experiment 1 with a slight procedural change. Rather than presenting each item twice on a trial (i.e., one time as response feedback), items were shown only once. Feedback was eliminated simply to ensure that the generation effect of Experiment 1 could not be attributed in some way to differential processing of the feedback among the conditions. It is possible, for example, that even when subjects successfully identified items on the generate trials, there was enough uncertainty about what was presented to support full processing of the item during its second presentation (see Cuddy & Jacoby, 1982; Jacoby, 1978). On read trials, in contrast, there may have been little, if any, processing of the second occurrence because subjects were certain about what was initially presented. In this case, the locus of the generation advantage would be placed in the processing of the feedback rather than as a manifestation of the data-driven identification process itself.

In addition to the feedback issue, Experiment 2 was also designed to examine the mnemonic value of the identification task on a more conceptually driven retention test. It is well known that recognition memory is sometimes sensitive to data-driven processing (e.g., Johnston, Dark, & Jacoby, 1985; Mandler, 1980), but this is less true, if at all, for recall (see Roediger & Blaxton, 1987); consequently, one might not expect to find a generation effect for perceptual identification, which is a data-driven task, on a truly conceptually driven test. Following the identification trials, one group of subjects in Experiment 2 was asked to recall all of the items that had been presented in the experiment, in any order of occurrence. In an effort to improve recall somewhat, these subjects were first shown a list of all the words that had been presented during the critical identification phase and were told to try and remember them. A second group of subjects received a recognition test that was identical to the one given in Experiment 1.

Method

Subjects and apparatus. The subjects were 36 undergraduates from the University of Texas who participated for credit in an introductory psychology course. Subjects were assigned to the different groups of the experiment based on their order of arrival in the laboratory. The apparatus was the same as in Experiment 1.

Materials. The stimulus materials were the ones used in Experiment 1.

Procedure. In most respects, Experiment 2 was an exact replication of Experiment 1. Subjects participated in three experimental phases: Phase 1 was designed to establish an exposure duration for the generate trials; Phase 2 contained the 24 critical read and generate trials, and Phase 3 was the surprise retention test. The identification trials differed from the preceding experiment in only one way: The number signs were not followed by a second occurrence of the just presented item. This change was accomplished by merely deleting one line in the controlling computer program; thus, exactly the same intertrial intervals were used in both Experiments 1 and 2.

Immediately following the second identification phase, each subject was asked to count backward from 100 to 0 by 3s prior to administration of the retention test. The group receiving the recognition test (n = 18) then received the same test used in Experiment 1. Each subject was told to circle only items that had occurred previously in the experiment. In the group receiving the recall test (n = 18), subjects were first presented with each of the 24 critical items from the second identification phase (12 read and 12 generate), along with four primacy and recency buffers, on the CRT screen. Items were presented individually in random order for 4.5 s, separated by a blank screen for 500 ms. Subjects were instructed to try and remember each item during its presentation. Following the last item on the list, subjects were asked once again to count backward from 100 by 3s and then to recall each of the presented items in any order of occurrence.

Results and Discussion

Collapsed across the two retention groups, subjects correctly named 100% of the items presented on the read trials. On generate trials, subjects correctly named .57 of the presented items (range = .25–.92, SD = .13); the two retention groups did not differ in identification performance on the generate trials (in fact, each group named .57 of the generate items correctly).

The recognition data, which are displayed in Table 1, show a replication of the generation effect found in Experiment 1. An overall ANOVA on the hit rates from the three conditions (G, GF, and R) revealed the presence of significant differences, $F(2, 34) = 82.22$, $MS_e = .026$, $p < .001$; a Neuman-Keuls analysis indicated that all three conditions differed significantly from one another ($p < .05$). For the critical G versus R comparison, 13 subjects recognized more of the correctly named items from the generate trials, 4 subjects showed greater recognition of the read items, and there was one tie. This pattern confirms the generation advantage of Experiment 1 and allows us to place its locus in the identification stage itself rather than in differential processing of response feedback. Although these data do not allow us to comment on the role that feedback might play in enhancing G versus R differences, it is clear that feedback is not a necessary com-

Table 1
Retention Performance: Experiment 2

	Trial type		
Measure	Generate	Generate Failure	Read
Recognition	.77	.10	.59
Recall	.36	.37	.36

ponent for obtaining a highly reliable generation effect in perceptual identification.

Next, as in Experiment 1, it is of interest to compare performance on generate trials as a function of whether the item was correctly or incorrectly named during its single presentation. As the data clearly show, when the surface attributes of a generation failure are not presented again as response feedback, subjects do a very poor job of recognizing the item's prior occurrence. In fact, there is no significant difference in recognition performance between items in Condition GF (.10) and the overall false alarm rate (.11). This result is not surprising given that the subject, in some respect, was never functionally presented with the item; still, above-chance performance would have been understandable if some of the surface attributes had been encoded during presentation.

The lower half of Table 1 shows mean proportion correct performance on the recall test as a function of the three major conditions. An overall ANOVA on these data revealed no significant differences among the conditions, $F(2, 34) < 1$. For the comparison of major interest, 9 subjects recalled more of the correctly named items from the generate trials; 8 subjects showed the reverse pattern, and there was one tie. In sum, there is no indication of a generation effect in these data. Even more striking is the level of performance for the items in Condition GF. Whereas these items were recognized no better than chance, they were recalled at levels identical to those obtained in Conditions G and R. This lack of sensitivity to condition differences is bothersome, and most likely reflects the manner in which the retention test was conducted. It is possible that presenting subjects with a list of to-be-recalled items prior to recall, along with instructions to remember the items, may have masked any differences that were present following the identification phase. One could argue, for example, that the operations used during the conscious efforts to memorize were the controlling operations of recall, and that this reduced the likelihood that the effects of the prior perceptual identification phase would transfer. Whatever the explanation, the data remain important because they provide a good test of the item-selection account introduced at the close of Experiment 1. If the items in Condition GF were simply more difficult to remember than the items in Condition G, then they should have been more difficult to learn as revealed by performance on the recall test. Because no differences were found in recall of the G and GF items, this reduces the plausibility of an item-selection account of the basic recognition findings.

Experiment 3

The rationale for Experiment 3 was similar to that given for the preceding experiment: Does one find an advantage for perceptual identification over reading when retention is assessed through a truly conceptually driven test? In Experiment 3, however, rather than giving everyone a list of the critical items prior to recall, subjects were simply asked to free recall all of the presented items following the last identification trials.

Method

Subjects and apparatus. The subjects were 36 undergraduates who participated for course credit. Stimulus events, as in the preceding experiments, were presented on a Televideo 950 CRT screen and were controlled by a Northstar Horizon microcomputer.

Materials. The to-be-named items were chosen from the same pool as the previous experiments. The only difference was that three, rather than two, groups of items were rotated through the read and generate trials across subjects; again, each item participated equally often in the read and generate conditions.

Procedure. Subjects participated in two identification phases followed by a surprise retention test. The first phase was designed to familiarize subjects with the two trial types and to establish a timing interval for the generate trials of the second identification phase. During this second critical identification phase, subjects received 12 read and 12 generate trials, randomly mixed, surrounded by recency and primacy buffers of the same trial types. No response feedback was given in either identification phase. Following the last identification trial, everyone was asked to count backward from 100 to 0 by 3s and then to recall as many of the presented items as possible in any order of occurrence. Subjects were given a total of 4 min for free recall and were encouraged to guess if they needed to.

Results and Discussion

On the critical generate trials, subjects correctly named .62 of the presented items (range = .42–.83, $SD = .13$) and no errors were made on the read trials.

For the free-recall test, subjects correctly recalled .14 of the items that were correctly identified on the generate trials, .13 of the items from the read trials, and only .01 of the items which were generation failures. An overall ANOVA on these data revealed the presence of significant differences, $F(2, 70) = 25.55$, $MS_e = .007$, $p < .001$; the Neuman-Keuls test indicated that Conditions G and R differed from Condition GF but not from each other. For the G versus R comparison, 19 subjects recalled more items from Condition G than from Condition R, 14 subjects showed the reverse pattern, and there were three ties. Consequently, as in the preceding experiment, there was no evidence of a generation effect in recall. One could argue that the present comparison is somewhat more convincing, however, because recall was not preceded by presentation of all of the to-be-recalled items; the latter may have masked condition differences in Experiment 2. As expected, subjects were unable to recall items that were not functionally presented (Condition GF).

Because of the low levels of recall in this experiment, it is necessary to consider the possibility of a floor effect. Perhaps a generation effect would have emerged in free recall if a higher level of performance had been obtained. To check on this possibility, the data were broken down into high and low responders, on the basis of the overall recall levels irrespective of condition. Specifically, subjects who recalled at least 3 of the 24 critical items were designated as high responders (range = 3–9; $N = 13$) and subjects who recalled 2 or 1 items were termed low responders ($N = 20$); there were 3 subjects who recalled nothing and these individuals were excluded from the analysis. For the high recallers, .25 of the items which were correctly identified on generate trials were recalled, compared with .22 of the read items. Although there is a small mean advantage for generation, the G versus R comparison did not approach significance 7 subjects recalled more items from Condition G and 6 subjects showed higher recall of items from the read trials. For the low recallers, .10 of the items were recalled in both Conditions G and R. Out of 20 subjects, 12 showed a generate advantage and 8 subjects showed a read advantage. Clearly, these data provide no suggestion of an advantage for perceptual identification over reading on the surprise recall test; in this sense, these results replicate the recall outcome of Experiment 2.

Experiment 4

Experiment 4 was designed to take another look at recall; this time, however, subjects were given a surprise cued-recall test after the identification trials. The recall cues were general category names, taken from the Battig and Montague (1969) norms. It was expected that these cues would increase the overall recall levels somewhat relative to Experiment 3; in addition, it was felt that these cues might help the subject access any general semantic information that could have been encoded as a consequence of identification. Rabinowitz and Craik (1986) argued, for example, that general semantic information can be enhanced as a consequence of generation, especially when the generation process is not tightly constrained by relational information. Because perceptual identification occurs in the absence of contextual guidance, it is possible that general semantic cues may selectively aid later retention of identified items. On the other hand, if the identification process is truly data driven, then no differences should be found between the identified and read items on a cued-recall test, which is conceptually driven.

Method

Subjects and apparatus. The subjects were 20 undergraduates who participated for course credit. The to-be-named stimulus events were presented on the equipment used in the previous experiments.

Materials. The stimulus materials were high-frequency nouns (A or AA according to Thorndike & Lorge, 1944) which were four to six letters in length. The items were taken from the Battig and Montague (1969) norms with the criteria that each needed to be high frequency and each could only be a moderately likely category response. The mean ranking for the items in the category norms was 5.95; the range was three to nine. At test, the subjects were given the category names as stated exactly in the Battig and Montague norms. Each item was rotated through the major conditions of the experiment, participating as a read item and a generate item across subjects.

Procedure. Subjects participated in two identification phases followed by the surprise cued-recall test. The first phase contained 40 trials (half read and half generate) to familiarize the subjects with the procedure and to establish individual timing intervals for the following phase; once again, the step-wise procedure described in Experiment 1 was used to estimate a timing interval for each subject that would yield about 50% correct performance on generate trials. During the second identification phase, subjects were presented with 10 generate and 10 read trials, along with four primacy and recency buffers. Trial types were randomly mixed and subjects were told simply to name each item aloud as it appeared on the screen. No response feedback was given.

Following the last identification trial, subjects were asked to count backward from 100 to 0 by 3s and then were given a sheet containing 20 category names. Half of these cues were relevant to items that had been presented on generate trials and the other half were related to read items. Subjects were told to use the cues as aids in recalling information from the second identification phase; that is, episodic retrieval instructions were given and subjects were told not to write down simply the first response that came to mind. Subjects were given as much time as needed to complete the test.

Results and Discussion

Subjects correctly named .57 of the generate items (range = .20–.90, $SD = .15$); no errors were made on the read trials. On the cued-recall test, subjects recalled .24 of the items that were identified correctly on the generate trials, .25 of the items from the read trials, and .08 of generation failures. An overall ANOVA revealed the presence of significant differences, $F(2, 38) = 8.94$, $MS_e = .02$, $p < .01$; the Neuman-Keuls test indicated that Conditions G and R differed significantly from Condition GF, but not from each other. On the comparison of major interest, 8 subjects recalled more items from Condition G than R, 8 subjects showed the reverse pattern, and there were four ties. This pattern replicates the one found in Experiment 3: there were no differences between perceptual identification and reading when recall was used as the retention index. This null effect remained despite the fact that the recall levels were somewhat higher than in the previous experiment, and cues which were presumably sensitive to general semantic information were used during recall. The results of Experiments 2–4 indicate that whatever processes might be involved in perceptual identification, they yield no retention advantage over reading on a conceptually driven retention test.

General Discussion

The preceding four experiments were conducted to investigate the idea that items presented quickly for identification might be remembered better than items presented slowly. This rather unusual prediction originated from the claim that perceptual identification is a kind of functional generation task; under rapid viewing conditions, it was argued, subjects may fail to extract enough visual features to allow for immediate resolution, requiring the initiation of a kind of data-driven generation process. Consistent with the claim, when surprise retention for quickly presented words was assessed, a generation advantage was found on tests sensitive to data-driven operations, but not on tests driven by conceptual means. In Experiments 1 and 2, a substantial recognition advantage was found for items from the generate trials; the generate–read difference was quite large, consistent across subjects, and did not depend on whether feedback was given about what had actually occurred. However, when a conceptually driven recall test was administered, no generation effect was found across three separate experiments.

Furthermore, it is unlikely that one can appeal to any simple methodological concern to explain the recognition memory pattern. One argument, for example, might dwell on the fact that generation failures were not included in the analysis of generate trials, thereby setting up different item pools on generate and read trials. One can assume that each item set, prior to the experimental session, contained items that were easy to recognize (or identify) and items that were difficult to recognize (or identify); elimination of the generation failures, then, may have created a set of items that were easy to recognize in the generate condition. This is actually a problem in any generation experiment in which conditionalized data are presented, but is particularly problematic here because of the high error rates. Although item selection arguments are notoriously difficult to dismiss, several features of the present experiments argue against such an interpretation: For example, an overall analysis of the data in Experiment 1, when no distinction was drawn between generation successes and failures, still revealed a highly reliable generation advantage. In addition, in recognition memory one would expect high-frequency words to be recognized less well than low-frequency words; yet, the perceptual identification literature clearly shows that the probability of successful word identification is positively correlated with word frequency (e.g., see Jacoby & Dallas, 1981). If the items in Condition G are then of higher frequency, they should have been recognized less well than the read items, not better. Finally, if the recognition memory results truly reflected idiosyncratic item differences, then one would have expected similar effects in recall and recognition; yet, very different results were obtained for the two retention measures. Collectively, then, these arguments provide persuasive evidence against a item-selection account of the recognition data.

Perceptual Identification as Generation

Although these results have been described throughout as indicating the presence or absence of generation effects, one might quarrel with the idea that perceptual identification is actually a generation task. Certainly from a nominal perspective, generate trials resembled read trials: In each case, the whole item was presented for the subject to identify, the only difference was that items were presented quickly in one case and relatively slowly in the other. Of course, the pertinent theoretical question then asks why the brief, difficult to perceive, presentation condition yielded the superior recognition performance. Such a finding, among other things, represents an extreme violation of the total-time hypothesis (see Cooper & Pantle, 1967). Perhaps if subjects had been aware that certain items were going to be presented quickly, one could argue, then they would have developed a set to pay more attention to the surface features of those items when they appeared. Yet, in the present experiments, the read and generate trials were randomly mixed throughout the session, so any initial spurt of attention should have occurred equally for each trial type.

It has been reported a number of times in the perceptual identification literature that intrusion errors tend to resemble the actual presented item visually (e.g., Jacoby & Dallas, 1981; Morton, 1964). A natural interpretation of this result would be to assume that subjects actually perceive only fragmentary information (or at least some kind of degraded visual form), from which a likely visual candidate is generated. In fact, relatively sophisticated models currently exist that describe in detail how subjects might accomplish this feat (e.g., see McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). In this sense, perceptual identification is thought to resemble a kind of highspeed fragment completion task, at least under viewing conditions similar to the ones used on the present generate trials. Further, the generation idea fits in well with what subjects actually report after the identification sessions. Subjects frequently claim that the majority of their responses were "guesses" and are often surprised that they produced many correct responses. Moreover, although no systematic observations were conducted, it was not uncommon for subjects to report fragments as identification responses (e.g., "I saw an A, or an AE") rather than complete words. Such comments mesh well with the claim that the identification process is generative.

If one accepts this idea that perceptual identification is a generation procedure, then the present results can be interpreted from within the frameworks of existing theoretical accounts of the generation effect. Most particularly, these data seem consistent with the transfer-appropriate processing view, which places the locus of the generation advantage in the match between the operations performed at study and test (e.g., see Glisky & Rabinowitz, 1985; Kolers & Roediger, 1984). To the extent that the operations engaged during study resemble the operations required by the testing environment, then large mnemonic transfer effects will be expected. It is generally acknowledged that the processes involved in perceptual identification are not semantically or conceptually based; when used as a retention measure, for example, perceptual identification performance shows little sensitivity to variables like the depth of processing that an item has received during its initial episodic exposure (e.g., Jacoby & Dallas, 1981). It follows, then, that when the identification task is used as an encoding vehicle, as in the present experiments, one might not expect the mnemonic attributes that are produced to transfer well to a retention measure which is typically guided by conceptual means.

Whereas recall is traditionally viewed as a "constructive" retention measure, relying in large part on interitem elaborations that are established during encoding, somewhat different processes are thought to govern recognition. Variables that affect the success of recognition memory performance, typically, induce operations that influence the distinctiveness of an item's representation in memory. (Jacoby, Craik, & Begg, 1979). Apparently, the process of perceptual identification leads the subject to highlight those aspects of the presented stimulus that make it distinctive or unique. These distinctive features then aid the recognition process, which can rely on intraitem distinctiveness, but have little impact on recall, which relies on interitem elaborations. Once again an analysis of this sort makes intuitive sense: Given that the item is presented quickly, and only some of the presented stimulus features are extracted, the subject is required to undergo a decision-making process to decide what item has actually been presented. Similar processing is not required in the read condition because the available information allows for an immediate resolution of what has occurred. Because the decision-making process for the rapidly presented item is based on the available surface attributes only, one would expect these operations to transfer effectively only to retention tests that are sensitive to the perceptual characteristics of the test item.

Regardless of one's interpretation of the data, it is clear that perceptual identification can represent a potent mnemonic technique for improving later recognition memory performance. Performance on perceptual identification tasks has become popular in recent years as a dependent variable, and it has frequently been shown to be especially sensitive to the effects of prior data-driven processing. Most relevant to the present discussion is Jacoby's (1983b) finding that conceptually driven generation tasks transfer less well to a later identification task than does reading; encoding of the latter, presumably, is controlled primarily by perceptual characteristics rather than semantic ones. The present results seem perfectly consistent with this outcome; however, perceptual identification has been used here as an encoding vehicle, and quickly presented items were argued to be driven more by perceptual characteristics than simple reading. Whether such a conclusion will apply to all comparisons between reading and identification is, of course, unknown at present.

References

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs, 80*(3, Pt. 2).

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 68,* 221–234.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior, 21,* 451–467.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effects in word identification. *Journal of Experimental Psychology: General, 112,* 309–346.

Gardiner, J. M., & Hampton, J. A. (1985). Semantic memory and the generation effect: Some tests of the lexical activation hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 732–741.

Glisky, E. L., & Rabinowitz, J. C. (1985). Enhancing the generation effect through repetition of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 193–205.

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

Jacoby, L. L. (1983a). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 21–38.

Jacoby, L. L. (1983b). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 485–508.

Jacoby, L. L., Craik, F. I. M., & Begg, I. (1979). Effects of decision difficulty on recognition and recall. *Journal of Verbal Learning and Verbal Behavior, 18,* 585–600.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Johnston, W. A., Dark, V. J., & Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 3–11.

Kolers, P. A., & Roediger, H. L. III. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23,* 425–449.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

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 and Verbal Behavior, 21,* 243–259.

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

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533.

Morton, J. (1964). The effects of context on the visual duration threshold for words. *British Journal of Psychology, 55,* 165–180.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstal, & H. Bouma (Eds.), *Processing visible language I* (pp. 259–268). New York: Plenum Press.

Nairne, J. S., Pusen, C. P., & Widner, R. L., Jr. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition, 13,* 183–191.

Nairne, J. S., & Widner, R. L., Jr. (1987). Generation effects with nonwords: The role of test appropriateness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 164–171.

Payne, D. G., Neely, J. H., & Burns, D. J. (1986). The generation effect: Further tests of the lexical activation hypothesis. *Memory & Cognition, 14,* 246–252.

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

Roediger, H. L. III., & Blaxton, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus centennial conference* (pp. 349–379). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Russo, J. E., & Wisher, R. A. (1976). Reprocessing as a recognition cue. *Memory & Cognition, 4,* 683–689.

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

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

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Teachers College, Columbia University.

Witherspoon, D., & Allan, L. G. (1985). The effect of prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition, 13,* 101–111.

Received December 9, 1986
Revision received March 16, 1987
Accepted March 27, 1987 ■