

Is there something special about memory for internally generated information?

CAROL L. RAYE

Barnard College, New York, New York 10027

and

MARCIA K. JOHNSON and THOMAS H. TAYLOR

State University of New York at Stony Brook, Stony Brook, New York 11794

In two experiments, the number of times subjects saw and generated words was varied, and then half the subjects were asked to judge presentation frequency and half were asked to judge generation frequency. People were more sensitive to the relative frequency of internally generated, compared with externally generated, events. Also, judgments of internal events were less affected by whether or not the subjects were given a restricted range of numbers for their judgments, and internally generated events produced more confusion in estimating external event frequency than vice versa. The results are discussed in terms of potential differences in the characteristics of memory representations originating in external, compared with internal, events and the way these might interact with the judgment process.

This paper reports two experiments in which memory for internally generated events was compared with memory for externally generated events. External events are those occurring in the world around us and entering our awareness through the senses. Internally generated events originate in mental processes such as thought, imagination, and so on. While a sharp distinction between these two classes of events is perhaps impossible, it seems reasonable to distinguish memories in terms of their primary source.

Memory research has focused almost exclusively on memory for external events. When internally generated information has been studied (as in research on organization, imagery, encoding factors, context effects, etc.), it has primarily been from the point of view of understanding the role of thought in remembering external events (e.g., Bower, 1972; Cofer, 1973; Craik & Lockhart, 1972; Kintsch, 1970; Mandler, 1967; Martin, 1968; Paivio, 1969; Tulving, 1968; Tulving & Thomson, 1971). Various experiments along these lines tend to highlight either the potential contribution of imaginative processes to veridical memory (e.g., Bransford & Johnson, 1973; Paivio, 1969; Tulving, 1968) or the potential contribution of imaginative processes to inaccurate memory (e.g., Bartlett, 1932; Johnson, Bransford, & Solomon, 1973; Postman, 1954; Sulin & Dooling, 1974; Underwood, 1965). Only recently have a

few investigators begun to examine internal events as the targets of remembering (e.g., Doshier & Russo, 1976; Johnson, Taylor, & Raye, 1977; Klee & Gardiner, 1976; Slamecka & Graf, 1978), but most (with the exception of Slamecka and Graf) have not attempted to compare the memories generated by external events with the memories generated by internal events. Thus, while it is generally acknowledged that thought plays a very critical role in both remembering and misremembering external events, we have little information about memory for internally generated information itself and the way in which it might differ from memory for external events.

The present experiments represent two basic strategies for comparing characteristics of memories for internally and externally generated events. One strategy is to attempt, insofar as possible, to compare the two directly. The other is to determine whether a particular manipulation has similar effects on both. The first experiment represents an example of the first strategy, and the second experiment an example of the second strategy. The basic design used in both experiments was to present items (words) varying numbers of times to subjects and have subjects generate, or "rerepresent," those items varying numbers of times. The procedure was designed to control as much as possible the occurrence and content of items originating both from external and internal sources. Later, subjects were asked to estimate either the number of times an item had been presented as an external event or the number of times they had generated the item. In Experiment 1, the sensitivity of subjects to the relative frequency of internally and externally generated events was compared directly. Experiment 2 gave an indirect comparison of characteristics of the two types of memories through the intro-

This research has received support from Grant BNS-7813054 from the National Science Foundation. Part of these results were initially presented at the annual meeting of the Eastern Psychological Association, March 1978. Requests for reprints should be sent to Carol L. Raye, Department of Psychology, Barnard College, 606 West 120th, New York, New York 10027.

duction of another independent variable intended to influence the criteria subjects used in making frequency estimates.

EXPERIMENT 1

Method

In the first phase of Experiment 1, words were presented two, five, or eight times and generated two, five, or eight times. Presentation and generation frequency were factorially combined within subjects. There were four items at each combination of presentation and generation frequency, for a total of 36 different items in all. On a presentation trial, subjects saw a cue and a to-be-remembered word (e.g., Color-BLUE). Subjects were told to try to learn the word that went with each cue so they could recall the word when the cue was given alone, and they were told some items would occur more often than others. On a generation trial, the cue (e.g., Color) was presented and the subject had to think of and write down the appropriate word (BLUE). Cues and words were initially selected from the Battig and Montague (1969) norms. Category names served as cues, and the responses were from the most popular instances of each category. Items were then tested with pilot subjects to assure that the cues would reliably produce the correct response. Consequently, in the present studies, as in Johnson et al. (1977), failures to generate the desired items were extremely rare. Blocks of 18 presentation trials alternated with blocks of 18 generation trials on successive pages of a mimeographed booklet, for 10 blocks of each type. No item occurred more than once in a block. Across the experiment, items were assigned randomly to presentation and generation positions with the restrictions that no item be generated before it was presented and that, across items, average position and lag between occurrences were approximately the same for presentation and generation conditions.

Subjects were given 3 sec on presentations to read each presented item and to make a check mark indicating that they had read it; they were given 3 sec on generations to read each cue and write down the appropriate to-be-remembered word. Rate was controlled by having the subject move a mask down the page in time with an auditory signal.

In Phase 2, which followed immediately, a random half of the subjects were asked to judge how many times each to-be-remembered word had been presented in the booklet. The other half of the subjects were asked to judge how many times they had generated each item. This was the first mention of the frequency judgment task. Instructions and materials for the judgment task were also presented in booklets.

Twenty female Barnard College students participated as subjects. All subjects were tested simultaneously.

Results

Since Experiment 1 was prompted by the results obtained in a previous experiment (Johnson, et al., 1977, Experiment 1), the relevant data from that experiment will be examined first.⁴ These statistical comparisons were not reported previously because they were not the major focus of the previous paper, and the conclusions that could be drawn from them were limited, as will be discussed below. The basic design of the previous study was like that of the present Experiment 1, and the results are shown in Figure 1. The left panel of Figure 1 shows judgments of presentation frequency. The slopes of the lines indicate that judged frequency increased as actual presentation frequency increased—as many others have found. The right panel of Figure 1 shows the corre-

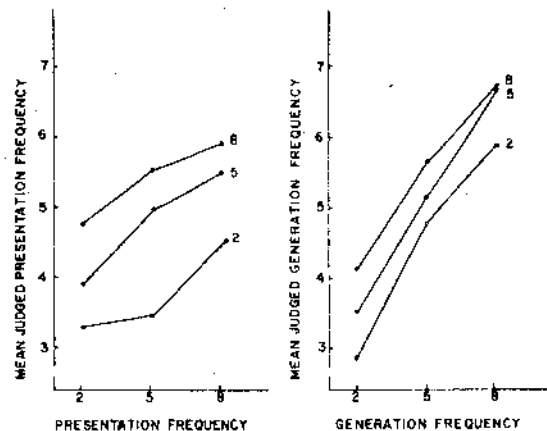


Figure 1. The left panel shows mean judged presentation frequency as a function of actual presentation frequency. Each line represents a different generation frequency. The right panel shows mean judged generation frequency as a function of actual generation frequency. Each line represents a different presentation frequency. Data are from Johnson, Taylor, and Raye (1977, Experiment 1).

sponding results for judgments of generation frequency. Here, the slopes of the lines represent subjects' sensitivity to the frequency with which they have generated an item. If you compare the slopes of the generation judgments (right panel) with the slopes of the presentation judgments (left panel), you will notice that the slopes are greater for judged generations. An analysis of variance was conducted in which one factor was type of judgment (presentation or generation) and another was frequency of the event that was appropriate to the judgment task. There was an interaction between type of judgment and frequency of the appropriate event [$F(2,76) = 19.71$, $MSe = 1.07$], indicating that increasing generation frequency increased generation judgments more than increasing presentation frequency increased presentation judgments. That is, subjects appeared more sensitive to the relative frequency of generations than to the relative frequency of presentations.

In the left panel of Figure 1, the presentation judgments are shown separately for items according to the frequency with which they were generated (the frequency of the inappropriate event). The separations between these lines show that judgments of presentation frequency were influenced by the number of times an item was generated. This shows the confusion of externally and internally generated information, which was the primary focus of the earlier study (Johnson et al., 1977, Experiment 1). Next, compare this confusion effect for presentation judgments with the corresponding effect for generation judgments, represented by the separation between the lines in the right panel of Figure 1. It appears that presentations had less influence on estimates of the number of generations of items than generations had on judgments about presentations. This was confirmed by an interaction between type of judgment

and frequency of the inappropriate event (generation frequency for presentation judgments and presentation frequency for generation judgments) [$F(2,76) = 3.41$, $MSe = 1.04$].

In general, then, judgments about the frequency of internal events or generations seemed more sensitive to relative frequency and more stable, in that they were less influenced by the opposite event. However, in the Johnson et al. (1977) study, more time was allotted to generations than to presentations in Phase 1, and this difference might account for the apparent superiority of the generation judgments. In addition, in this previous study, subjects were forced to respond to each item on generation but not on presentation trials, and we were concerned that greater control might have been exerted over generation than over presentation frequencies. Thus, the current study was an attempt to see if the same pattern of results would emerge when presentation and generation rates were equal and when subjects were required to respond to each item on presentation trials with a check to indicate they had read each item.

The results of the present experiment are shown in Figure 2. Presentation judgments are shown in the left panel and generation judgments in the right panel. Judgments were increased by increases in the number of both the appropriate [$F(2,36) = 59.84$, $MSe = 1.43$] and the inappropriate [$F(2,36) = 76.35$, $MSe = .51$] event. These results replicate those reported by Johnson et al. (1977), indicating that people are sensitive to the relative frequency of internally generated as well as externally generated events and demonstrating confusion between the two sources of information.

More important for the present purposes, the two interactions reported above were also replicated. Subjects were more sensitive to the relative frequency of

internally generated events than they were to the relative frequency of externally generated events [$F(2,36) = 4.41$]. Also, presentation judgments were affected more by the number of generations than generation judgments were by the number of presentations [$F(2,36) = 3.45$]. Therefore, with presentation and generation rates equated and attention given to each item, judgments about generations again appeared in general to be more sensitive and more stable.

We conducted an additional analysis, in which the factors were judgment condition, frequency, and appropriateness of events for the judgment required. This allows us to compare the impact of increases in frequency of appropriate events with increases in frequency of inappropriate events on the same judgment. There was both an Appropriateness by Frequency interaction [$F(2,36) = 7.16$, $MSe = .22$] and an Appropriateness by Frequency by Judgment interaction [$F(2,36) = 9.39$]. These analyses reflect the fact that generation judgments were affected more by generation frequency than by presentation frequency and that both types of events had approximately equal impact on presentation judgments. This difference in the relative contribution of the two types of events in the two judgment conditions supports the assumption that subjects were indeed trying to estimate the frequency of two different types of events in the two judgment conditions. The pattern here is similar to that found by Johnson et al. (1977).

In both Johnson et al. (1977) and Experiment 1 here (and, in fact, in Experiment 2 to be described below), there was no overall effect of the type of judgment. One might predict, for example, that the frequency estimates given for generated memories would on the average be greater than the frequency estimates given for externally derived memories, based on prior findings of superior retention for internally generated memories (Slamecka & Graf, 1978; Johnson & Raye, Note 1). We have previously discussed (Johnson, Raye, Wang, & Taylor, 1979) how subjects might reasonably be expected to initially determine an "anchor" for their judgments based on such information as estimates of the number of different items or concepts experienced, the rate of occurrence of events (regardless of whether they were externally or internally generated), and the total time in the situation. If subjects in the two judgment conditions use a similar average anchor point, the overall magnitude of their judgments might not differ. For these reasons, the slope of the judgment function seems a more responsive measure of potential differences in characteristics of memories that might be expected to influence how memories enter into frequency estimates.

EXPERIMENT 2

The second experiment involved the same design and general procedure as the first. That is, words were presented two, five, or eight times and generated two, five,

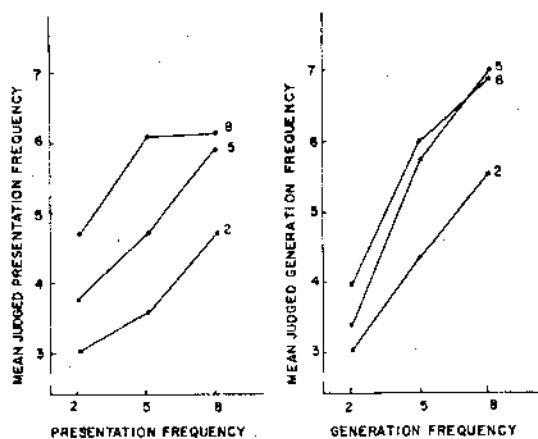


Figure 2. The left panel shows mean judged presentation frequency as a function of actual presentation frequency. Each line represents a different generation frequency. The right panel shows mean judged generation frequency as a function of actual generation frequency. Each line represents a different presentation frequency (Experiment 1).

or eight times, and later, half of the subjects were asked to judge the number of times each item was presented by the experimenter and half of the subjects were asked to judge the number of times they had generated each item. In addition, however, half of the people judging each type of event were given a limit for their estimates: They were asked to use numbers between 0 and 10. The other half of the subjects were not given any limit or range for their estimates. We expected that this manipulation would affect the criteria subjects used for including memories in an "event count." Of special interest was whether the two types of judgments would be affected similarly. If they were, it would emphasize the similarity between externally and internally derived memories. However, if the two were affected differently, this would point to differences in the memories or judgment processes in the two cases. For example, if memories for one type of information are more stable and faithful, they might be less susceptible to variations in judgment criteria. In addition, most of our previous studies were run with the limit in order to provide a relatively stringent estimate of confusion between externally and internally generated memories, assuming that subjects would be less likely to include "marginally appropriate" events (that is, events from the inappropriate class) when estimating with a limit. Experiment 2 provided a test of this assumption that is, itself, based on the assumption that the two classes of memories differ and that this difference provides a basis for distinguishing them. If so, there should be some evidence that making the subjects more cautious results in a selective "gating out" of inappropriate memories, rather than simply an overall suppression in frequency estimates.

Method

The subjects were male and female undergraduate volunteers at the State University of New York at Stony Brook. The 32 subjects were assigned randomly to conditions and tested in groups of from one to four people; testing sessions were assigned to conditions in randomized blocks of the experimental treatments. Frequency judgment instructions were given just prior to the frequency judgment task and did or did not include a limit on the range of numbers to be used, depending on the subject's condition. As in Experiment 1, prior to these instructions, no mention of the frequency judgment task had been made.

Results

An analysis of variance was conducted that included the between-subjects factors of type of judgment (presentations or generations) and type of instructions (with the limit or without the limit) and the within-subjects factors of number of events appropriate to the judgment task (two, five, or eight) and number of events inappropriate to the judgment task (two, five, or eight).

The contribution of appropriate events can be seen in Figure 3, in which presentation judgments are in the left panel and generation judgments are in the right panel. As in previous studies, frequency estimates increased with increases in the actual frequency of the appropriate

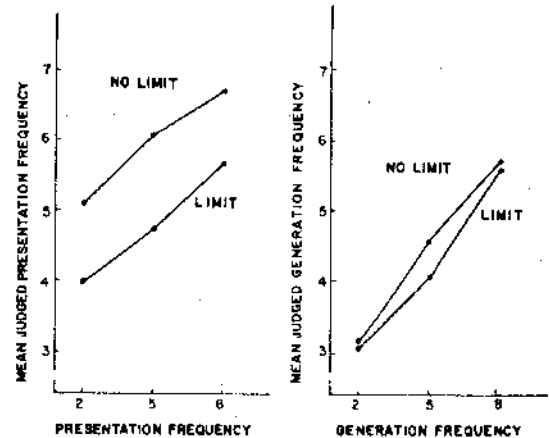


Figure 3. The left panel shows mean judged presentation frequency as a function of actual presentation frequency. The two lines represent different test conditions. The data are collapsed across generation frequency. The right panel shows mean judged generation frequency as a function of actual generation frequency. The two lines represent different test conditions. The data are collapsed across presentation frequency (Experiment 2).

events [$F(2,56) = 99.39$, $MSe = 1.08$]. In both cases, the data have been collapsed across the frequency of the inappropriate event. By comparing the slopes of the lines in the left panel with those in the right panel, you can see the interaction of type of judgment with the frequency of the appropriate event [$F(2,56) = 4.18$]. Again, the slopes of the lines relating judged generation frequency to actual generation frequency are steeper than the slopes of the lines relating judged presentation frequency to actual presentation frequency. In addition, while judgments about the number of presentations appeared to be higher without the limit than with the limit, the instructional manipulation had no effect on judgments about generations.

Figure 4 gives a better indication of how the limit influenced presentation judgments. This is a figure showing confusion between externally and internally generated information. Presentation judgments are on the left, and generation judgments are on the right. Each is represented as a function of the inappropriate event. (The data have been collapsed across the frequency of the appropriate event.) The functions in each panel would be flat if frequency judgments were independent of the frequency of the inappropriate event, but as is apparent, inappropriate events contributed to frequency estimates [$F(2,56) = 32.52$, $MSe = 1.48$]. Once more, presentations had less effect on judgments about generations than vice versa [$F(2,56) = 5.69$]; there was less confusion when subjects were judging generations. More importantly, there was a triple-order interaction of instruction, judgment, and inappropriate frequency [$F(2,56) = 4.48$]. Without the limit, there was an even greater effect of generations on judged presentations. Given the limit, there was less confusion in making

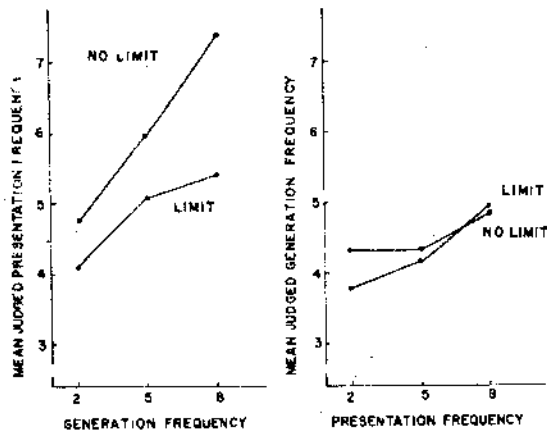


Figure 4. This figure shows judgments as a function of the inappropriate event. The left panel shows mean judged presentation frequency as a function of generation frequency. Each line represents a different test condition. The data are collapsed across presentation frequency. The right panel shows mean judged generation frequency as a function of presentation frequency, collapsed across generation frequency (Experiment 2).

presentation judgments. In contrast, the limit/no-limit instruction had no effect on generation judgments.

A final analysis included the factors of judgment, limit instruction, frequency, and appropriateness to the designated judgment, for the purpose of comparing the contribution of appropriate and inappropriate events to the same judgments. There was an appropriateness by frequency interaction [$F(2,56) = 5.55$, $MSe = .36$], an interaction among appropriateness, frequency, and judgment [$F(2,56) = 11.06$], and an interaction of these three factors with the limit instruction [$F(2,56) = 3.61$]. While this sounds complex, the pattern accounting for these interactions can be described easily: Appropriate events affected judgments more than inappropriate events when subjects were judging either presentations or generations under the limit instructions. Appropriate events also influenced judgments more when subjects were judging generation frequency without the limit. In contrast to these three cases, inappropriate events had the greater impact when subjects were judging presentation frequency without the limit. This analysis serves to highlight the pattern discussed previously: The number of times subjects generated items had a large effect on their judgments of presentation frequency, actually contributing more than presentations when subjects made judgments without the limit. However, the impact of the inappropriate traces was greatly reduced by introducing a manipulation that presumably made the subjects more cautious in their judgments.

DISCUSSION

To summarize the results of these two experiments: (1) People seemed to be more sensitive to the relative frequency of internally generated events compared with

the relative frequency of externally generated events; (2) the number of times an item was presented had less effect on estimates of generation frequency than the reverse confusion effect; and (3) judgments of generation frequency were less affected by the test conditions. These results also are consistent with Slamecka and Graf's (1978) findings of superior recall and recognition of internally generated memories compared with comparable externally generated memories and, along with their results, indicate that there is indeed something special about the memories produced from internally generated information. Therefore, two questions clearly seem worth pursuing: (1) Are memories for internally generated information always "superior," or are there circumstances in which externally derived information produces better memory? (2) What are the differences in the nature of the memory representations produced as a consequence of perceptual vs. internally generated activity? Which of these differences might account for the apparent superiority of self-generated information in the current context?

It is not so surprising that generating should augment recall, since recall obviously includes a generation process, and in generating, a subject may be practicing part of the recall processes. Why generated information should have any advantage for recognition or frequency judgments is not so obvious. One possibility is that recognition and frequency judgments in these tasks are based on these same recall processes (e.g., see Mandler, 1972; Tversky & Kahneman, 1973). Alternatively, the advantage for subject-generated information in tasks such as judging relative frequency may not depend on prior recall practice, but rather, on differences in actual properties of the memories generated from external and internal sources. For example, Raye, Johnson, and Wang (Note 2) have found that small variations in features of external events improve judgments of the relative frequency of external events. This suggests that if the memories of internally generated representations of the same event were, on the average, more variable than the memories of externally derived representations, the former should be more unique and/or discriminable and thus perhaps easier to "count" in a frequency judgment task. This would predict that there should be some circumstances, namely, those in which representations of internal events are less variable than recurrences of external events, in which judgments of relative frequency should be better for the externally derived events.

While recurrences of which type of event are more variable probably depends on the nature of the material, it seems reasonable to suppose that there are probably some more general and pervasive differences in the characteristics of internally and externally generated memories. One classic notion is that internally generated information is typically weaker or less vivid than information derived from perceptual experience. However, as already discussed, available evidence suggests that under

some circumstances, and by some measures, internally generated memories are stronger, not weaker, than externally generated memories (e.g., Bobrow & Bower, 1969; Doshier & Russo, 1976; Slamecka & Graf, 1978). Therefore, it seems more promising to pursue the possibility that internally and externally generated memories differ in some way or ways that cannot easily be encompassed under the single notion of strength. Johnson and Raye (Note 1, Note 3) have described a working model of the process of discriminating memories of past perceptions from memories of past imaginations (or "reality monitoring"; see Johnson, 1977) that proposes that, as general classes, externally generated and internally generated memories differ in several ways (e.g., traces derived from external sources typically should have more sensory information, whereas internally generated traces are proposed to include more information regarding cognitive operations that took place at the time the trace was established). Furthermore, the specific dimensions and criteria used to identify the origin of memories may depend on the particular events involved.

The current finding of more confusion from internal events when judging external event frequency than from external events when judging internal event frequency could be produced by a tendency to include a wider range of traces in external event counts than in internal event counts. This could result simply from a less stringent criterion along a single dimension in the former case. Alternatively and more probably, the specific criteria used to select traces may actually depend on which event is being estimated. Typically, when judging the frequency of external events, selection may be based largely on certain sensory features, whereas when judging the frequency of internal events, selection may be based largely on the availability of information about cognitive operations. However, in the present situation, the external events were words. Words, especially when read, are probably rather "impoverished" sensory events compared with other external events. Thus in the current case, sensory information may not have been a very good discriminator between externally and internally generated representations (especially if, as some investigators have proposed, there is a phonological recoding component occurring on presentation trials, e.g., Spoehr & Smith, 1975, which also would be expected to occur on generation trials). In addition, visual input occurred on both presentation and generation trials, since subjects wrote the responses on generation trials. Also, the fact that a memory representation contained some information regarding cognitive operations would not necessarily be used as a reason for rejecting it as part of an external occurrence, since even given the current task, the memory of a word might be accompanied by the memory of a thought that occurred during its presentation. On the other hand, information about a cognitive operation (e.g., the recall process) would be a

critical determinant in judging internal occurrences. Thus, the subject would have been exercising a stricter criterion for internal than for external occurrences, producing the result of less confusion during judgments of internally generated occurrences than during judgments of presentations.

The idea that traces originating in external and internal sources differ is further supported by the results of Experiment 2. In Experiment 2, the limit did not affect the magnitude of confusion when people were judging the frequency of internally generated events. But when they were judging external event frequency, imposing the limit reduced apparent confusion. This suggests that the limit produced a more stringent criterion for including a memory trace in a frequency estimate, but only for judgments of external frequency. That internal judgments were not similarly affected is consistent with the notion that different criteria can be used in the two judgment conditions, implying that the two classes of memories are distinguishable. Furthermore, when the limit was imposed, the more stringent criterion for judging external frequency fell selectively on internally generated traces and did not affect the contribution of externally generated traces to judgments of presentation frequency. That is, the limit reduced the slope of the function relating generation frequency to judged presentation frequency, but it did not affect the slope of the function relating presentation frequency to judged presentation frequency. This fact indicated that these two types of traces differ in some way that allows the subject to discriminate between them.

One explanation, then, of the outcome of the limit manipulation (which would be consistent with the relative magnitudes of the confusion effects discussed above) is that the more stringent criterion produced by imposing a limit on judgments induced subjects judging presentation frequency to reject memories containing information about cognitive operations that would have been admitted otherwise, thus excluding more of the inappropriate traces. The limit would not be expected to have the same effect on judgments of internally generated events, since information about cognitive operations is, in either test condition, the major criterion used to identify internal traces.

Along with previously reported results (Johnson et al., 1979), the present data argue strongly that all occurrences of a concept do not simply increment the same concept counter, but that specific information is preserved (Hintzman, Block, & Summers, 1973). Otherwise, subjects would not show differential sensitivity to relative frequency of events from internal and external sources, and they would not be able selectively to exclude traces of one type. We have suggested that subjects are particularly sensitive to relative frequency of self-generated, as compared with other, events, and that the presence of cognitive operation information regarding generating is preserved in memory traces and

can be used as a discriminative cue. However, "generation" is obviously a much more complex concept than we have acknowledged in the above discussion. Generation can involve primarily central acts (e.g., thought), more peripheral motor components such as speech or writing with their consequent kinesthetic feedback, and/or feedback into the distance senses (hearing, vision) from the products of speech, writing, or other motor acts.

All of these can be associated with external as well as internal events, but our model assumes that some are more characteristic of events originating externally (e.g., sensory information) and some are more characteristic of self-generated events (e.g., cognitive operations). The procedure used in the present experiments does not permit an assessment of the relative effects of central, motor, and sensory components of self-generated information. Because subjects wrote their generated items, the internally generated traces included, in addition to thought, motor and sensory information produced by the subjects, whereas the externally generated traces did not contain the same subject-produced motor and sensory information. However, this difference between external and internal events cannot alone account for the entire pattern of results obtained in the present experiments. We previously reported that overt, as compared with covert, production of items on generation trials increased estimated presentation frequency (Johnson et al., 1977, Experiment 2). This finding, we suggested, indicated that some aspects of overt generations made those traces seem more like presentations; a likely candidate is the larger amount of sensory information from overt compared with covert generations, along with the special association of sensory information with external events. The large impact of generations on presentation judgments in the present experiments might be partially accounted for by this subject-produced sensory information. However, if sensory information were the only characteristic of memories taken into account, one would expect the magnitude of confusion to be the same when subjects judged generation frequency as when they judged presentation frequency, and clearly, it was smaller in the former case. Also this would not explain, at the same time, the greater discriminability of relative frequency among internally generated traces. As discussed above, the present results are more easily accounted for by postulating that one class of information predominated when external judgments are made and another class predominated when internal judgments are made. Furthermore, Experiment 2 suggests that what information is used is somewhat under the subject's control.

It is possible that the information that usually dominates judgments about internally generated events is not limited to memories of cognitive acts of generation, as we imply above, but also involves the memory record of the motor components of the overt responses. We have two other findings that bear on this issue. One experi-

ment was very similar to Experiment 1, except that subjects pronounced items aloud on both presentation trials and generation trials (Taylor, Johnson, Birnbaum, & Raye, Note 4). Thus, both subject-generated and experimenter-generated events should have included motor and consequent auditory components. Again, subjects were significantly more sensitive to generation frequency than to presentation frequency, and confusion from internal events was greater than that produced by external events. In another experiment, free recall was obtained for events that had occurred a single time, either internally or externally generated (Johnson & Raye, Note 1). As Slamecka and Graf (1978) found, subject-generated items had the advantage over experimenter-generated items. We found, in addition, that the advantage of subject-generated items was of the same magnitude whether externally and internally generated items had both been written at acquisition or had simply been heard and thought, respectively. These two findings argue that internally generated information has an advantage over externally generated information that is not dependent on overt production. In summary, it appears to us that the major potential cue about the source of memories for generated verbal materials originates in the thought and not the deed; however, the relative importance of these aspects of self-generated events in other situations may well be different.

REFERENCE NOTES

1. Johnson, M. K., & Raye, C. L. *Reality monitoring*. Manuscript submitted for publication, 1980.
2. Raye, C. L., Johnson, M. K., & Wang, A. Y. *Event variability and judgments of relative frequency*. Manuscript in preparation.
3. Johnson, M. K., & Raye, C. L. *A working model of reality monitoring*. Paper presented at the annual meeting of the Psychonomic Society, San Antonio, Texas, 1978.
4. Taylor, T. H., Johnson, M. K., Birnbaum, I. M., & Raye, C. L. *Alcohol and reality monitoring*. Manuscript submitted for publication, 1980.

REFERENCES

- BARTLETT, F. C. *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press, 1932.
- BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 1969, 80(3, Pt. 2).
- BOBBROW, S. A., & BOWER, G. H. Comprehension and recall of sentences. *Journal of Experimental Psychology*, 1969, 80(3, Pt. 1), 455-461.
- BOWER, G. H. A selective review of organizational factors in memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- BRANSFORD, J. D., & JOHNSON, M. K. Considerations of some problems of comprehension. In W. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- COFER, C. N. Constructive processes in memory. *American Scientist*, 1973, 61, 537-543.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing:

- A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- DOSHER, B. A., & RUSSO, J. E. Memory for internally generated stimuli. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 633-640.
- HINTZMAN, D. L., BLOCK, R. A., & SUMMERS, J. J. Modality tags and memory for repetitions: Locus of the spacing effect. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 229-238.
- JOHNSON, M. K. What is being counted none the less? In I. M. Birnbaum & E. S. Parker (Eds.), *Alcohol and human memory*. Hillsdale, N.J.: Erlbaum, 1977.
- JOHNSON, M. K., BRANSFORD, J. D., & SOLOMON, S. K. Memory for tacit implications of sentences. *Journal of Experimental Psychology*, 1973, 98, 203-205.
- JOHNSON, M. K., RAYE, C. L., WANG, A. Y., & TAYLOR, T. H. Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 229-240.
- JOHNSON, M. K., TAYLOR, T. H., & RAYE, C. L. Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition*, 1977, 5, 116-122.
- KINTSCH, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- KLEE, H., & GARDINER, J. M. Memory for remembered events: Contrasting recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1976, 15, 471-478.
- MANDLER, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press, 1967.
- MANDLER, G. Organization and recognition. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- MARTIN, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 1968, 75, 421-441.
- PAIVIO, A. Mental imagery in associative learning and memory. *Psychological Review*, 1969, 76, 241-263.
- POSTMAN, L. Learned principles of organization in memory. *Psychological Monographs*, 1954, 68(374).
- SLAMECKA, N. J., & GRAF, P. The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 592-604.
- SPOEHR, K. T., & SMITH, E. E. The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, 1, 21-34.
- SULIN, R. A., & DOOLING, D. J. Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology*, 1974, 103, 255-262.
- TULVING, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- TULVING, E., & THOMSON, D. M. Retrieval processes in recognition memory: Effect of associative context. *Journal of Experimental Psychology*, 1971, 87, 116-124.
- TVERSKY, A., & KAHNEMAN, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.
- UNDERWOOD, B. J. False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, 1965, 70, 122-129.

NOTE

1. Reported F values are significant at the .05 level or better.

(Received for publication August 8, 1978;
revision accepted November 26, 1979.)