

The accuracy of beliefs about retrieval cues

PAULA T. HERTEL, LINDA J. ANOOSHIAN, and PATRICIA ASHBROOK
Trinity University, San Antonio, Texas

We investigated the accuracy of predictions about semantic, environmental, and phonological cues for remembering. Subjects rated the pleasantness of 10 words in each of four rooms, predicted the number of words that they would recall with and without one of the three types of cues, and then were tested for free or cued recall. Consistent with their predictions, subjects who received semantic cues recalled more words than did subjects in the free-recall group. The subjects in the other cuing conditions did not benefit from the cues; furthermore, they overestimated the value of phonological cues, and they believed that environmental cues were ineffective. Finally, confidence ratings for cued-recall predictions did not reflect the pattern of cued-recall performance. Subjects were least confident about their predictions for semantic cuing and most confident about their predictions for recall to be cued phonologically.

Every day, rememberers encounter a variety of situations in which they are successful at remembering and a similarly diverse set of circumstances in which they fail to remember. How well do they understand the factors that underlie their successes and failures? The present research examined subjects' beliefs about the effectiveness of particular types of retrieval cues in the context of actual cued-recall performance. In contemplating this research, we were struck with the extent to which anecdotal evidence about certain kinds of retrieval cues seemed to be discrepant with the evidence from memory research. Although memory research has tended to stress the diversity of cues that could facilitate retrieval, anecdotal evidence seems to suggest an over-reliance on particular types of cues and ignorance of others. For example, in the now popular game of Trivial Pursuit, partners often attempt to provide each other with phonological cues, such as "It sounds like..." or "I think the name starts with..." Similarly, students (and pursuers of trivia) seem to acknowledge the effectiveness of various kinds of semantic cues, such as category names (or associated trivia). In contrast, students in our psychology classes often find it incredible that environmental cues can be effective—that performance on an exam, for example, might be negatively affected by taking the exam in a different classroom.

In many of our everyday cognitive activities, we naturally invoke semantic memory in our attempts to make sense of the information to which we are exposed. In this study, therefore, subjects performed a judgment task to incidentally learn the words they would later attempt to remember. Subsequently, we asked for predictions about how many words they could remember, first without help

and then with the help of one of three types of retrieval cues. These included semantic cues as well as two other types which were, as in many of our everyday memory attempts, less prominent during the original encoding: phonological cues (see Bransford, Franks, Morris, & Stein, 1979; Nelson, Wheeler, Borden, & Brooks, 1974) and cues regarding the environmental context during the original encoding (see Smith, 1979, 1982; Smith, Glensberg, & Bjork, 1978). The learning task was designed so that an environmental reinstatement effect was likely (see Smith, 1982): subjects learned the words in four different rooms.

Our aim was to investigate subjects' beliefs about this variety of retrieval cues and the accuracy of those beliefs. Would subjects think that, in addition to semantic cues, phonological and environmental cues could affect their cued-recall performance?

METHOD

Subjects

A total of 64 students participated as subjects in exchange for extra-credit points in their lower-division psychology courses. They were randomly assigned to four groups ($n = 16$ each) that differed according to the recall test: environmental, semantic, phonological, or free recall.

Materials

We selected 40 low-frequency words, each having both a unique phonological cue and a unique semantic cue. When compared to other words on the list, each word was unique in terms of the sound of the first two letters and was the only instance of a more general concept or category name. The sound of the first two letters served as the phonological cue, and the category name served as the semantic cue. Target words, followed by their corresponding semantic cues, included: ginger, seasoning; radiator, car part; lettuce, vegetable; tortoise, reptile; barracuda, fish; wrench, tool; zoology, science; and daffodil, flower.

The 40 words were randomly ordered and assigned to one of four rooms for the incidental learning task (10 words per room). In each of these rooms, seven large envelopes were placed on an otherwise blank wall to form a 7-point Likert scale; the ends were labeled "pleasant" and "unpleasant." Each word was printed on a

We thank Penny Latham and Allison Waldron for assistance in testing subjects and Tom Nelson for comments on an earlier manuscript.

Requests for reprints should be sent to P. T. Hertel, Department of Psychology, Trinity University, 715 Stadium Drive, San Antonio, TX 78284.

card, so that the rating task could be carried out by placing the word into the appropriate envelope.

Materials for the phonological and semantic cued-recall tasks were presented on tape. Each set of 40 cues was randomly arranged and then split into four equal blocks. The cues were then tape-recorded in a manner that corresponded to the timing of subjects' recall in the environmental group. Each block of 10 cues was presented over 3 min with a 45-sec interval between blocks. An additional version of each tape was prepared in which the 40 words were presented in reversed order.

Procedure

Each subject was individually tested by a single experimenter. To prepare the subject for pleasantness ratings in different rooms, the experimenter told the subject that several graduate students, completing a common assignment for a research methodology class, had set up rating scales in different areas.

The subject and experimenter then went to each of the four rooms for the pleasantness-rating task. The rooms had been selected for their distinctiveness (laboratory rooms set up for different kinds of studies), with the goal of ensuring that environmental context would be effective in aiding retrieval. Subjects followed an unfamiliar route, walking through an interior service corridor to get from one room to another.

As they walked the designated route, the experimenter made sure that the walking time between rooms was approximately 45 sec for all subjects. Two orders for traveling from one room to another (Rooms 1 to 4 or Rooms 4 to 1) were counterbalanced with the four conditions for recall. Upon arriving at each room, the subject rated the pleasantness of 10 words. Then the experimenter allowed 90 sec for travel to a fifth "neutral" room that was clearly different from any of the four exposure rooms. In this room, all subjects were asked to provide two predictions about recall. First, regardless of their assigned condition for actual recall, they predicted the number of words they could (freely) recall if asked to recall them at their present location. Second, they indicated the number of words they thought they could recall if they could do one of the following: (1) follow the same route and recall the words inside the previously visited rooms (if they were assigned to the environmental condition for recall), (2) hear the sound of the first two letters for each of the words (if they were assigned to the phonological condition), or (3) hear a more general word that was meaningfully related to each of the target words (in the semantic condition). Examples of semantic or phonological cues were provided, according to the assigned condition. In the semantic condition, the word *city* was provided as a more general word to cue memory for *San Antonio*. In the phonological condition, the sound *wə* was provided as an example to cue memory for *workshop*.

So that all subjects had comparable experiences, the experimenter made sure that those who were assigned to the free-recall condition also made cued-recall predictions, following their free-recall predictions. For this purpose, they were randomly assigned one of the three types of retrieval cues. Notice, however, that these subjects actually recalled without cues, even though they had predicted recall under one of the cuing conditions.

Following each prediction for free and cued recall, subjects rated their confidence in its accuracy using a 7-point Likert scale. In addition, they rated their confidence in the difference between their predictions for free versus cued recall.

Finally, subjects were asked to recall the 40 words in the recall condition to which they had been assigned. Two orders of cues were used per condition, with an equal number of subjects receiving each order. In the phonological and semantic groups, subjects heard the cues in one of the two orders described above. Similarly, for the environmental group, subjects revisited the four rooms in one of two directions (Rooms 1 to 4 or Rooms 4 to 1).

For subjects in the environmental group, the experimenter allowed 90 sec for travel to the first room; she escorted the subjects in the remaining groups out of the neutral room for a 90-sec break prior to recall in that same room. Travel times for the environmental

group, as well as the timing in the recorded tapes for the phonological and semantic groups, allowed for 3 min for each of four blocks of 10 words separated by 45 sec between blocks. Hence, for subjects in the free-recall group, the experimenter called for 45-sec breaks between four separate 3-min intervals for free recall. During the 45-sec breaks for the semantic, phonological, and free-recall groups, the subjects moved to different areas within the same room for the four different recall attempts. (We told them that these procedures were necessary to ensure that everyone had the same time and opportunities for recall.)

RESULTS AND DISCUSSION

Our approach to data analyses was first to examine the effectiveness of the various types of retrieval cues (i.e., whether subjects in a particular cued-recall group outperformed those assigned to free recall). With this information in hand, we then examined subjects' predictions of the number of words recalled with and without cues, in each of the three cued-recall groups. Again, our purpose was to determine whether the predictions differed according to the type of cue and to compare this pattern to the pattern of recall data. The significance level for all analyses was set at .05.

Recall

The mean number of words correctly recalled by each recall group can be seen in the first row of Table 1. A one-way analysis of variance revealed that recall varied according to the type of recall test [$F(3,60) = 10.21$, $MSe = 27.90$]. The subjects who received semantic cues recalled more words than did subjects in the other three groups. Cued recall was comparable for the phonological and environmental groups; however, both types of cues were ineffective relative to free recall.

The lack of a cuing effect for phonological and environmental groups must be evaluated in the context of the conditions for free recall. Although we attempted to equate the timing and conditions for recall across the four groups, subjects in the free-recall group had a clear advantage. That is, although we mandated breaks between 3-min recall attempts, subjects had completed most of their free recall well before those in the cued-recall groups had received many of their cues. Therefore, the functional retention interval was somewhat longer for cued recall.

Predictions

The mean number of words predicted by each recall group is also reported in Table 1. Row 2 contains the mean number of words predicted without the aid of cues,

Table 1
Mean Number of Words Recalled and Predicted

Measure	Recall Test*			
	Semantic	Phonological	Environmental	Free
Recall	19.19	12.06	12.94	11.69
Predictions for:				
Free Recall	13.94	13.00	12.75	12.75
Cued Recall	19.81	18.63	13.81	†

* $n = 16$ for each recall group. †This mean is not meaningful because different subjects considered different types of cues.

and Row 3 reports the mean number predicted if one of the three types of cues was provided. A comparison of these rows of means reveals that subjects predicted a cued-recall advantage if provided with either phonological or semantic cues.

The predictions by the three cued-recall groups were submitted to a mixed-design analysis of variance with a between-subjects factor for the cue type (environmental, semantic, or phonological) and a within-subjects factor for class of prediction (free vs. cued recall). Only the data from the three cued-recall groups were included in order to maintain the same groups for analyses of actual and predicted cued recall. However, analyses that included the recall predictions of the free-recall group yielded comparable results.¹

A significant main effect for class of prediction [$F(1,45) = 32.92$, $MSe = 12.78$] was interpreted in the context of a significant interaction [$F(2,45) = 4.59$]. The subjects in the semantic or phonological groups predicted an advantage for those cues (relative to free recall), whereas subjects in the environmental group did not predict a similar advantage. This outcome was supported by a reliable interaction contrast (class of prediction \times environmental vs. the other two types of cues) that accounted for 99% of the variance in the overall interaction [$F(1,45) = 9.17$]. The cuing advantage predicted by the subjects in the semantic group did not reliably differ from that predicted in the phonological group.

As discussed earlier, the functional retention interval for the free-recall group was shorter on the average than the retention interval for the other groups. Hence, evidence that recall cued by phonological cues was no better than free recall is not, by itself, convincing evidence that subjects overestimated the number of words that could be recalled with phonological cues. However, such overestimation is further indicated by evidence that subjects made comparable cued-recall predictions for phonological and semantic cues, but semantic cues were much more effective in aiding recall. Furthermore, despite evidence that cued-recall was quite comparable given phonological or environmental cues, subjects predicted much greater recall when considering phonological cues. In fact, in the environmental group, subjects failed to give higher predictions for cued than free recall. They clearly believed that environmental cues were irrelevant and ineffective as retrieval cues.

There is, of course, the possibility that the levels of recall performance could be attributed to subjects' motivations to make their predictions come true. Did subjects stop recalling words when they had achieved the level of recall that they had just predicted? This seemed unlikely for the phonological group, since cued-recall predictions were much higher than cued-recall performance. It also seemed unlikely for the semantic group, because recall was paced by providing a different cue for each of the 40 words. Presumably, it would be difficult for subjects to keep track of the number of correctly recalled words while hearing additional cues. However, the low recall

by subjects in both the environmental and free-recall groups could reflect attempts to limit recall such that it was consistent with prior predictions. We assessed this possibility by determining the proportion of subjects who recalled more words than they had predicted. These proportions are inconsistent with the operation of such a strategy: environmental, .38; semantic, .44; phonological, .12; free recall, .44.

Confidence in Predictions

Subjects had rated their confidence in free-recall and cued-recall predictions as well as their confidence in the difference between those predictions. Each set of confidence ratings was separately submitted to Kruskal-Wallis tests with a factor for cue type.

Only the analysis of the confidence ratings for cued-recall predictions revealed significant group differences [$\chi^2 = 7.05$, corrected for ties]. Median confidence ratings for predictions about environmental cues ($Mdn = 3.44$) and phonological cues ($Mdn = 3.63$) were higher than ratings for semantic cues ($Mdn = 2.75$) [Mann-Whitney $U = 152$, $Z = 2.00$, for the analysis of semantic vs. environmental cues; $U = 133.5$, $z = 2.46$, for the analysis of semantic vs. phonological cues]. Median ratings for phonological and environmental cues were not significantly different. Subjects were least confident when they predicted the consequences of semantic cuing, the most effective type of cuing for recall. The subjects who predicted recall with phonological cues were moderately confident in what would turn out to be inaccurate predictions.

GENERAL DISCUSSION

Beliefs about retrieval cues were quite accurate when subjects considered cues that were most compatible with the original encoding. The average prediction for semantic cues was quite close to the average cued-recall score.² Yet metamemory typically fell somewhat short with regard to the diversity of cues that can play a significant role in attempts at remembering. Compared to beliefs about semantic cues, beliefs about both environmental and phonological cues were less accurate.

Subjects in our environmental-cue groups viewed environmental cues as irrelevant to remembering. Cued-recall predictions were considerably lower for environmental cues than for the other types, as well as quite comparable to their free-recall predictions. In this experiment, the environmental reinstatement effect was not obtained. These subjects, therefore, could be viewed as having made somewhat accurate predictions as a group. However, in a second experiment, in which multiple settings in the same room were substituted for multiple rooms, the reinstatement effect was obtained, and subjects again predicted that returning to the "rating" room would not be useful in aiding recall. Although further research on the topic of environmental cuing is obviously necessary before much can be concluded about such beliefs, our results do

suggest that the effectiveness of environmental cues might be underestimated. Furthermore, such an outcome can be interpreted within the framework provided by Hasher and Zacks (1979), who have argued that information about spatial location is encoded automatically rather than effortfully. Indeed, other research has confirmed that spatial attributes of events are encoded with considerably less effort than are other types of attributes (Anooshian & Siegel, 1985; Lovelace & Southall, 1983; Park & James, 1983). If the processing of certain attributes involves little effort, awareness, or conscious attention, it seems likely that subjects would be similarly unaware that these attributes could be reinstated to aid retrieval during deliberate attempts to remember (see Kellogg, 1982).

Perhaps the most interesting outcome of this experiment was that in making predictions about phonological cues, subjects overestimated their effectiveness. Following an encoding task that stressed semantic processing, it was not surprising that phonological cues were considerably less effective than were semantic cues. However, at first blush, it was surprising that cued-recall predictions in this context were similar in the two groups. In a similarly interesting comparison, phonological cues were no more effective than were environmental cues, but subjects thought that they would be considerably more effective. How can the inaccurate predictions about phonological cues be understood?

First, in considering our everyday experiences in which sounds cue our memories, it seems to us that phonological cues are helpful when the search set has already been limited to certain domains of semantic or episodic memory. This is certainly true of Trivial Pursuit, for example, when the question has been asked. In other words, people do not commonly declare, "I'm thinking of a word that begins with *qu* that you encountered about 30 minutes ago; what is it?" We did just that for 40 words, probably for the first time in our subjects' experience. Hence, it is perhaps not surprising that these cues failed or that subjects did not predict that failure.

An extension of this reasoning also emphasizes the role of everyday experience with phonological cues when semantic information is at hand. We sometimes experience feelings of knowing when we have retrieved the semantic referent for a word, but not its phonological label (see Krinsky & Nelson, 1985; Nelson, Gerler, & Narens, 1984). Under those conditions, a beginning sound clearly would be very useful. Thus, inaccurate beliefs about phonological cues may reflect that individuals form generalized beliefs about the benefits of such cues, rather than considering the particular circumstances of the original encoding or their retrieval attempts.

Our subjects, therefore, may have made cued-recall predictions based on overgeneralizations from past experiences with situations that are considerably different. In a sense, they were correct in recognizing that phonological cues can be quite effective as memory aids, for example, when the search set has been limited, perhaps by accurate retrieval of the semantic referent, or when

phonological processing is emphasized in the original encoding task (see Bransford et al., 1979). What they failed to recognize was that neither our original encoding task nor our task of recalling on the basis of phonological cues alone was designed such that phonological cues would be maximally effective. This outcome is particularly interesting in light of subjects' abilities to fairly accurately differentiate among semantic and phonological orienting tasks in determining the amount of processing or degree of difficulty they require (see Seamon & Virostek, 1978). The outcome suggests that some characteristics of encoding tasks are more easily judged than are processes involved in retrieval, or that college students' experiences have taught them the following lesson: Learning is a complex state of affairs, but remembering merely requires a jostle provided by the sound of a word or meaningfully related information.

In concluding our discussion, it seems rather striking to us that, although subjects were clearly most accurate in their beliefs about semantic cues, they were most confident of their beliefs about phonological and environmental cues. This misplaced confidence has important implications for how rememberers interpret their everyday successes and failures (see Martin & Jones, 1984; Morris, 1984). As Morris has said, we rarely attend to commonplace memory failures that we can readily explain (e.g., "No one remembers infrequently used phone numbers"). However, if one overestimates the effectiveness of phonological cues, the failure to remember when given a specific phonological cue would not be easily explained and might, therefore, be alarming, especially for those of us who fear memory deficits. Expectations regarding the effects of environmental cues are considerably different. Most people would be surprised and quite pleased with themselves if, upon visiting an old classroom, they remembered the name of their fifth grade teacher. Yet, if one loses familiar environmental cues (upon moving to a new town, for example, or into a home for the elderly) and does not understand why remembering has become more difficult, one's failures might again be cause for alarm (see Anooshian, Ashbrook, & Hertel, 1985; Hertel, 1985). Uneducated metamemory can contribute to deficits (and occasional enhancements) in self-esteem.

REFERENCES

- ANOOSHIAN, L. J., ASHBROOK, P., & HERTEL, P. T. (1985). Self-esteem and beliefs about memory in environmentally stable and relocated students. *Journal of Research in Personality*, *19*, 457-471.
- ANOOSHIAN, L. J., & SIEGEL, A. W. (1985). From cognitive to procedural mapping. In C. J. Brainerd & M. Pressley (Eds.), *Basic processes in memory development: Progress in cognitive development research* (pp. 47-101). New York: Springer-Verlag.
- BRANSFORD, J. D., FRANKS, J. J., MORRIS, C. D., & STEIN, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory*. Hillsdale, NJ: Erlbaum.
- HASHER, L., & ZACKS, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356-388.
- HERTEL, P. T. (1985, May). *What was that? Expectations for interac-*

- tive cueing*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.
- KELLOGG, R. T. (1982). When can we introspect accurately about mental processes? *Memory & Cognition*, **10**, 141-144.
- KRINSKY, R., & NELSON, T. O. (1985). The feeling of knowing for different types of retrieval failure. *Acta Psychologica*, **58**, 141-158.
- LOVELACE, E. A., & SOUTHALL, S. D. (1983). Memory for words in prose and their locations on the page. *Memory & Cognition*, **11**, 429-434.
- MARTIN, M., & JONES, G. V. (1984). Cognitive failures in everyday life. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory actions and absent-mindedness* (pp. 173-190). New York: Academic Press.
- MORRIS, P. E. (1984). The validity of subjective reports on memory. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory actions and absent-mindedness* (pp. 153-172). New York: Academic Press.
- NELSON, D. L., WHEELER, J. W., BORDEN, R. C., & BROOKS, D. H. (1974). Levels of processing and cuing: Sensory vs. meaning features. *Journal of Experimental Psychology*, **103**, 971-977.
- NELSON, T. O., GERLER, D., & NARENS, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General*, **113**, 282-300.
- PARK, D. D., & JAMES, C. Q. (1983). Effect of encoding instructions on children's spatial and color memory: Is there evidence for automaticity? *Child Development*, **54**, 61-68.
- SEAMON, J. G., & VIROSTEK, S. (1978). Memory performance and subject-defined depth of processing. *Memory & Cognition*, **6**, 283-287.
- SMITH, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 460-471.
- SMITH, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory & Cognition*, **10**, 405-412.
- SMITH, S. M., GLENBERG, A. M., & BJORK, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, **6**, 342-353.

NOTES

1. For these analyses, data from each cued-recall group were combined with data from subjects in the free-recall group who had made the same cued-recall predictions. The combined means for free- versus cued-recall predictions were 13.18 versus 19.05 for semantic cues ($n = 22$), 12.38 versus 17.24 for phonological cues ($n = 21$), and 13.76 versus 13.10 for environmental cues ($n = 21$). Both the main effect for class of prediction [$F(1,61) = 19.48$, $MSe = 18.88$] and the interaction [$F(2,61) = 6.96$] were reliable.
2. Also, the correlation between predicted cued recall and actual recall was reliable only for subjects who received semantic cues [$r(16) = .53$].

(Manuscript received August 15, 1985;
revision accepted for publication December 23, 1985.)