

Classification and aural coding in short-term memory¹

DOUGLAS L. HINTZMAN
STANFORD UNIVERSITY

Abstract

Analysis of errors in a short-term memory task indicates that Ss adopted two possible coding strategies: digit vs. letter categorization and subvocal or aural rehearsals. White noise had no effect on types of errors made or on over-all performance, but did bring out the usually covert rehearsal process. Evidence from errors and effects of noise point to a reinterpretation of "auditory" coding in terms of kinesthetic feedback produced by subvocal rehearsals.

Problem

Conrad (1964) has shown that confusions among visually presented letters in a short-term memory task are highly correlated with perceptual confusions that Ss make when the same letters are read aloud with a noise background. In other words, the letters most likely to be confused in short-term memory (STM) are those which sound most alike. This strongly suggests that STM is primarily maintained through a coding process closely associated with the auditory modality. Nevertheless, it seems unlikely that such an aural coding system is the only means of maintaining a temporary memory trace. Various classification schemes should be available when meaningful stimuli are used—for example, if letters and digits are presented alternately, Ss should be able to remember this fact and use it to more accurately reconstruct the sequence when asked to recall.

It is possible to infer the particular coding strategies an S has employed from the errors he makes in recall (and also to some extent from verbal reports at the end of an experimental session). This technique was used in the present study to determine (a) to what degree aural coding is evident when another means of coding is available, and (b) whether loud random ("white") noise during presentation and recall can interfere with STM and thus force S to abandon aural coding for a non-aural classification strategy.

Procedure

Five digits and five letters were selected so that for each digit there was a letter with a similar sounding name. The digits were: 2, 3, 5, 6, and 8; the corresponding matched letters were Q, T, Y, X, H. On each trial 8 of these 10 characters—four digits and their corresponding letters—were used to generate a randomly ordered sequence (e.g., 6, H, X, 5, Y, Q, 8, 2). The items were presented on a memory drum, one at a time at a 1-sec. rate. Immediately after presentation of the eighth item S wrote down as much of the sequence as he could remember. Instructions stressed that every item had to appear on the answer sheet in its proper position and that

Ss were not to guess unless reasonably certain. There were no restrictions on recall order or the amount of time allowed for recall.

Each S wore headphones during the entire experimental session. During half the trials white noise was presented over the headphones from a nearby tape recorder. The noise level was high enough so that S could not hear his own voice when speaking, but not so high as to be unpleasant after a short adaptation period. Eighty different sequences were presented, with noise and non-noise alternating in blocks of five trials. The Ss were fully informed of this procedure and always warned before the noise was turned on or off. Twenty Ss served as part of the laboratory requirement for an introductory psychology course.

Results and Discussion

An item analysis was done on all errors, providing a count of the number of times 3 was given in the position of H, and so on for all possible substitutions. As no directionality was evident in these confusions, "3 substituted for H" and "H for 3" were combined to provide a total 3-H confusion index. Omissions were also tallied. Extra-task intrusions (only seven in all) were counted as omissions for simplicity.

It has been shown that most intrusions in a task of this sort come from within the sequence presented—that is, some items are recalled but their positions apparently forgotten (Conrad, 1959). Under the assumption that all confusions come from within the 8-item sequence and that all such confusions are equally likely, we can predict that 3/7 of the substitution errors should be intra-class errors (i.e. digit-digit and letter-letter substitutions), 1/7 should be interclass aural confusions (e.g., 2-Q, 8-H, etc.), and 3/7 should cross class boundaries but be non-aural (e.g., 3-H, 5-X, etc.).

The numbers of omission and substitution errors and a comparison of substitution error percentages with the above chance predictions are presented in Table 1. Differences between noise and non-noise error distri-

TABLE I
BREAKDOWN OF ERRORS AND
COMPARISON WITH CHANCE EXPECTATIONS

Type of error	Observed		Chance
Omission	Noise (n)	927	903
Substitution	Non-noise (n)	1637	1681
Intraclass		46.5%	45.3%
Interclass			42.7%
aural		19.4%	21.8%
non-aural		34.1%	33.0%
			14.6%
			42.7%

butions produced a χ^2 of 2.84, which is not significant ($p > .50$, df = 4). Combining noise and non-noise errors and comparing with the chance expectations yielded a χ^2 of 166.33, which is very highly significant ($p < .001$, df = 2).

The directions of deviation from chance indicate both a high frequency of aural confusions and a high frequency of intraclass confusions. It appears that when Ss have an opportunity they will code in several ways. In the present case aural coding is very evident, but classification coding also shows up strongly. How classification coding could be attempted is illustrated in one S's report that he had "tried to make mathematical formulas" out of the sequences. This S made few aural confusions. Thus, it appears that the types of errors made depends in part upon the particular coding strategy S adopts. Most Ss made both aural and classification errors with high frequency.

Several confusions occurred with noticeably higher frequencies than others of the same types. Of intraclass confusions, X-H, Q-T, and 2-3 were most common. Of non-aural interclass confusions, the most frequent were 2-T, 3-Q, and 6-H. In fact, it appears that 2-3-Q-T is a cluster of highly confusable items both within and across class boundaries. Another, less effective cluster is 6-8-X-H. A possible posthoc rationalization for these clusters lies in articulatory similarities. Thus, while 3 and Q do not sound much alike, their names have a common consonant-vowel configuration. Similarly, the other cluster has a common vowel-consonant articulation. This interpretation suggests that what have been called "aural" or "auditory" confusions are really kinesthetic confusions, arising from similar muscular feedback patterns produced by subvocal rehearsal. Since the majority of acoustically similar items are also articulatorily similar, virtually all the data showing acoustic confusions in STM can be accounted for by a kinesthetic feedback hypothesis.

Lending credence to this interpretation are two outcomes of the white noise condition in the present experiment. First, it did not have the effects upon error dis-

tributions or over-all performance on the STM task which might be expected if coding were auditory. Second, it produced overt rehearsal in a few Ss (in the present study, all males). This effect was probably due to the elimination of auditory feedback as a means of monitoring (and thus suppressing) overt speech. Ss questioned indicated that they had not been aware of "thinking aloud." Since information from kinesthetic feedback was available to these Ss, we can safely conclude that the same kinesthetic feedback must have been present on non-noise trials as well—presumably produced by small, covert muscular responses in the vocal apparatus. This interpretation fits in nicely with the common experience that thinking (and in this case STM) involves "talking to oneself" without producing overt vocalizations.

It should be possible, if the kinesthetic feedback hypothesis is correct, to find certain predictable discrepancies between acoustic confusions and memory confusions. One test can be derived from Conrad's (1964) data. The name of the letter P is articulationally very similar to B, but is acoustically more similar to T, as P and T both begin with an "unvoiced stop." It can therefore be predicted that the ratio of P-B and P-T confusions should differ accordingly from the listening to the recall data. Conrad's data produce the following frequencies:

Confusion	Listening	Recall
P-T	786	103
P-B	237	164

The differences between listening and recall is highly significant ($\chi^2 = 144.6$, $p < .001$, df = 1) and in the predicted direction.

References

- Conrad, R. Errors in immediate memory. *Brit. J. Psychol.*, 1959, 50, 349-359.
Conrad, R. Acoustic confusions in immediate memory. *Brit. J. Psychol.*, 1964, 55, 75-84.

Note

1. This research was done while the author was a graduate fellow of the National Science Foundation. He wishes to thank Drs. Gordon H. Bower and Leonard M. Horowitz for their helpful suggestions.