

Instructions and processing of right vs. wrong items in multiple-item recognition learning

DONALD H. KAUSLER and RUTH E. DALEZMAN
University of Missouri, Columbia, Missouri 65201

Recognition of individual items after one, two, or four study trials on paired right (R) and wrong (W) items was contrasted for subjects expecting a standard paired R-W item test and subjects expecting an individual item test. Expecting to be tested on individual items resulted in less disparity between R and W items in terms of their recognizability as prior study, or old, items and in more false recognitions of W items as having functioned as prior R items. Informing subjects about the precise nature of their individual item test presumably diminishes the distinction between R and W items as nominally relevant and irrelevant sources of informational input, respectively.

The individual-item test procedure (Kausler, Paur, & Yadrick, 1975) provides a means of contrasting the processing given to right (R) vs. wrong (W) items during practice on a multiple-item recognition, or verbal discrimination, learning task. Recent research with this procedure (Kausler & Yadrick, 1976, 1977) has indicated that the processing of the two types of items differs both quantitatively and qualitatively. The quantitative disparity reflects the greater number of response-based frequency units accruing to R items than to W items, as postulated by frequency theory (Ekstrand, Wallace, & Underwood, 1966) and the qualitative disparity seemingly reflects the more elaborative rehearsal (Craik & Tulving, 1975) received by R items than by W items.

Variations in task conditions presumably affect multiple-item recognition learning by altering the pattern of differential processing between R and W items. Insights into the nature of this alteration may be obtained through the use of the individual-item test procedure. That is, the processing changes produced by a task variation should be manifested in hit rates both for recognizing prior R and W items as old and for identifying the prior study list functions of these old items. For example, Kausler and Remisovsky (1976) employed this procedure to determine locus of the adverse effects on multiple-item recognition learning produced by the overt pronunciation of both R and W items during study trials (e.g., Kausler & Sardello, 1967). Their results revealed that overt pronunciation had little effect on the recognition of W items, whereas it reduced markedly the recognition of R items in terms of both their oldness and their prior pairwise function. The probable changes in the processing of intrapair items produced by overt pronunciation are discussed fully by Kausler and Remisovsky (1976).

This study was supported in part by National Science Foundation Grant BMS75-05007 awarded to the first author.

Another task variable of considerable importance concerns instructions about the test subjects are to receive after their study trials are completed. The instructions for all of the previous studies employing the individual-item test procedure conformed to the standard format of multiple-item recognition learning. That is, all subjects were told, incorrectly, that they would receive a pairwise test in which they were to select each R item within the context offered by its previously paired W item. The intent, of course, was to elicit the pattern of processing that normally occurs during the course of multiple-item recognition learning. Of general interest is the extent to which this pattern departs from that obtained when subjects are informed about the nature of the individual-item test they will receive at the end of pairwise study trials. As a means of evaluating these differences in item processing, the present study compared item recognition scores after one, two, or four study trials under standard (i.e., pairwise test expectancy) instructions and informed (i.e., individual item test expectancy) instructions. The distinction between R and W items as nominally relevant and irrelevant sources of information, respectively, that is indigenous to a standard multiple-item recognition learning task should be greatly attenuated when subjects know in advance that they will be tested on both types of items.

METHOD

Subjects and Design

The subjects were 96 students in general psychology classes who fulfilled a course requirement by their participation in this study. They were assigned in randomized blocks to the six groups formed by a 2 by 3 factorial design, with instructions (standard or informed) and number of study trials, (one, two, or four) preceding the single individual-item test trial as between-groups variables. In addition, the 16 subjects of each group were further assigned to four subgroups that differed only in specific item content. The effects attributable to item content were subsequently found to be negligible. Consequently, the subgroups were combined for all analyses reported later.

Lists

Each study list contained 23 paired R and W items, 18 of which were experimental pairs and 5 of which were buffer pairs filling primacy and recency positions. Four list variants were employed, two composed of one set of words and two of a different set. The intraset variants were constructed by having one member of each experimental pair serve as the R item for one variant and as the W item for the other variant. The test list contained the 18 R items and the 18 W items, together with 18 new items. The 18 new items were words randomly selected from the alternate study list set. These lists were identical to those used by Kausler and Remisovsky (1976) (cf. Kausler and Remisovsky for a description of the overall word pool from which the lists were constructed).

Procedure

The procedure for subjects who were given standard instructions closely replicated that of earlier studies (e.g., Kausler et al., 1975) involving individual-item test trials. That is, these subjects were told about the pairwise format of multiple-item recognition learning tasks, and they practiced on a short list that conformed to this format. Informed subjects were told about the precise nature of the individual-item test they would receive after completing their designated number of study trials. In addition, all subjects practiced on a short list that conformed to the paired-item study phase/individual-item test phase procedure. For all subjects the R member of each study pair was identified by being underlined. Spatial position and serial order effects for experimental pairs were controlled across the designated number of study trials. The rate of exposure was 3 sec/pair in each study trial and 6 sec/item in the test trial, and the intertrial interval, where appropriate, was 18 sec. Old-new and R-W decisions during the test trial were recorded by each subject on an answer sheet.

RESULTS

Summary statistics for the hit rates of recognizing both R and W items as old are given in Table 1. A 2 by 3 by 2 mixed analysis of variance (arc sine transformation) with instructions (standard vs. informed) and study trials (one, two, or four) as between-groups variables and pairwise function (R vs. W) as the within-groups variable, yielded a significant main effect for study trials [$F(2,90) = 38.05, p < .001$] and for function [$F(1,90) = 44.31, p < .001$], but not for instructions [$F(1,90) < 1$]. The Instructions by Function interaction was significant [$F(1,90) = 5.24, p < .05$], as was the Instructions by Function by Study Trials

interaction [$F(2,90) = 3.72, p < .05$]. However, the Function by Study Trials interaction was clearly not significant [$F(2,90) < 1$].

The main effect for study trials resulted from increasing hit rates for old items, whether R or W, over increasing study trials, and the main effect for function resulted from the overall greater hit rate for R items than for W items. Both of these effects replicate those found in earlier studies (e.g., Kausler & Yadrick, 1977). Our present concern, however, is with the effects of instructions on old-item recognition. The significant interaction with item function reflects the greater disparity in hit rates between R and W items (favoring R items, of course) under standard instructions than under informed instructions. In addition, the significant second-order interaction with study trials indicates that the differential disparity between the two instructional conditions was most pronounced after only one study trial. From Table 1 it may be seen that the disparity in hit rates between R and W items remained relatively invariant at approximately .06 across study trials under informed instructions. On the other hand, the disparity decreased from .156, after one study trial, to .087, after four study trials, under standard instructions. This decrease in disparity most likely resulted from a ceiling effect for the recognition of prior R items as old after multiple study trials.

Summary statistics for the absolute hit rates of identifying prior R items as right in function are given in Table 2, together with the absolute false alarm rates of identifying prior W items as right in function. Considering R items first, a 2 by 3 between-groups analysis of variance (arc sine transformation) revealed a significant main effect for study trials [$F(2,90) = 14.69, p < .001$] but not for instructions [$F(1,90) = 1.97, p > .05$]. The increasing hit rates over increasing study trials again replicates an effect found in earlier studies (e.g., Kausler & Yadrick, 1977). More importantly, the variation in instructions had little effect on identifying prior R items as right. By contrast, a 2 by 3 analysis of variance revealed a significant main effect for instructions on the false alarm rates for W items [$F(1,90) = 4.62, p < .05$], as well as a significant main effect for study trials [$F(2,90) = 5.72, p < .01$]. From Table 2 it may be seen that more false alarms occurred under informed instructions than under standard instructions and that the overall false alarm rate decreased progressively as the number of study trials increased. The latter finding also replicates earlier studies involving multiple study trials (e.g., Kausler & Yadrick, 1977). Finally, for both analyses the Instructions by Study Trials interaction fell far short of statistical significance [$F(2,90) < 1$].

False alarm rates for new items misidentified as old were also determined. Means and standard deviations were .264 and .167, .149 and .101, and .097 and .104 after one, two, and four study trials, respectively, under standard instructions and .233 and .133, .191 and .137,

Table 1
Summary Statistics for Hit Rates of Recognizing
Prior Study Items as Old

Condition	Study Trials					
	1		2		4	
	Mean	SD	Mean	SD	Mean	SD
Standard						
R Items	.826	.101	.924	.064	.972	.045
W Items	.670	.143	.753	.203	.885	.098
Informed						
R Items	.785	.086	.910	.093	.948	.082
W Items	.722	.137	.837	.072	.889	.109

Table 2
Summary Statistics for Hit Rates (HR) of Identifying Prior R Items as Right and False Alarm Rates (FAR) of Misidentifying Prior W Items as Right

Condition	Study Trials					
	1		2		4	
	Mean	SD	Mean	SD	Mean	SD
Standard						
HR, R Items	.497	.192	.684	.219	.746	.177
FAR, W Items	.208	.122	.136	.111	.122	.149
Informed						
HR, R Items	.483	.151	.590	.156	.715	.159
FAR, W Items	.236	.122	.236	.127	.125	.085

and .045 and .089 under informed instructions. Only the main effect for study trials attained statistical significance [$F(1,90) = 20.33, p < .001$].

DISCUSSION

The present results support the hypothesis that the disparity in processing between R and W items is reduced when the usual distinction between the two types of items in terms of their informational relevance to the task at hand is diminished. The reduction presumably results from an increase in the number of rehearsal responses directed at W items when these items are no longer perceived as nominally irrelevant sources of informational input. The net effect is to increase the number of frequency units accruing to W items, thereby increasing their recognizability as old items and making their hit rates more comparable to those of R items. On the other hand, increased responding to W items decreases the discriminability between R and W items in terms of their functional roles during prior study trials. This decrease is apparent from the increment in false recognitions of W items as prior R items when subjects know in advance that they will be tested on individual items. The reliability of both frequency cues and nonfrequency cues in mediating functional identifications (Kausler & Yadrack, 1977) is likely to be reduced considerably under such conditions.

Reducing the disparity in the number of responses to R and W items should also make standard intrapair discriminations between R and W items more difficult. That is, the frequency advantage that normally occurs to R items under standard practice conditions (Ekstrand et al., 1966) should be attenuated, thereby reducing the efficacy of frequency cues to mediate

intrapair selections of R items. To test this prediction, the present experiment was replicated with one important exception, namely, the use of a paired-item test trial rather than an individual-item test trial. Thus, subjects receiving standard instructions were tested in the manner they expected to be tested. On the other hand, informed subjects received an unexpected test in which R and W items continued to be paired and they were required to select the R member of each pair. Means and standard deviations for the number of correct selections were 14.56 and 2.03, 16.56 and 1.50, and 17.00 and 1.32 after one, two, and four study trials, respectively, under standard instructions. Comparable values under informed instructions were 12.81 and 2.83, 14.00 and 2.10, and 17.19 and 1.14. As expected, fewer correct selections were made under informed instructions than under standard instructions [$F(1,90) = 12.50, p < .001$]. In addition, the main effect for study trials was significant [$F(2,90) = 25.61, p < .001$], as was the Instructions by Study Trials interaction [$F(2,90) = 4.40, p < .05$].

REFERENCES

- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, **104**, 268-294.
- EKSTRAND, B. R., WALLACE, W. P., & UNDERWOOD, B. J. A frequency theory of verbal discrimination learning. *Psychological Review*, 1966, **73**, 566-578.
- KAUSLER, D. H., PAVUR, E. J., JR., & YADRICK, R. M. Single-item recognition following a verbal discrimination study trial. *Memory & Cognition*, 1975, **3**, 135-139.
- KAUSLER, D. H., & REMISOVSKY, J. E. Pronunciation and individual item identifications in multiple-item recognition learning. *Bulletin of the Psychonomic Society*, 1976, **8**, 224-226.
- KAUSLER, D. H., & SARDELLO, R. J. Item recall in verbal discrimination learning as related to pronunciation and degree of practice. *Psychonomic Science*, 1967, **7**, 285-286.
- KAUSLER, D. H., & YADRICK, R. M. Differential processing of bilingual right and wrong items in multiple-item recognition learning. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 284-292.
- KAUSLER, D. H., & YADRICK, R. M. Item identifications following varying study trials on a multiple-item recognition learning task. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, **3**, 203-210.

(Received for publication June 13, 1977.)