Memory & Cognition
1980, Vol. 8 (5), 468-475

An examination of orienting task relationships in a proactive interference paradigm

CHARLES P. BIRD and RENAE ROBERTS
Ohio State University, Mansfield, Ohio 44906

Three experiments investigating release from proactive interference were conducted, in which orienting tasks were employed to bias encoding. Following earlier experiments by Bird (1976, 1977), it was expected that release would be observed when tasks were changed after several trials, but only to the extent that the tasks required different processing. Two obviously related nonsemantic tasks were compared in Experiment 1, and no release was obtained. Experiment 2 was a comparison of part-of-speech classification, considered by some to be a nonsemantic task, and of judgments of word pleasantness. The release obtained was sufficiently low to suggest that part-of-speech decisions involve substantial semantic processing. Finally, Experiment 3 employed four tasks, in order to address various questions about task relationships raised in the earlier experiments. Based on the levels of release observed across experiments and the finding that some tasks led to less proactive interference than others, a tentative categorization of tasks was proposed.

Since Craik and Lockhart (1972) introduced their levels-of-processing framework, theoretical development in memory research has tended to stress a processing orientation over the multistore models that dominated thinking for several years (e.g., Atkinson & Shiffrin, 1968). One of the critical points in the levels framework was that word retention should be a function of processing depth. That is, recall of words should be higher following the processing of word meaning (a deep level of processing) than following the processing of auditory or structural information. Although there is considerable evidence to support the depth notion, recent experimental work has indicated that variables other than depth may be critical determinants of performance. For example, it has been shown that the retention of a "shallow" encoding may be higher than that of a "deep" encoding if the retention test requires attention to shallow word features (Morris, Bransford, & Franks, 1977) or if the deeper features are subject to heavy proactive interference (PI) (Bird, 1976, 1977).

As a result of increasing difficulties with the depth-of-processing concept, the most recent trend seems to be to stress a broad notion of encoding distinctiveness (e.g., Jacoby & Craik, 1979). Distinctiveness may be defined as the extent to which a particular processing strategy differentiates an item or a set of items from other information in memory. As developed by Jacoby and Craik, distinctiveness and depth are not unrelated, since encodings that stress word meaning should have greater potential for developing distinct codes than should encodings that stress structural or auditory features. However, distinctiveness also depends on previously encoded events, the elaborateness of the code developed, and the cues or demands present at retrieval. Thus, distinctiveness is always relative to some particular set of conditions, and an encoding that is distinctive in one retrieval context may not be in another.

Experimental tests of both the depth and distinctiveness ideas generally require control over encoding or retrieval processes. One popular method for controlling encoding involves having subjects make a decision about each word as it is presented. These decisions are intended to orient attention toward some features of the words and, presumably, away from other features. For example, requiring subjects to judge each word as pleasant or unpleasant (pleasantness judgments) may direct attention to semantic features, whereas requiring subjects to estimate the number of letters in each word (letter estimates) may stress nonsemantic features.

Although the use of orienting tasks is central to tests of important theoretical issues, there has been relatively little research on the tasks themselves or on the relationships among them. Following the work of Jenkins and his colleagues (see Jenkins, 1974, for a review), it has been common to call any task that is followed by high recall semantic and any followed by poor recall nonsemantic. However, there are problems with the semantic-nonsemantic distinction, as well as with the initial assumptions about the effect of tasks on processing.

One potential problem is that the semantic-nonsemantic distinction may miss important differences between tasks in one category or the other (cf. Postman & Kruesi, 1977). A second, perhaps more crucial, problem is that the basis of classification was essentially

This research was supported by a Mansfield Campus Professional Development Grant and by an OSU Small Research Grant. We thank Peggy Burton for assistance in the data collection for Experiment 3. Requests for reprints should be sent to Charles Bird, Ohio State University, Mansfield Campus, 1680 University Drive, Mansfield, Ohio 44906.

  0090-502X/80/050468-08$01.05

a combination of intuition and the assumption that high retention is associated with semantic processing. The assumption seems to be weak, given the problems with a simple depth-of-processing notion, and it has not always been obvious why certain semantic tasks should stress meaning more than some nonsemantic tasks (cf. Postman, 1975).

If the application of orienting tasks in memory research is going to continue, it would seem to be prudent to develop a more sophisticated understanding of their effects on performance. If these effects are complex, as we suspect, then researchers should be aware of the complexity when selecting tasks and types of retention tests. The present study was an attempt to learn more about how orienting tasks affect performance on a recall test and to specify at least some aspects of the relationships among tasks. A better understanding of these relationships might allow the development of a classification scheme that avoids the problem of assuming that high recall follows only semantic (deep) processing and that is more consistent with the current theoretical stress on encoding distinctiveness.

The method was a variant on the release-from-PI procedure developed by Wickens (see Wickens, 1972, for a review of the development and initial applications of this procedure). The release procedure involves a number of trials (usually four), during each of which several to-be-remembered (critical) items are presented, followed by some distractor task and a test for recall of the original items. Typically, the items presented on the first several trials share some attribute (e.g., they might be names of animals), and recall tends to decrease from trial to trial (a build-up of PI). After several trials, there is a shift so that the critical items share some different attribute (e.g., they might be names of trees), and, for a number of dimensions that have been tested, recall increases dramatically.

Although there has been considerable controversy concerning the locus of PI effects, it would seem that PI and PI release are rather easily explained by the concept of encoding distinctiveness. As the subject proceeds from trial to trial, with no change in item organization and with relatively trivial context changes, the distinctiveness of the encodings produced will become increasingly low, decreasing the likelihood of successful recall. When the organization of items is changed after several trials, there is an increase in distinctiveness (i.e., the encoded attribute serves as a more effective retrieval cue) and recall increases.

Orienting tasks can be employed in a release-from-PI procedure by requiring one task on all words for several trials and then either continuing with the same task on the final trial (the no-release control condition) or changing to a different task (the release condition). Any observed improvement in the release condition on the final trial can provide a measure of endoding distinctiveness.[1]

The proposed measure, percentage release, is based only on the distinctiveness between two tasks, but it can serve a useful purpose. For example, it will provide an empirical basis for grouping similar tasks (i.e., tasks that do not produce a high percentage release in relation to each other but that do produce high release in relation to tasks in other groups) that goes beyond recall totals. In addition, since encoding distinctiveness is a relative condition, depending on the total set of processing events, an empirical grouping will allow researchers to manipulate distinctiveness by selecting orienting tasks that provide a known level of distinctiveness in relation to each other.

The potential usefulness of combining orienting tasks and a release procedure was demonstrated in three experiments by Bird (1976, Experiment 1; 1977). In two of the experiments, a semantic task was followed by a nonsemantic task or vice versa. In one of these, the task was either modifiers (the subject generated an adjective to modify a presented noun or a noun that was appropriately modified by a presented adjective) or rhymes (the subject generated a word that rhymed with the presented word); in the other comparison, the task was either pleasantness judgments or letter estimates. In both experiments, the percentage release from PI following a task change was high (80% in the first experiment and 81% in the second). The third experiment was a comparison of two semantic tasks, pleasantness judgments and judgments of whether each word was active or passive (active-passive judgments). Since a change from one semantic task to another should provide less increase in distinctiveness than a change from a semantic to a nonsemantic task, less release was predicted. In fact, only 41% was obtained, suggesting the release procedure can be used to discriminate closely related tasks from those that allow considerable distinctiveness.

Of course, the percentage release produced by a change from one task to another will not necessarily equal that of a change in the opposite direction, so both directions of change were examined. With the semantic-semantic comparison, release was symmetrical (44% for a change from active-passive judgments to pleasantness judgments and 38% for a change in the opposite direction), but both semantic-nonsemantic changes were asymmetrical (109% and 100% for the semantic-to-nonsemantic change and 58% and 57% for the nonsemantic-to-semantic change in the modifiers-rhymes study and in the pleasantness-letter estimates study, respectively). It is apparent that greater release was obtained when changing to a task for which total recall is generally poor than when changing to a task that typically produces high recall. This finding requires some discussion before proceeding with the present study.

An important point in considering asymmetry is that the control group used to calculate percentage release for a one-direction change performs the same task on all trials that the experimental group performs on only the final trial. The asymmetry, therefore, may be due to subjects in the experimental group performing both the earlier task and the new task. Such an event could improve recall on the final trial when changing to a nonsemantic task, relative to the control group, or it could retard recall when changing to a semantic task, distorting the release percentages. This effect would be cancelled out when looking at combined release percentages, making combined release a more stable measure of distinctiveness.

The asymmetry could also be due to extremely high or low PI development in the control group. For example, if there is very little reduction in recall over the first few trials, then even a small increase in recall on the final trial will produce a large percentage release. In fact, Bird (1977) did find that less PI developed with letter estimates than with pleasantness judgments, although Bird (1976) found comparable levels of PI with rhymes and modifiers. Again, this potential cause of asymmetry should be somewhat compensated for by looking at combined release, since differences in PI sensitivity would be eliminated by the counterbalancing. In any case, there was considerable interest in the present study in performance over trials, as well as in release effects. Examining PI buildup should provide an important additional perspective on task differences.

In the present study, three additional release-from-PI experiments were conducted in an attempt to specify relationships between a number of orienting tasks. Experiment 1 was a comparison of two nonsemantic tasks, a type of comparison not previously made; these tasks were selected because they appeared to require very similar processing. The tasks required subjects to count the number of es in each word (E-checking) or to count the number of gs (G-checking). We were interested in whether the apparent similarity of processing would be confirmed by the finding of little if any release following a task change and whether any release observed would be symmetrical. In addition, since PI apparently developed more slowly with letter estimates than with pleasantness judgments in an earlier study, we were interested in the pattern of recall over the first few trials. Finally, if the expected trivial level of release was observed, the usefulness of the release procedure would be confirmed once again.

EXPERIMENT 1

Method

Design. The design was a 2 by 2 by 4 factorial, in which the factors were experimental condition (release vs. no-release), counterbalancing on orienting task, and the within-subjects trials. Release conditions were produced with a change in orienting task on Trial 4 (E-checking/G-checking and G-checking/E-checking conditions), whereas no-release conditions required the same task on all trials (E-checking/E-checking and G-checking/G-checking conditions).

Materials and Procedure. The experimental items were 20 nouns and 20 adjectives assigned to groups of five critical items such that there were 3 nouns and 2 adjectives or 3 adjectives and 2 nouns in each group. The eight groups of words thus selected were divided into two sets of four groups, so that there were two independent sets of experimental items. The presentation order of the groups of words within a set was determined by a 4 by 4 Latin square.

The words were printed in block letters on 2 x 2 in. slides. The interval between presentation of the words and recall was filled with a digit-reading task involving randomly chosen digits (0-9) printed in the form of a 5 by 5 array on 2 x 2 in. slides. All materials were presented by a timer-controlled Kodak Carousel projector, and the recall intervals were timed with a stopwatch.

Instructions were read by the subject and were the same in all conditions, so that all subjects were prepared to perform both processing operations. If the operation was "E-checking," then as each word appeared on the screen the subject was to count the number of es and say the number aloud. If the operation was "G-checking," then the subject was to count and say aloud the number of gs. No suggestions were made that the orienting tasks should be used to assist the subject in remembering the words, but the subject knew, of course, that there would be a recall test at the end of the retention interval.

Each trial began with a 3-sec presentation of a slide, with either "E-checking" or "G-checking" printed in letters larger than those used for the experimental items. The five words for that trial were then presented one at a time, at a 3-sec rate (onset to onset). While each word was on the screen, the subject gave an appropriate response aloud. Following the five words, four number arrays were presented for a total of 20 sec (at a 5-sec rate); during this time, the subject read the numbers aloud as quickly as possible from left to right across each row. Finally, a slide with three question marks was presented for 20 sec, and the subject tried to recall the five words aloud. The procedure continued in this manner for four trials.

Subjects. Students at the Ohio State University served as subjects in partial fulfillment of a requirement of the general psychology course. There were 32 subjects in each of the four conditions, making a total of 128 subjects. Subjects were tested individually and were assigned to conditions randomly, with the restrictions that all subjects receiving a particular word order within a particular word set be tested before the order was changed and that there be equal proportions of males to females in the four experimental conditions.

Results and Discussion

Responses were scored as correct when they were given within the recall interval of the trial on which they first occurred, without regard to the original order of the words. Table 1 presents the percentage of correct responses on each trial for the four between-subjects and for the two release and two no-release conditions combined. There was a modest decrease in performance over the first three trials, indicating a buildup of PI, but

Table 1
Percentage of Correct Responses, Experiment 1

Condition	Trial 1	2	3	4
E-G	63	47	42	41
G-G	56	56	49	51
G-E	59	50	51	53
E-E	56	49	50	51
Release Combined	61	48	47	47
No Release Combined	56	52	50	51

Note—E refers to E-checking and G to G-checking. The label combination refers to the Trials 1-3 and Trial 4 tasks, in that order.

there was no release from PI on Trial 4 for the release condition. In fact, performance on Trial 4 was somewhat better in the no-release than in the release condition.

A 2 by 2 by 3 analysis of variance was calculated for Trials 1-3; it revealed only a significant main effect of trials [$F(2,248) = 8.93$, $p < .001$, MSe = 1.07]. A separate 2 by 2 analysis calculated for Trial 4 revealed no significant differences. Thus, although significant PI developed in this experiment, there was no release with the apparently trivial shift from E-checking to G-checking or vice versa. Since the decrease in performance across Trials 1-3 was slight in the control (no-release) conditions, Wickens' (1972) release formula did not provide a very accurate description of the data. The percentage release was −80% overall and was grossly asymmetrical (48% for the change from G-checking to E-checking and −200% for the E-checking to G-checking change). Obviously, these percentages do not provide a meaningful comparison with earlier studies. It appears that PI did not build as dramatically with E-checking and G-checking as it did with other tasks and that the asymmetry problem was magnified as a result.

EXPERIMENT 2

Since the effectiveness of the release procedure as a means of measuring relative distinctiveness has been established through four experiments, the procedure was employed in Experiment 2 to examine a task that is not obviously semantic or nonsemantic. The task, called part of speech, requires the subject to classify each word into a syntactic category. In the present study, the items were either nouns or adjectives, and the comparison task was pleasantness judgments. Although some researchers have found relatively poor recall following a part-of-speech task and, therefore, have classified it as a nonsemantic task (e.g., Hyde & Jenkins, 1973), others have pointed out that semantic processing must be an important part of performing the necessary operations (e.g., Postman, 1975). In the present experiment, it should be possible to determine how part of speech compares with tasks previously combined with pleasantness judgments.

Method

Design, Materials, and Procedure. All details of this experiment were identical to those of Experiment 1, except that subjects either judged an item as pleasant or unpleasant or classified it as a noun or an adjective (in all cases, giving their responses aloud). As before, there were four trials, and the required task was identified by a slide presented at the beginning of each trial.

Subjects. There were 128 subjects, drawn from the same pool and assigned to conditions in the same manner as before.

Results and Discussion

Table 2 presents the percentage of correct responses on each trial for the four between-subjects conditions and for release and no-release conditions collapsed over the orienting task required. Performance was more consistent across conditions than in Experiment 1, and there was a clear increase in performance on Trial 4 in each release condition relative to the appropriate no-release condition.

Table 2
Percentage of Correct Responses, Experiment 2

Condition	Trial 1	2	3	4
PJ-PS	74	59	45	54
PS-PS	67	54	48	47
PS-PJ	73	59	46	63
PJ-PJ	79	62	51	54
Release Combined	73	59	45	58
No Release Combined	73	58	49	50

Note–PJ refers to pleasantness judgments and PS to part of speech. The label combination refers to the Trials 1-3 and Trial 4 tasks, in that order.

An analysis of variance calculated for Trials 1-3 revealed only a significant trials factor [$F(2,248) = 54.03$, $p < .001$, MSe = 1.00]. The part-of-speech and pleasantness judgments tasks did not produce reliably different performance [$F(1,124) = 2.00$, $p > .10$, MSe = 1.69]. Analysis of Trial 4 data revealed both that performance in the release condition was superior to that in the no-release condition and that performance was better following pleasantness judgments than following part of speech [$F(1,124) = 4.34$ in both cases, $p < .05$, MSe = 1.22]. There was no interaction of task with release condition.

The PI buildup in control conditions was much more substantial in Experiment 2 than in Experiment 1. Wickens' (1972) formula indicated a moderate level of release on Trial 4, with nearly perfect symmetry for the two directions of change (35% for the combined conditions, 35% for the pleasantness/part-of-speech change, and 36% for the part-of-speech/pleasantness change). Performance in Experiment 2 was very similar to performance in the comparison of pleasantness judgments and active-passive judgments by Bird (1977, Experiment 2). The level of release is comparable (35% here and 41% in the earlier study) and symmetrical in both. In the present study, there was some advantage with pleasantness judgments, and there was a similar but nonsignificant advantage with pleasantness judgments in the earlier study. Thus, the relationship between pleasantness judgments and part of speech is comparable to that between pleasantness judgments and active-passive judgments, and both relationships were closer than that between pleasantness judgments and letter estimates. Part of speech was also included in Experiment 3, but the indication here is that it may be at least as close to tasks usually considered to be semantic as to tasks considered to be nonsemantic.

EXPERIMENT 3

Experiment 3 was conducted to compare four tasks that are often considered to require only superficial processing. Three of these tasks (letter estimates, part of speech, and rhymes) were previously employed in a comparison with some task requiring attention to meaning, and the fourth task, E-checking was compared with the equally superficial task of G-checking in Experiment 1. It was expected that the four tasks employed would require rather different processing activity, with the exception that E-checking and letter estimates should be closely related. In addition, there was some special interest in comparing the part-of-speech task with tasks that more clearly do not require attention to meaning. Finally, we were interested in examining the rate of PI buildup for these tasks, since in previous studies there appeared to be less PI over trials with superficial tasks than with those that stress meaning.

Method

Design, Materials, and Procedure. The design was a 4 by 4 by 4 factorial, in which the factors were the task performed on Trials 1-3, the task performed on Trial 4, and the within-subjects trials. The materials and procedure were the same as in the first two experiments, except that all subjects were instructed on the requirements of four tasks instead of two, so that each had equal expectations. When the task was letter estimates, subjects gave an estimate of the number of letters in each word. When it was rhymes, they generated a rhyme. When it was E-checking, they counted the number of es in each word. When it was part of speech, they classified each word as a noun or an adjective. All responses were given aloud.

Subjects. There were 32 participants in each of the 16 between-subjects conditions, making a total of 512 subjects. They were drawn from the same pool and assigned to conditions in the same manner as before.

Results and Discussion

Trials 1-3. Performance on Trials 1-3 for each of the four tasks required is presented in Table 3. It may be seen that recall decreased over trials with all tasks, although recall differences from trial to trial varied considerably. A 4 by 4 by 3 analysis of variance indicated that the main effect of the orienting task required was reliable [$F(3,496) = 10.66$, $p < .001$, MSe = 1.79], as was the trials factor [$F(2,992) = 123.90$, $p < .001$, MSe = .92]. In addition, there was an interaction of task and trials [$F(6,992) = 6.36$, $p < .001$, MSe = .92]. No other factors were significant. In particular, the four groups of subjects performing a specific task did not differ prior to receiving the Trial 4 requirement.

A Newman-Keuls test was calculated for the significant main effect of orienting task; it revealed that each condition differed from the other three ($p < .05$ or $p < .01$). The order of the tasks, from the highest to the lowest performance, was part of speech, E-checking, letter estimates, and rhymes. A Newman-Keuls test calculated on the interaction of task with trials indicated that performance on Trial 3 was lower than that on Trial 1 for all tasks ($p < .01$) and that performance decreased significantly from one trial to the next in all cases ($p < .05$ or $p < .01$), except for Trial 1 to Trial 2 with E-checking.

Table 3
Percentage Correct Recall on Trials 1-3, Experiment 3

Task	Trial 1	Trial 2	Trial 3	Total
E-Checking	58	54	48	53
Letter Estimates	58	53	43	51
Rhymes	61	46	35	48
Part of Speech	72	54	48	58

Table 4
Percentage Correct Recall on Trial 4, Experiment 3

Trial 4 Task	Task on Trials 1-3: E	LE	R	PS
E	44	51	49	62
LE	50	51	49	49
R	58	61	38	58
PS	64	64	63	50

Note–E refers to E-checking, LE to letter estimates, R to rhymes, and PS to part of speech. Performance in the control groups (same task on all trials) is shown on the diagonal beginning in the upper left-hand corner.

The results for Trials 1-3 fit well with the earlier experiments. First, recall following part of speech is best, reinforcing the idea that it may actually be closer to the "semantic" tasks. Second, only modest, although reliable, PI developed with the E-checking task, which is consistent with Experiment 1 and may account for the higher overall performance with E-checking than with letter estimates or rhymes. An interesting final point is that performance on Trials 2 and 3 was very poor in the rhymes condition, indicating that the rhymes task is quite sensitive to PI. This fits with the finding by Bird (1976, Experiment 1) that rhymes produced as much PI as the semantic modifiers task. Hence, it is not the case that all nonsemantic tasks produce little PI.

Trial 4. The percentage of correct responses on Trial 4 is presented in Table 4 for each of the four tasks required, broken down according to the task required on Trials 1-3. A 4 by 4 analysis of variance indicated that both the main effect of the task on Trial 4 and the interaction of the task on Trial 4 with the task on Trials 1-3 were reliable [$F(3,496) = 5.78$ and $F(9,496) = 4.01$, respectively, ps $< .001$, MSe = 1.20].

A Newman-Keuls test calculated for the main effect indicated that recall was highest again following part of speech ($p < .05$ in comparison with rhymes and $p < .01$ in comparison with E-checking and letter estimates), but the other three tasks did not differ reliably. Although the differences were not significant, the order of performance was rhymes, E-checking, and letter estimates, which places rhymes in the same position as on Trial 1, but ahead of its position for Trials 1-3 combined.

A Newman-Keuls test calculated for the interaction requires a separate description for each condition. The appropriate comparison is to consider each group that performed the same task on all four trials to be a control group and to compare performance following that task on Trial 4 with performance on the same task when preceded by each of the other three.

When the Trial 4 task was letter estimates, there were no differences between the control group and the other groups. When the task was rhymes, all groups were significantly better than the control ($p < .01$). When the task was part of speech, all groups were better than the control ($p < .05$), but the magnitude of the difference was clearly less than that with rhymes. Finally, when the task was E-checking, having part of speech as the preceding task led to significantly better performance than that of the control group ($p < .01$), but having letter estimates or rhymes as the preceding task did not improve performance relative to the control group. Discussion of these differences will be presented in the context of release from PI.

Release from PI. Table 5 presents a matrix of the percentage release from PI for each pair of tasks. Table 6 presents the matrix for the two directions of change separately to allow an examination of symmetry.

It may be seen in Table 5 that a change between letter estimates and E-checking produced only a low level of release. This was expected to be the closest relationship among the comparisons, and it fits well with Experiment 1, in which E-checking and G-checking were even closer. A change between letter estimates and part of speech produced moderate release, and all other changes produced a high percentage of release. Thus, except for the letter estimates/E-checking relationship, it appears that the four tasks examined are not particularly close to each other in processing requirements and that a change from one to another allows for substantial differentiation.

Examination of Table 6 will reveal a lack of symmetry for the release effect. Performing rhymes following any other task led to greatly improved performance, but changing to letter estimates consistently led to slightly poorer performance than that in the control group and, thus, to negative release. Changing to E-checking did not greatly improve performance (see Table 4), but since E-checking is relatively insensitive to PI, any improvement can lead to a high percentage of release. Changing to part of speech consistently led to a high level of release, but not with percentages above 100 as with rhymes. Except for E-checking following part of speech, the release percentages after performing any particular task on Trials 1-3 seem to be extremely consistent. However, as mentioned in the introduction, the actual percentages obtained reflect the amount of PI developed in the control group, as well as improvement in performance following a task change, so the combined percentages (Table 5) probably are a more reliable indicator of task relationships.

Table 5
Combined Percentage Release from PI

Task	Task: LE	R	PS
E	23	63	80
LE		71	40
R			85

Note—E refers to E-checking, LE to letter estimates, R to rhymes, and PS to part of speech.

Table 6
Percentage Release from PI

Trial 4 Task	Task on Trials 1-3: E	LE	R	PS
E		39	28	100
LE	−10		−20	−20
R	100	115		100
PS	67	67	62	

Note—E refers to E-checking, LE to letter estimates, R to rhymes, and PS to part of speech.

In summary, the release obtained suggests that E-checking and letter estimates are closely related and that rhymes and part of speech differ both from each other and from the "letter" tasks. The percentages recorded in Table 6 make it clear that percentage release is not sufficiently precise as a measure to make fine discriminations among tasks. Even so, the general task relationships are straightforward, and we suggest that the anomalies reflect inconsistency in performance following the most superficial tasks. (Note, there are similar irregularities with E-checking and G-checking in Experiment 1.).

GENERAL DISCUSSION

Through a series of six release-from-PI experiments, including those reported by Bird (1976, 1977), a total of eight orienting tasks have been tested at least once. Two principal objectives of this research were to obtain new information about orienting task effects on recall and to develop a potential measure of task relatedness. The PI-release method appears to have met both objectives.

Obviously, other measures of task relatedness are important, and the release procedure allows the comparison of only two tasks at a time. Direct conclusions, therefore, will have to be limited to the tasks actually compared. Nevertheless, the tasks we examined have been used frequently and seem to be representative of a wide variety of potential tasks. In the discussion, we consider the direct implications of the release studies for the tasks tested, the general implications that apply to any experiment involving orienting tasks, and the relationship of the results to the concept of encoding distinctiveness.

As expected, the tasks examined were sensitive to the release procedure. It was entirely possible, a priori, that all task changes would lead to 100% or to 0% release. Instead, we obtained a range of values for various comparisons, providing an empirical basis for task classification. In this sense, the present research is much like Wickens' (1972) work on encoding dimensions in short-term memory.

Except for the conclusion that part of speech is as closely related to semantic tasks as to nonsemantic tasks, the present findings are in good agreement with the work of Jenkins (1974) and his colleagues. However, we propose that the category of nonsemantic tasks be divided into categories of auditory and structural tasks. This division reflects the fact that all nonsemantic tasks do not require the same processing, even though they all tend to produce poor free recall. More important, however, we found significant differences between the rhymes task and the structural tasks (E-checking, G-checking, and letter estimates) in sensitivity to PI. Further, rhymes produced high release in comparison with both semantic and structural tasks.

The structural tasks produced low (sometimes negative) release in comparison with each other and high, asymmetrical release when compared with tasks outside the group. Within the structural group, percentage release was not strikingly symmetrical, but examination of actual performance reveals that the asymmetry may be an artifact due to the low level of PI developed over Trials 1-3. No shift from one structural task to another led to an impressive increase in recall.

We are not at all convinced that the four semantic tasks (pleasantness judgments, active-passive judgments, modifiers, and part of speech) are closely related to each other. The percentage release produced in semantic-semantic comparisons is higher than that produced within the group of structural tasks, and pleasantness judgments tended to produce higher recall than did either active-passive judgments or part of speech (which is consistent with data reported on pleasantness judgments by Packman & Battig, 1978, and by Postman & Kruesi, 1977). In addition, the part-of-speech/letter estimates release percentage was only 40%, suggesting that part of speech might be placed on a continuum between pleasantness judgments and letter estimates. However, with the one exception, release percentages were more moderate for within-group comparisons than for cross-group comparisons, and the within-group comparisons were unusually symmetrical. Development of a task continuum seems plausible, but sufficient data to specify points are lacking.

The present research has implications for understanding orienting tasks that go beyond the specific tasks tested. One major point follows from task differences in sensitivity to PI. If researchers are using a particular task because of its effect on performance, they should be aware that similar levels of total recall do not necessarily reflect a similar pattern of recall over trials. Consider, as an example, recall following rhymes and E-checking over Trials 1-3 in Experiment 3. Rhymes led to rather high recall on Trial 1, but for the combined trials, E-checking produced better total recall (because relatively little PI developed).

We are especially concerned about the practice of labeling as nonsemantic all tasks that produce poor total recall. Since we now know that nonsemantic tasks can lead to higher performance than follows semantic tasks under certain conditions, recall differences alone do not provide an adequate basis for classification. Further, structural and auditory tasks may direct attention to nonsemantic features, but they clearly do not eliminate semantic processing (cf. Nelson, Walling, & McEvoy, 1979; Postman, Thompkins, & Gray, 1978). In fact, the less consistent performance and weak PI developed with structural tasks lead us to agree with others that superficial tasks often produce poor recall simply because they interfere with usual processing more than semantic tasks do. This point of view contrasts with that of Eysenck (1979), who suggested that shallow encodings are followed by poor recall in part because they are more susceptible to PI than are deep encodings.

We view the release experiments as lending modest support to the developing concept of encoding distinctiveness. In the context of PI research, however, distinctiveness refers only to the relationship between pairs of tasks. This necessarily overlooks the effect of different retention tests or the distinctiveness a particular encoding produces in relation to the total set of information in memory. To describe the relationship between pairs of tasks, it may be useful to employ a notion of psychological distance.

Distance was used previously by Wickens, Dalezman, and Eggemeier (1976) to account for varying levels of release obtained with a shift in word category. With regard to orienting tasks, we define distance in terms of the release from PI obtained with a task change. Two tasks that are very distant could be said to produce mutually distinct encodings.

Obviously, the asymmetry observed with some task comparisons is a problem for a concept of distance. However, except for the modifiers-rhymes comparison (Bird, 1976, Experiment 1), asymmetry occurred only when one or both of the tasks were structural. This type of asymmetry has already been attributed to weak PI development in structural control groups and its effect on the percentage-release formula. We do not want to minimize the importance of asymmetry, but it may be caused more by variability in performance, especially with the structural tasks, than by any theoretically crucial events.

In conclusion, we recommend that researchers exercise caution in employing orienting tasks. Equivalent levels of total recall do not necessarily reflect equivalent patterns of performance or a close relationship between encodings. As with many memory procedures, the effect of orienting tasks is not as simple as it first appeared.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2). New York: Academic Press, 1968.

Bird, C. P. On the role of processing requirements in short-term memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 234-243.

BIRD, C. P. Proactive inhibition as a function of orienting task characteristics. *Memory & Cognition*, 1977, **5**, 27-31.

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.

EYSENCK, M. W. Depth, elaboration, and distinctiveness. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J: Erlbaum, 1979.

HYDE, T. S., & JENKINS, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 471-480.

JACOBY, L. L., & CRAIK, F. I. M. Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J: Erlbaum, 1979.

JENKINS, J. J. Can we have a theory of meaningful memory? In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*. Potomac, Md: Erlbaum, 1974.

MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 519-533.

NELSON, D. L., WALLING, J. R., & MCEVOY, C. L. Doubts about depth. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, **5**, 24-44.

PACKMAN, J. L., & BATTIG, W. F. Effects of different kinds of semantic processing on memory for words. *Memory & Cognition*, 1978, **6**, 502-508.

POSTMAN, L. Verbal learning and memory. *Annual Review of Psychology*, 1975, **26**, 291-335.

POSTMAN, L., & KRUESI, E. The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 353-369.

POSTMAN, L., THOMPKINS, B. A., & GRAY, W. D. The interpretation of encoding effects in retention. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 681-705.

WICKENS, D. D. Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C: Winston, 1972.

WICKENS, D. D., DALEZMAN, R. E., & EGGEMEIER, F. T. Multiple encoding of word attributes in memory. *Memory & Cognition*, 1976, **4**, 307-310.

NOTE

1. We measured PI release using Wickens' (1972) formula: Percentage release equals the difference between percentages recalled on the final trial in the release and no-release conditions divided by the total drop over trials in percentage recalled in the no-release condition multiplied by 100.

(Received for publication August 9, 1979;
revision accepted March 31, 1980.)