27

JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 8, 622–632 (1969)

Compositional Structure in Recall

WILLIAM W. ROZEBOOM

University of Alberta, Edmonton, Alberta, Canada

Subjects trained on complexly structured stimulus displays were tested on fragments of these structures in search of recall phenomena which differentiate between associative (part–part) and redintegrative (part–whole) modes of memory. Spatial relations among display components and intradisplay redundancy were both found to exert appreciable effects on recall.

Although the concept of "structure" has become Big Magic for increasingly large sectors of contemporary psychology, efforts to make clear just what structure is and does have been remarkably scanty. Where the concept had received any attempted explication at all (e.g., Peak, 1958; Garner, 1960; Deese, 1965; Mandler, 1967; Pollio, 1968), it has been construed almost entirely as *associative structure*; i.e., a system of covariational linkages among stimulus and response variables or their central (ideational) counterparts. This is, to be sure, one perfectly good type of psychological "structure," but there also exists a second kind, fully as important as the first, of which we have as yet virtually no technical understanding despite the fervent but fumbling efforts of the Gestalt movement[1] to get hold of it and the pioneer work of Selz[2] currently being rediscovered by the psychology of thought. I refer to *compositional structure*, namely, the pattern by which a complex whole is composed of an ensemblage of parts standing in specific relations to one another in the way, e.g., the U.S. Constitution and *Peyton Place* are different constructions from the same stock of English phonemes or a newsprint photograph is a distinctive configuration of tiny dots.

For purposes of this report, it will not be necessary to discuss the generic nature and psychonomic significance of compositional structure, desirable in its own right as such an analysis would be. It suffices to speak briefly about the (compositional) structure of a compound stimulus.

It is well known that a person who has repeated joint experiences of two stimuli A and B eventually becomes able to reproduce B (or its motor equivalent) when presented with A alone as a cue for recall. Traditionally, this is explained by the hypothesis that co-experiencing A and B causes an "association" to be learned between them (more precisely, between their ideational counterparts, or from the first as stimulus to the motor equivalent of the second as response), in virtue of which stimulation by A elicits the idea or motor equivalent of B. However, this account ignores the fact that unless stimuli A and B affect different sensory modalities (and perhaps even if they do), the perceptible relationship between A and B on a given joint presentation is much richer than bare co-occurrence. Thus if A and B are visual stimuli, A may be to the left of, or above, or super-

[1] As my own grasp of psychological "structure" has matured (cf. Rozeboom, 1960, 1961, 1965, 1967), I find myself increasingly sympathetic to Gestalt outlooks such as Wertheimer (1959) and Asch (1969). Unfortunately, however, the Gestalt literature has verbalized its crucial insights far too obscurely, often with an active antipathy for analytic approaches, to provide an intellectually workable account of them.

[2] See Humphrey, 1951, pp. 132–142; Mandler and Mandler, 1964, pp. 225–234; de Groot, 1965, pp. 52–72. Although the experiment reported here was designed before I learned of Selz's work, it is strongly in the spirit of the latter.

imposed upon, or darker, or smaller, or more intense than *B*, etc. That is, a person who jointly experiences items *A* and *B* actually receives a stimulus compound $R(A, B)$ comprising elements *A* and *B* standing in some relation or set of relations *R* which inevitably includes considerably more than just temporal contiguity.

Insomuch as the particular fashion, *R*, in which elements are coupled in stimulus complex $R(A, B)$ is a phenomenally prominent feature of the perceiver's total experience, it would be remarkable if the distinctive character of this structure did not contribute significantly to the experience's causal consequences for, *inter alia*, learning and recall. Yet to date the only aspect of stimulus structure explicitly studied by learning theory is that of temporal contiguity, the sole recognized effect of which in turn is formation of evocation bonds. An outstanding and still largely virginal research problem is thus the following: What effect does experiencing a stimulus complex $R(A, B)$ have which is *more* than mere strengthening of a part-part association $A \rightarrow B$ through which subsequent stimulation by *A* alone tends to arouse the idea or motor equivalent of *B*?

One particularly important such effect, for example, may well be that experiences of $R(A, B)$ lay down a "memory trace" in virtue of which stimulation by *A* redintegrates the idea of the entire structured complex $R(A, B)$ rather than of *B* alone—i.e., that what *S* learns is better symbolized as "$A \rightarrow R(A, B)$" than as "$A \rightarrow B$," a part–whole rather than part–part mode of recall (cf. Horowitz & Prytulak, in press). Although verbal learning data are on record which clearly favor a part–whole interpretation of memory over the traditional part–part model (e.g., Rozeboom, 1967), a head-on empirical confrontation between these two alternatives proves surprisingly difficult to arrange. Reported here is an attempt to explore the problem of redintegration through study of a related question: If *S* is trained on a set of complex stimuli, each having a structure wherein one element *A* stands in relation R_1 to an element *B* as well as in another relation R_2 to another element *C*, what phenomena will appear when *S* is presented with a partial complex $R_1(A, \_)$ or $R_2(A, \_)$ and asked to remember which element stood in that relation to *A*? That is, what happens when *S*'s cue for recalling part of an experience is not just another detached component thereof, but that component embedded in the relational framework which previously united those elements? The present experiment was designed not to test any particular hypothesis about such phenomena, but simply to provide a setting wherein the patterns of recall errors produced by the structural features of past and present experience reveal something about the role of compositional structure in memory.

Method

Subjects. Two hundred forty male and 144 female University of Alberta undergraduates participated in this experiment to satisfy a course requirement. The *S*s were tested in batches of roughly 50 *S*s per batch. Each *S* was issued a packet of experimental materials which he manipulated throughout the session in accord with paced instructions. The procedure was identical for all *S*s except that the materials received by the various *S*s were of 48 different kinds, eight alternative training sets combined with six alternative orders of recall. For six of the eight training sets, each recall was administered to an equal number of male and female *S*s, whereas *S*s receiving the other two training sets were all males.

Task. Experimental materials consisted of one training booklet and nine sequentially numbered test booklets, each containing 12 pages. Each page of the training booklet displayed a distinctive stimulus configuration, while the *i*th ($i = 1, \ldots, 12$) page of each test booklet contained a portion of the stimulus display on page *i* of the training booklet together with a dotted blank in place of one of the missing parts. On each of the experiment's nine training-and-test series, *S* first studied each successive page in the training booklet for 8 sec timed by *E*, and then had 150 sec (unpaced) to attempt filling each dotted blank in the test booklet for that series with the item which occurred in that position in the training booklet. Subjects were urged to guess at the missing items when they felt unsure of the correct answer, and to respond to each page of the test booklet in the order of its occurrence.

Stimulus Materials. Each page of *S*'s training booklet displayed a 42 × 30-mm black rectangular border, each corner of the rectangle so framed containing a familiar male or female name in 3-mm high black type (see Fig. 1). The booklet paper was sufficiently heavy to mask stimuli on pages below the one exposed. The 12 stimulus displays in *S*'s training series were dichotomized along two orthogonal dimensions yielding three training displays in each of four categories, 1*C-N*, 2*C-N*, 1*C-R*, and 2*C-R*. The 1*C* (1-cue) vs. 2*C* (2-cue) distinction concerns the number of stimulus components presented to *S* on test trials of that display: *S*'s recall cue for each test of a 1*C* display consisted of the rectangular frame, the upper-left stimulus item, and a dotted blank in one of the three empty corners; whereas for each test of a 2*C* display, the recall cues were the frame, the upper-left stimulus item, a second item in the corner of its training occurrence, and a dotted blank in one of the two remaining empty corners. The *N* vs. *R* (nonredundant vs. redundant) distinction concerns duplication of stimulus elements within a display: Whereas the items in the four corners of each *N*-display were all distinct, one item appeared twice in each of the *R*-displays. The upper-left corner of an *R*-display never contained the redundant item, while each of the three ways in which a redundant item could occupy two of the three remaining corners of a display occurred once in each category 1*C-R* and 2*C-R*. Although stimulus items were thus sometimes repeated within a display, two different displays in *S*'s training set never contained any common items. Training and test stimuli in each of these categories are illustrated in Fig. 1.

The training set (i.e., booklet of 12 stimulus configurations) for each *S* was constructed from the names Art, Bill, Bob, Carl, Dave, Dan, Earl, Fred, Guy, Hugh, Ian, John, Jim, Luke, Neal, Paul, Roy, Sam, Tom, Vic, Walt, Abby, Ann, Bess, Dot, Edna, Eve, Gwen, Hope, Ida, Jane, Judy, Kate, Lisa, Lucy, Mary, Nora, Peg, Rita, Ruth, Sara, Sue. (In data analysis, spelling variants of these names were scored as equivalent to their training prototypes.) To counterbalance possible specific-stimulus effects, eight different training sets were formed by different arrangements of these names. The 1*C-N* and 2*C-R* displays were constructed exclusively of male names and the 1*C-R* and 2*C-N* displays exclusively of female names for half the training sets, while the reverse was true of the remainder. The serial order (page in training booklet) of displays in each category 1*C-N*, 1*C-R*, 2*C-N*, and 2*C-R* were varied among training sets in such fashion that across all *S*s, each category occurred with the same frequency in each serial position.

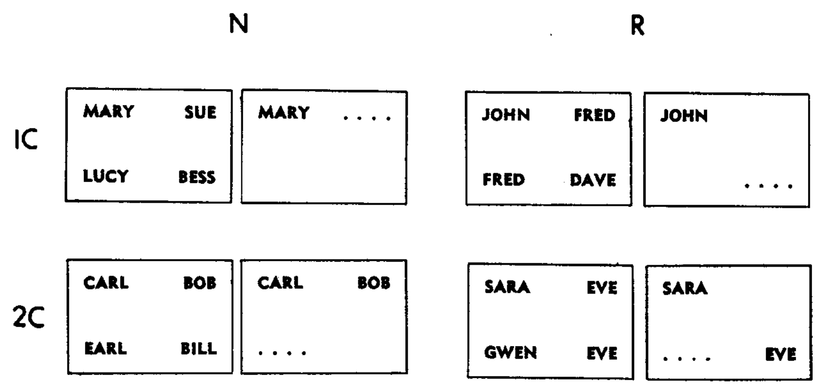

FIG. 1. Photographic reproductions of a training display and test stimulus in each of the four display categories.

Since the item in the upper-left corner of each training display was always included in the recall cues for that display, there were three components in each display whose recall could be tested. Test series were accordingly organized in blocks of three so that within each block, all display components other than the fixed cues were tested once. All six possible orders of the three test series in a block occurred with equal frequency across all *S*s, and within each test series (i.e., one test each of the 12 displays) one test of each position was made in each of the four display categories. For categories 2*C-N* and 2*C-R*, the second cue was also counterbalanced across position in each test series.

RESULTS

Although it was anticipated that *S*s would find memorizing 12 four-element displays a difficult task, the total percentages of correct responses on successive blocks of three test trials were 34%, 72%, and 87%, respectively.

TABLE 1

THE SEVEN STRUCTURALLY DISTINCT TEST TYPES[a]

Test type	Training-display category	Description of test	Example	
			Training	Test
1	1*C-N*	One cue; test of nonredundant item	A B D C	A …
2	2*C-N*	Two cues; test of nonredundant item	A B D C	A B …
3	1*C-R*	One cue; test of redundant item	A B C C	A …
4	1*C-R*	One cue; test of nonredundant item	A B C C	A …
5	2*C-R*	Two cues, one of redundant item; test of redundant item	A B C C	A … C
6	2*C-R*	Two cues, one of redundant item; test of nonredundant item	A B C C	A … C
7	2*C-R*	Two cues, neither of redundant item; test of redundant item	A B C C	A B …

[a] Within each test type, the recall blank and the second recall cue for 2*C*-displays occurred equally often in all positions except the upper left.

Male *S*s made 30% more errors than did females, a difference statistically significant far beyond the .01 level. Differences in error rates among the alternative training sets, on the other hand, were barely significant at the .05 level.

The design of this experiment yields seven structurally distinct types of recall tests, the respective natures of which are detailed in Table 1. The types of responses possible on these tests in turn are as follows.

A: Correct recall of nonredundant (A_N) or redundant (A_R) item.

B: Intrusion of nonredundant (B_N) or redundant (B_R) item from same display.

C: Intrusion of nonredundant item from *same position* in other same-sex display.

D: Intrusion of nonredundant item (other than fixed cue) from *different position* in other same-sex display.

E: Intrusion of fixed cue from other same-sex display.

F: Intrusion of redundant item from other same-sex display.

G: Repetition of fixed cue (same display).

H: Incorrect repetition of nonredundant (H_N) or redundant (H_R) variable cue (same display).

X: Intrusion from other-sex display.

Y: Outside intrusion.

Z: Omission.

An incorrect response is a "repetition error" if it duplicates a cue present on that test trial, and is an "intrusion error" otherwise. Intrusion errors are distinguished according to where *S* previously saw the reproduced item in the training series (unless it was not in the training series at all, in which case it is an "outside intrusion"). An intrusion from "same display" is an item which was in the same training display as the correct item for that test but in a different position from it; an intrusion from a "same-sex display" is an item of the same gender (i.e., male vs. female name) as the test cue; and an intrusion from the "same position" is an item which occupied the position being tested (upper right, lower right, or lower left) in another training display. By "fixed cue" is meant any item which occurred in the upper-left position of some training display, while the "variable cue" on a 2-cue test is the one additional to the fixed cue on that test. And as already explained, a "redundant" item is one which occurred in two positions in its training display.

The total number of responses obtained from all *S*s in each error category on each test type is shown in Table 2. The rapidity of *S*s' learning precluded detailed analysis of shifts in the error distribution as a function of training, but on the whole, the relative error frequencies were remarkably constant across successive test blocks. The only major exceptions are that (*a*) on all test types, error type *B* became conspicuously more prominent as training progressed; (*b*) the same was true, though more weakly and not quite so consistently, for type-*C* errors; and (*c*) a disproportionately large number of the type-*H* errors on test type 6 occurred in the first two test blocks, whereas the exact opposite was true of error type *H* on test types 2 and 7.

Before interpreting the error propensities revealed by these data, it is necessary to make two adjustments on the raw error frequencies in Table 2. The first corrects for the fact that the number of alternative responses qualifying as an instance of each error type is not a constant. On a type-1 or type-2 test, for example, there are three different ways to make a type-*C* error (i.e., three different items which so qualify) but six different ways to make a type-*D* error; hence even if an item's position in its training display does not affect its intrusion probability, type-*C* errors should occur twice as often on these tests as do type-*D* errors. To correct for this inequality, each raw frequency in Table 2 must be divided by the number of different response items which count for that entry, yielding the corresponding "per-unit" error frequency. Expressed as a proportion of total responses for its test type, the per-unit rate of error type *i* on test type *j* estimates the probability that a particular item is the response to a type-*j* test when that item qualifies as a type-*i* error. Within-test contrasts between per-item error frequencies thus reveal the relative strengths by which the structural properties defining the various error alternatives potentiate recall on tests of a given type. *Between*-test comparisons of per-unit error rates cannot be interpreted this simply, however, for the incidence of response type *i* on test type *j* is determined not only by the intrinsic strength with which type-*j* cues arouse type-*i* responses but also by competition from responses of types other than *i*. Thus the probability of response type *i* could be higher on type-*j* than on type-*k* tests either because type-*j* cues potentiate type-*i* responses more strongly than do type-*k* cues *or* because the latter predispose alternative responses more strongly than do the former. For between-test comparisons, therefore, it is most insightful to express the per-unit error rates on each test type as a multiple of the rate for some relatively neutral error type which can rationally be expected to have about the same intrinsic strength on all test types. For reasons discussed below, the average of error types *D* and *E* appears to be the best standard for such comparisons, and the result of dividing the per-unit response frequency corresponding to each entry in Table 2 by the mean of the per-unit frequencies of error types *D* and *E* for that column is given in Table 3. Comparisons among Table 3 entries then show how the various determinants of recall differ in the strengths by which they exceed the background level of "noise" intrusions.

Assessing the statistical significance of the contrasts in these data poses something of a problem. The null hypothesis of no-true-difference can be appraised for any two

TABLE 2

TOTAL NUMBER OF RESPONSES FROM ALL 384 *S*s IN EACH RESPONSE CATEGORY ON EACH TEST TYPE[a]

Error type	Test type						
	1	2	3	4	5	6	7
A	6050 (1)	6607 (1)	4459 (1)	2020 (1)	2940 (1)	2143 (1)	2361 (1)
B	580 (2)	482 (1)	186 (1)	210 (1)	69 (1)	—	—
C	986 (3)	670 (3)	656 (4)	266 (3)	106 (4)	180 (3)	230 (4)
D	702 (6)	637 (6)	450 (7)	305 (8)	85 (7)	254 (8)	178 (7)
E	511 (5)	457 (5)	286 (5)	172 (5)	81 (5)	186 (5)	140 (5)
F	666 (3)	414 (3)	306 (2)	165 (2)	55 (2)	57 (2)	132 (2)
G	137 (1)	72 (1)	117 (1)	60 (1)	10 (1)	27 (1)	42 (1)
H	—	316 (1)	—	—	—	448 (1)	133 (1)
X	84 (21)	74 (21)	59 (21)	23 (21)	16 (21)	15 (21)	18 (21)
Y	223	249	132	94	30	58	78
Z	429	390	261	141	64	88	144

[a] The number in parentheses after each entry is the number of different items which qualified as an instance of that error type on that test type. Dividing each entry by its parenthesized companion yields the corresponding "per-unit" error frequency. For computing the last two rows of Table 3, error types *Y* and *Z* are treated as though having a parenthesized divisor of unity.

TABLE 3

BASELINE EQUATED PER-UNIT ERROR STRENGTHS[a]

Error type	Test type						
	1	2	3	4	5	6	7
A_R	—	—	73.4	—	207.5	—	88.4
A_N	55.2	66.9	—	55.7	—	62.2	—
B_R	—	—	—	5.79	—	—	—
B_N	2.64	4.88	3.06	—	4.87	—	—
C	3.00	2.26	2.70	2.44	1.87	1.73	2.15
D	1.07	1.08	1.06	1.05	.86	.92	.95
E	.93	.93	.94	.94	1.14	1.08	1.04
F	2.03	1.40	2.52	2.28	1.94	.83	2.47
G	1.25	.73	1.93	1.66	.74	.78	1.57
H_R	—	—	—	—	—	12.99	—
H_N	—	3.20	—	—	—	—	4.98
X	.037	.036	.046	.030	.054	.021	.032
Y	2.04	2.52	2.17	2.59	2.12	1.68	2.92
Z	3.91	3.94	4.30	3.89	4.52	2.55	5.39

[a] Each entry equals the corresponding per-unit frequency in Table 2, divided by the mean per-unit frequency of error types *D* and *E* on that test type.

per-unit error rates on the same test type by computing the corresponding difference for each S separately and testing the mean of these differences by the t statistic. Concluding that there is a significant difference between two per-unit error rates on tests of different type, however, does not tell whether this is due to a difference in the inherent strengths of these two error types' specific sources, or whether it is because they suffer unequal degrees of competition on their respective test types. To analyze the statistical significance of between-test differences with competition inequalities partialled out, we would need to start with the equivalent of Table 3 for each S individually —the practical difficulty with which being that the raw per-S frequency of baseline errors is too low to yield meaningful individual-S error ratios. However, the magnitude of within-test comparisons in Table 3 which are statistically significant may also be taken as a rough guide to which between-test contrasts are large enough to be taken seriously. And since Table-3 ratios greater than 3:2 are nearly all significant at the .01 level for the within-test comparisons,[3] between-test comparison ratios on the order of 2:1 may be considered reasonably secure statistically.

Discussion

Two main classes of phenomena are manifest in these data, (*a*) effects involving the display positions of test blank and response item, and (*b*) effects of redundancy in the training displays.

Position Effects. Most striking of all the phenomena here observed is simply that most Ss are, in fact, able with surprising rapidity to achieve nearly errorless performance on the test trials. This is almost—though not quite—conclusive evidence that S's recall in this setting is predominantly part–whole rather than part–part in character. It will be worth taking some pains to explain why this is so, for a major reason why the severe limitations of association theory are seldom adequately appreciated by its partisans is that simple redintegrative explanations of data are often proposed in loose terms which *seem* to be merely associative.

For example, when in paired-associate research it is said that S learns *which items were associated* (i.e., paired) or that what S recalls on a test trial is an *association*, what S is thereby construed to acquire from experience is something which permits retrieval not merely of disconnected elements—which is all that associations yield—but of their previously experienced relations as well (Rozeboom, 1965, esp. pp. 349 f., 365). Thus if S has learned merely part–part associations $A \rightarrow B$ and $C \rightarrow D$ from past pairings of A with B and C with D, *joint* perception of A and B will arouse C and D without, however, giving S any clue as to how these simultaneously active ideas A, B, C, and D were previously linked. On the other hand, S can easily tell what went with what if A and C simultaneously arouse integrated complexes AB and CD, but these evoked structures AB and CD—which are also called "associations" by ordinary language—are not at all the same kind of thing as the association-theoretical $A \rightarrow B$ and $C \rightarrow D$. These two senses of the word "association" have been repeatedly confounded throughout the history of association theory.

This general point about simultaneous arousal is critical for why correct recall in the present experiment cannot be explained by hypothesizing that S learns to associate—*in the association-theoretic sense of the term*—each item with its position in the display. Let the frame cues in the various corners of a given display be designated P_1, P_2, P_3, P_4, and suppose that in this display items A, B, C, and D were in these respective positions. How might position associations enable S to reproduce C, rather than B or D, when shown a test configuration comprising the frame, item A in position P_1, and the test blank in position P_3? To begin, let us provisionally grant (*a*) that training has developed position associations $P_2 \rightarrow B$, $P_3 \rightarrow C$, and

[3] Further details of these significance tests together with a somewhat finer-grained analysis of the data are contained in an earlier draft of this article, available upon request from the author.

$P_4 \rightarrow D$ in *S* but *not* $P_3 \rightarrow B$, $P_3 \rightarrow D$, etc. (The position cues will also be associated with items on the other training displays, but for simplicity we shall assume that cue A makes these other associations irrelevant, i.e., that on the test in question, position cue P_i elicits only items which are associated both with A and P_i.) Then if the test configuration consisted just of A, P_3, and the test blank, associative arousal would yield response C as desired. But in fact, the test configuration contains position cues P_2 and P_4 as well—as becomes evident upon reflection that the only difference between a test of C and a test of B or D with fixed cue A lies in the test blank's relation to the other cues. Hence if the only determinants of response to this test were associations of form $x \rightarrow y$ where y is a single item and x is a cue or complex of cues previously experienced with y, the test stimulus would simultaneously arouse B, C, and D with no indication of where these were previously positioned. To make the blank's relation to the other cues a relevant factor, an associative explanation must appeal to selective perception and assume that *S* attends just to the position P_3 containing the blank while suppressing the sensory impact of P_2 and P_4. (Whether such selectivity could be an associative process will not be disputed here, though that is certainly problematic.) It seems unreasonable to assume that such suppression would be complete, however, so there should be at least *some* arousal of B and D through associations $P_2 \rightarrow B$ and $P_4 \rightarrow D$ even if dominated by arousal of C through $P_3 \rightarrow C$. Moreover, supposition (*a*) is not very plausible, for during training each position cue is co-experienced with every item in the display—the fact that P_3 is *spatially* closer to C than to B and D does not keep it from being fully as contiguous *temporally* with B and D as with C. To argue that association $P_3 \rightarrow C$ becomes stronger than $P_3 \rightarrow B$ and $P_3 \rightarrow D$, we must either invoke principles of association learning other than bare co-occurrence or conjecture that the pattern of *S*'s eye movements results in greater temporal contiguity between adjacent elements than between ones more spatially remote from each other.

Although position associations are not the only possibility for associative explanations here, their fundamental difficulty is common to all: At best, a strictly associative account can yield only a slightly stronger test-trial arousal of the correct response than of the other items previously experienced by *S* in that configuration, while conversely, comparative strength of arousal is *S*'s only clue to an item's correctness for the position tested. Moreover, there is no associative mechanism by which *S* can scan the various ideas evoked in him and report the most intense one with near-perfect accuracy even if "stronger arousal" in this context means (as it may not) a consistent difference in degree of activation rather than a higher probability of arousal. It must be concluded that elementwise associations alone can at best account for only a slight dominance of correct over incorrect recall in this experiment. On the other hand, if the test stimulus redintegrates the total configuration previously experienced by *S*, or at least evokes partial structures in which each recalled item is integrated with a distinctive portion of the frame, then *S* can match his recall against the test display and read out the item which, in the former, bears the same relation to the frame as does the test blank in the latter.

That configural position plays an important role in recall *somehow* is amply demonstrated in Table 3 by the data for error types *C* and *D*. On each of the seven test types, the error rate for extradisplay intrusions of nonredundant items is about twice as great or greater when the intrusion is from the position marked by the test blank as when it is from a different position. (A similar position phenomenon has previously been noted by Asch, Hay, & Diamond, 1960.) Unless association learning is a function of spatial as well as temporal contiguity, this establishes that the relation of one component x to another y in a stimulus configuration (here the relation of test blank to frame) can serve as a cue to recall of other items previously experienced by *S* as standing in that same relation to y.

Comparing the strength of same-position extradisplay intrusions to that of intradisplay intrusions brings out a curious effect of the number of test cues. Contrary to original expectation, presenting *S* with two cues on a test trial rather than one does not appreciably facilitate recall (cf. test types 1 and 4 vs. 2 and 6 for row A_N in Table 3). What little superiority there is of 2-cue over 1-cue recall could be an artifact of the additional training *S* received from exposure to the 2-cue test displays. Increasing the number of test cues does, however, alter the balance between intradisplay and extradisplay intrusions. Specifically, inspection of error types *B*, *C*, *D*, and *E* on test types 1, 2, 3, and 5 in Table 3 reveals that nonredundant-item intrusions from other positions in the same display vs. the same position in other displays vs. other positions in other displays have strength ratios of roughly 3:3:1 on 1-cue tests but about 5:2:1 on 2-cue tests. Possibly this increased dominance of intradisplay over extradisplay arousal is no more than enhanced redintegration on 2-cue tests, akin to the strongly increasing incidence of intradisplay intrusions on the later test trials. But that does not explain why this enrichment of recall appears as a significantly augmented arousal of items from positions other than the one tested without a commensurate facilitation of the correctly positioned item. (As the enhanced-integration interpretation would predict, the relative increase of intradisplay intrusions as a function of training is accompanied by an increase in the correct-response: intradisplay-intrusions ratio as well. In contrast, the increase in intradisplay intrusions effected by adding a second test cue is accompanied by a marked decrease in the correct-response:intradisplay-intrusions ratio.) It may be that detachable components of an experienced configuration tend to form elementwise associations more strongly than they do holistic traces, and that responsibility for redintegrative memory falls disproportionately upon the relational aspects of the recall stimulus. If so, increasing the content-to-structure balance of a test stimulus should have the effect here observed.

Redundancy Effects. Since the remaining analysis concerns between-test comparisons, the adjustment in Table 3 for competition inequalities deserves further explanation. This was accomplished, it will be recalled, by expressing each per-item error rate as a proportion of its test type's average per-item rate for certain baseline errors which presumably have about the same intrinsic propensity on all tests. The errors chosen to define this baseline (*a*) should be "background noise" in the sense that a test's distinctive features ought not to bias their arousal, and (*b*) should occur with sufficient frequency to insure statistical stability. Error types *X* (other-sex intrusions), *Y* (outside intrusions), and perhaps *Z* (omissions[4]) are ideal from standpoint (*a*) but are precarious on grounds (*b*). Error types *D* and *E* (extradisplay intrusions of nonredundant items from other positions) are scarcely less satisfactory regarding (*a*), however, and constituting as they do almost 30% of all errors, are optimal by (*b*). The scale values for error types *X*, *Y*, and *Z* which result from the latter choice of baseline test the validity of this adjustment, and it is gratifying to observe in Table 3 that these values are, as they should be, highly constant across all seven test types. Were a composite of error types *X*, *Y*, and *Z* rather than of *D* and *E* taken for the baseline, only between-test comparisons involving test type 6 would be altered more than trivially and even those not enough to affect their qualitative interpretations.

In what follows, the Table-3 strength of error type j ($j = A, \ldots, H$) on test type i ($i = 1, \ldots, 7$) will be denoted by the phrase "datum ij."

Scanning Table 3 for major departures from the between-test norm on each error type reveals two enormous singularities, one for test type 5 on correct recall (datum 5*A*) and the other for test type 6 on error type *H*. Both of these tests present *S* with two cues, one of which was redundant in the training display; their difference is that on tests of type 5 the position tested also contained the redundant item in training so that the correct response is the same as the variable test cue, whereas test type 6 calls for the remaining nonredundant item. It is evident from data 5*A* and 6*H* that when confronted with a previously redundant item on a 2-cue test, *S* strongly tends to repeat the redundant item whether it is correct for the position tested or

[4] The present calculations may be viewed as an application of the Bradley-Terry-Luce model for decomposing response probabilities into a measure of the underlying strengths of competing sources (Luce, 1959). In terms of this model, the probability of omissions can be construed to reflect a response threshold which behaves mathematically like any other response source.

not. This phenomenon is by far the most spectacular of all the error effects in the present data. Why it occurs, moreover, is illuminated by some of the less robust contrasts in Table 3.

It is first of all clear from comparison of datum $6H_R$ with data $2H_N$ and $7H_N$ that the effect just noted is genuinely an effect of past redundancy, not just a propensity for S to repeat the variable cue regardless of its specific history for him. The subject somehow recognizes which test cues were redundant in his past experience and is thereby incited to make those cues redundant in his recall reproduction as well. Moreover, if the observed superiority of datum $7H_N$ over datum $2H_N$ is statistically genuine, a second factor operative here is that repetition of a test cue is more strongly disposed by tests of redundant displays than by tests of nonredundant ones even when the redundant item is not included in the test stimulus. The subject apparently tends to recall whether test cues not themselves redundant were part of a display containing redundancies, while such recall of co-experienced redundancy induces S to put redundancy into his test reproduction even when this is done incorrectly. The between-test contrasts for error type G (repetition of fixed cue), though statistically precarious, also support this interpretation; for among the five test types with no redundant material in the test stimulus, the three which test redundant displays (types 3, 4, and 7) all show stronger fixed-cue repetition than do the two testing non-redundant displays (types 1 and 2). The low type-G errors for test types 5 and 6 may then also be explained by the not-implausible interaction conjecture that S's propensity to repeat previously redundant test cues also acts as prepotent channel for whatever additional unfocused tendency S may have to reproduce redundancy on a test trial.

Finally, the contrast between data $3B_N$ and $4B_R$ shows that test cues arouse a previously co-experienced item more strongly when that item was redundant in the original display than when it was not—which would also explain why recall is more accurate on tests of redundant items than on tests of nonredundant items even when the tested redundancy is not included in the test stimulus (cf. data $3A_R$ and $7A_R$ vs. $1A_N$, $2A_N$, $4A_N$, and $6A_N$). This is specifically a redintegrative or associative phenomenon rather than an uncued "response availability" difference due to the redundant items' higher training frequencies, for comparison of error type F with error types C and D in Table 3 exhibits little if any tendency for extradisplay intrusions to favor redundant over nonredundant items, nor are type-F errors any more prominent on the first block of test trials than on the later ones. To explain this difference on associative grounds, it might be argued that a training display in which A is the fixed cue, B is single, and C is redundant pairs A twice as often with C as with B and should hence develop correspondingly greater strength of $A \rightarrow C$ than of $A \rightarrow B$. That this effect should also occur in redintegrative recall is intuitively evident, but the explicit theory behind that intuition is complex and will not be discussed here.

To summarize the observed redundancy effects, then: (*a*) S's recall behavior sharply discriminates test cues which were redundant in prior experience from those which were not. (*b*) Test cues previously co-experienced with redundancies are similarly though more weakly distinguished, even when the specific redundancy is not itself recalled, from cues which appeared only in nonredundant displays. (*c*) The redundant components of a prior experience are more strongly evoked by another element of that experience than are its nonredundant components.

OVERVIEW

It is clear from these data not only that the compositional structure of past and present experience plays a significant role in recall, but also that there exists an abundance of structural phenomena awaiting identification and analysis. It would be senseless any longer to shun redintegrative models of memory, for traditional part–part associative concepts alone simply lack the formal potency needed to account adequately for such phenomena. But it is also important to emphasize that the present findings give little comfort to a simplistic holism which views recall as an entirely seamless reinstatement of past experience. For all the effects noted above, even to some extent correct recall, reflect part-processes of one sort or another; not merely partial recall of content detached from structure, but also components of structure functioning as cues or reproduced in recall independent of specific content. The significant task ahead is not to debate which interpretation of memory—associative vs. redintegrative—is the correct one, but to determine what mechanisms at various positions along a part–part/part–whole continuum contribute to learning and recall, how they interact to produce a composite outcome, and what parameters govern the respective strengths of their contributions on particular occasions.

References

ASCH, S. E. A reformulation of the problem of associations. *American Psychologist*, 1969, **24**, 92–102.

ASCH, S. E., HAY, J., & DIAMOND, R. M. Perceptual organization in serial rote-learning. *American Journal of Psychology*, 1960, **73**, 177–198.

DEESE, J. *The structure of associations in language and thought*. Baltimore: Johns Hopkins University Press, 1965.

DE GROOT, A. D. *Thought and choice in chess*. The Hague: Mouton, 1965.

GARNER, W. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.

HOROWITZ, L. M., & PRYTULAK, L. S. Redintegrative memory. *Psychological Review*, in press.

HUMPHREY, G. *Thinking*. London: Methuen, 1951.

LUCE, R. D. *Individual choice behavior: A theoretical analysis*. New York: Wiley, 1959.

MANDLER, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*, Vol. 1. New York: Academic Press, 1967.

MANDLER, J. M., & MANDLER, G. *Thinking: From association to gestalt*. New York: Wiley, 1964.

PEAK, H. Psychological structure and psychological activity. *Psychological Review*, 1958, **65**, 325–347.

POLLIO, H. R. Associative structure and verbal behavior. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

ROZEBOOM, W. W. Do stimuli elicit behavior?—a study in the logical foundations of behavioristics. *Philosophy of Science*, 1960, **27**, 159–170.

ROZEBOOM, W. W. Formal analysis and the language of behavior theory. In H. Feigl & G. Maxwell (Eds.), *Current issues in the philosophy of science*. New York: Holt, 1961.

ROZEBOOM, W. W. The concept of "memory." *Psychological Record*, 1965, **15**, 329–368.

ROZEBOOM, W. W. Conditioned generalization, cognitive set, and the structure of human learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 491–500.

WERTHEIMER, M. On discrimination experiments: I. Two logical structures. *Psychological Review*, 1959, **66**, 252–266.

(Received March 31, 1969)