Journal of Experimental Psychology: General
1986, Vol. 115, No. 1, 26-38

0096-3445/86/$00.75

Elaboration, Organization, and the Self-Reference Effect in Memory

Stanley B. Klein
Harvard University

John F. Kihlstrom
University of Wisconsin

Relating information to the self (self-referent encoding) has been shown to produce better recall than purely semantic encoding. This finding has been interpreted as demonstrating that self-reference produces a more elaborate memory trace than semantic encoding, and it has been cited frequently as evidence that the self is one of the most highly elaborated structures in memory. The experiments reported in this article challenge this interpretation of the self-reference effect by demonstrating that self-referent and semantic encodings produce virtually identical free recall levels if they are first equated for the amount of organization they encourage. On the basis of our findings we conclude the following: (a) Organization, not elaboration, is responsible for the superior recall performance obtained when information is encoded self-referentially, and (b) organization is not a necessary component of self-referent encoding and can be orthogonally varied within self-referent and semantic encoding tasks. Finally, we discuss how a single-factor theory based on organization can account for many of the self-referent recall findings reported in the literature.

One of the most influential approaches to the study of human memory has been the Depth of Processing (DOP) framework proposed by Craik and Lockhart (1972). These investigators suggested that retention of a memory trace is determined by the nature of encoding operations carried out on stimulus material. Deep, meaningful analyses such as those prompted by semantic encoding tasks allow formation of a more durable trace than do shallow, structural analyses of the sound or appearance of stimuli. Until 1977, semantic encoding was commonly considered the optimal way of achieving good retention (e.g., Craik, 1973; Craik, 1977; Craik & Tulving, 1975; Hyde & Jenkins, 1973; Moscovitch & Craik, 1976). However, in an extension of the DOP framework, Rogers, Kuiper, and Kriker (1977) demonstrated that judging stimulus material for its personal descriptiveness (self-referent encoding) produced even higher levels of recall than semantic encoding. This Self-Reference Effect (SRE) has been replicated in a number of studies (Bellezza, 1984; Bower & Gilligan, 1979; Friedman & Pullyblank, 1982; Ganellen & Carver, 1985; Halpin, Puff, Mason, & Martson, 1984; Kendzierski, 1980; Klein & Kihlstrom, 1984; Kuiper & Rogers, 1979; Lord, 1980; McCaul & Maki, 1984; Mross & Kihlstrom, 1985; Rogers et al., 1977; Warren, Chattin, Thompson, & Tomsky, 1983). Its nature is not yet completely understood, however.

Several cognitive models have been advanced to explain the SRE (see Greenwald & Pratkanis, 1984; Kihlstrom, 1981; Kihlstrom & Cantor, 1984; Rogers, 1981, for recent reviews). The most popular of these, the elaboration model, is based on Craik and Tulving's (1975) proposal that the DOP results can be understood in terms of the idea that retention is an increasing function of the amount of elaboration that stimulus material receives during encoding: Encoding tasks that are said to promote deeper processing of the stimulus material can be viewed as encouraging the subject to engage in greater trace elaboration (see also J. R. Anderson, 1976, 1983b; J. R. Anderson & Reder, 1979). The clear recall superiority of self-referent over structural and semantic encoding tasks has been interpreted as indicating that self-referent encoding is unique among these tasks in its potential for creating a richly elaborated memory trace (Keenan & Baillet, 1980; Kendzierski, 1980; Rogers et al., 1977). Accordingly, a number of SRE investigators have suggested that self-referent trace elaboration can be explained by thinking of the self as a highly elaborate memory structure that, when activated, can form many links between the stimulus and preexisting information about the self in memory (e.g., Ingram, Smith, & Brehm, 1983; Keenan & Baillet, 1980; Markus & Smith, 1981; Rogers et al., 1977).

An alternative explanation of the SRE is that the evaluative aspect, rather than the elaborative potential, of the self-reference task leads to better retention (e.g., Ferguson, Rule, & Carlson, 1983). In the typical SRE study, evaluation is confounded with self-reference: Given the tendency of most subjects to perceive themselves in socially desirable terms, self-referent judgments tend to have an evaluative connotation (e.g., "Does this word describe you?"), whereas semantic encoding tasks are typically nonevaluative (e.g., "Does the word fit in the following sentence?"). Ferguson et al. (1983) found that a semantic task that required evaluative judgments ("Is this word a desirable characteristic?") enhanced memory to the same degree as a self-referent task, suggesting that no special mnemonic properties need be attributed to self-referent encoding to account for the results obtained from the SRE paradigm. Other data, however, indicate that evaluation does not invariably lead to recall levels equaling

This article is based in part on a dissertation submitted to Harvard University by the first author and partially supported by Grant MH-35856 from the National Institute of Mental Health, United States Public Health Service to the second author.

We would like to thank Francis Bellezza, Roger Brown, David Buss, David Funder, Reid Hastie, Carolyn Phinny, Michael Watkins, and especially Judith Loftus and Miriam Schustack for their helpful comments on earlier versions of this article. We gratefully acknowledge Martin Murphy for his valuable suggestions regarding the clustering index developed in this paper.

Correspondence concerning this article should be addressed to Stanley B. Klein, who is now at the Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820.

those obtained through self-reference (McCaul & Maki, 1984; Rogers, 1981). To date, then, there has been no unifying theoretical account of the SRE, and some theorists (Bellezza, 1984; Greenwald & Pratkanis, 1984) question whether self-referent recall superiority can be explained by any single-factor theory.

We take issue with this conclusion. We propose that the effects of self-reference on free recall in the DOP paradigm can be explained by a single principle: the well-documented finding that organization of stimulus material improves recall (e.g., Bower, Clark, Lesgold, & Winzenz, 1969; G. Mandler, 1967; Puff, 1970; Underwood, 1964). Our hypothesis rests on the observation that two factors, self-referent encoding and categorical organization, have been confounded in every study that has demonstrated self-referent recall superiority using the DOP framework. To date, self-referent encoding tasks, unlike semantic and structural encoding tasks, have been so designed that their successful completion required organization of the stimulus material into two categories: words that do and do not describe the subject. The purpose of the present experiments was to examine the hypothesis that organization, not elaboration or evaluation, is responsible for the high levels of recall found by relating words to the self.

Organization and Elaboration

We will evaluate categorical organization and elaborative encoding as explanations for the SRE. Elaboration refers to encoding operations performed on a single word, whereas organization involves associating the list words together. Specifically, elaborating a word transforms it *independently* of other list words by forming multiple connections between it and related, extralist material in memory (e.g., Bellezza, Cheesman, & Reddy, 1977; Ellis & Hunt, 1983; Ellis, Thomas, & Rodriguez, 1984). The word *doctor,* for example, may prompt the subject to encode multiple propositions such as "Doctors work in hospitals" and "Doctors wear white coats." Relating a word to extralist material in memory increases its recall probability by providing multiple routes for subsequent retrieval (J. R. Anderson, 1983a, 1983b, 1985).

In contrast, categorical organization, defined by G. Mandler (1977) as the process of grouping words that "go together" according to some semantic criterion, results in the encoding of relations *between* list words: direct relations based on word-to-word associations, and mediated relations that emerge from the fact that the words share a common category (e.g., Allen, Puff, & Weist, 1968; Crowder, 1976; Wallace, 1970). Organizing a list of words into categories should augment recall in two ways: (a) It encourages the encoding of interitem associations that typically exist among members of the same category (e.g., Cofer, 1966, 1967), thereby establishing associative paths in memory that can be used during retrieval (e.g., Hastie, 1980; Srull, 1981, 1983), and (b) associations between each of the members of a category and the category label itself are encoded (e.g., Bower et al., 1969; Underwood, 1964), enabling the category label to act as a cue for category members during recall (e.g., Barsalou, 1983; Bower, 1970; Tulving & Pearlstone, 1966).

Depth of Processing and Organization

To appreciate how organization might be confounded with self-referent encoding tasks, it will help first to examine the way in which organizational factors are conceptualized in the DOP framework. Most studies conducted within this paradigm minimize or ignore the influence of organization in promoting good recall. This lack of attention is a direct consequence of the methodological procedures employed. The typical DOP experiment (e.g., Craik & Tulving, 1975, Experiment 1) involves an incidental learning situation in which the subject is presented with a list of words, one at a time, and is asked to consider each one only in relation to its specific orienting question. Care is taken to make sure that the list does not have any obvious taxonomic structure, and subjects are not encouraged to search for relations between the words. As Battig and Bellezza (1979) note, these conditions inhibit the discovery or use of organizational strategies during acquisition. It is not surprising, therefore, that the DOP paradigm emphasizes semantic elaboration as a major determinant of recall.

A small set of DOP studies, however, indicates that organization can have a powerful effect on recall even when the list lacks an obvious taxonomic structure. Bellezza and his associates (Bellezza et al., 1977; Bellezza, Richards, & Geiselman, 1976) have demonstrated that tasks that involve associating the list words together in some way (organizational processing) lead to better free recall than semantic encoding tasks that encourage subjects to elaborate each word independently of other list words (individual-item processing). A series of studies by Einstein and Hunt (1980; Hunt & Einstein, 1981) reported similar findings and showed that organization was not simply a variant of semantic encoding but was functionally distinct.

Einstein and Hunt's experimental design (1980, Experiment 2) has important implications for our arguments concerning self-reference and organization. As part of a larger study, half of their subjects were shown a list of words (familiar nouns) whose relations were obscure, but that could be organized into several ad hoc categories (e.g., things that make noise) if subjects were made aware of the appropriate category labels. Without this provision, subjects reported that they perceived the list as a collection of unrelated nouns. Subjects were assigned either to a semantic encoding condition that encouraged individual-item processing or to an organization condition. Semantic encoding of the list consisted of rating each noun for pleasantness on a 5-point scale; in the organization condition subjects were presented with category labels and asked to place each noun in its appropriate category. Although pleasantness ratings have been shown to produce better retention than a variety of other semantic encoding tasks (Packman & Battig, 1978; Postman & Kruesi, 1977), Einstein and Hunt found that organization was associated with significantly higher recall than pleasantness ratings. Further analysis of the recall protocols indicated significantly greater clustering in the organization condition than in the semantic condition, suggesting that categorizing words encourages the encoding of relations between them to a greater degree than does rating their pleasantness.

A recent study (cited in the work of J. M. Mandler, 1984) demonstrates that the same pattern of results can be obtained even when the categories in the list are easily identifiable (e.g., Animals). Subjects performing a semantic encoding task that encouraged individual-item processing recalled significantly fewer words than subjects who performed a task that drew attention to the list's categorical structure. An important implication of

these two studies is that the encoding task must direct attention to relations between list words, and an effort must be made to encode them if the effects of organization are to be observed in recall performance (see also Bower et al., 1969; Hudson, 1968; G. Mandler, 1979; J. M. Mandler, 1979; Srull & Brand, 1983; Wood & Underwood, 1967).

Self-Reference and Organization

When the DOP paradigm requires that words be encoded structurally (e.g., "Big letters?"), phonemically (e.g., "Rhymes with *happy*?") or semantically (e.g., "Fits in the following sentence?"), each word is judged individually in reference to its particular orienting question. Under these conditions the usual DOP rank ordering of mnemonic effectiveness emerges: Semantic encoding is superior to nonsemantic encoding. In the self-reference task, however, the words are not treated as unrelated items. The words presented (often a list of trait adjectives) have the potential to be organized into categories. Furthermore, their implicit categorical structure is made salient because the self-referent orienting question ("Describes you?") is functionally equivalent to providing the category labels "words that describe me" and "words that do not describe me." As Einstein and Hunt (1980) demonstrated, when potential relations exist among list words, and the subjects are explicitly alerted to them by the appropriate orienting questions, recall is better than when the list is treated as a set of unrelated words.

The idea that self-reference improves recall by promoting the organization of stimulus material has also been suggested by Greenwald (1981) and Hamilton (1981). We take issue, however, with their suggestion that self-reference automatically produces organization. If this were so, then our proposal would modify the cognitive explanation of the SRE without challenging the claim that self-referent encoding per se has special mnemonic properties: Instead of producing the most richly elaborated memory traces, it would promote the best organization of list words. However, it is possible to show that organization is separable from self-reference and can be manipulated independently. When self-referent and semantic encoding tasks are equated for the amount of organization they encourage, words judged for self-reference should not be better recalled than words encoded semantically. To test this hypothesis, the present experiments orthogonally varied both type of encoding task (semantic and self) and amount of organization (organized and unorganized).

The organization hypothesis makes three predictions that can be tested using this experimental design:

1. The strongest statement of the hypothesis that organization, not type of encoding, is responsible for the SRE, is that organization will affect recall independently of encoding task, and that the encoding task will have no effect, either by itself or in interaction with organization.

2. Consistent with the typical SRE findings, self-referent encoding will lead to better recall than semantic encoding when self-reference encourages organization but semantic judgments do not.

3. However, when semantic encoding encourages organization, and self-referent encoding minimizes it, there will be a reversal of the usual SRE, such that semantic encoding results in better recall than self-reference.

Taken together, these three predictions imply that self-referent encodings produce no more elaboration or organization than do semantic encodings.

The alternative hypothesis, that the SRE occurs because self-reference produces highly elaborate encodings, predicts that self-referent recall will be superior to semantic recall when organization is controlled. It is possible that when both self-referent and semantic encoding tasks encourage organization, the organizational effects may be powerful enough to mask any effects of self-referent elaboration on recall. However, since elaboration does not depend on the discovery or encoding of relations between list words (Battig & Bellezza, 1979), the effects of self-referent elaboration should certainly emerge when both encoding tasks minimize list organization.

Experiment 1

This experiment had two purposes: We wanted to replicate the original Rogers et al. (1977) findings by comparing self-referent, semantic, and structural encoding tasks, and we wanted to examine clustering differences between these encoding tasks during free recall. Our hypothesis that self-referent tasks encourage organization suggests that words so encoded should be output in clusters. Furthermore, because semantic and structural encoding tasks encourage individual-item processing, significantly less clustering should be found in these conditions.

Method

Subjects. Twenty-two undergraduates from Harvard University served as paid subjects. They were tested individually in sessions lasting approximately 30 min.

Materials. The target words were 48 trait adjectives from N. H. Anderson's (1968) norms, chosen to represent a broad range of personality characteristics. Six additional trait adjectives served as buffer items and were not included in any of the recall analyses. Three were placed at the beginning of the list and three at the end.

Two supplementary lists of 54 trait adjectives were constructed using Roget's *Thesaurus* to serve as synonyms and nonsynonyms for target words in the semantic encoding task (described below).

Design. Subjects performed three types of tasks during the encoding phase. (a) For the structural task, subjects decided whether the target word was presented in small or capital letters ("Is the word printed in capital letters?"); (b) in the semantic task, subjects decided whether the target word was synonymous with a paired trait adjective ("Does the word mean the same as XXXX?"); (c) in the self-referent task, subjects judged each target word for its personal descriptiveness ("Does this word describe you?").

The experimental design was a 3×2 factorial with repeated measures on both factors—encoding task (structural, semantic, and self-referent) and response (*yes* and *no*).

All subjects saw the same list of 48 target words. The assignment of words to the three encoding tasks (16 words per task) and the order of presentation of encoding tasks were randomized across subjects. To ensure equal numbers of *yes* and *no* responses during the structural and semantic tasks, half the target words were randomly paired with orienting questions designed to elicit a *yes* response and half with questions eliciting a *no* response. In the self-reference task, however, *yes* and *no* responses could not be counterbalanced, because subjects' self-opinions determined which trait words they endorsed.

Procedure. An incidental learning situation was established by telling subjects that the study was investigating the speed and accuracy with which they could make decisions about the stimulus material.

All stimuli were presented on a Radio Shack TRS 80 Model I computer, which also recorded the subject's responses and response latencies. Each orienting question was presented for 4 s, followed immediately by presentation of the target word. A timer started with the appearance of the target word and stopped when the subject responded by pushing either a "Yes" or "No" key located on opposite sides of the keyboard. After a 1-s blank interval the next trial began. Subjects were instructed to rest their index fingers lightly on the keys while waiting to respond. The pairing of *yes* and *no* responses with the right and left keys was balanced across subjects.

Upon completion of the rating tasks a surprise free recall test was administered. Subjects were read 1 min of instructions describing the recall procedure and then were given 5 min to write, in any order, as many of the target words as they could remember. So that order of recall could be examined, subjects were asked to write only one target word per line.

Results and Discussion

Recall performance. Words receiving *yes* ratings are often better recalled than words receiving *no* ratings (Craik & Tulving, 1975; Schulman, 1974). Because *yes* and *no* responses in the self-reference encoding condition were not under experimental control, it is possible that the *yes/no* response distribution may have influenced recall. Therefore, subjects' *yes* and *no* recall scores were changed to adjusted proportions: The number of *yes*-rated words recalled for a given encoding task was divided by the total number of *yes* responses made while doing that task to form a score representing the proportion of words recalled that the subject rated *yes*, and the *no* responses were adjusted in the same manner. Although the adjusted recall score frequently is employed as a measure of retention in the SRE paradigm (e.g., Lord, 1980; Maki & McCaul, 1985; Rogers et al., 1977) interpretive difficulties arise because different absolute recall levels can yield identical adjusted scores. A two-way analysis of variance (ANOVA) on the adjusted recall scores showed no main effect for *yes/no* response, nor did the interaction reach significance (similar findings are reported in the work of Rogers et al., 1977). Therefore, recall scores were collapsed over *yes/no* responses to yield overall proportions whose means are presented in the top line of Table 1. An ANOVA on these proportions revealed a significant main effect for encoding task, $F(2, 42) = 29.31$, $p < .0001$. Subsequent tests (Tukey) confirmed the pattern of recall obtained in previous SRE studies: Self-reference led to significantly better recall than semantic encoding ($p < .01$), which, in turn, produced reliably better recall than structural encoding ($p < .01$).

Response latency. Although a large body of research suggests that retention is directly proportional to the time spent processing the information (for a review, see Cooper & Pantle, 1967), some studies indicate that this relation may not hold in the DOP paradigm (Craik & Tulving, 1975; Eysenck & Eysenck, 1979; Gardiner, 1974). Nonetheless, it would be reassuring if we could demonstrate that differences in recall were not simply the result of differences in the amount of processing time devoted to the encoding tasks. Response latency—the time between the onset of the target word and a subject's response—is a measure of the time spent processing each word (e.g., Craik, 1973; Craik & Tulving, 1975). The middle item of Table 1 shows the mean response latencies for each encoding task. A one-way ANOVA showed a significant effect of encoding task, $F(2, 42) = 31.92$, $p < .0001$; Tukey comparisons revealed that although structural decisions took significantly less time than semantic and self-referent decisions ($p < .01$), the latter two conditions did not differ reliably. Before dismissing a processing-time explanation, however, we computed mean response latency as a function of whether a word was recalled. These means are shown in the bottom lines of Table 1 (mean response latencies were not computed for the structural task because half of the subjects recalled no words in this condition). A 2 × 2 ANOVA with the factors encoding task (self-reference and semantic) and recall status (recalled words and nonrecalled words) revealed no significant effects. Taken together, the data in the middle and bottom lines of Table 1 offer no support for a processing time explanation of the recall results.

Table 1
Proportion of Words Recalled and Mean Response Latency (RL) As a Function of Encoding Task (Experiment 1)

Recall and response latency	Encoding task		
	Structural	Semantic	Self
Unconditional recall	.05	.13	.27
Overall Mean RL (in ms)	917	1,564	1,355
Mean RL (in ms)			
Recalled words		1,404	1,299
Nonrecalled words		1,577	1,365

Clustering analysis. An analysis was performed to determine whether words encoded self-referentially exhibited greater clustering in recall than words encoded in the other two conditions. In the only study to examine clustering in the SRE paradigm, Kuiper & Rogers (1979) used an index that, although providing a measure of total clustering, did not measure clustering for individual encoding tasks represented within the recall protocol. We therefore developed an index that allowed us to compare clustering differences among encoding tasks performed by each subject. The formula for our measure is as follows:

$$\frac{R_i - E(R_i)}{\text{Max } R_i - E(R_i)},$$

where R_i = the number of observed repetitions from encoding task i (i.e., the number of times a word from Task i follows a word from the same task); Max R_i = the maximum possible number of repetitions from encoding Task i; and $E(R_i)$ = the expected number of chance repetitions from encoding Task i, and is computed using Hunt's (1971) corrections for Bousfield and Bousfield's (1966) equation for expected number of chance repetitions.

This index, similar to the adjusted ratio of clustering proposed by Roenker, Thompson, and Brown (1971), sets chance clustering at zero and perfect clustering at unity. Although designed to provide a measure of clustering that is independent of amount recalled, it cannot be applied to an encoding task unless at least two words from that task are recalled (i.e., the minimum number of words necessary for clustering to take place). Therefore, the structural task, which averaged less than one word recalled, was

excluded from the clustering analysis; and Yates' procedure for replacing missing data (see Cochran & Cox, 1957) was used to estimate clustering scores in the 8 cases (out of 44) in which semantic or self-reference recall performance was less than two. Clustering analysis supported the hypothesis: Self-referent encoding showed significantly greater clustering than semantic encoding, $.32 > -.09$, $t(21) = 3.18$, $p < .01$.

In sum, there is evidence to suggest that self-reference tasks encourage organization to a greater degree than do semantic tasks, leaving open the possibility that organization, not encoding task, is responsible for self-referent recall enhancement. The next experiment examined this hypothesis directly.

Experiment 2

In Experiment 1, encoding task was confounded with organization. To permit organization and encoding task to be varied orthogonally, it was necessary to find a list of words that could be subdivided into easily identifiable categories under both semantic and self-referent encoding conditions. McGuire and Padawer-Singer (1976) report that in response to the open-ended question "Tell us about yourself," physical characteristics are mentioned almost as frequently as personality characteristics. This finding is consistent with the claim that body parts constitute an important part of a person's self-representation (Keenan & Baillet, 1980) and suggests that a task requiring self-referent encoding of a list of body part words should produce recall enhancement similar to that found for trait terms. Unlike trait adjectives, which are not easily grouped into well-defined categories by semantic encoding tasks, words for body parts can be organized into stable and identifiable subsets by both self-referent and semantic tasks.

Method

Subjects. Sixty-four Harvard undergraduates participated for pay. They were tested individually in sessions lasting approximately 30 min.

Materials. All subjects received the same list of target words, 28 nouns chosen from the Battig and Montague (1969) norms from the category "A Part of the Human Body." Half of the words were external body parts (e.g., hand, nose), half were internal body parts (e.g., heart, stomach), and all were capable of being used in trait-descriptive idioms (e.g., "to give a hand"; "to turn one's stomach"). Six additional body part words satisfying the same criteria served as primacy and recency buffers and were not included in any statistical analyses.

Design. The experimental design was a $2 \times 2 \times 2 \times 2$ mixed factorial, with two between-subjects factors, encoding task (self and semantic) and organization (organized and unorganized), and two within-subjects factors, trial and response (*yes* and *no*). Sixteen subjects were randomly assigned to each of the four experimental conditions representing the factorial combination of the two types of encoding tasks with the two levels of organization.

In the organized conditions, organization was encouraged by having target words all share a common orienting question that focused attention on the fact that the words could be divided into two categories. In the unorganized conditions each target word was associated with its own orienting question, unrelated to any other orienting question in the set, making it likely that subjects would treat "each orienting question–target word" pair as an independent unit and not actively organize the target words into categories. Examples of the orienting questions used in each condition are presented in Table 2.

Table 2
Examples of Stimulus Materials Used in Experiment 2

Condition	Orienting question	Target word
Semantic/ unorganized	Does this word fit in the sentence? "The young woman had very fair _____ "	Skin
Self/unorganized	Does this describe you? "I would stick my _____ out for a friend"	Neck
Semantic/ organized	Is this an external body part?	Heart
Self/organized	Can you think of an incident in which you had an injury or an illness associated with your?	Leg

In the semantic/organized condition, the orienting question asked subjects to decide whether the target word was an external body part. Half the target words elicited a *yes* response and half a *no* response. This task required subjects to divide the list into two categories (external body parts and internal body parts) but demanded only a nonevaluative, semantic analysis of the stimulus material.

Subjects in the self/organized condition saw an orienting question that asked them to decide if they had ever had an injury or an illness associated with the presented target word. This required a self-referent decision and also sorted the list into two categories: body parts associated with a personal illness or injury and body parts not so associated. Although this task requires the subject to search autobiographical memory for an incident similar to the one described by each orienting question rather than judge each question for its personal descriptiveness, strong self-reference effects have been obtained under these conditions using both nouns (Bower & Gilligan, 1979, Experiment 1; Warren et al., 1983) and traits (Bellezza, 1984, Experiment 1; Bower & Gilligan, 1979, Experiment 2) as target words.

In both organized conditions the order in which the target words were presented was randomized across subjects.

In the semantic/unorganized condition, the orienting questions were sentence frames each missing a single word (the target word). After reading each sentence frame subjects were presented with the target word and asked to determine whether it fit meaningfully into the sentence. Two different sentence frames were constructed for each of the 28 target words and were designed so that the associated target word fit meaningfully into one of them but not the other (e.g., "The soldier preferred to keep his _____ short: Hair?"; "The _____ got rusty after being left out in the rain: Hair?"). To ensure equal numbers of *yes* and *no* responses, half of the target words were randomly paired with frames that elicited a *yes* response and half with frames eliciting a *no* response. Order of presentation of the frames was randomized across subjects.

Finally, subjects in the self/unorganized condition read self-descriptive sentence frames each missing the target word and after viewing the target word decided whether the completed sentence described them. All the sentences began with the word *I* and used body part target words in an idiomatic trait-descriptive sense (e.g., "I always try to keep a civil _____: Tongue?"). The order of presentation of the frames was randomized across subjects, but *yes* and *no* responses could not be balanced because the subjects' personal opinions and recollections determined which stimulus items they endorsed.

Procedure. All aspects of the procedure were identical to Experiment 1 except that subjects received a second recall trial one minute after their recall protocols from the first trial were collected. Examination of output consistency across recall trials permitted measurement of subjective organization.

Results and Discussion

Recall performance. A four-way ANOVA on the adjusted recall scores showed no main effect for response, $F(1, 60) = .67$, $p > .40$ (the mean proportion of words recalled that received a *yes* rating was .69 and that of words rated *no* was .68); nor did response interact with any of the other variables. Therefore, recall scores were collapsed over *yes/no* responses, and the mean overall proportions are shown in Table 3. A 2 × 2 × 2 (Encoding Task × Organization × Trial) mixed ANOVA on these recall scores revealed a highly significant main effect of organization, $F(1, 60) = 136.21$, $p < .0001$, indicating superior recall when organization is encouraged, and a significant main effect of trial, $F(1, 60) = 57.98$, $p < .0001$, due to an 8% increase in recall on Trial 2. When organization was controlled, recall of self-referentially encoded words ($M = .64$ and .71 for Trials 1 and 2, respectively) was virtually the same as recall of semantically encoded words ($M = .64$ and .73 for Trials 1 and 2, respectively), $F < 1$ for encoding task.[1] None of the interactions reached significance.

These analyses cast serious doubt on the argument that self-reference produces a more elaborate trace than semantic encoding of the same stimuli. Even if we entertain the hypothesis that organization is powerful enough to mask any contribution of elaboration, the effect of self-referent elaboration on recall should be observed when organization is minimized. No such effect emerged; self/unorganized and semantic/unorganized showed almost identical recall performance on Trial 1 and on Trial 2. The amount of organization, not the type of encoding task, appears to be the primary determinant of recall.

It is important to note that the failure to find a significant main effect of self-referent encoding is not subject to some of the usual criticisms concerning null findings (e.g., Greenwald, 1975). In fact, a strong replication of the SRE is evident in the top half of Table 3 (recall on Trial 1). Self-referent encoding is clearly superior to semantic encoding when the self is given an organizational advantage, .78 > .49, $t(30) = 7.51$, $p < .0001$. However, when the organizational advantage is shifted to semantic encoding, there is a complete reversal of the SRE: Semantic encoding results in much better recall than self-referent encoding, .79 > .51, $t(30) = 8.23$, $p < .0001$. Identical patterns of significant differences are found in the recall performance on Trial 2.

Finally, an explanation based solely on the amount of evaluation induced by the different conditions cannot account for these recall findings. Certainly the self/unorganized task (e.g., "Does this describe you: 'I think it is wise to get things off your _____': Chest?") required a greater degree of evaluation than the semantic/unorganized task (e.g., "Does this word fit in the sentence: 'The bloodhound's _____ was very sensitive': *nose?*"), yet recall in these two conditions did not differ.

Organization. An analysis of organization was performed to evaluate the claim that the two conditions we have labeled "organized" encourage more organization than the unorganized conditions. Most of the techniques for measuring organization in free recall provide an index reflecting the degree to which a subject's clustering of stimulus material conforms to a set of experimenter-established categories. Unfortunately, we were unable to specify categories for all conditions (e.g., it is not clear what, if any, organizational strategies subjects in the semantic/unorganized condition might use). We chose, therefore, the bi-directional paired frequency (PF) measure of subjective organization proposed by Sternberg and Tulving (1977), which permits an analysis of subject-imposed organization with material lacking preselected categories by providing a measure of output consistency across successive recall trials.

Table 3
Proportion of Words Recalled, Mean PF Scores, and Mean Response Latency as a Function of Encoding Task and Level of Organization, Experiment 2

Encoding task	Level of organization: Unorganized	Level of organization: Organized	*M*
	Recall Trial 1		
Semantic	.49	.79	.64
PF	.76	3.35	
Self	.51	.78	.64
PF	.84	4.36	
M	.50	.78	
	Recall Trial 2		
Semantic	.60	.85	.73
Self	.59	.84	.71
M	.59	.85	
	Mean response latency (ms)		
Semantic	1,062	1,531	1,269
Self	1,175	2,001	1,588
M	1,118	1,766	

Note. PF = paired frequency.

The mean values for the PF measure for the four Encoding × Organization conditions are presented in Table 3. An inspection of these means reveals that subjects in the two organized conditions showed much greater subjective organization than subjects in the two unorganized conditions. This observation was confirmed by a 2 × 2 analysis of variance (Encoding Task × Organization) on the PF scores. The only effect to reach significance was a main effect of Organization, $F(1, 60) = 56.99$, $p < .0001$, offering strong support for our assumptions about the relative

[1] Because *yes* and *no* responses could not be balanced in the self/organized condition, the size of the categories they defined varied between subjects. The possibility existed, therefore, that the uneveness of the *yes/no* response distribution in this condition may have lessened the effectiveness of categorical organization. Specifically, it has been shown that as a category label comes to subsume more and more members, its effectiveness as a retrieval cue for any particular category member declines (e.g., Cohen, 1966; Tulving & Pearlstone, 1966; Watkins, 1979). Perhaps recall in the self/organized condition would have exceeded the semantic/organized condition if it also constrained subjects to divide the list words into two equal-sized categories. A Spearman rank-order correlation coefficient, calculated between the number of items recalled by subjects in the self/organized condition and their degree of deviation from a 50/50 response distribution, revealed that uneveness of the distribution had little effect on the amount recalled, $r = .16$, *ns*. This analysis suggests that unequal category sizes did not significantly reduce amount recalled in the self/organized condition.

effectiveness of the different conditions in encouraging organization of the target words.

However, PF values have been shown to be correlated with the number of items recalled (Murphy, 1979), raising the possibility that the larger PF values for the organized tasks are an artifact of the high levels of recall found in these conditions. The problem may be corrected by calculating the ratio of the obtained PF to the maximum PF possible given the number of items recalled (Kihlstrom & Wilson, 1984; Murphy & Puff, 1982; Pellegrino, 1971). Although the correction is controversial (Sternberg & Tulving, 1977), it seemed appropriate to reanalyze the data using the adjusted ratio measure. A two-way ANOVA on the adjusted PF scores replicated the findings of the raw PF analysis: The only effect to reach significance was organization, $F(1, 60) = 21.21$, $p < .0001$.

The fact that organization, as inferred from both recall performance and PF score analysis, can be orthogonally varied between encoding tasks suggests that it is not a necessary property of self-referent encoding.

Response latency. The bottom panel of Table 3 presents the mean response latencies for each experimental condition. A 2 × 2 ANOVA variance (Encoding Task × Organization) revealed a single significant main effect of organization, $F(1, 60) = 17.35$, $p < .001$, indicating that the organized tasks required more processing time than the unorganized tasks. To examine the relative contributions of organization and processing time to recall performance, we repeated our recall analysis with response latency as a covariate. The analysis of covariance revealed only two significant results, a main effect for organization, $F(1, 59) = 90.52$, $p < .0001$, and a main effect for trial, $F(1, 60) = 59.00$, $p < .0001$, thus replicating the pattern of findings from the ANOVA. It appears that organization, not processing time, determines the number of target words recalled within a trial.

Experiment 3

Experiment 2 supported the prediction that semantic and self-referent encodings would produce comparable recall levels if they were equated for the amount of organization they encouraged. However, a potential confound in the design needs consideration: in the self/unorganized encoding condition, the focus was on idiomatic meanings of the target words (e.g., "To keep a civil tongue"), while in the other three conditions core meanings of the words were encoded (e.g., tongue as a body part). Perhaps memory performance was relatively poor in the self/unorganized condition because idiomatic meanings of words, which are largely determined by context (e.g., Ortony, Schallert, Reynolds, & Antos, 1978), are more difficult to recall than core meanings. Evidence can be found supporting such an interpretation (Hashtroudi, 1983). One purpose of Experiment 3, therefore, was to see if the pattern of results in Experiment 2 could be replicated when all of the experimental conditions directed encoding to the core aspect of the target word's meaning.

A second purpose of Experiment 3 was to test our claim that recall performance in the organized conditions resulted from establishing connections between list words during encoding. Alternately, our findings could be explained by assuming that recall superiority in the organized conditions depended on processes localized primarily at the retrieval stage. That is, subjects in our organized conditions may have perceived the list's categorical structure but stored the traces of the individual category instances as independent units (e.g., Slamecka, 1968, 1972). At the time of recall, subjects could retrieve the category labels and use them to guide the search through the unorganized network to create organization and improve recall. Subjects in our unorganized conditions, however, having failed to note the category structure of the list, would not have category labels available to guide their search (postexperimental interviews found only 16% of the subjects in these conditions were aware that all words on the list were body part words).

To evaluate these two alternatives, we included a cued recall condition in Experiment 3. Half the subjects in our unorganized conditions were informed at recall that the words they had rated belonged to two categories, internal and external body parts. We expected cuing to enhance recall, since the presentation of category label retrieval cues has been shown to facilitate recall even when subjects are unaware of categorical structure in a list during encoding (Hunt & Seta, 1984). However, if encoding plays a crucial role in organization, cued recall levels would still be significantly less than those found in the organized groups. On the other hand, if the mechanisms underlying the organizational advantage reside primarily at retrieval, then we would expect the cued and organized conditions to show approximately equal memory performance.

Method

Subjects. Subjects were 96 Harvard undergraduates who were paid for their participation. They were tested individually in sessions lasting approximately 20 min.

Materials and design. The materials were the same as those used in Experiment 2 with the following exception: The 28 sentence frames in the self/unorganized condition were identical to the 28 frames that elicited a *yes* response in the semantic/unorganized condition except that all other-referent terms (e.g., "the boy," "his") were changed to self-referent terms (e.g., "I," "My"). For example, if the semantic/unorganized frame was "The soldier preferred to keep his _____ short: Hair?", then subjects in the self/unorganized condition were shown "I prefer to keep my ____ short: Hair?" and asked to decide if that was true for them. This allowed the self/unorganized condition to use the body part sense of the target words.

The design was similar to that of Experiment 2 with two exceptions: (a) Only a single free recall trial was used, and (b) two cued recall conditions were added. Subjects in the semantic/unorganized cued group received the same encoding procedure as subjects in the semantic/unorganized condition, while subjects in the self/unorganized cued group were treated the same as subjects in the self/unorganized condition. Before beginning recall, however, subjects in the two cued groups were informed that half of the words they had just seen were internal body parts and half were external body parts. These changes resulted in a 3 × 2 × 2 factorial with organization (organized, unorganized, and unorganized plus cue) and encoding test (self and semantic) being varied between subjects, and response (*yes* and *no*) being varied within subjects. Sixteen subjects were randomly assigned to each of the six experimental conditions.

Procedure. Except for the provision of category label retrieval cues in the two cued conditions, the procedure was identical to Experiment 1.

Results and Discussion

In the present experiment, and in the two that follow, both adjusted and overall proportion recalled were analyzed. However,

because adjusted recall produced the same pattern of results as the overall recall analyses, only the latter are presented here. The interested reader can find a complete analysis of adjusted recall scores in the work of Klein (1985).

Table 4 presents the mean proportions recalled for each experimental condition. An ANOVA on these scores yielded only a main effect of organization, $F(1, 90) = 57.36$, $p < .0001$. Post hoc Tukey tests revealed that more words were recalled in the organized condition (.78) than in either the cued (.58) or unorganized (.51) conditions ($p < .01$); the difference between the latter two conditions was not statistically significant. The absence of a reliable difference in recall between the cued and unorganized conditions suggests that the beneficial effects of organization depend to some degree on associations established during encoding; the question of whether this advantage results from an interaction between encoding operations and retrieval conditions (e.g., Tulving, 1979; Tulving & Thomson, 1973) cannot be answered by the present design. However, any model proposing that the major contribution of organization to recall is the greater availability of category cues during retrieval must explain why the provision of category labels at recall failed to reliably increase memory over uncued levels.

Finally, examination of the mean recall proportions in Tables 3 and 4 reveals that the pattern of recall in Experiment 2 was replicated even when all the experimental conditions used core meanings of the target words.

Experiment 4

Our interpretation of the results of Experiments 2 and 3 depends on the assumption that the crucial difference between organized and unorganized encoding tasks is the amount of organization they engender. There is, however, another difference between these tasks that merits consideration: The number of target words associated with a particular orienting question is confounded with the organizational factor. Since all of the target words in the organized conditions are judged with reference to the same orienting question, the repeated presentations of the question should require minimal processing, allowing encoding efforts to be focused on each target word. In the unorganized conditions, however, subjects must consider each target word's meaning in the context of its own unique sentence frame. Therefore, they must actively process both sentence frame and target word, necessarily devoting less processing capacity to the target words at the time of encoding.

In Experiment 4, a single semantic encoding question that required sorting the target words into two categories was compared with a single semantic encoding question that did not encourage such organization. If organization was the primary factor contributing to the pattern of recall in Experiments 2 and 3, then subjects in the single-question semantic/organized group should remember more words than subjects in the single-question semantic/unorganized group. If, however, the use of a single orienting question was responsible for recall performance in the organized conditions of Experiments 2 and 3, the two groups in the present experiment should show comparable levels of recall.

Method

Subjects. Thirty Harvard undergraduates were tested individually in sessions lasting approximately 20 min and paid for their participation.

Table 4
Proportion of Words Recalled As a Function of Encoding Task and Level of Organization, Experiment 3

	Level of organization			
Encoding task	Unorganized	Cued	Organized	*M*
Semantic	.50	.56	.78	.62
Self	.51	.60	.77	.63
M	.51	.58	.78	

Materials and design. Two groups of subjects, 15 per condition, were compared. The first group received a semantic encoding task that encouraged organization, and the second group performed a semantic task that encouraged individual-item processing. Both encoding tasks controlled for the number of target words per question by requiring all of the words to be processed with reference to a single orienting question. In the organized condition, the subjects saw the question "Is this an external body part?"; in the unorganized condition, subjects were asked "Does a second common meaning of this word come to mind?" Subjects in the latter condition were informed in advance that all target words were body part words and that they should decide quickly if they could think of a second, nonanatomical meaning for each word. For example, the target word *head* might evoke the secondary meaning "to be in command" or "to go in a specified direction." Deciding which list words have a second common meaning provides little basis for categorization and is not likely, therefore, to encourage organization. Target words were the same as those used in Experiments 2 and 3.

Procedure. The procedure followed was identical to that of Experiment 1.

Results and Discussion

The use of single orienting questions ensured that target words in both the organized and unorganized conditions would be the focus of encoding during input. Despite this, the semantic/organized group still recalled significantly more words than the semantic/unorganized group, .80 > .68, $t(28) = 2.94$, $p < .01$. It must be noted, however, that more words were recalled in the unorganized condition of this experiment than in the semantic/unorganized conditions of Experiments 2 and 3 (average recall for Trial 1 = .50). A possible explanation is that although few subjects in the unorganized conditions of Experiments 2 and 3 indicated they were aware of relations between target words, subjects in Experiment 4 were alerted during encoding to the fact that all target words were body part words. This knowledge may have prompted some organization during encoding, but recall was still inferior to that found in the semantic/organized condition because the encoding task encouraged individual-item processing.

Experiment 5

This experiment was designed to provide a replication of Experiment 2 and to extend the generality of our findings by demonstrating the same pattern of recall using a different set of encoding tasks and using target words chosen from a different category.

Method

Subjects. Sixty Harvard undergraduates served as paid subjects. They were tested individually in sessions that lasted approximately 20 minutes.

Materials and design. Career aspirations form a significant part of a person's self-concept (Gordon, 1968; McGuire & Padawer-Singer, 1976). Therefore, 32 nouns, chosen from the Battig and Montague norms from the category "An Occupation or Profession," served as target words for Experiment 5. Half of the nouns were occupations that required a college education (e.g., judge, chemist) and half were occupations for which a college education was not necessary (e.g., janitor, grocer). Six additional occupation words were selected to serve as primacy and recency buffers and were not included in any statistical analyses.

The experimental design was a 2 × 2 factorial with both encoding task (self and semantic) and organization (organized and unorganized) being varied between subjects. Fifteen subjects were randomly assigned to each of the four Encoding × Organization conditions.

Subjects in the semantic/organized condition were asked to decide whether each occupation required a college education (e.g., "Does this job require a college education _____: Butcher?"). In addition to requiring a semantic judgment, this task encouraged the subjects to divide the list into two categories: occupations that do and do not require a college education. In the self/organized condition subjects were asked if they had ever aspired to a career in each of the presented occupations (e.g., "Have you ever wanted to be a _____: Biologist?"). This task required a self-referent decision and encouraged sorting the list words into the categories "occupations I would like" and "occupations I would not like."

In the semantic/unorganized condition, the orienting questions were single sentence descriptions of job responsibilities. After viewing each description, subjects were shown a target word and asked to decide if that occupation entailed performing the duty described by the orienting question (e.g., "Does this person perform operations: Surgeon?") For each subject, half of the target words were randomly paired with job descriptions designed to elicit a *yes* response, and half with descriptions eliciting a *no* response. This condition focused attention on judging each target word for compatibility with its associated job description without drawing attention to relations among the target words.

In the self/unorganized condition, subjects were shown self-descriptive sentence frames that were each missing the target word. After reading each frame, they were shown the target word and asked to decide if the completed sentence described their opinions or feelings. All sentences began with the word "I" and asked a personal question (e.g., "I place complete trust in my _____: Doctor?", "I often look like I could use the services of a good _____: Barber?"). The individual frames were designed to be as unrelated to each other as possible, encouraging subjects to treat each "sentence frame-target word" pairing as an independent unit.

In all other aspects the design of Experiment 5 was identical to Experiment 2.

Procedure. The procedure was the same as Experiment 1.

Results and Discussion

Table 5 presents the mean proportions recalled for each experimental condition. A two-way ANOVA (Encoding Task × Organization) on these scores yielded only a main effect of organization, $F(1, 56) = 16.59$, $p < .0005$, indicating that organized encoding tasks resulted in significantly better recall than unorganized tasks, .56 > .48.

Although the magnitude of the recall superiority associated with organization was smaller in Experiment 5 than that found in the experiments using words from the category "body parts" as target words, the same pattern of recall results was obtained: (a) The SRE was clearly in evidence when organized, self-referent encoding was compared with unorganized, semantic encoding, .55 > .46, $t(28) = 3.29$, $p < .005$; (b) a reversal of the SRE occurred when the organizational advantage was given to the semantic encoding task, .56 > .49, $t(28) = 2.53$, $p < .05$; (c) the SRE was eliminated when self and semantic encoding tasks were equated for organization (no main effect of Encoding Task, $F < 1$); (d) even allowing for the possibility that organization masked the effects of elaboration in the organized conditions, the comparison of self/unorganized (.49) and semantic/unorganized (.46) recall offered little evidence that self-referent encoding is more elaborate than semantic encoding, $t(28) = .95$, $p > .30$.

Table 5
Proportion of Words Recalled As a Function of Encoding Task and Level of Organization, Experiment 5

	Level of organization		
Encoding task	Unorganized	Organized	*M*
Semantic	.46	.56	.51
Self	.49	.55	.52
M	.48	.56	

Once again, the hypothesis that the evaluative nature of the self-referent decision is responsible for the SRE receives no support. The two self-referent tasks were clearly more evaluative then the two semantic tasks, yet when amount of organization was controlled, self and semantic encodings showed nearly identical recall scores.

In summary, Experiment 5 demonstrates the generality of our findings by showing that the pattern of recall results found in Experiment 2 is not restricted to a specific category of target words or a particular set of self-referent and semantic encoding tasks.

General Discussion

The present series of studies provides converging evidence that the amount of organization, not the nature of the encoding task, is responsible for what has been called the self-reference effect. Experiment 1 showed that in addition to producing better recall, self-referent encoding tasks encouraged significantly greater clustering than semantic tasks, suggesting that self-referent recall enhancement may be due to a confounding of organization with self-referent encoding. Experiment 2 put the hypothesis to a direct test by independently manipulating both encoding task and organization. The results clearly showed that when organization was controlled, the pattern of recall produced by self-referent and semantic encoding tasks was virtually identical. These findings provide no support for the hypothesis that self-referent encoding produces a more durable or more elaborate memory trace than does semantic encoding.

The claim that amount of organization determined memory performance is pivotal for our explanation of the SRE. Therefore, several alternative interpretations of our recall findings were examined. An analysis of response latency (Experiments 1 and 2) failed to support a processing-time explanation of recall performance. Experiment 3 showed that enhanced memory in the organized conditions was not just the result of category cue availability at recall, but depended on organization taking place at the time of encoding. Finally, Experiment 4 demonstrated that recall was not simply a function of the degree to which target

words in the organized conditions were made the focus of encoding operations.

Our research also establishes that organization is not a necessary property of self-reference and that it can be orthogonally varied within self-referent and semantic encoding tasks. Trait words, body part words, and occupation words are easily associated with the category labels "things that describe self" and "things that do not describe self," but this should not be considered evidence that there is something unique about self-reference as a means of organizing stimulus material; we have shown that the same stimuli may be organized just as effectively by a semantic encoding task if it provides suitable category labels (e.g., "external body parts" and "internal body parts"). Furthermore, regardless of whether semantic or self-referent judgments are required, the task must first direct attention to relations between list words if the effects of organization are to be observed in recall performance. For example, a complete reversal of the SRE is found when semantic encoding tasks encourage organization and self-referent encoding tasks emphasize individual-item processing.

To date, no single principle has satisfactorily explained and integrated the recall findings from the various procedural modifications of the DOP paradigm that have been used to explore the SRE. For example, some researchers (e.g., Ferguson et al., 1983) contend that self-reference improves recall by focusing encoding on the evaluative aspects of words, whereas others (e.g., Bower & Gilligan, 1979) view self-reference as a special case of the general principle that memory for words is promoted by using any well-known person (e.g., mother) as a referent during encoding. Although a single explanatory principle is not necessary, the results of the experiments reported in this article provide the basis for a reinterpretation of the effects of self-reference on recall in terms of a single factor, the amount of organization encouraged by the encoding task. In the next two sections we will suggest how this organization hypothesis can accommodate the findings from both evaluation and person-reference studies of the SRE.

Self-Reference and Evaluation

Experiments 2, 3, and 5 found no recall advantage for evaluative judgments: evaluative, self-referent decisions and nonevaluative, semantic decisions led to identical recall provided organization was equated between conditions. Ferguson et al. (1983), however, found that an evaluative semantic task ("Is this word a desirable characteristic?") not only produced significantly greater recall than three nonevaluative semantic tasks (judging words for their meaningfulness, familiarity, or imageability) but equaled self-referent free recall levels. This led them to propose that the inherently evaluative nature of most self-referent judgments might be responsible for the SRE. However, a careful examination of their stimulus materials suggests that organization and evaluation may have been confounded in their desirability judgment condition. Their stimuli were 40 trait adjectives chosen such that half were high in desirability and half low in desirability according to Kirby and Gardner's (1972) norms, but no attempt was made to structure the list similarly along the dimensions of meaningfulness, familiarity, or imageability. Each of the four semantic encoding tasks used by Ferguson et al. can be conceptualized as requiring subjects to sort the list words into two categories defined by the presence or absence of the dimension stressed by the particular task (desirability, meaningfulness, familiarity, and imageability). The fact that the list comprised adjectives that clearly fell into two groups along the desirability dimension, but was not structured comparably on the other three dimensions, implies that the desirability dimension would be the most conducive to categorical organization. Thus, Ferguson et al.'s results can be interpreted as an illustration of the facilitating effects of categorical organization on recall.

Self-Reference and Person-Reference

In their original exploration of the mnemonic effects of self-reference, Rogers et al. (1977) concluded that their recall results demonstrated its uniqueness as a means of producing an elaborate encoding. They neglected, however, to control for the possibility that *any* person-referent encoding task might produce recall levels equivalent to those found with self-referent encoding. Subsequent tests of this hypothesis indicated that the degree of familiarity with the person used as the referent during encoding is a primary factor determining recall. When judgments about trait adjectives are made with reference to a well-known other (e.g., "Describes your mother?"), recall levels are comparable to those attained in self-reference tasks (Bower & Gilligan, 1979, Experiment 2; Friedman & Pullyblank, 1982, Experiment 1; Kuiper & Rogers, 1979, Experiments 4 and 5; but see Ferguson et al., 1983 and Lord, 1980, Experiment 1 for conflicting findings), whereas the use of unfamiliar others (e.g., "Describes Walter Cronkite?") fails to match self-reference (Bower & Gilligan, 1979, Experiment 2; Ferguson et al., 1983; Kuiper & Rogers, 1979, Experiment 1; Lord, 1980, Experiment 1).

An organizational account for the observed dependence of recall on referent familiarity relies on the fact that judging trait adjectives for their referent descriptiveness encourages the organization of the words into the categories "describes referent" and "does not describe referent." There is evidence indicating a positive relation between the degree of familiarity with a category and the strength of its category-to-member associations (Barsalou, 1983). This suggests that the ease of forming associations between trait adjectives and the category labels "describes referent" and "does not describe referent," as well as the effectiveness of these labels in activating category members during recall, should be an increasing function of how well the subject knows the composition of these categories (i.e., how familiar the subject is with the person being judged). Although clearly post hoc, this model provides an explanation of the general phenomenon of person-referent recall, whether the person is unfamiliar, well-known, or the self.

It should be noted that the organization hypothesis is only intended to provide an alternative for DOP explanations of the SRE. Greenwald and his associates (e.g., Banaji & Greenwald, 1984) have explored the SRE using "generation tasks" that require subjects to construct sentences containing the target words. These studies generally demonstrate superior recall for sentences describing personal experiences, but there is some question concerning whether generation task effects depend on the same memory processes that are involved in the standard DOP experiment (Slamecka & Graf, 1978).

Organization and Depth of Processing

One finding in the DOP literature appears to be at odds with our demonstration of the memorial superiority of encoding tasks

that encourage categorical organization. Moscovitch and Craik (1976, Experiments 2 and 3), using conditions quite similar to our organized and unorganized semantic tasks, found that retention was better when each target word was paired with a unique orienting question than when groups of target words shared a single orienting question. In light of this, it is surprising that our findings consistently showed that the use of a single orienting question produced better recall than providing unique orienting questions for each word. An explanation for these conflicting results is suggested by an examination of the recall conditions used in the two studies. Moscovitch and Craik used a task in which the original orienting questions were presented as retrieval cues. Since the probability of recalling any single word associated with a retrieval cue diminishes as the number of words associated with that cue increases (e.g., Cohen, 1966; Tulving & Pearlstone, 1966; Watkins, 1979), memory should be better when each word is cued with a unique question than with a shared one. However, different results may be expected from the recall conditions we employed. Without the provision of orienting question retrieval cues, the usual beneficial effects of categorical organization on free recall (e.g., G. Mandler, 1967) give the single question encoding task an advantage over the unique question task.

Conclusion

Much of the initial interest in the SRE reflected the claim that the phenomenon was informative about the nature of the self-concept as a memory structure. Given the elaboration account of the DOP results, it seemed as if self-referent encoding produced many links between the stimulus and preexisting information about the self in memory (e.g., Ingram et al., 1983). These links, in turn, formed a richly elaborated memory trace of the stimulus, and afforded many paths to its successful retrieval. Thus, it was inferred that the self was a highly elaborate memory structure (Ingram et al., 1983; Keenan & Baillet, 1980; Kendzierski, 1980; Lord, 1980; Markus, 1980; Markus & Smith, 1981; Mills, 1983; Rogers et al., 1977)—an inference that was generalized to memory structures representing highly familiar other people as well (Bower & Gilligan, 1979; Kuiper & Rogers, 1979). The experiments reported here, however, show that the effects of self-referent encoding are produced by organization rather than elaboration and that organization is not an inherent property of self-referent encoding. Although it may well be that the self is a highly elaborate memory structure, the SRE does not give evidence to support this conclusion, and other paradigms must be used to address this issue.

On a more positive note, these experiments clearly indicate that elaboration accounts of memory functioning will have to be supplemented by an appeal to organizational principles. It appears that encoding tasks that only encourage the processing of individual items, no matter how elaborate that processing may be, will not produce optimal retention in the DOP paradigm (Bellezza et al., 1977; Bellezza et al., 1976). Retention will be maximal only when some degree of interitem organization takes place as well as elaboration. Organizational principles once were of central importance in theories of memory (e.g., Bower, 1970; Mandler, 1967; Tulving & Donaldson, 1972), but their importance has been overshadowed by current emphases on elaboration, and in fact many current paradigms actively prevent organizational effects from being observed (Battig & Bellezza, 1979). Just as Ebbinghaus's exclusion of meaningful stimulus material from his experiments prevented him from studying the effects of a variable central to remembering—"effort after meaning"—so the exclusion of organizational factors from the DOP paradigm impedes our understanding of the functioning of one of the most significant variables in natural memory processes: the search for relations among stimuli.

References

Allen, M., Puff, C. R., & Weist, R. (1968). The effects of associative and coding processes on organization in free recall. *Journal of Verbal Learning and Verbal Behavior, 7,* 531–538.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983a). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

Anderson, J. R. (1983b). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). San Francisco: W. H. Freeman and Company.

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385–403). Hillsdale, NJ: Erlbaum.

Anderson, N. H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9,* 272–279.

Banaji, M. R., & Greenwald, A. G. (1984, April). *When does self-reference facilitate recall?* Paper presented at the meeting of the Eastern Psychological Association, Baltimore.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211–227.

Battig, W. F., & Bellezza, F. S. (1979). Organization and levels of processing. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 321–346). New York: Academic Press.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80* (3, Pt. 2).

Bellezza, F. S. (1984). The self as a mnemonic device: The role of internal cues. *Journal of Personality and Social Psychology, 47,* 506–516.

Bellezza, F. S., Cheesman, F. L., & Reddy, B. G. (1977). Organization and semantic elaboration in free recall. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 539–550.

Bellezza, F. S., Richards, D. L., & Geiselman, R. E. (1976). Semantic processing and organization in free recall. *Memory & Cognition, 4,* 415–421.

Bousfield, A. K., & Bousfield, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports, 19,* 935–942.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology, 1,* 18–46.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior, 8,* 323–343.

Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. *Journal of Research in Personality, 13,* 420–432.

Cochran, W. G., & Cox, G. M. (1957). *Experimental designs* (2nd ed.). New York: Wiley.

Cofer, C. N. (1966). Some evidence for coding processes derived from clustering in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 188–192.

Cofer, C. N. (1967). Does conceptual organization influence the amount retained in immediate free recall? In B. Kleinmuntz (Ed.), *Concepts and the structure of memory* (pp. 181–214). New York: Wiley.

Cohen, B. H. (1966). Some-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior, 5,* 182–187.

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 68,* 221–234.

Craik, F. I. M. (1973). A "levels of analysis" view of memory. In P. Pilner, L. Krames, & T. M. Alloway (Eds.), *Communication and affect: Language and thought* (pp. 45–65). London: Academic Press.

Craik, F. I. M. (1977). Depth of processing in recall and recognition. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 679–697). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 11,* 268–294.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 588–598.

Ellis, H. C., & Hunt, R. R. (1983). *Fundamentals of human memory and cognition.* Dubuque, IA: Wm. C. Brown Company.

Ellis, H. C., Thomas, R. L., & Rodriguez, I. A. (1984). Emotional mood states and memory: Elaborative encoding, semantic processing, and cognitive effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 470–482.

Eysenck, M. W., & Eysenck, C. E. (1979). Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 472–484.

Ferguson, T. J., Rule, G. R., & Carlson, D. (1983). Memory for personally relevant information. *Journal of Personality and Social Psychology, 44,* 251–261.

Friedman, A., & Pullyblank, J. (1982, November). *Remembering information about oneself and others: The role of distinctiveness.* Paper presented at the meeting of the Psychonomic Society, Minneapolis.

Ganellen, R. J., & Carver, C. S. (1985). Why does self-reference promote incidental encoding? *Journal of Experimental Social Psychology, 21,* 284–300.

Gardiner, J. M. (1974). Levels of processing in word recognition and subsequent free recall. *Journal of Experimental Psychology, 102,* 101–105.

Gordon, C. (1968). Self-conceptions: Configurations of content. In C. Gordon & K. J. Gergen (Eds.), *The self in social interaction* (pp. 115–136). New York: Wiley.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82,* 1–20.

Greenwald, A. G. (1981). Self and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 15, pp. 201–236). New York: Academic Press.

Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition: Vol. 3* (pp. 129–178). Hillsdale, NJ: Erlbaum.

Halpin, J. A., Puff, C. R., Mason, H. F., & Martson, S. P. (1984). Self-reference and incidental recall by children. *Bulletin of the Psychonomic Society, 22,* 87–89.

Hamilton, D. L. (1981). Cognitive representations of persons. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium: Vol. 1* (pp. 135–159). Hillsdale, NJ: Erlbaum.

Hashtroudi, S. (1983). Type of semantic elaboration and recall. *Memory & Cognition, 11,* 476–484.

Hastie, R. (1980). Memory for information that confirms or contradicts a general impression. In R. Hastie, T. M. Ostrom, E. B. Ebbesen, R. S. Wyer, Jr., D. L. Hamilton, & D. E. Carlston (Eds.), *Person memory: The cognitive basis of social perception* (pp. 155–177). Hillsdale, NJ: Erlbaum.

Hudson, R. L. (1968). Category clustering as a function of level of information and number of stimulus presentations. *Journal of Verbal Learning and Verbal Behavior, 7,* 1106–1108.

Hunt, K. (1971). Measuring clustering differences among categories in the same list. *Behavior Research Methods and Instrumentation, 3,* 67–69.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 497–514.

Hunt, R. R., & Seta, C. E. (1984). Category size effects in recall: The roles of relational and individual item information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 454–464.

Hyde, T. S., & Jenkins J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior, 12,* 471–480.

Ingram, R. E., Smith, T. W., & Brehm, S. S. (1983). Depression and information processing: Self-schemata and the encoding of self-referent information. *Journal of Personality and Social Psychology, 45,* 412–420.

Keenan, J. M., & Baillet, S. D. (1980). Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and performance* (Vol. 8, pp. 651–669). Hillsdale, NJ: Erlbaum.

Kendzierski, D. (1980). Self-schemata and scripts: The recall of self-referent and scriptal information. *Personality and Social Psychology Bulletin, 6,* 23–29.

Kihlstrom, J. F. (1981). On personality and memory. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 123–149). Hillsdale, NJ: Erlbaum.

Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 1–47). New York: Academic Press.

Kihlstrom, J. F., & Wilson, L. (1984). Temporal organization of recall during posthypnotic amnesia. *Journal of Abnormal Psychology, 93,* 200–208.

Kirby, D. M., & Gardner, R. C. (1972). Ethnic stereotypes: Norms on 208 words typically used in their assessment. *Canadian Journal of Psychology, 26,* 140–154.

Klein, S. B. (1985). The self-reference effect in human memory. *Dissertation Abstracts International, 46,* 07B. (University Microfilms No. 85-20229)

Klein, S. B., & Kihlstrom, J. F. (1984). *Self-reference and hypermnesia.* Unpublished manuscript.

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self–other differences. *Journal of Personality and Social Psychology, 37,* 499–514.

Lord, C. G. (1980). Schemas and images as memory aids: Two modes of processing social information. *Journal of Personality and Social Psychology, 38,* 257–269.

Maki, R. H., & McCaul, K. D. (1985). The effects of self-reference versus other reference on the recall of traits and nouns. *Bulletin of the Psychonomic Society, 23,* 169–172.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 327–372). New York: Academic Press.

Mandler, G. (1977). Commentary on "Organization and Memory." In G. Bower (Ed.), *Human memory: Basic processes* (pp. 297–308). New York: Academic Press.

Mandler, G. (1979). Organization, memory, and mental structures. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 303–319). New York: Academic Press.

Mandler, J. M. (1979). Categorical and schematic organization in memory. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 259–299). New York: Academic Press.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: Erlbaum.

Markus, H. (1980). The self in thought and memory. In D. M. Wegner & R. R. Vallacher (Eds.), *The self in social psychology* (pp. 102–130). New York: Oxford University Press.

Markus, H., & Smith, J. (1981). The influence of self-schema on the perception of others. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 233–262). Hillsdale, NJ: Erlbaum.

McCaul, K. D., & Makl, R. H. (1984). Self-reference versus desirability ratings and memory for traits. *Journal of Personality and Social Psychology, 47,* 953–955.

McGuire, W. J., & Padawer-Singer, A. (1976). Trait salience in the spontaneous self-concept. *Journal of Personality and Social Psychology, 33,* 743–754.

Mills, C. J. (1983). Sex-typing and self-schemata effects on memory and response latency. *Journal of Personality and Social Psychology, 45,* 163–172.

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior, 15,* 447–458.

Mross, E. F., & Kihlstrom, J. F. (1985). *Levels of processing and levels of recall in hypermnesia.* Unpublished manuscript.

Murphy, M. D. (1979). Measurement of category clustering in free recall. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 51–83). New York: Academic Press.

Murphy, M. D., & Puff, C. R. (1982). Free recall: Basic methodology and analyses. In C. R. Puff (Ed.), *Handbook of research methods in memory and cognition* (pp. 99–128). New York: Academic Press.

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17,* 465–477.

Packman, J. L., & Battig, W. F. (1978). Effects of different kinds of semantic processing on memory for words. *Memory & Cognition, 6,* 502–508.

Pellegrino, J. W. (1971). A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavior Research Methods and Instrumentation, 3,* 241–246.

Postman, L., & Kruesi, E. (1977). The influence of orienting task on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior, 16,* 353–369.

Puff, C. R. (1970). Role of clustering in free recall. *Journal of Experimental Psychology, 86,* 384–386.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76,* 45–48.

Rogers, T. B. (1981). A model of the self as an aspect of the human information processing system. In N. Cantor and J. F. Kihlstrom (Eds.), *Personality, cognition and social interaction* (pp. 193–214). Hillsdale, NJ: Erlbaum.

Rogers, T. B., Kuiper, N. A., & Kriker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35,* 677–688.

Schulman, A. I. (1974). Memory for words recently classified. *Memory & Cognition, 2,* 47–52.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76,* 504–513.

Slamecka, N. J. (1972). The question of associative growth in the learning of categorized material. *Journal of Verbal Learning and Verbal Behavior, 11,* 324–332.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.

Srull, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 440–463.

Srull, T. K. (1983). Organizational and retrieval processes in person memory: An examination of processing objectives, presentation format, and the possible role of self-generated retrieval cues. *Journal of Personality and Social Psychology, 44,* 1157–1170.

Srull, T. K., & Brand, J. F. (1983). Memory for information about persons: The effect of encoding operations on subsequent retrieval. *Journal of Verbal Learning and Verbal Behavior, 22,* 219–230.

Sternberg, R. J., & Tulving, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin, 84,* 539–556.

Tulving, E. (1979). Relation between encoding specificity and levels of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 405–428). Hillsdale, NJ: Erlbaum.

Tulving, E., & Donaldson, W. (Eds.). (1972). *Organization of memory.* New York: Academic Press.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Underwood, B. J. (1964). The representativeness of rote verbal learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 47–78). New York: Academic Press.

Wallace, W. P. (1970). Consistency of emission order in free recall. *Journal of Verbal Learning and Verbal Behavior, 9,* 58–68.

Warren, M. W., Chattin, D., Thompson, D. D., & Tomsky, M. T. (1983). The effects of autobiographical elaboration on noun recall. *Memory & Cognition, 11,* 445–455.

Watkins, M. J. (1979). Engrams as cuegrams and forgetting as cue overload: A cueing approach to the structure of memory. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 347–372). New York: Academic Press.

Wood, G., & Underwood, B. J. (1967). Implicit responses and conceptual similarity. *Journal of Verbal Learning and Verbal Behavior, 6,* 1–10.

Received February 19, 1985
Revision received July 12, 1985 ■