The Set Size Effect in Person Memory

Stephen G. Gilligan & Gordon H. Bower
Stanford University

Rare is the occasion when we do not form or use personality impressions while interacting with another person. These impressions are revealed through the evaluative descriptions which frequently saturate discussion of another person, such as "he's a nice guy", 'she's as smart as a whip", "he's so darn stubborn", etc. How do these generalizations operate as cognitive structures? One obvious function is simplification: not only do such generalizations make it easer to communicate information about an individual's perceived behavioral patterns, they also enable performance of the common and quite important cognitive task of evaluating and predicting someone's behavior. From the point of view of learning theory, these generalizations serve as reference structures which can be used to assimilate new information into prior information in memory. This assimilation process, which can operate in a variety of ways (see Bower and Hilgard, 1980), facilitates the learning of new information.

The cognitive nature of personality structures has long interested researchers in social and personality psychology (e.g., Asch, 1946, 1952; Heider, 1958; Anderson, 1962, 1974; Schneider, 1973). In the recently developed firld of social cognition (e.g., Higgins, Herman, and Zanna, 1981), where investigators have borrowed heavily from the models and methodologies of cognitive psychology, the area of person memory occupies a central position (see Hastie, Ostrom, Ebbesen, Wyer, Hamilton, and Carlston, 1980). A central question in person memory research is how personality structures affect how behavioral information is learned and remembered.

One line of research in this area indicates that using personality structures to encode behavioral information enhances overall learning. For example, Hamilton (1981) presented subjects with a set of behavioral information about a signal character. Some subjects were instructed to form a coherent personality impression from the behaviors; others were simply told to memorize the material. The impression formation group later recalled more of the behaviors than the memory set group. Hamilton (1981) interpreted this finding, which has been replicated by Gilligan and Bower (Note 1) and Wyer and Gordon (1982), as evidence that schematic organizations enhance overall learning rates.

Research by Cantor and Mischel (1979) supports this conclusion. They provided subjects with trait descriptions about various characters. Some descriptions were consistent with a salient personality prototype ("extroversion" or "introversion") while others were unrelated to any prototype. Subjects later recalled more about the "prototype consistent" characters. Cantor and Mischel (1979) explained this result, which is analogous to Chase and Simon's (1973) "chess effect", in terms of the learning enhancement provided by available personality structures.

A related line of research has investigated whether personality impressions bias learning processes. Many of these studies have manipulated the personality stereotype attributed to a character before presenting behaviors which vary in their relation to the stereotype. For example, Zadny and Gerard (1974) had subjects watch a skit of a student trying to enroll in classes. They varied the subject major (music, chemistry, or psychology) ascribed to the student before presenting the skit, which contained information (e.g., books, remarks) relevant to each of the majors. Subjects later remembered more of the information congruent with the major they believed the actor to be. This finding of sterreotype-relevant information being remembered better than stereotype-irrelevant information has been reported by numerous other investigators (e.g., Sulin and Dooling, 1974; Marcus, 1977; Lingle and Ostrom, 1979; Rothbart, Evans, and Fulero, 1979; Snyder, 1981; Cohen, 1981).

As Hastie (1980) has cogently argued, however, most of these studies have only examined memory rates for relevant vs. irrelevant information, ignoring the fact that a behavior can also be incongruent to a personality schema. When this third type of relation is included, a more complex picture emerges. While irrelevant information is generally recalled less than relevant or incongruent information (for review, see Hastie, 1980), it is not clear which of the latter two types of behaviors is remembered better. Some studies (Hastie and Kumar, 1978; Rothbart, Evans, and Fulero, 1979; Hastie and Mazer, 1980; Wyer and Gordon, 1982) show superior memory for incongruent behaviors, while others (Snyder and Cantor, 1979; Snyder and Cambell, 1980) suggest the opposite pattern. In short, it appears that a variety of different factors affect how these different types of information are remembered (for discussions, see Hastie, 1980; Gilligan, Note 2; Srull, 1980).

A third line of research has examined how the complexity of elaborateness of a personality structure affects learning. An influential study in this area is by Rogers, Kuipers, and Kirker (1977), who used a modified depth-of-processing paradigm (Craik and Lockhart, 1972) in which subjects were shown trait adjectives and asked to judge some of the words for their surface appearance, ("Is this word printed in capital letters?"), some for phonetic features ("Does it rhyme with 'toy'?"), some for semantic meaning ("Does it mean the same as 'assertive'?"), and some for their self-relevance ("Does it describe you?"). Words judged in the self-relevance condition were recalled best. Rogers *et al.*, suggested that this finding, replicated by Lord (1980), demonstrates some special property of "self-schemas".

We challenged this interpretation of the results in an earlier paper

(Bower and Gilligan, 1979). In that paper we also reported experiments which first replicated Rogers et al's (1977) results, and then showed (1) equally high memory rates for self-rated traits and for traits used to retrieve episodes either from one's own life or from one's mother's life, but (2) poor memory for words rated for their relatedness to a less familiar person (Walter Cronkite). We argued that these data clearly indicate that there is nothing special about the self-schema as a mnemonic peg, as Rogers et al. (1977) and Lord (1980) have claimed, but that good memory depends on relating inputs on a well-differentiated memory structure.

The experiments described below tested this letter hypothesis in a more direct fashion. The major variable manipulated was the complexity of the personality structure ascribed to a character whose behavioral activities subjects read about and classified in relation to the personality structure. The general procedure used in the experiments involved having subjects first study personality descriptions of one or more characters, with a given character being described by two to ten traits. Subjects were then directed to use these personality trait lists to classify and explain a set of trait-related behaviors ostensibly enacted by the character(s), with an equal number of behaviors ascribed to each character. Thus, the number of trait categories available was varied while the number of behaviors was held constant, such that a given trait would be linked to one, two, or four behaviors. Our general prediction was that this within-subject manipulation of behavioral set size would affect memorability of the behaviors. Specifically, we expected surprise tests of free recall and cued recall (with the traits used as cues) to show that the probability of recalling a behavior increases with the richness (i.e., set size) of the trait description available for encoding the behavior.

Experiment 1

Subjects

There were 28 subjects in the experiment; all were volunteers from an introductory psychology course at California State University at Sacramento. Each subject received extra credit for his participation.

Materials and Design

The major experimental manipulation was the within S factor of trait set size. To develop the materials, fourteen trait adjectives were first selected from Anderson's (1968) normative ratings of 555 trait adjectives. Each adjective that met the criteria of examples are outgoing, practical, boastful, and courteous. For each of the 14 traits, 4 exemplary behaviors were generated. For example, "outgoing" behaviors included yelling at a clerk in the store, telling jokes at a party, and giving a rousing impromptu speech at a meeting. Each behavior was developed into a short vignette containing a situation and action pair. For example, a vignette exemplifying intelligence was: "Ann had just received her semester grades. As usual, she had earned excellent marks along with superb evaluations from her instructors."

To check the goodness-of-fit of these vignettes to their corresponding traits, 6 Stanford graduate students were given both the vignettes and the trait adjectives and asked to select, for each of the 56 vignettes, which of the 14 traits best explained such behavior. With few exceptions, the judges' choices matched perfectly.

To counterbalance the experimental groups of traits, four different orderings of the trait adjective list were randomly generated. As Figure 1 illustrates, each list was then partitioned into 3 different-sized trait descriptions of characters by pairing the first two traits with the name of a target character ("Joe", "Bill", or "Ted"), the next four traits with a different character, and the final eight with the remaining character. Eight vignettes were ascribed to each character. For the characters with two traits, each of the four vignettes originally generated for those traits was used.

Insert Figure 1 about here

For the four-trait characters, two vignettes from each of the four traits were randomly selected. For the eight-trait character, a single behavior corresponding to each trait was randomly chosen. Thus, depending on the condition, each trait was linked to one, two, or four vignettes.

Four different forms of test materials, corresponding to the counterbalancing trait groupings described above, were constructed. Each form contained a personality description sheet and a vignette booklet. The personality description sheet described Joe, Bill, and Ted as normal college students whose personality traits had been ascertained from a battery of personality tests and interviews. The names of the three characters were typed in capital letters beneath this introduction, along with their respective lists of two, four, or eight personality traits. The booklets contained 24 target vignettes corresponding to one of the trait description sheets, as well as 3 practice examples. Each vignette was typed on a separate half sheet of paper, with the question, "Which of (the character's) traits best explain this behavior?" appearing beneath each vignette. The order of the 24 target vignettes was randomized for each booklet.

Other test materials included a distractor task, a free recall test, and a cued recall test. The distractor task was simply a sheet of arithmetic problems that the subject was to perform; the free recall form instructed subjects to write down the gist of every vignette they had read earlier, trying to be as specific as possible, and the cued recall test--which was given after the subject had been handed back the trait description sheet--noted essentially the same recall instructions.

Subjects were run in groups of 7 to 12. The experimenter introduced himself as "a Stanford psychology researcher interested in personality processes", and vaguely described the study as one investigating how people form and use personality impressions. Subjects were then handed both a vignette booklet, which they were asked not to examine until instructed, and a corresponding personality description sheet, which they were directed to study to "get a good idea of each person's personality". After five minutes, subjects were informed that the booklets described various behaviors enacted by the three characters, with each page detailing a single episode from one of the character's lives. They were told that their task would be to read through the booklet, page by page, and decide for each behavior ascribed to a character which of his traits best explained that behavior. They were also told to use these behaviors to form a coherent personality impression of each character. It was noted that their progress would be paced with a stop-watch by the experimenter: they would have 9 seconds to read a vignette, after which the experimenter would say "all right", whereupon they would have five seconds to read the trait list of the character (being described) and select and write down the most explanatory trait. After the allotted five seconds, the experimenter would call out "Turn!", thereby signalling that the page was to be turned to the next vignette. This procedure--9 seconds to read, 5 seconds to select and write down the trait--was to be repeated for each page of the booklet. No mention of any memory tests was made.

To ensure that subjects understood the task, they were guided through the three practice examples, with instructions repeated and clarified after each example. Subjects were then paced through the target vignettes. Following completion of the task, the booklets and trait descriptions were collected, and the arithmetic distractor task administered for ten minutes. Subjects were then given a surprise free rcall test in which they asked to write down the gist of every behavioral vignette they could remember having read. Emphasis was placed on trying to include the character, situation, and basic action involved in each vignette. After fifteen minutes, the recall sheets were collected. Subjects were then handed back their personality description sheets and asked to recall the vignettes again--including those that they had already written down on the preceeding test--but to this time use the traits as reminder cues. Fifteen minutes was also given for this cued recall test, after which subjects were debriefed and dismissed.

Results

When subjects read through the vignette booklet, they wrote down on each page the trait that best explained that particular behavior. We first examined the degree to which their choices agreed with the original experimental groupings. Averaging across subjects, the predicted trait adjective was selected 96% of the time, with no subject miscategorizing more than 2 of the 24 behaviors. Thus, subjects usually paired a trait with its expected set of 1, 2, or 4 behaviors, thereby indicating that our set size manipulation was indeed operative.

We next analyzed the recall protocols. An item was scored as having been recalled if the character and basic action were correctly identified. The proportions of vignettes recalled across the three set size categories are shown in Figure 2. Looking at the free recall scores first, we find our predicted effect:

Insert Figure 2 about here

when only two trait categories were available, an average of 2.16 (27% of eight behaviors) was recalled; when four categories were available, the mean rose to 2.48 (31%); and when eight categories were used, 3.36 (42%) of the behaviors were recalled. This overall effect of set size on memory is reliable ($F(2,52) = 8.36$, $p<.01$). Post-hoc tests on the differences between means revealed that the memory increase was most prominent when the number of available traits rose from 4 to 8 (31 vs. .42; $t(26) = 2.44$, $p<.05$); the difference between the 2-trait and 4-trait conditions is not substantial (.27 vs. .31; $t(26) = 0.89$, ns).

Analysis of cued recall protocols yielded similar results. As Figure 2 illustrates, the recall means in the 2, 4, and 8 trait conditions were 2.64 (33%), 3.52 (44%), and 4.72 (59%), respectively. The overall effect is significant ($F(2,52) = 24.0$, $p<.001$), as are post-hoc test differences between the 2-trait and 4-trait conditions ($t(26) = 2.44$, $p<.05$) and the 4-trait and 8-trait conditions ($t(26) = 7.17$, $p<.01$).

Discussion

The major result of the experiment was that memory for a character's behaviors increased as a function of the number of trait categories available for encoding these behaviors. These results support our hypothesis of a set size effect in person memory, i.e., the richer the available personality structure, the easier it is to learn and later remember behaviors relevant to that structure. The basic idea is that a person attempts to understand another person's behavior by relating it to what he already knows and believes about that person. In other words, incoming behaviors are integrated into previously-developed personality impressions. Later retrieval of those behaviors is affected by the personality structure: a well-elaborated personality structure enhances memory by providing many cues for encoding and later remembering behaviors, while an impoverished personality struture tends to produce the Watkins and Watkins (1975) cue-overload effect whereby recall is inhibited by many items being associated to the same cue.

In competition with this general interpretation of the results are several alternate explanations based on experimental artifacts. The first and perhaps most obvious is that the effect resulted from variations in the target items. Specifically, while the number of behaviors in each trait condition was kept constant, the relations among the behaviors varied. The eight behaviors in the 2-trait condition represented only several types of behavior, while the eight behaviors in the 8-trait condition exemplified eight different traits. Thus, behaviors in the former condition were generally more similar to each other than those appearing in the latter condition. Since inter-item similarity tends to increase interference effects in memory (Baddeley and Dale, 1966; Wicken, 1972), the results may thus owe to this confound of item similarity and set size in the experimental design.

Another possible explanation of the results was suggested by the research of Ross (19), who showed that memory for an item increased as a function of the number of different questions asked about that item during the learning phase. In our experiment, subjects were directed to ask themselves, for each behavior in the vignette booklet, which of the character's traits best explained that behavior. In such a procedure, tyhe number of questions asked about a given behavior would increase with the number of traits available as explanatory candidates. This may have been why the different memory rates obtained.

Experiment 2

To test whether these alternative explanations could account for the findings of Experiment 1, we ran a second experiment whose design differed in several major ways from the first experiment. First, inter-item similarity was controlled by using the same set of behaviors in all conditions. To manipulate the set size, we varied the number of traits available to explain those behaviors, such that a given trait was paired with either one or four behaviors. And second, to test the number-of-questions hypothesis we manipulated as a between-subjects factor the number of characters to whom the behaviors were ascribed. For half the subjects the target behaviors were split evenly between two characters--one with two trait, the other with eight traits--while for the other half of the subjects the behaviors were all ascribed to a single character with ten traits (with eight traits having a 1-1 trait-behavior ration, and two traits bearing a 1-4 trait-behavior ratio). The number-of-questions hypothesis would expect recall scores to be higher in the one-character condition, since the greater number of traits ascribed to the single character would lead to more questions being asked during the behavior categorization task. The set size hypothesis, on the other hand, prerdicts (1) no differences between the one- and two-character conditions but (2) reliable differences between the 1-1 and 1-4 trait behavior conditions, with the former ratio producing much better memory than the latter.

Method

Subjects

The sixty subjects who volunteered for the experiment were all students in a large introductory psychology course at Santa Clara University in California. For their participation they received extra credit in their course.

Materials and Design

The experiment employed a 2 x 2 factorial design which manipulated the between subjects factor of number of characters (one vs. two) and the within subjects factor of number of behaviors ascribed to a trait (one vs. four). To develop the materials, we first used Rosenberg, Nelson, and Vivekonanthan's (1968) clustering analysis of personality traits to generate four separate groups of traits:

intelligent/imaginative/scientific/industrious/sociable/popular/
humour/helpful/unhappy/moody/pessimistic/humorless, and
dependent/conforming/submissive/impressionable. Each trait group represented

a different quadrant of the two-dimensional clustering space obtained by Rosenberg et al.'s analysis; members of each group were chosen on the basis of their close proximity to each other in their respective quadrants.

As in Experiment 1, we next generated for each individual trait a situation-action behavioral vignette exemplifying that trait. For example, the "intelligent" vignette described the character receiving all A's on a report card; the "scientific" vignette described the successful completion of a series of complex chemistry experiments; the "industrious" vignette detailed the character resourcefully completing a series of extra-credit assignments; and the "imaginative" vignette depicted the character painting a series of novel and creative pictures. Four graduate students were given the set of vignettes along with the list of sixteen traits and asked to select the best-fitting trait for each vignette. They did this with near-perfect accuracy, thus confirming our intuitions about the goodness of fit between each trait and its correspodning behavior.

These sixteen behavioral vignettes and sixteen traits were then used to form various conditions of the experimental design, as schematized in Figure 3. As can be seen in the figure, the same set of sixteen behaviors was used across all conditions. The within subject factor varied the number of traits available (one vs.

Insert Figure 3 about here

---

four) for explaining each of the four quartets of behavors. For example, assume that behaviors B through the B in the figure represent the intelligent" trait group--intelligent, imaginative, scientific, and industrious. In the condition on the left side of the diagram, only one trait--intelligent--would be available to explain all four behaviors, thus ensuring that all of these behaviors would be linked to the single trait during the learning procedure. In the right hand condition, however, all four traits would be available, thereby enabling a one-to-one mapping from each behavior to its exemplifying trait. As Figure 3 further illustrates, half the behaviors presented to a subject would be linked in a four-to-one mapping with the traits, while the other eight behaviors would form one-to-one links with traits. The assignment of each quartet of trait-related behaviors to a particular mapping condition was conducted in a counter-balanced fashion across subjects.

Figure 3 also shows how the between subjects factor was manipulated. For half the subjects, all of the behaviors were ascribed to one character having ten traits; for the other half of the subjects, eight behaviors from two randomly selected behavior quartettes were assigned to one character having two traits, while the eight behaviors from the remaining two quartets were ascribed to a second character bearing eight traits.

The experimental materials provided to each subjects were thus similar to those used in Experiment 1. Each subject received a personality trait list describing either one or two characters, along with a corresponding vignette booklet containing three practice examples followed by a random ordering of the sixteen behaviors ascribed to the character(s) and three practice examples. In addition, an arithmetic distractor task, a free recall test, and a cued test were used; these materials were essentially identical to those used in the first experiment.

<u>The procedure was also essentially identical to the one followed in</u> Experiment 1. Subjects were run in four groups ranging in size from six to fifteen. The experiment was introduced as a study of how people form and use personality impressions of other individuals. Each subject was then handed a personality trait sheet and a corresponding vignette booklet. Attention was focused first on studying the personality description(s) of the ostensibly

real individual(s) whose traits had been ascertained through extensive personality tests. After five minutes, subjects were told that the vignette booklet described behaviors expressed by the character(s) introduced on the trait sheet. Their general task was identified as reading through the vignettes to develop a good personality impression of the person(s); their specific task was described as deciding and then writing down, for the single vignette listed on each page of the booklet, that trait from their personality description sheet which best explained the character's behavior. It was also explained that their progress would be experimenter-paced: after nine seconds of reading a vignette, they owuld be signalled (with the cue phrase "All right") to select and write down the most explanatory trait; after five more seconds, they would be again cued ("Turn") to turn the page and begin to read the next vignette.

The instructions were reiterated and any questions answered as subjects were through the three practice examples at the beginning of the booklet. After then pacing them through the vignettes, the experimenter collected the booklets and trait description sheets, and administered the distractor task for about five minutes. Next given were the surprise memory tests--free recall and cued recall, with the latter using the personality trait sheets to cue memory. Ten minutes was allowed for each test. Finally, subjects were debriefed, thanked for their participation, and dismissed.

Results and Discussion

As in the first experiment, we first examined subjects' trait selections for explaining the various behaviors. Again, subjects' choices were almost always in agreement with our original pairings; the predicted trait adjective was selected 94% of the time on the average. Thus, our experimental manipulation of set size seemed to be in effect.

In next analyzing recall protocols, credit was given for a recalled item if the character and essential gist of the action were included. The data from both the free recall and cued recall tests, shown in Figure 4 in terms of the proportion of behaviors recalled in the various experimental conditions, were then subjected to 2 x 2 ANOVA tests. (Since there were no reliable differences among the 4 test forms used for

---

Insert Figure 4 about here

---

item randomization, they will not be discussed further here.) Looking first at the free recall results depicted in Panel A, several conclusions become apparent. First, the category size effect was again significant: behaviors were remembered better overall in the one-to-one pairing condition ($F(1,58) = 58.8$, $p<.01$). Second, the number of characters manipulation obviously had no effect ($F(1,58) = 0.39, p>.10$). Finally, there was no interaction between the two factors ($F(1,58) = 0.33$, $p>.10$).

The same pattern of results obtained for the cued recall scores represented in Panel B. The within-subjects factor of set size was again significant ($F(1,58) = 102.23$, $p<.01$), whereas there were no reliable differences in the between subjects groups ($F(1,58) = 0.33$, $p>.10$) or in the interaction scores ($F(1,58) = .01$, $p>.10$).

The results essentially replicate the findings of Experiment 1. They show that the more behaviors attached to a particular trait, the more difficult it becomes to recall one of those behaviors. They also constitute strong evidence against the number-of-question-asked hypothesis suggested as an alternative explanation in Experiment 1. Varying the number of traits the subject had to consider in explaining a behavior did not affect memory; what <u>did</u> have a large effect, though, was varying the number of times a particular trait was actually selected. In other words, of importance was the number of behaviors linked to a trait, <u>not</u> the number of traits ascribed to a character.

The results also provide apparent grounds for rejecting the other alternate explanation offered for the effect in Fxperiment 1, that of differences in similarities among behaviors used in the various set size conditions. We controlled for this in the present experiment by using the same behavior clusters in all experimental conditions, and still found the memory differences. However, this experiment contains a similar and possibly important confounding variable: the goodness of fit between a behavior and the trait selected to explain it. In the condition where four behaviors were mapped onto one trait, the associative strength of a given behavior to that trait might have been weaker than in the one-to-one condition, where the greater number of relevant traits might have enabled selection of a better-fitting trait. For example, one of the vignettes in the <u>unhappy</u>/<u>moody</u>/<u>pessimistic</u>/<u>humorless</u> trait group depicted the character scoffing skeptically while reading in the morining newspaper about various political reform programs. For half the subjects, this behavior appeared in a four-behaviors-to-one-trait condition in which <u>unhappy</u> was the only relevant explanatory candidate; for the other subjects, the behavior was included in a one-to-one mapping condition where all four "unhappy" traits were available, thus enabling <u>pessimistic</u> to be chosen to explain the behavior. Because <u>pessimistic</u> fits better than <u>unhappy</u> in explaining the behavior, it may have served as a better retrieval cue when later recalling the behaviors. If this was the case, superior recall in the one-to-one mapping conditions may have resulted from better trait-behavior pairings, <u>not</u> the number of items linked to a given trait.

Fortunately, our data provided a means to examine this possibility. Specifically, the same trait in a quartette was used every time that the quartette was assigned to the one-to-four mapping condition; e.g., when the "intelligent" behaviors were in the multiple mapping condition, <u>intelligent</u>

was always the single trait available to explain all four behaviors. Thus, we examined memory for the particular behavior in each quartette for which the single trait used in the one-to-four condition was actually the best-fitting explanatory candidate (e.g., the behavior which subjects explained as "intelligent', regardless of whether the other traits in the "intelligent" quartette were available). If our differential memory rates arose from the goodness-of-fit confound, then this particular behavior in each quartette should be recalled equally well in both mapping conditions.

This turned out not to be the case. Collapsing across sixty subjects, the proportions of critical items recalled were .45 (54 of 120 items) in the one-to-four mapping condition and .58 (70 to 120) in the one-to-one condition. These proportions are essentially no different from those obtained for all items in the two mapping conditions (.45 vs .43; .58 vs .61). The same pattern occurred when each critical item was individually compared across mapping conditions, and when the cued recall data was examined. Thus, it seems safe to conclude that the set size effect cannot be attributed to this experimental confound.

Experiment 3

The major result of Experiment 1 and 2 was that memory for a given behavior decreased as the number of other behaviors attached to the same trait increased. This general relationship between set size and memorability is certainly not a novel finding, having been demonstrated previously in a variety of studies in cognitive psychology. Early investigations (e.g., Mathews, 1954; Dallet, 1964; Tulving and Pearlstone, 1966; Mandler, 1967) used free and cued recall tests in showing that memory for a given word in a category decreased as its category size grew larger. Later studies by Anderson and Bower (1973; Anderson, 1974) found that reaction time to verify a proposition was inversely related to the number of related propositions, a result they termed "the fan effect". More recently, variations of the set size effect have been demonstrated by Watkins and Watkins (1975) and Mosocovitch and Craik (1976).

Recently, however, the validity of some of these studies has been challenged. For example, Smith, Adams, and Schorr (1979) and Moesser (1979) pointed out that the far effect--i.e., reaction time slows as number of facts grows--seemed inconsistent with the cognitive processes of an "expert" in a field, who has many related facts stored in memory yet can evaluate a given fact in a relatively quick and accurate fashion. Smith et al. (1979) rectified this apparent "expert's paradox" by showing that the fan effect did not hold when relevant facts could be integrated into a coherent organization. Relatedly, Moesser (1979) showed that the fan effect disappeared under repeated testing. In addition, Moesser found a recognition memory fan effect when materials were presented in a mixed design (the design used in the Anderson and Bower studies), but not when blocked presentations were used. This general result was also reported in free recall studies by Bower (1972), Brown and Underwood (1974), and Postman, Thompkins, an Gray (1978).

One general explanation for this disparity between mixed and blocked design results is that later-item organizations can be formed among category members in blocked presentations, where related items all appear together, but not in mixed presentations, where category members are scattered throughout the learning materials. Because such inter-item integrations can overcome the typical interference and unlearning effects found in larger-sized categories (Smith et al, 1979), the set size (or fan) effect occurs in the mixed condition but is overcome in the blocked condition. This general explanation is discussed more extensively by Smith et al (1979) and by Moesser (1979); other explanations involving reduction of sensory loads and increased probability of displaced rehearsal have been advanced by Postman et al (1978).

Of major importance here is the possibility that the set size effect obtained in the first two experiments is conditional on the way the materials are presented. If this is true, it raises grave doubts about validly generalizing our findings to naturalistic situations of person memory, which generaly involve the use of well-learned and well-integrated personality structures. As such, Experiment 3 tested the possibility that our findings would not replicate under a blocked presentation design. The experiment followed the same design and procedure used in the one-character condition of Expeirment 2, with the exception that half the subjects received the materials in a blocked condition, and half in a mixed condition. Thus, subjects would study ten traits ascribed to a hypothetical person, and then categorize 16 behaviors ostensibly enacted by the character. Half the behaviors formed one-to-one mappings with eight of the traits, the other eight behaviors were best matched in four-to-one mappings to the remaining two traits. Afterwards, subjects were unexpectedly asked to recall the behaviors.

Method

Subjects

Thirty six Stanford University undergraduates participated in the experiment in partial fulfillment of their course requirements for an Introductory Psychology class.

The materials were essentially the same as those used in Experiment 2. There were four quartettes of related traits and sixteen behavioral vignettes, with each vignette best exemplifying one trait in particular. In addition, the materials included an arithmetic distractor task, a free recall test, and a cued recall test.

The two major variablès in the design were set size and type of presentation. As in Experiment 2, the within subjects factor of set size was manipulated by varying the number of traits (one vs. four) available to explain the four behaviors in a quartette. The between subjects factor of type of presentation (mixed vs. blocked) was orthogonally manipulated by giving half the subjects vignette booklets containing a blocked-by-quarteete ordering of the 16 behaviors (i.e., the first four behaviors were all from one trait quartette, the next four from a different quartette, etc.), and half the subjects booklets containing a random ordering of the behaviors. To emphasize the possible inter-item organizations, all subjects received instructions that some of the behaviors were very related, and that they should try to interrelated as many of them as possible.

To ensure proper counterbalancing four test forms were developed. As in the previous experiments, each test form consisted of a personality description sheet, which briefly introduced the character and then listed ten traits purportedly characterizing his personality, and a vignette booklet. Eight of the ten traits on the personality description sheet of a given test form were taken from two randomly-selected trait quartettes; the other two traits were the most general members of the remaining two quartettes. This ensured that half the behaviors would be encoded in one-ton-one mappings with traits, and half in four-to-one mappings. The vignette booklets contained three practice examples followed by the 16 target behaviors, with each of the 19 vignettes appearing on a separate page. In the blocked presentation condition, the behaviors were partitioned into their four respective quartettes, with presentation order of the quartettes randomized across subjects; in the mixed condition, presentation order of the 16 behaviors was randomized for each subject.

Procedure

Subjects were tested in groups of six; test groups were assigned in random alternation to the blocked and mixed presentation conditions. The general procedure was identical to that of Experiment 2. Subjects were given the personality impression cover-story, handed the trait description sheets to study for five minutes, and then paced through the vignette booklets. This time, though, 20 (rather than 15) seconds was given for each page of the booklet (15 to read the vignette, 5 to select and write down the best-explaining trait). In addition, special instructions to try to interrelate the behaviors were emphasized. (These changes in procedure were undertaken to encourage the use of inter-item organizational strategies.) After reading the booklets, subjects were given n succession a five minute arithmetic distractor task, a ten minute surprise free recall test of the 16 behaviors, and a cued recall test involving using the handed-back trait sheets to cue memory for behaviors. Subjects were then debriefed and dismissed.

Results and Discussion

As in the previous experiments, we first examined the degree to which subjects' choices of the best-fitting trait for each vignette agreed with our expectations. The experimental pairings again proved highly reliable: no subject chose an unexpected trait more than once.

We next scored the recall perotocols, giving credit for an item if the basic gist of the situation and action was listed. A 2 (presentation style) x 4 (different test forms) x 2 (within subject set size) was used to analyze the recall data. As expected, the test forms had no significant effects, and thus will not be discussed further. Figure 5 illustrates how recall performance fared under the major factorical conditions of set size and presentation style. Looking first at the free recall data in the left panel of the figure,

---

Insert Table 5 about here

---

an immediate observation is that set size had a major effect: one the average, subjects recalled .45 of the behaviors in the four-to-one mapping condition (.42 in the Blocked condition and .48 in the Mixed condition), and .57

behaviors in the unique mapping condition (.56 in the Blocked condition and .59 in the Mixed condition). This difference between mapping conditions is highly reliable ($F(1,30) = 18.9$, $p<.01$). However, presentation style had no significant impact: subjects recalled about the same amount in both groups (.49 in the Blocked condition vs $F(1,30) = 1.43$, ns), and there was no interaction between set size and presentation style ($F(1,30) = .45$, ns). As the right panel of Figure 5 indicates, this pattern of results was even more pronounced in cued recall: there was a whopping effect of set size ($F(1,30) = 122.72$, $<.001$), no effect of presentation style ($F(1,30) = .96$, ns), and no interaction effect ($F(1,30) = 0.29$, ns).

The results indicate that recall of behavioral information is an increasing function of the number of relevant trait categories available for encoding the behaviors. This suggests that having a well-developed personality structure about a person provides, among other benefits, an organizational structure that greatly facilitates our encoding and retrieval of that person's actions. This conclusion is made cautiously, since the procedure used in the above experiments was restrictive and artificial. One limitation was that the behaviors were always relatable to the trait set, whereas in real life there are many behaviors which are either irrelevant or incongruent with a personality impression. We would expect that the facilitating effects of the personality organization would vary according to the relatability of the behavioral items (cf. Hastie, 1981). Another restruction was that each behavior was fitted to only one trait, whereas typically an action may be categorized along multiple dimensions (e.g., returning a wallet may indicate either <u>kindness</u> or <u>honesty</u>). However, there is no reason to expect that under the latter conditions the effect would not have obtained. Thus, the artificial conditions of the study do not provide sufficient grounds for challenging the validity of the general result. We therefore turn now to an explanation and discussion of the obtained effect.

General Discussion

A Network Theory Explanation of the Results

The major result in the three experiments was that memory for social behaviors increased with the number of traits available for categorizing the behaviors. This "set size" effect is consistent with the well-established finding in cognitive psychology that memory for a category word decreases with an increase in category size (e.g., Tulving and Pearlstone, 1966; Mandler, 1967). It is also consistent with Bower and Gilligan's (1979) earlier work showing that personality traits encoded in terms of a well-differentiated person schemata (e.g., self, mother) were remembered better than traits learned in reference to poorly-differentiated schemata (e.g., Walter Cronkite). It thus appears that the complexity of the personality structure used to organize and udnerstand another person's actions can affect the degree to which we can later recall those actions.

This finding can be understood in terms of contemporary semantic network models such as HAM (Anderson and Bower, 1973) or ACT (Anderson, 1976). The basic elements in such models are nodes which represent concepts and links which indicate associative relationships between concepts. The basic units of thought are subject-predicate propositions. These propositions can express general concepts as well as specific episodic information. For example, Figure 4(?) illustrates a portion of a network structure subjects may have formed to represent one of the characters studied in the experiment. Note how the representation expresses both general statements about the person's personality (e.g., "Bob is kind") and specific behaviors enacted by the person (e.g., "Bob helped the old lady across the street"). Links connecting the general trait categories and specific behaviors are developed as the latter are defined as instances of the former. Thus, "helping the old lady across the street" is associated both with the node representing the actor's name ("Bill") and the best-fitting trait category ("kindness").

In network models, the links between nodes are critical for retrieval for information. Specifically, recall is presumed to involve a process whereby several entry points into the network are first activated. Activation of an ode spreads automatically down its connective links to associate nodes. When a node's activation level crosses a specified threshold level, it "pops" into the person's conscious awareness (i.e., it is recalled). Because the amount of activation is conscious awareness (i.e., it is recalled). Because the amount of activation is limited, spreading activation received by a given neighboring node decreases as the number of other nodes associated to the originally activated node increases. That is, more associated nodes means the spreading activation is mroe dissipated, which decreases the probability that a given associate node receives activation sufficient to prime it into awareness. This is a partial explanation of the "set size" effect discussed above.

This activation process need not be a completely passive one. A person may strengthen or renew the activation level of a node by, say, repeating its name over and over again in his head; e.g., a subject might repeatedly say to hismelf something like, "I know there was something about kindness on that list", thereby continuing to spread activation from the "kindness" and "list" nodes. Similarly, the person might reduce activation by deliberately forgetting about some concept node.

This brief account of network models provides a basis for expalining the results of the present experiments. During the study phase, subjects were given sets of traits describing characters and asked to form personality impressions of the characters. They were then asked to encode behaviors in terms of the developed personality structures. We presume that this resulted in a representation like that illustrated by Figure ___. Thus, a given behavior would become associated to the "list" node, a trait node, and the character's name node. Because of the varying number of traits available for encoding a constant number of behaviors, some traits became linked to only one behavior while others developed connections to multiple behaviors. To explain why this difference resulted in the botained recall differences, the network theory assumes that subjects would initiate recall of the behaviors by activating various entry points into the network representation--e.g., the "list" node and the names of the experimental characters and/or the trait categories--and thus spread activation down their associative pathways. An intersection of activated pathways emanating from the source nodes would activate a target behavior and enable it to be identified and overtly recalled as having been on the list.

We assume for various reasons that the trait names were superior to character names as retrieval cues and were thus relied on more heavily by subjects during recall. First, there were more trait categories than person categories available to the subject. Thus, the number of behaviors connected to a trait was, on the average, less than the number connected to a character. This smaller category size would result in lesser interference and hence better memory. Second, the connections between the character and behaviors were arbitrary, while the trait/behavior linkages were based on meaningful associations (e.g., "returning a wallet" is commonly interpreted as indicating honesty). Third, subjects were instructed to study and elaborate on the trait descriptions before being presented with the behaviors, thus ensuring that the traits were well-learned and available as retrieval cues at time of recall. For these reasons, the traits constituted more powerful retrieval cues.

Given the importance of the traits as recall cues, it is not surprising that varying their number affected recall. specifically, the more behaviors attached to a trait, the less spreading activation each behavior would receive form the activated trait node. Also, increasing the number of pathways emanating from a node tends to reduce the probability that a subject would search down a given pathway. Thus, increasing the number of behaviors linked to a trait node would lessen the changes that a given behavior would be recalled.

<u>General Comments on Person Memory Structures</u>

The network model of person memory assumes that knowledge of another person is structured in memory in temrs of a unified representation (cf. Asch, 1946). Such a representation would undoubtedly be more complex than the simple example illustrated in Figure 4; it probably would include other information such as physical appearance, interests, hobbies, acquaintances, occupation, etc. It seems likely that such representations might also include references to general personality prototypes or stereotypes (Cantor and Mischel, 1977). For example, I might decide that Bill is a typical extrovert, and thus link a pointer from my representation of him to my general prototype representation of an extroverted personality.

The above experiments suggest that such complex organizational structures enable behavioral information to be more easily understood and remembered. Besides this major value, person memory structures seem to have other important functions (see Taylor and Crocker, 1981). For example, they permit organization of a large body of information; they enable prediction of the person's behavior; they enable inferences and evaluations to be made; they suggest appropriate response strategies; etc. But such structures also have potential disadvantages. They may bias interpretation of behavior (Carlston, 1980; Higgins, Rholes, and Jones, 1977; Bower, 1977); they may sustain stereotypes (Hamilton, 1979; Hamilton and Gifford, 1976; Bilig and Tajfel, 1973); they may lead people to makme what Ross (1977) calls the fundamental attribution error of attributing someone's behavior to underlying dispositional qualities (rather than situational variables); and they may distort memory of a person's actions ( ). Thus, person memory structures seem to be an invaluable but double-edged tool for interacting with other people.

This is not to say that person memory structures will always be used to organize experience about another person's behavior. One situation when they are not used is when the person is unfamiliar to us (Ostrom, Pryor, and Simpson, 1981). Another situation may be when situational factors appear to dictate a person's behavior (Bem, 1972; Jones and Davis, 1965; KMelley, 1967; but see Ross, 1977). A third situation is when the explicit processing goal is something other than impression formation. For example, Hamilton (1981) found that recall of behaviors about a person was clustered in terms of trait-related behaviors when Impression Formation instructions were given, but <u>not</u> when Memory Set instructions were issued; similarly, Ostrom, Pryor, and Simpson (1981) found that recall of information about multiple persons was clustered according to individual persons when subjects were forming impressions of the people but not when they were simply trying to memorize the presented information. A fourth situation in which person memory structures may not be used is when the behavioral information is not relevant to our available structures. For instance, Markus (1977) found that subjects who did not "define themselves" in terms of certain traits--i.e., they considered such traits irrelevant to their representations of themselves--seemed to process behaviors related to such traits independently from their "self-schemas".

Even when the personality representation of a specific person is available and appropriate, other types of schematic structures might be used in combination or as alternatives (Taylor and Crocker, 1981). These additional social schemata might include <u>general person prototypes</u> (Cantor and

Mischel, 1977), _self-schemata_ (Markus, 1977; Bower and Gilligan, 1979), _event schemata_ which specify wello-practiced behavioral scripts (Schank and Abelson, 1977; Bower, Black, and Turner, 1979), _role schemata_ for various occupations (e.g., fireman, psychologist) or social roles (wife, parent, radical, etc.), and _causual schemata_ for understanding the motives, intentions, and goals underlying behavioral actions (cf. Schank, 1975). In the complex activity of social perception, it seems likely that all of these types of structures are used to some extent.

The point to be made from this is that person memory structures occupy a critical but not exclusive role in how we understand another person's behavior. The present experiments demonstrated that the complexity of the structure used to encode behaviors was related positively to how well the behaviors were later recalled. While this finding is certainly not startling, it does provide an explanation for earlier results by Bower and Gilligan (1979), Lord (1981), and Rogers, Kuipers, and Kirker (1977) showing that information related to the self (or, in the Bower and Gilligan experiment, to one's representation of one's mother) was remembered far better than information encoded learned in relation to lesser- _differentiated_ representations. Further research is needed to determine if the obtained effect can be generalized to other types of social cognition structures as well.

Reference Notes

1. The figure provides only rudimentary information. A full network representation would include other information such as the time and location of the encoded event and contextual information (e.g., the person's emotional state, thoughts during the experimental task, instructions, etc. For a more complete representation, see Anderson, 1976.)