G. A. RADVANSKY

PH D

1992

This is to certify that the

dissertation entitled

Recognition, Recall, and Mental Models

presented by

Gabriel Allen Radvansky

has been accepted towards fulfillment
of the requirements for

PhD degree in Psychology

Rose T. Zacks
Major professor

Date August 6, 1992

RECOGNITION, RECALL, AND MENTAL MODELS

By

Gabriel Allen Radvansky

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1992

ABSTRACT

RECOGNITION, RECALL, AND MENTAL MODELS

By

Gabriel Allen Radvansky

Previous experiments investigating the influence of mental model organizations on memory retrieval have concentrated on recognition methods. The experiments reported in this dissertation assess the utility of recall methods for investigating such organizations. Experiments 1 and 2 found that recognition showed differential organization patterns based on information content (location-based for object-location facts, such as "The potted palm is in the high school, and person-based for person-small location facts, such as "The farmer is in the telephone booth") whereas free recall consistently showed only location-based organization. Experiments 3-5 explored factors that could contribute to the recall-recognition dissociation. Experiment 3 used cued recognition and found partial evidence in support of the earlier recognition results. Experiment 4 found a location-based bias in explicit organization. Finally, in Experiment 5, cued recall supported the recognition data. Voice onset times showed evidence of the presumed mental model organizations, although production times were not definitive. In general, Experiments 3-5 suggest that memory measures that use direct memory access more accurately reflect the underlying mental representation's structure than tasks that are prone to reflect the structure of a retrieval plan.

ACKNOWLEDGMENTS

I would like to extend my thanks to those faculty members at Michigan State University, especially Barbara Abbott, Galen Bodenhausen, Tom Carr, Rich Hall, Bill Hartmann, Les Hyman, Tony Nunez, Brad Rakerd, and Jim Zacks, who have helped me progress towards becoming a competent researcher, with special appreciation to Rose Zacks for being an outstanding advisor throughout my five years at Michigan State. I would also like to thank my graduate student colleagues for their discussions on various topics. Much appreciation also goes to Zenzi Griffin, Dan Spieler, Christy Stewart, and Debbie Van Wieren for their assistance in the data collection. Finally, many, many thanks go to Laura Carlson Radvansky for her support, understanding, and patience throughout the progress of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The work reported in this dissertation extends a series of studies investigating the role of mental models (Johnson-Laird, 1983) in memory retrieval. Specifically, it concentrates on whether recall measures can be used to assess the mental model theory of knowledge organization, and retrieval. The structure of this dissertation is as follows: First, a brief review of the mental model literature is presented, with an emphasis on those aspects relevant to the current work. Following this, a more in-depth review of completed experiments investigating mental models in fact retrieval is given to illustrate the successes of the mental model view in predicting how facts are organized in memory and how this organization affects retrieval. Next, the SAM model of memory (e.g., Raaijmakers & Shiffrin, 1981) is briefly described and adopted as the framework for considering the processes involved in the recall and recognition of information stored in mental models. After this exposition, a series of experiments exploring the access of information in mental models through recall and recognition is presented[1].

[1] Although the thrust of this dissertation is a comparison of the differential properties of recall and recognition with respect to the retrieval of information from mental models, this was not always the case. Originally the focus of this dissertation was to test whether the organization of mental models is guided by properties other than spatial

Mental Models

Mental models (Johnson-Laird, 1983) are representations of specific situations involving real or possible states of the world. Each situational representation is assumed to derive from, and to represent, the functional relations between elements in the world such that "its structure corresponds to the structure of the situation that it represents" (Johnson-Laird, 1989, p. 488). A mental model is an analogue to certain aspects of the world in the head serving as a "high grade simulation" (Johnson-Laird, 1983, p. 4). Mental models are updated as subsequent information about a situation is gained. This includes either the integration of new information into the mental model (e.g., Ehrlich & Johnson-Laird, 1982) or the removal of information that is no longer a part of the situation (e.g., Glenberg, Meyer, & Lindem, 1987). A person's general world knowledge serves to guide this updating process. Any new information that is encountered that cannot be integrated into, or is irrelevant with respect to, any existing mental model will be used to create a new and separate mental model.

Although a complete definition of exactly what a mental model is has not been fully developed (Glenberg, et al., 1987;

---

relations, in particular, temporal and abstract relations, such as ownership. This was to be done within the general framework of situations being interpretable as either states-of-affairs or courses-of-events (Barwise & Perry, 1983). The method for testing these notions was to be free recall in which the order of item recall was to reflect the internal organization of the mental model. In particular, course-of-events situations were expected to be more unidirectionally structured, as though being based on a timeline. It was during the initial stages of this investigation that it was discovered that recall and recognition demonstrated basic differences in the observed organizational patterns. As a result, the course of this dissertation was altered to pursue this difference.

Johnson-Laird, 1989), the mental model construct is useful in helping to understand how information is organized in memory. This theoretical construct can then be used to make predictions about how the mental model organization affects various cognitive processes such as knowledge integration (Ehrlich & Johnson-Laird, 1982), inference making (Byrne & Johnson-Laird, 1989), and fact retrieval (Radvansky & Zacks, 1991).

Research on Mental Models

A great deal of mental model research in memory and language comprehension has centered on how the relational structure of the mental model affects learning, memory and retrieval. For example, I (Radvansky, Spieler, & Zacks, in press; Radvansky & Zacks, 1991) have studied how general aspects of organization affect the access of information from mental models containing spatial information. Other researchers investigating mental models have been concerned with the access of information based on the described spatial position of a protagonist in a passage (Albrecht & O'Brien, 1991; Bower & Morrow, 1990; Franklin & Tversky, 1990) or the distance between different objects mentioned in the text (Glenberg, et al., 1987; Morrow, Bower, & Greenspan, 1988; Morrow, Greenspan, & Bower, 1987; Wilson, Rinck, McNamara, Bower & Morrow, 1992). In general, mental models do show spatial organization effects which incorporate properties such as distance, environmental structures, such as rooms in a building, and a protagonist's path of travel, so long as the measurement task encourages taking the perspective of the protagonist (Wilson et al. 1992).

Other studies have looked at how the arrangement of objects in a description influences the creation of mental models and the access of information from them. In general, different types of descriptions of the same relational structure, such as a route or survey description of a fictional town, lead to equivalently structured mental models (Perrig & Kintsch, 1985; Taylor & Tversky, 1992). However, the amount of effort needed to construct a mental model is mediated by the degree of referential continuity between successive statements in a description (Ehrlich & Johnson-Laird, 1982; Mani & Johnson-Laird, 1982), the number of possible arrangements of objects (Byrne & Johnson-Laird, 1989) and the correspondence with an external abstract reference, such as a diagram (Glenberg & Langston, 1992).

Once the mental model has been created, the functional relations between the protagonist and the other objects can serve to guide information access from either a single perspective (Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990; Sharp & McNamara, 1990), or an alternative perspective when a switch in point of view is needed (Franklin, Tversky, & Coon, in press). For instance, people are able to access information about which object is directly in front of a story protagonist faster than information about which object is to the left of the protagonist. This is because a person is more likely to interact with things that are directly in front of him/her, than with some other object.

Differing relations among entities can guide not only the creation of spatial mental models, but also temporal mental models. There has been some research on how temporal distances affect the structure or organization of information (Anderson, Garrod & Sanford, 1983; Sanford & Garrod, 1981). In particular, people are more likely to consider that a large time change is consistent with more than one situation, and that a short time change is consistent with different aspects of a single situation. However, mental models conveying temporal relations may not necessarily incorporate spatial aspects as well. Franklin et al. (in press) presented subjects with narratives about a person being in a location at two different times (either 1 or 6 hours apart). After reading a narrative, subjects were asked to indicate which object was in a given direction from the person at a certain time. Previous experiments that did not use different time periods (e.g., Franklin & Tversky, 1990) demonstrated spatial direction effects. Subjects in this experiment identified objects from the first time period faster than the second, indicating a temporal ordering; however, this occurred in the absence of spatial direction effects, suggesting that a single, temporally organized mental model was created which lacked spatial characteristics.

A major claim of the mental model view is that mental models represent what a text is about, its gist, rather than the text itself. For example, when people attempt to select between a previously encountered sentence and a distractor which describes the same situation, they select the distractor more often, relative to conditions in which similarly altered distractors describe different situations (Bransford, Barclay, & Franks, 1972; Garnham, 1981; 1987; Radvansky, Gerard, Zacks, & Hasher, 1990). For example, people are more likely to confuse sentences such as

"The judge got his contact lenses from the optician" and "The judge got his contact lenses at the optician's", than sentence pairs such as "The judge received a telegram from the optician" and "The judge received a telegram at the optician's". This is because people are unlikely to receive telegrams from other people when they share a common location as part of a single situation.

In general, this research has provided evidence for mental models representations which capture the functional relations among entities in various types of situations. The structure and organization of these representations are separate from the structure and form of the original input information. This can be seen by the variety of ways in which information access is mediated by mental model structure. What is of central concern for this dissertation is how these mental model organizations are revealed in the process of information retrieval for recall and recognition.

Previous Experiments on Mental Models and Fact Retrieval

The experiments described in this dissertation stem from a series of studies investigating the role of mental models in fact retrieval (Radvansky et al., in press; Radvansky & Zacks, 1991; Zacks, Radvansky, & Hasher, in progress). The initial impetus of this investigation was a consideration of whether the structure of the representation used in fact retrieval more closely resembled a mental model or a propositional network, such as in the ACT* model of memory (Anderson, 1983), and how the memory organization affected the retrieval process. This comparison of propositional and mental model representations is a common theme in a number of discussions of the plausibility and benefit of considering mental models as a representational form (Ehrlich & Johnson-Laird, 1982; Johnson-Laird, Hermann, & Chaffin, 1984; Mani & Johnson-Laird, 1982; Perrig & Kintsch, 1985; Sharp & McNamara, 1990). In general, the data suggest that people prefer to organize the information into mental models and to use these representations during fact retrieval rather than build and use propositional networks.

The paradigm used for the comparison of propositional network and mental models views in this series of experiments is the fan effect. This paradigm was initially used because it has been a main source of supportive evidence for the propositional network view. The fan effect itself is an increase in retrieval time or error rate on a recognition memory test accompanying an increase in the number of associations with a concept in the memory probe (Anderson, 1974). To illustrate, consider the following set of sentences:

1. The potted palm is in the hotel.
2. The potted palm is in the barber shop.
3. The potted palm is in the public library.
4. The cola machine is in the airport.
5. The welcome mat is in the airport.
6. The fire extinguisher is in the airport.
7. The bulletin board is in the city hall.

This set of sentences can be divided up into three cases. In the first case (Sentences 1 to 3) there are three locations associated with a single object (the potted palm). This case is referred to as the multiple location (ML) condition. In the second case (Sentences 4 to 6) there are three objects associated with a single location (the airport). This case is referred to as the single location (SL) condition. In the third case (Sentence 7) there is only a single associate for each concept. This case is consistent with both the SL and ML conditions and can be used as a baseline for comparison.

According to the propositional network view, retrieval time for any of the facts from either of the first two conditions should be slower than for the third. This is because three associations need to be searched for the first two cases, compared to only one for the third. The propositional network view predicts no fan effect differences between the SL and ML conditions. A fan effect should occur as long as there are no pre-experimental relationships among the items, such as thematic grouping, that would cause them to be organized together (e.g., Smith, Adams, & Schorr, 1978). In those cases where there is a thematic grouping, the fan effect would be attenuated or eliminated. The basic notion is that information which can be integrated into a single unit does not produce retrieval interference; interference is produced when there are several distinct and related groups of information competing for retrieval (Moeser, 1979). The present experiments rely on the notion that integrated representations do not show a fan effect whereas independently stored representations do.

For the mental model view, the basis for organizing a set of facts into an integrated representation is the situation.

Consistent with the Smith et al. (1978) and Moeser (1979) positions, the mental model view predicts a fan effect only when related information cannot be organized and integrated into a single representation. In the ML condition, it is unlikely that a potted palm in several locations constitutes a single situation. Instead, each sentence seems to describe a different situation. So, a separate mental model would be constructed for each of the locations associated with the potted palm. While these mental models are separate, they are all related because they all contain the potted palm. During fact retrieval, the appropriate mental model needs to be distinguished from those mental models that are related to it. Consequently, the related and irrelevant mental models interfere with the retrieval of the appropriate one, thus producing increases in retrieval time and errors. This is the fan effect.

Again, consistent with the Smith et al. (1978) and Moeser (1979) positions, the mental model view predicts no fan effect when information can be integrated into a single representation, provided there is no need to check the individual members of the representation (Reder & Anderson, 1980). In the SL condition, a situation in which several objects are in the airport can be easily conceived, so these facts would be integrated into a single mental model. During fact retrieval, there are no other related mental models because all of the concepts are contained in the one, so there is little or no interference, and retrieval times and error rates should remain constant regardless of the number of items incorporated in the model. There is no fan effect.

A question that might be asked is why one would expect no fan effect, rather than a smaller fan effect, for the SL condition? After all, a more complex mental model should take longer to search than a simpler one. At this point in time there is no evidence directly related to this issue, but a few hypotheses can be put forward as a preliminary explanation. One hypothesis is that the mental models created in these sorts of experiments are not very complex, containing four or fewer items. As a result, a search of any particular model can be accomplished rather rapidly.

Another hypothesis has to do with the fact that subjects are presented with sentences in which the content words (the subject and predicate) vary but the relation among them does not. As a result, only the content information is needed for accurate retrieval. Several researchers have found that content information is available prior to relational information (Anderson, 1975; Dosher, 1983; Dyne, Humphreys, Bain, & Pike, 1990; Gronlund & Ratcliff, 1989; Ratcliff & McKoon, 1982; 1989). For example, Ratcliff and McKoon presented subjects with sets of simple sentences, such as "Bill hit Mary". They found that if subjects in a recognition test were given a response deadline, at short deadlines subjects were just as likely to accept sentences such as "Mary hit Bill" in which the content concepts remained the same and only the relational information was wrong. Only at a longer deadline (after 700 ms), were subjects able to accurately differentiate between the two different sentence types.

It could be that during the retrieval of information from mental models, only the earlier available content information would be needed to verify the fact. If both concepts were part of the same representation, a positive response could be made early because it could be safely assumed that the relation was the appropriate one. Thus, no internal search of the mental model would be required and no fan effect would be produced in the SL condition. If the relations were varied, then a fan effect of some magnitude would arise for the SL condition. However, it is the difference in the fan effects for the ML and SL conditions that is the critical comparison in this dissertation.

This difference between SL and ML fan effects has been demonstrated repeatedly, suggesting the use of mental models (Radvansky, et al. in press; Radvansky & Zacks, 1991; Zacks, et al., in progress). In the most basic experiment (i.e., Radvansky & Zacks, 1991, Exp. 1) each subject memorized a list of sentences about objects in locations, such as "The cola machine is in the public library". They were then given a speeded recognition test. The results have shown that a fan effect occurs for the ML condition but not for the SL condition. Presumably, the locations defined the situations. Several factors could contribute to the location-based organization. The first is the animacy of the sentence subject. The sentence subjects used were inanimate objects which typically do not move from place to place, making it implausible that they would be the basis for organizing mental models. The second is the nature of the predicates, which were locations which provide a plausible source for organizing the information into mental models. The third is the relation between the object and the location, which was one of containment rather than the object being shipped to the location, missing from the location, or some other relation. This reinforces the suggestion that location-based mental models are to be created because the location readily provides the framework for the containment.

The tendency to organize mental models around locations is very reliable. It occurs despite instructions to organize by other means (Radvansky & Zacks, 1991, Exp. 2), with the use of either definite and indefinite articles for the object or location concepts (Radvansky et al., in press), in both younger and older adults (Zacks et al., in progress), and when the sentences are presented with the location serving as either the sentence subject or predicate (Radvansky & Zacks, 1991, Exps. 1 & 3).

Despite the large tendency to create location-based mental models, this is not the exclusive organizational pattern found. If other situation types are possible, other types of mental models may be created. One simple distinction between situation types is between states-of-affairs and courses-of-events (Barwise & Perry, 1983). A state-of-affairs is a situation that occurs at one spatial-temporal location and is static within that location. As long as there are no changes to the individuals' properties and relations to one another, the state-of-affairs remains the same. A course-of-events is a set of states-of-affairs that include changes in portions of the situation which extend in a sequence across several time periods, provided that other aspects of the situation remain invariant. Such an invariant could be a foregrounded individual who has a common purpose throughout the sequence. For example, descriptions of a person going to the bank, then to the cleaner's, and finally to the grocery store could all be construed as part of a course-of-events involved in a "running errands" situation.

Radvansky et al. (in press) found that when people (e.g., "The lawyer") were paired with locations (like museums, hotels, barber shops, etc.), there was no clear organizational preference. Instead, there were fan effects for both SL and ML conditions. This presumably occurred because subjects were forming mental models that were location-based some of the time and person-based the rest of the time. For example, consider the sentences "The lawyer is in the hotel", "The lawyer is in the barber shop", and "The lawyer is in the museum." For a location-based organization, each sentence could be a separate state-of-affairs, leading to the creation of separate mental models. For a person-based organization, these sentences could be integrated into a course-of-events of the lawyer going from one place to another, such as starting out at the hotel, then going to the barber shop (for a haircut) and finally ending up at the car dealership as part of his travels for the day. So, a set of facts about people being in locations could be organized and represented in two different ways.

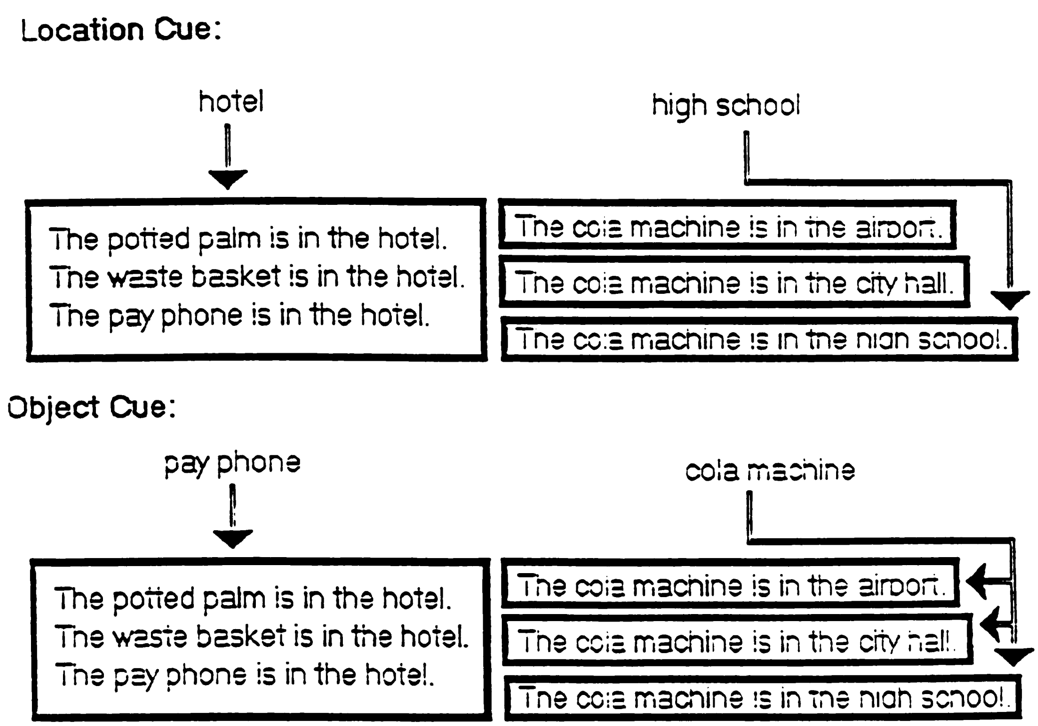
In a second experiment by Radvansky et al. (in press), the study sentences were modified so that the locations were small ones that typically contain only a single person, such as phone booths, confessionals, and bathrooms on Greyhound buses. Through this manipulation, location-based organizations were rendered less plausible, because several people do not tend to occupy these locations as part of a single situation. This leaves the person-based organizations as the more viable choice. In this study, a fan effect was found for the SL condition but not for the ML condition, reflecting a person-based organization.

Finally, there has also been some investigation of whether mental models can be selectively retrieved (Radvansky & Zacks, Exp. 3, 1991). For object-location materials, if subjects are precued on the recognition test with the location concept 1s prior to the probe sentence, then the fan effect in the ML condition is eliminated. The assumption is that the location precue selects the mental model needed to verify the probe sentence, thus facilitating the subjects' reaction times. Object concept and neutral precues (the word "READY") were unable to provide this sort of benefit because they do not select a specific mental model.

For example, consider the situation illustrated by Figure 1. The case where the probe sentence is "The pay phone is in the hotel" is an SL condition trial. Providing either "hotel" or "pay phone" as a precue does not change the pattern of data because there is already little to no interference in this condition. The case where the probe sentence is "The cola machine is in the high school" is an ML condition trial. Providing "high school" as a precue allows for the early selection of the high school mental model before the city hall and airport mental models begin to interfere. Providing "cola machine", however, does not allow for the selection of a single mental model prior to the presentation

Figure 1

Selective activation of location-mental models for single and multiple location conditions for both location and object cue conditions.



of the probe and interference effects are still seen.

Insert Figure 1 about here

In summary, these experiments suggest that mental models are used in fact retrieval when the information can be organized into situational representations and that the fan effect is a reliable method for assessing this organization. Therefore, the fan effect may now also be used to help discover the conditions under which mental model organizations are created, rather than just to discriminate between propositional network and mental model theories.

Recall and Recognition

Although recognition and recall are both considered direct (as opposed to indirect) memory tasks (e.g., Richardson-Klavehn & Bjork, 1988) because they require conscious recollection, they do exhibit clear differences. Several models of human memory might be considered to account for these differences (for a review, see Raaijmakers & Shiffrin, 1992); however, the framework of the SAM (Search of Associative Memory) (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981) model is adopted to provide an account for how information is retrieved from mental models in recall and recognition.

Assumptions of SAM that are applicable in the current setting are that (a) information is stored in separate traces, (b) recall and recognition involve different sets of processes, (c) recognition is a complex direct access process that involves the activation of large sets of traces, (d) recall involves the creation of a retrieval plan, as well as (e) a memory search. This set of assumptions of the SAM model allows for the most direct interpretation of a model of memory storage and retrieval for the mental model approach.

The separate trace assumption is appealing because each mental model can be considered a separate trace. The composition of each trace is assumed to be the set of features which characterize the individual event. Although a mental model does not necessarily reflect an actual episode in a person's experience, it represents a particular situation. Therefore, a mental model trace would be comprised of features representing the elements of a situation and the relations among them. Each trace is assumed to be stored as a semi-independent unit -- semi-independent because although they can be clearly identified as distinct units, there also are assumed to be various associative relations among similar mental models. These associative relations among models can be structured as networks, hierarchies, or other types of relational structures. Finally, the memory representation that is returned to working memory is an exact copy of the trace stored in memory, rather than a composite of several traces, as is assumed in other memory models (e.g., Hintzman, 1986).

In SAM, recognition and recall rely on the same general retrieval mechanism, although in different ways. Recognition involves the direct access to the appropriate memory trace based on the strength of the association between the recognition probe and the memory trace. Recall, however, involves a sequential probabilistic search through various memory traces that demonstrate a reasonably large probe-to-trace association strength. The probability that an individual trace is retrieved is a function of not only the associative strength between the retrieval probe and the memory trace, but also the set of features comprising the memory probe and the presence and strength of the intertrace associations. In both recognition and recall, the activation of individual traces occurs through probe-trace associative strengths as well as intertrace associations, such that related traces are activated as they would be in a network model.

Recognition in SAM is accomplished through a memory probe that is composed of target features. This memory probe samples long-term memory in a content addressable fashion. Contextual features are also encoded into the memory probe, but are ignored because only minimal contextual variation is assumed here since the subject is in the same experimental setting and only a single list of items is being used. The memory probe activates the entire set of memory traces which share features in common with the probe. As such, SAM employs a global matching process with the memory probe being globally compared to all traces in long-term memory. When more than one memory trace is activated, the appropriate trace must be differentiated from the others. As the number of related traces increases, it becomes more difficult to select the appropriate trace amidst the competing activation of other related traces. In cases where the activation level of the target trace is too close to threshold to make an unambiguous decision, a search process similar to recall may be engaged (e.g., Gillund & Shiffrin, 1984; Atkinson & Juola, 1973). Although SAM does not address the issue directly, it can be assumed that increased numbers of competing traces would lengthen the retrieval time.

Recall in SAM is a more involved process. First, a retrieval plan needs to be created or accessed from earlier recalls. This retrieval plan guides the search of long-term memory, limiting it to different parcels of long-term memory at different portions of the search. The retrieval plan organization may or may not correspond to the organizational structure of traces in long-term memory. The memory probes generated by the retrieval plan are used to activate the relevant traces in long-term memory. All of the related traces are activated by the probes, sequentially sampled and evaluated for relevance to the task. The sampling process is probabilistic, so any given trace may or may not be retrieved during a given memory search, and certain traces may be sampled more than once. Once a trace has been successfully recalled, the probe-to-trace association strength is increased, making it more likely that the trace will be retrieved by similar probes later.

Recall, Recognition, and Mental Models

In terms of mental models, recognition is a direct access process in which the memory probe directly activates the relevant mental models. The time it takes to retrieve the appropriate mental model is affected by whether there are other related mental models which are also activated by the memory probe, and, hence, compete for retrieval. If there were, this would bring about the fan effect seen in previous recognition experiments (e.g., Radvansky & Zacks, 1991). Recognition should be unaffected by the need to create a retrieval plan or a search process that would involve the sequentially sampling and retrieving of mental models until the appropriate one has been retrieved.

For recall, however, a retrieval plan needs to be created. The structure of the retrieval plan could correspond to the structure and organization of the stored mental models, although this is not necessarily true. For instance, the structure of the retrieval plan could be guided by the structure of the original presentation of the information, such as the text structure. In one study, Perrig and Kintsch (1985) found that although different types of descriptions of the layout of a fictitious town did not affect recognition, recall was superior when given a route description rather than a survey description. This presumably occurred because the route description possessed the more coherent temporal structure. In any case, the retrieval plan would serve to guide the retrieval of the mental models from long-term memory.

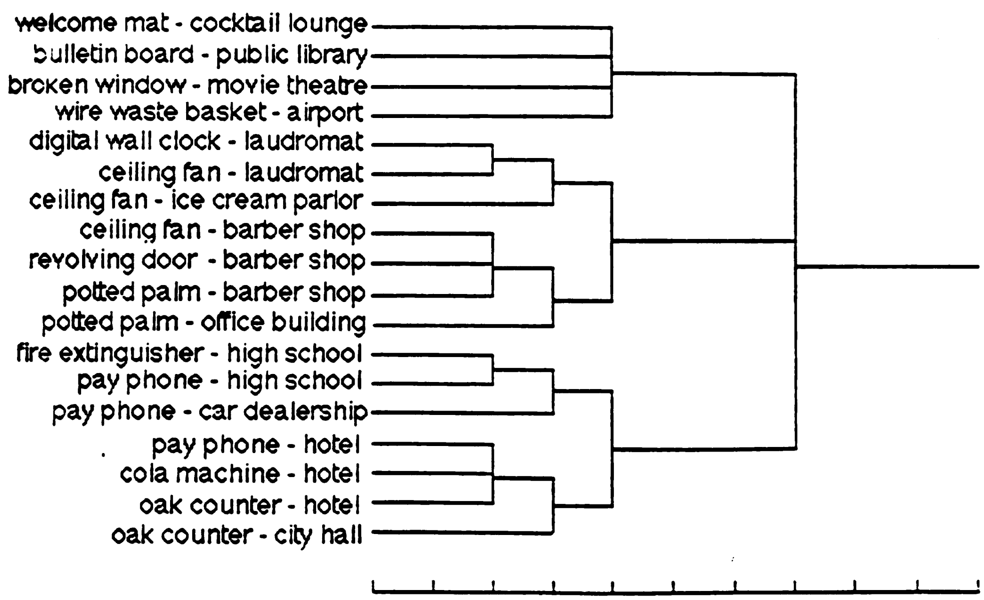
Aside from the structure of the retrieval plan, there are consequences of mental model organizations on the order in which information is recalled that can be anticipated. Mental model organizations could have two influences on free recall information output order. First, for mental models representing the integration of several facts, all of the information contained in that mental model would be expected to be output once it has been retrieved into working memory. For example, consider that a person created a mental model of the barber shop with a revolving door, potted palm, and ceiling fan in it. During recall, when the barber shop mental model is retrieved, all of the facts about the objects contained in the model would be output at that time.

The second influence of mental model organization on recall is based on the notion that retrieval probes activate a complex of memory traces rather than just a single one. After a particular mental model has been retrieved, those mental models that were related to the previously retrieved one would be at a higher activation level for the next memory sampling, and would be more likely to be retrieved. For example, if the office building mental model also contained the potted palm, it would be at a higher activation level following the retrieval of the barber shop mental model, and would be more likely to be retrieved next. This would result in a recall order pattern organized in terms of individual mental model organizations at the lowest level, and in terms of related mental models at the next highest level.

Figure 2 presents an illustration of the sort of pattern of recall organization that can be expected given the above description. The assumption here is that object-location facts are being recalled and that the facts have been stored in location-based mental models. At the lowest level of the hierarchy are sentences which refer to common locations. This would presumably occur because all of the facts represented by a single mental model would be output when that mental model has been retrieved. At the next level in the hierarchy, facts which

Figure 2

Potential hierarchical structure from the free recall of object-location facts stored in location-based mental models.



refer to common objects are also grouped together. It might be noted that Figure 2 also reflects the assumption that facts from cases where there is only a single associate for both the subject and predicate are also be grouped together simply on the basis in their similarity in the number of associations for each of the facts. Finally, at the highest level of organization, all of these larger units are structured together since they come from the same set.

Insert Figure 2 about here

In summary, the differences in recall and recognition of information stored in mental models should reflect differences in the way that information is located and retrieved from long-term memory. Recognition is a direct access process which shows interference effects only from related memory models, not from the structural complexity within a mental model. Recall is a search process in which output order is guided by mental model organizations with information from within the same mental model being reported together at the lowest level and information from similar mental models being recalled together at the next highest level.

Current Experiments

All of the previous experiments in the present series relied on recognition to assess mental model organization. The present experiments tested whether the mental model organizations revealed by recognition are also found with recall. This evaluation is done within the framework of the SAM memory model. In general, it is expected that recognition will show retrieval interference only when there are several competing mental models, but not from within mental model complexity. Recall is expected to show patterns of organization with the lowest level of organization being from within mental models, and the next highest level being from sets of related mental models.

Experiment 1 compared recognition and recall measures with facts about objects being in locations (e.g., The ceiling fan is in the barber shop) and demonstrated what initially appeared to be consistent evidence of mental model organization for both retrieval tasks. The recall measure was an evaluation of consistencies in item order across repeated free recalls of materials involving objects and locations.

Experiment 2 attempted to replicate Experiment 1 with facts about people being in small locations (e.g., The farmer is in the greyhound bus's bathroom). Previous experiments have shown that materials of this type elicit person-based organizations. To anticipate the results, there was an organizational dissociation between recognition and recall in Experiment 2. In particular, recognition supported a person-based organization, whereas free recall and, to a lesser extent, a card sorting task, showed a location-based organization.

Experiments 3-5 assessed which, if either, of these retrieval processes revealed the underlying organizations, which are assumed to be mental models in the present case, and tried to provide an explanation of the dissociation. Experiment 3 attempted to reinforce the integrity of the recognition test results through cued recognition. Experiment 4 assessed organizational preference in the absence of memorization and memory tests. Experiment 5 tested the notion that the organizations in free recall are due to the organization of the retrieval plan rather than the underlying representation. The presentation of these experiments is followed by a discussion of their implications, along with suggestions for further studies to complete this assessment of the differences between the recall and recognition of information stored in mental models.

CHAPTER II

EXPERIMENT 1

Experiment 1 tested whether mental model organizations could be detected through repeated free recalls. There is a long history of the use of free recall to assess how people organize information apart from any organization that may have been present when the information was initially learned, such as presentation order (e.g., Bousfield, 1953). The basic idea is that items grouped together in memory tend to occur in close proximity during recall.

There are two general classes of organization measurement in free recall data: categorical clustering and subjective organization. For categorical clustering, the experimenter defines the organizational form. This organizational form could encompass such things as including items from different categories in the study list and then assessing the degree to which a subject's recall conforms to the selected categories. Categorical clustering analyses require only a single trial to provide an estimate of the degree to which the organization under investigation is adopted. The present experiments rely on the ARC measure (Roenker, Thompson, & Brown, 1971) because it is the one of the most widely used and accurate methods of measuring categorical clustering (White & Kelly, 1978). The ARC measure is based on the following formula: ARC = (O - E) / (maxR - E), where O is the number of observed consecutive repetitions of items from within the same category, E is the number of item repetitions that can be expected by chance, and maxR is the maximum number of possible item repetitions. ARC values closer to 1.00 reflect greater amounts of categorical clustering, values closer to 0.00 reflect less clustering, and values less than 0.00 reflect less than chance clustering. Mental model organizations are considered here to be categories. In those cases where different organizations are possible, different scores were calculated. For example, in Experiment 1, different ARC scores were calculated for location- and object-based organizations (i.e., ARCloc and ARCobj).

For subjective organization, the experimenter does not select the organizational form, but instead attempts to reveal it by relying on consistencies in a subject's recall across a number of trials (Shuell, 1969). Specifically, items which tend to be consistently recalled together are assumed to be stored together in memory. The greater the degree of consistency across recalls, the greater the subjective organization. The present experiments employ two measures of subjective organization. The first is an adaptation of the ARC score developed for the assessment of sequential consistency and is referred to as ARC′ (Pellegrino, 1971). ARC′ essentially provides a measurement of the degree to which the pattern of recall on trial *t* is matched on trial *t+1*. The basic formula for ARC′ is the same as for the ARC score, except that the O, E, and maxR values are defined in terms of sequential consistencies from one trial to the next rather than the number of categorical repetitions within a single trial. The second measure is hierarchical clustering analysis, using a single linkage method of cluster determination. The dependent measure for cluster analysis is the average recall order distance between pairs of items across a set of trials which is used to determine the hierarchical relations among the individual items.

Previous research on free recall organization has shown that over the course of several recall trials, subjects develop a consistent output order, even when the lists are composed of random items (Bousfield, Puff, & Cowan, 1964; Mandler & Dean, 1969; Tulving, 1962). For instance, Tulving found that during list learning (in his case, a list of nouns), the lists were not only subjectively organized, but also that this organization increased systematically over trials, with a positive correlation between recall organization and performance, until perfect recall was achieved and the organization stabilized. Basically, when subjects are able to categorize lists of items to aid recall, they will do so (Mandler, 1967). Therefore, an accurate assessment of mental organization would best be done after the facts are well learned and recall is perfect. Under those circumstances, one can be more certain that the organization has been developed. So, in the current experiments, only trials in which all of the items are recalled were considered.

Method

Subjects. Twenty-four native English speakers were recruited from the subject pool at Michigan State University and given partial course credit for their participation. An additional 3 subjects were replaced, 2 for not being be able to do the recall test accurately, and 1 due to equipment failure.

Materials. The materials for each subject were a set of 18 sentences of either the form "The *object* is in the *location*." or "The *location* has the *object*." Half of the subjects received facts in which the object was the sentence subject (OS group) and the other half in which the location was the sentence subject (LS group). Both forms were included because although previous findings (Radvansky & Zacks, 1991) indicated that both sentence forms produce similar results, these previous experiments relied only on recognition not recall. There is a tendency for forward associations to be stronger than backward associations during recall (e.g., Wolford, 1971). This could influence recall orders by leading subjects to report all of the associates with the first concept mentioned in a sentence just after recalling it. Presenting the sentences in both forms should balance out any such bias.

The objects and locations used were the same as those used in previous object-location experiments (e.g., Radvansky & Zacks, 1991). These materials were used because the object-location combinations had been normed for sensibility, and because they reliably produced the SL and ML fan effect differences in recognition, therefore serving as a reliable basis for assessing recall. The combination of objects and locations is presented in Table 1. The lower case letters refer to specific objects and the upper case letters refer to specific locations. The design is such that there are cases where there are specific objects and locations that have 1 to 3 associations with each of them. The SL and ML superscripts refer to items that were included in the SL and ML conditions. Because the items in the 1-1 condition provided a baseline, half of them were arbitrarily assigned to the SL condition, and half to the ML condition. The remaining items that are not marked with an SL or ML were presented and tested but not included in the analysis. A different random assignment of objects to the *a* to *l* values and locations to the *A* to *L* values was used for each subject. This is the same study list design used in previous experiments. A complete listing of the objects and locations for Experiment 1 is presented in Appendix A.

Insert Table 1 about here

Procedure. Subjects were tested individually in a single session lasting approximately 2 hours. A similar study-test memorization procedure was used for all experiments. Each subject was first presented with a study list consisting of 18 sentences with the instruction to memorize the sentences as efficiently as possible. The sentences were displayed one at a time, for 7 s each, on a monochrome (white) screen controlled by an IBM compatible computer. The sentences appeared half-way down the screen beginning on the left-hand edge. A 40-column presentation mode was used. A different random order was used for each list presentation. After a list had been presented, a set of test questions were given. For Experiment 1, the test questions were

Table 1

Design used for combining subjects and predicates in the generation of the study lists. Items comprising the SL (single location) and ML (multiple location) conditions are marked with corresponding superscripts.

		Predicate Fan 1	Predicate Fan 2	Predicate Fan 3
Subject Fan	1	aA^{SL} bB^{SL}	eE^{SL}	gG^{SL}
		cC^{ML} dD^{ML}	fF^{SL}	hH^{SL}
	2	iI^{ML}	--	iG
		jJ^{ML}	--	jH
	3	kK^{ML}	kE	kG
		lL^{ML}	lF	lH

"Where is the *object?*" and "What is in the *location?*" for each of the objects and locations, respectively. The test questions were also randomly ordered on each cycle. The experimenter kept track of responses on a score sheet. After answering all of the questions the subjects returned to the study portion. The criterion for memorization was the ability to correctly answer all of the test questions twice in a row. Subjects in Experiment 1 required an average of 5.6 cycles to memorize the information. There was no difference between the OS and LS groups, $t < 1$. (Unless otherwise mentioned, the criterion for statistical significance for all tests was set at $p < .05$.)

Recognition. Once the sentences were memorized, recognition and recall tests were given. The order in which these tests were given was counterbalanced with half of the subjects receiving each order. The recognition test was timed and administered on the computer. Both studied and nonstudied facts were presented. Subjects pressed the left button on a computer mouse to indicate a studied fact, and the right button to indicate a nonstudied fact. Subjects were encouraged to be as fast and as accurate as possible in their decisions. A set of 18 practice trials was provided to familiarize the subjects with using the mouse buttons. On the practice trials the computer either displayed "SENTENCE STUDIED" or "SENTENCE NOT STUDIED", and the subject had to press the appropriate button. On the recognition test itself, studied probes were sentences from the study list and nonstudied probes were generated from repairings of the object and location concepts from within a cell of the design. For example, if the studied sentences from the same cell were sentences 8 and 9, the nonstudied sentences would be

10 and 11.

8. The pay phone is in the public library.

9. The ceiling fan is in the hotel.

10. The pay phone is in the hotel.

11. The ceiling fan is in the public library.

The order of probe presentation in the recognition test was randomized within each of 8 blocks. Those sentences from cells in which several objects were associated with several locations, although they were presented, were not entered into the analysis. This is because the data from these cells do not simply and directly address the question asked in these experiments: whether there are fan effects in the SL and ML conditions. Each sentence was presented 8 times, yielding a total of 288 recognition test trials, 192 of which were included in the analysis. Subjects were allowed 3 self-timed breaks at the end of each quarter of the recognition test. The computer recorded RTs and errors.

If a subject responded incorrectly on a trial they received immediate feedback. The feedback consisted of the presentation of a line that read either "*ERROR* SENTENCE STUDIED" or "*ERROR* SENTENCE NOT STUDIED", whichever was appropriate. This feedback was presented for 1 s. For the purposes of analysis, errors also included trials for which the RTs were shorter than 500 ms or longer than 10,000 ms. Also for purposes of analysis, for each subject, those RTs that were 2 standard deviations from the mean of a given cell of the design were dropped from the analysis, though they were not counted as errors. This trimming of extreme scores eliminated 5.1% of the data in Experiment 1.

<u>Recall</u>. For the recall test, subjects were asked to orally recall the memorized lists and to report the entire sentences.

The experimenter recorded recall orders on a tally sheet, as well as on a tape recorder to ensure that the tally was correct. There were 18 recall trials for each subject. At the beginning of each trial, one of the study sentences was given and the subject was asked to recall the list starting with that sentence as if it were the first one that they themselves had recalled. This was done to prevent subjects from adopting a stereotyped recall order that could potentially obscure mental model organizations. The order of the starting sentences was randomized for each subject. Any incomplete or erroneous recalls were repeated at the end of the recall test so that all starting sentence trials were present in the data. Aside from this error replacement procedure, a particular starting sentence was never repeated more than once.

Card sorting and questionnaire. After being presented with both memory tests, subjects were given 18 4 in. X 5 in. index cards with each sentence that they had memorized printed in the center. Subjects were asked to sort the cards based on what was written on them. They were told that they could make as many piles as they liked, so long as there was more than one pile. Once the subjects had sorted the cards to their satisfaction, they were given a short questionnaire to fill out about the strategies they used to memorize the sentences. Discussion of the results of the questionnaire is delayed until after all of the experiments have been presented.

Results

RECOGNITION TEST

Reaction times. The reaction time (RT) and error rate data are summarized in Figure 3. Responses to probes in the ML condition showed an increase in RT with an increase in the number of locations associated with an object, whereas in the SL condition, there was no increase in RT with an increase in the number of objects associated with a location. In addition, the SL probes were responded to faster (1720 ms) than ML probes (1817 ms), and there was a general fan effect (fan level 1 = 1696; 2 = 1796; 3 = 1812 ms). These findings are consistent with the location-based organization of mental models found in previous experiments.

Insert Figure 3 about here

Other differences were that responses to studied probes were faster (1721 ms) than to nonstudied probes (1816 ms), and the size of the SL-ML difference was larger for nonstudied probes (SL = 1739; ML = 1892 ms) than for studied probes (SL = 1700; ML = 1741 ms). Finally, there was an interaction of the order of memory tests with studied-nonstudied probes, the SL-ML difference, and the fan effect. A breakdown of these data is presented in Table 2. The same pattern of fan effects for the SL and ML conditions were found for both order groups across both studied and nonstudied probes. The cause of this interaction is the fluctuation in the patterns of retrieval times at the different

Figure 3

Reaction time and error rate results for Experiment 1.

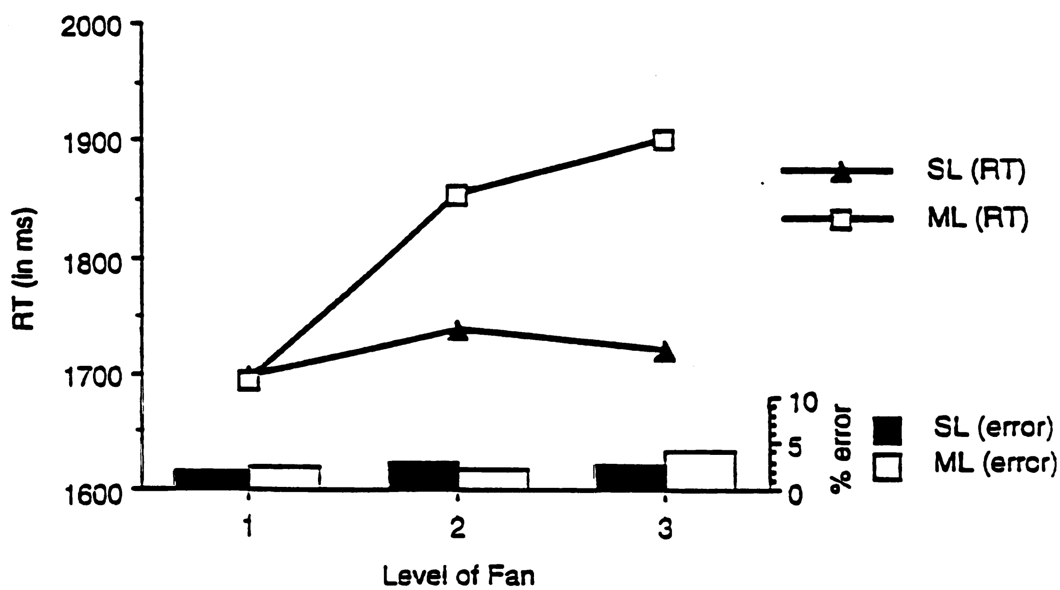


Table 2

Reaction time results (in ms) for Experiment 1 for recognition test 1st and 2nd groups. Error percentages are presented in parentheses.

Fan Level		1	2	3
Recognition Test 1st				
Studied	SL	1828 (1.04)	1759 (1.04)	1685 (2.60)
	ML	1726 (2.60)	1847 (2.60)	1792 (2.60)
Nonstudied	SL	1770 (1.56)	1735 (1.56)	1592 (3.65)
	ML	1727 (2.08)	2010 (2.08)	2189 (3.13)
Recognition Test 2nd				
Studied	SL	1601 (1.56)	1734 (7.81)	1592 (1.56)
	ML	1551 (1.04)	1750 (2.08)	1782 (6.25)
Nonstudied	SL	1594 (3.13)	1728 (1.56)	1756 (2.60)
	ML	1770 (3.65)	1808 (1.56)	1849 (4.69)

levels of fans in each of the different cases.

Insert Table 2 about here

The RT data were submitted to a 2 (Type of sentence) X 2 (Order of memory tests) X 2 (Studied-Nonstudied) X 2 (SL-ML) X 3 (Fan) mixed ANOVA. The first 2 variables were between subjects and the rest within subjects. Of central interest, there was a significant main effect of SL-ML, $F(1,20) = 8.05$, $MSe = 84605$, and a marginally significant main effect of Fan, $F(2,40) = 2.92$, $MSe = 130516$, $p < .07$, as well as a significant SL-ML X Fan interaction, $F(2,40) = 3.24$, $MSe = 66174$.

Also significant were the main effect of Studied-Nonstudied, $F(1,20) = 10.22$, $MSe = 63886$, and the Studied-Nonstudied X SL-ML interaction, $F(1,20) = 4.68$, $MSe = 47441$. Simple effects tests showed that the SL-ML difference was significant for the nonstudied probes, $F(1,20) = 12.05$, $MSe = 69741$, but not for the studied probes, $F = 1.01$. Finally, there was a marginally significant Order X Studied-Nonstudied X SL-ML X Fan interaction, $F(2,40) = 3.11$, $MSe = 29520$, $p < .06$.

Errors. Subjects in Experiment 1 had 2.7% errors on average for the recognition test. There was a noticeable difference in the fan effects for studied (1 = 1.6; 2 = 3.4; 3 = 3.3%) and nonstudied probes (1 = 2.6; 2 = 1.7; 3 = 3.5%), $F(2,40) = 3.29$, $MSe = 14$, $p < .06$, which was further qualified by a difference between whether the recognition test came first or second, $F(2,40) = 3.12$, $MSe = 14$, $p < .07$. Finally, the SL and ML fan effects differed, depending on whether the recognition test came first or second, $F(2,40) = 3.58$, $MSe = 24$. Neither of these interactions lends itself to easy interpretation. Given the low percentage of errors overall, it is most likely that these effects are spurious.

RECALL TEST

Subjects had 0 to 8 ($M = 4.1$) recall trials which were repeated at the end of the test due to errors. There was no difference for either memory test order or sentence types, $Fs < 1$.

Cluster Analysis. A hierarchical cluster analysis was performed on the recall data based on the mean distance in recall order among items for each subject. The results of this analysis are presented in Figure 4. As a reminder, the lower case letters refer to objects and the upper case letters refer to locations. As can be seen, there was a clear location-based organization. Items referring to common locations were grouped together more-so than items referring to common objects. In addition, locations containing the same objects were grouped together, although this organization was not as strong as the location-based organization.

The scale beneath each the hierarchical tree provides an index of the proximity between various clusters. This proximity index reflects the similarity distance threshold at which different items or subset of items are identified as belonging to the same cluster. For the recall data, the smaller numbers reflect smaller recall order pair distances, whereas for the card sorting data to be seen later, larger numbers reflect larger numbers of subjects grouping pairs of items together. The absolute magnitude of the proximity values can vary for a number of reasons. For example, a larger number of subjects in one condition could produce larger pair grouping values although no difference in structure is actually evident. These proximity values are given to provide a sense of scale to the hierarchical structures and to ease the comparison across experiments where the basis of the proximity value is relatively stable.

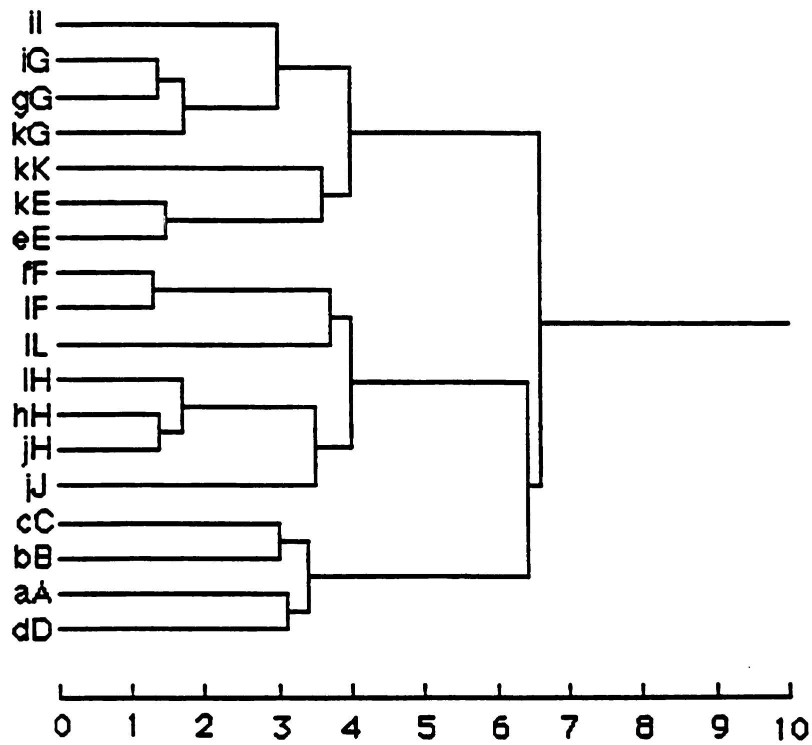
The recall order distances (not the scale values, but the average recall order distances between a pair of items for each subject) for all item pairs in the four memory test order - sentence type groups were highly correlated ($.88 \leq r \leq .94$), indicating that the organizations of these groups do not differ from one another in the general organization. The only other notable finding was that sentences in which both concepts had only one associate were also grouped together.

Insert Figure 4 about here

ARC. The ARC and ARC′ scores for all of the experiments are listed in Table 3. Separate ARC scores were calculated for location-based (ARCloc) and object-based organizations (ARCobj). These scores were submitted to a 2 (Type of sentence) X 2 (Order of memory tests) X 2 (Grouping: location or object) mixed ANOVA. The first two variables were between subjects and the third was within subjects. There was a greater degree of location-based than object-based organization as revealed by the main effect of Grouping, $F(1,20) = 74.54$, $MSe = .026$. In addition, the Type X Grouping interaction was significant, $F(1,20) = 5.064$, $MSe = .026$.

Figure 4

Results of a hierarchical cluster analysis for the recall data of Experiment 1 derived from the recall order pair distances.



The preference for a location-based organization was stronger for the LS group (ARCloc = .92; ARCobj = .41) than the OS group (ARCloc = .82; ARCobj = .53). Simple effects tests showed that the effect of Grouping was significant for both OS, $F(1,11) = 13.61$, $MSe = .039$, and LS groups, $F(1,11) = 67.94$, $MSe = .023$. Finally, there was a marginally significant Order X Grouping interaction, $F(1,20) = 3.93$, $MSe = .026$, $p < .07$. There was a tendency for a greater location-based organization when the recall test was second (ARCloc = .90; ARCobj = .41) rather than first (ARCloc = .84; ARCobj = .53). Simple effects tests showed that Grouping was significant both when the recall test came first, $F(1,11) = 12.87$, $MSe = .045$, as well as second, $F(1,11) = 74.73$, $MSe = .020$.

The average recall order pair distances were compared, and a strong negative correlation was found between the ARCloc and ARCobj scores ($r = -.69$). This would suggest that as a subject demonstrated greater preference for one organizational scheme, preference for the other diminished. There was no correlation with the SL-ML difference (an index of recognition organization) for either the ARCloc ($r = .02$) or ARCobj scores ($r = -.22$). This raises the possibility that one or both of these measures of organization unreliably assesses the underlying memory organization.

Insert Table 3 about here

Table 3

Results of the ARC. and ARC' analyses for the recall tests and card sorting tasks.

	ARCloc	ARCobj/per	ARC'
Recall:			
Experiment 1	.87	.47	.25
Experiment 2	.82	.55	.31
Card Sorting:			
Experiment 1	.90	.28	---
Experiment 2	.72	.57	---
Experiment 3	.73	.16	---
Experiment 4	.64	.39	---
Experiment 5	.76	.17	---

ARC'. The ARC' scores were submitted to a 2 (Type of sentence) X 2 (Order of memory tests) ANOVA. There were no significant differences. The ARC' scores were not correlated with the ARCloc scores (r = -.02) but were positively correlated with the ARCobj scores (r = .43). Most of the sequential consistency of the free recalls was in the order of the items referring to a common object. The ARC' score and SL-ML difference were not correlated (r = .19).

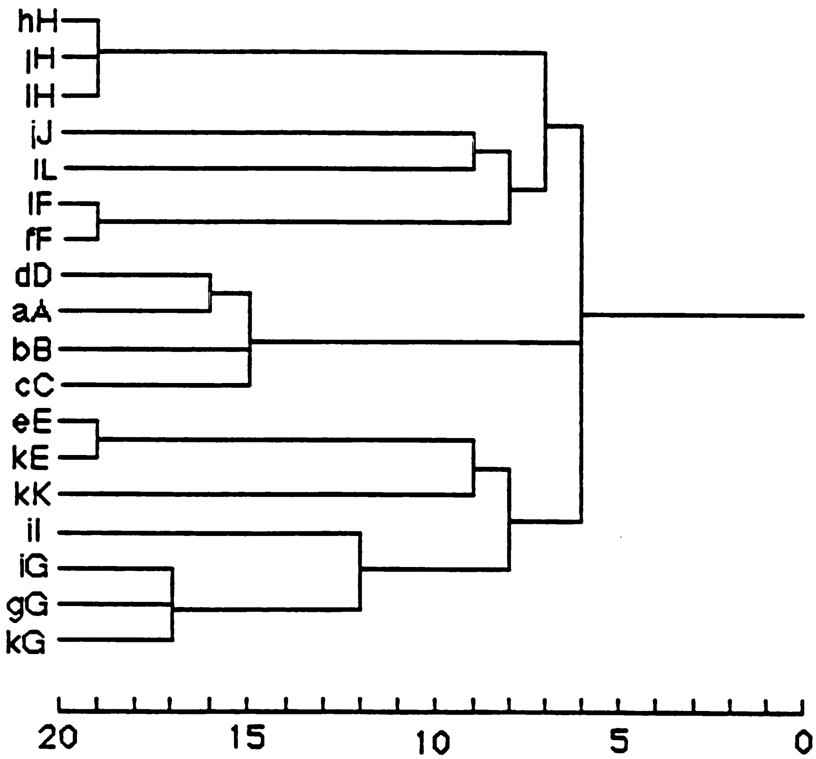
CARD SORTING

The data from the card sorting task were submitted to cluster and ARC analyses similar to those for the recall data. Three subjects were not given the card sorting task due to an experimenter oversight. The number of times a pair of items was placed in the same group was the dependent measure. ARC scores were calculated based on group membership. So, a group containing three items from the same category were counted as two categorical repetitions, even though the card sorting was not ordered.

The results of the cluster analysis are presented in Figure 5. As can be seen, there was a clear location-based organization. Items containing common locations were organized closer than items containing common objects. This result is consistent with the pattern found in the recognition and recall tests. The ARCloc scores were higher than ARCobj scores, $F(1,19) = 24.23$, MSe = .165. There was no significant difference between OS and LS sentence groups, $F < 1$.

Figure 5

Results of a hierarchical cluster analysis for the card sorting data of Experiment 1 derived from the number of times pairs of items were grouping together.



Insert Figure 5 about here

As can be seen in Table 4, the number of times a pairs of items were grouped together for the LS and OS groups was positively correlated. Furthermore, the ARCloc and ARCobj scores were negatively correlated, although not significantly so, suggesting, along with the results of the recall test, that as preference for one organizational scheme increased, preference for the other decreased. When compared with the results of the recall test, the card sort ARCloc score did not correlate with the recall ARCloc (r = .21) scores; however, the card sort ARCobj scores were positively correlated with the ARCobj scores (r = .44). This suggests that the recall and card sorting tasks used the same organizational structure.

Insert Table 4 about here

Discussion

For Experiment 1, the results of the recognition and recall tests, as well as the card sorting task, were in agreement with each other and previous experiments in showing a location-based organization of the mental models. There were also some effects involving the order of memory tests and sentence type for the different memory tests. For RT and error rates, there were variations in the SL and ML fan effects for studied and nonstudied probes depending on the order of the memory tests. However, this

Table 4

Correlation of item groupings in sorting task for all experiments. The LS-OS/PS correlation reflects a comparison of the degree two which items were grouped together for subjects receiving different sentence types. The ARCloc-ARCobj/per correlation reflects a comparison of the degree of location-based versus object/person-based organization for subjects within an experiment.

	LS-OS/PS	ARCloc-ARCobj/per
Experiment 1	.92*	-.22
Experiment 2	.92*	-.46*
Experiment 3	.86*	-.91*
Experiment 4	.79*	-.88*
Experiment 5	.87*	-.57*

Note: * indicates $p < .05$.

pattern of differences is not easily interpretable, and is likely to be spurious. Like previous recognition experiments, there were no effects of sentence type.

For the recall data, the organization more closely followed a location-based organization when the recall test was second rather than first. This is not surprising considering that the subjects had more exposure to the information when the recall test was given second, and the information is therefore likely to be more organized (e.g., Tulving, 1962). It is important to note the difference between the order of memory test groups is one of degree rather than direction. Also, the stronger location-based organization for the LS group could be an effect of locations being placed first in the sentences (Wolford, 1971). Again, the location-based organization for the OS group, while weaker than the LS group, is still quite strong.

It appears from Experiment 1 that recall and recognition reflect the location-based organization of mental model representations equally well. During recognition the mental models were directly accessed with no interference during retrieval, except when there were several mental models that were related to the probe. During recall, output organization was consistent with the location-based mental model organization. This organizational pattern is apparent in the results of the hierarchical cluster analysis. It would seem reasonable at this time to assume that the retrieval plan used to guide the retrieval of the mental models was organized to reflect the structure of the underlying memory traces. This organizational bias was also reflected in the card sorting task where subjects were explicitly asked to organize the facts.

However, the mental models used in Experiment 1 were of only one type, namely, location-based mental models. It is unclear whether the same results will be found for other types. Experiment 2 attempted to replicate Experiment 1 using facts about people being in small places which have been shown to produce person-based mental models (Radvansky et al., in press).

CHAPTER III

EXPERIMENT 2

Method

Subjects. Twenty-four native English speakers were recruited from the subject pool at Michigan State University and given partial course credit for their participation. Half of the subjects received PS (Person Subject) sentences and half LS (Location Subject) sentences. An additional 5 subjects were replaced, 3 for not completing the experiment, and 2 for not accurately recalling the facts.

Materials and Procedure. The sentences used in Experiment 2 were composed of random pairings of people and small locations. The materials were drawn from those used by Radvansky et al. (in press) which have been previously rated for sensibility. Like Experiment 1, half of the sentences were of the form "The person is in/on the location." and the other half were of the form "In/On the location is the person." Sentences 12 and 13 are examples of the first form and Sentences 14 and 15 are corresponding sentences in the second form.

12. The banker is in the telephone booth.
13. The carpenter is on the operating table.
14. In the telephone booth is the banker.
15. On the operating table is the carpenter.

The subjects and predicates used in Experiment 2 are listed in Appendix B. The memorization and testing procedures were like those of Experiment 1. Subjects in Experiment 2 took an average of 5 study-test cycles to learn the facts with no differences between PS and LS groups, $\underline{t} < 1$. In addition, the trimming procedure eliminated 4.7% of the recognition test data.

Results

RECOGNITION TEST

<u>Reaction times</u>. The RT and error rate results for Experiment 2 are presented in Figure 6. A fan effect occurred in the SL condition, but not the ML condition. This person-based organization is consistent with previous research using people and small places (Radvansky et al., in press). In addition, the SL probes were responded to more slowly (1749 ms) than the ML probes (1591 ms), and an overall fan effect was produced (fan level 1 = 1575; 2 = 1676; 3 = 1760 ms).

Insert Figure 6 about here

In addition to the organization results, studied probes were responded to faster (1628 ms) than nonstudied probes (1712 ms). Furthermore, the SL-ML difference was larger for nonstudied probes (SL = 1828; ML = 1596 ms) than for studied probes (SL = 1671; ML = 1586 ms). Finally, the SL and ML fan effect patterns differed depending on whether the recognition test came first or second. This is the only effect involving memory test order in this experiment, and, as can be seen in the summary in Table 5, this

Figure 6

Reaction time and error rate results for Experiment 2.

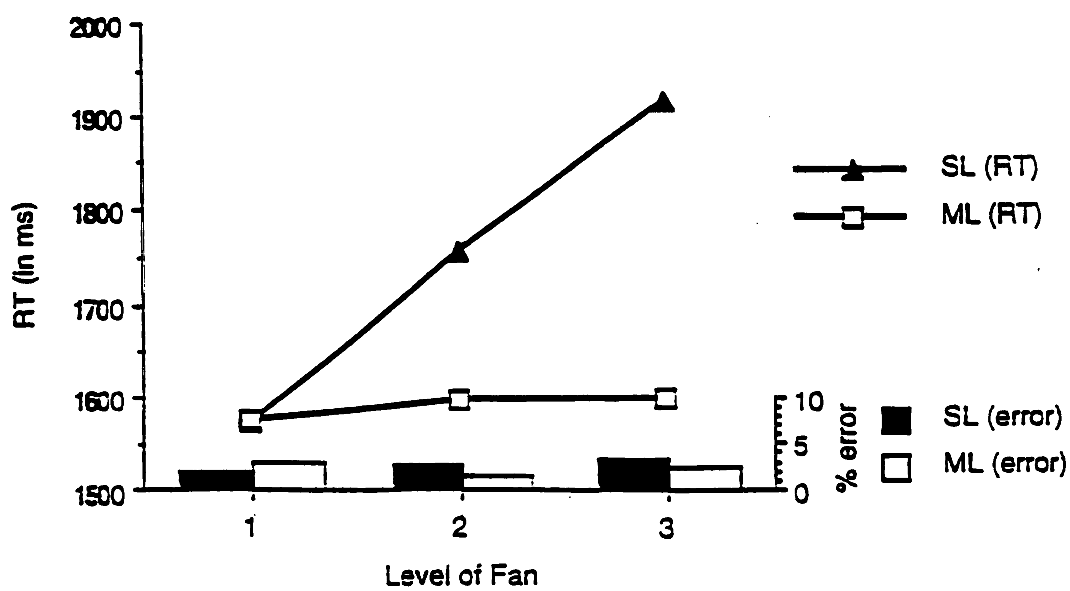


Table 5

Reaction time results in ms (and error percentages) for SL and ML fan effects for recognition test 1st and 2nd groups in Experiment 2.

Fan Level	1	2	3
Recognition Test 1st			
SL	1618 (2.08)	1768 (2.08)	1954 (3.39)
ML	1494 (1.82)	1722 (1.04)	1624 (1.56)
Recognition Test 2nd			
SL	1531 (1.56)	1741 (3.39)	1884 (3.39)
ML	1657 (3.91)	1475 (2.09)	1577 (3.39)

was due to fluctuations in the different cells rather than differences in the organizational pattern. It is unlikely that this fluctuation is due to any systematic differences attributable to the order of memory tests.

Insert Table 5 about here

The RT data were submitted to a 2 (Type of sentence) X 2 (Order of memory tests) X 2 (Studied-Nonstudied) X 2 (SL-ML) X 3 (Fan) mixed ANOVA. The first two variables were between subjects and the rest were within. There was a significant main effect of SL-ML, $F(1,20) = 14.60$, $MSe = 122869$, and Fan, $F(2,40) = 5.41$, $MSe = 152588$, as well as a significant SL-ML X Fan interaction, $F(2,40) = 9.20$, $MSe = 66339$. Simple effects test showed that the fan effect was significant for the SL condition, $F(2,40) = 8.60$, $MSe = 165835$, but not for the ML condition, $F < 1$.

There was also a significant main effect of Studied-Nonstudied, $F(1,20) = 11.94$, $MSe = 42300$, as well as a significant Studied-Nonstudied X SL-ML interaction, $F(1,20) = 10.42$, $MSe = 37605$. Simple effects tests showed that the SL-ML difference was significant for both studied, $F(1,20) = 3.40$, $MSe = 74802$, and nonstudied probes, $F(1,20) = 22.54$, $MSe = 85671$. Finally, there was a significant Order X SL-ML X Fan interaction, $F(2,40) = 4.99$, $MSe = 66340$. The SL-ML X Fan interaction was significant both when the recognition test came first, $F(2,20) = 3.79$, $MSe = 67990$, as well as second, $F(2,20) = 10.57$, $MSe = 64688$.

Errors. Subjects in Experiment 2 averaged 2.5% errors on the recognition test. The error rate data were analyzed in an ANOVA similar to that for the RTs. The only significant difference was a greater number of errors for studied probes (3.2%) than nonstudied probes (1.8%), $F(1,20) = 9.48$, $MSe = 15$.

RECALL TEST

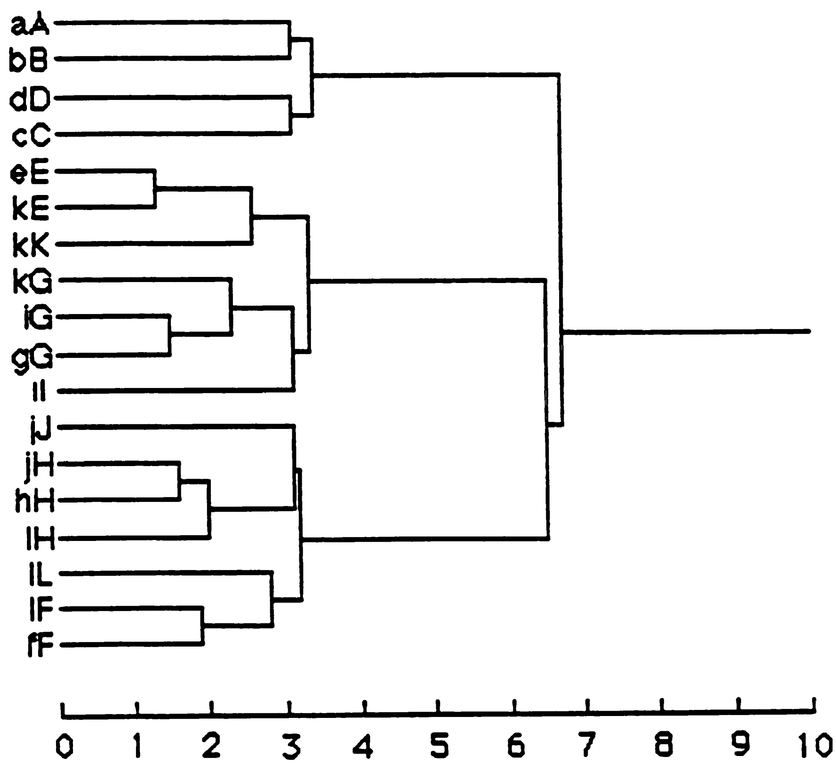
Recall errors were made on an average of 2.8 trials in Experiment 2. There were no differences based on the order of the memory tests, $F < 1$; however, there were more recall errors for the PS group (3.8) than the LS group (1.9), $F(1,20) = 3.72$, $MSe = 5.4$, $p < .07$.

Cluster Analysis. A cluster analysis was performed as in Experiment 1, and the results are presented in Figure 7 (lower case letters = people; upper case letters = locations). As can be seen, there was generally a location-based organization. This organization is inconsistent with the recognition data which showed a person-based organization. Like Experiment 1, the recall distances for the four memory test order - sentence type groups were highly correlated ($.85 \leq r \leq .93$) indicating that the organizations for those groups did not differ. In a comparison with the corresponding items in Experiment 1, the mean recall distances of all possible item pairs showed a very high positive correlation, $r = .97$. This indicates that the organizational structure of the two data sets was nearly identical. Finally, like Experiment 1, fan level 1 sentences were grouped together.

Insert Figure 7 about here

Figure 7

Results of a hierarchical cluster analysis for the recall data of Experiment 2 derived from the recall order pair distances.



ARC. The ARC scores for Experiment 2 were submitted to a 2 (Type of sentence) X 2 (Order of memory tests) X 2 (Grouping of facts) mixed ANOVA. The first two variables were between subjects and the third was within. Only the main effect of Grouping was significant, *F*(1,20) = 7.72, *MSe* = .112. As can be seen in Table 2, the ARCloc scores were greater than the ARCper scores. This is inconsistent with the notion that person-based mental models were formed.

There was a strong negative correlation between the ARCloc and ARCper scores (*r* = -.74). Again, this is interpreted as indicating that greater preferences for one organizational scheme are accompanied by smaller preferences for the other organizational scheme. Furthermore, the ARCloc scores were negatively correlated with the SL-ML RT difference (*r* = -.44), whereas the ARCper scores were positively correlated with this difference (*r* = .28), although not significantly so.

ARC'. For the ARC' data, no effects of sentence type or order of memory tests were found. Like Experiment 1, the ARC' score for Experiment 2 was negatively correlated with ARCloc (*r* = -.36), but positively correlated with ARCper (*r* = .36). This would suggest that a larger portion of the sequential consistency in subjects' recalls were in repetitions of items containing common people rather than common locations. The ARC' scores also showed a small nonsignificant correlation with the size of the SL-ML difference (*r* = .28).

CARD SORTING

The result of the card sorting task differed from Experiment

1 in that it did not show a clear organizational preference, as can be seen in Figure 8, although there was still a general preference for location-based organization. In Experiment 2, the person- and location-based organizations were of equivalent strength (see Table 3). The difference between the ARCloc and ARCper scores was not significant, $F < 1$. Like Experiment 1, there were no differences involving sentence type, $Fs < 1$.

The pair groupings for the PS and LS groups were highly correlated (see Table 4) suggesting equivalent organizational strategies in the two groups. Furthermore, the ARCloc scores and ARCper scores were negatively correlated. This suggests that increased preference from one organizational strategy was accompanied by a decreased preference for the other, rather than a random adoption of strategies.

Insert Figure 8 about here

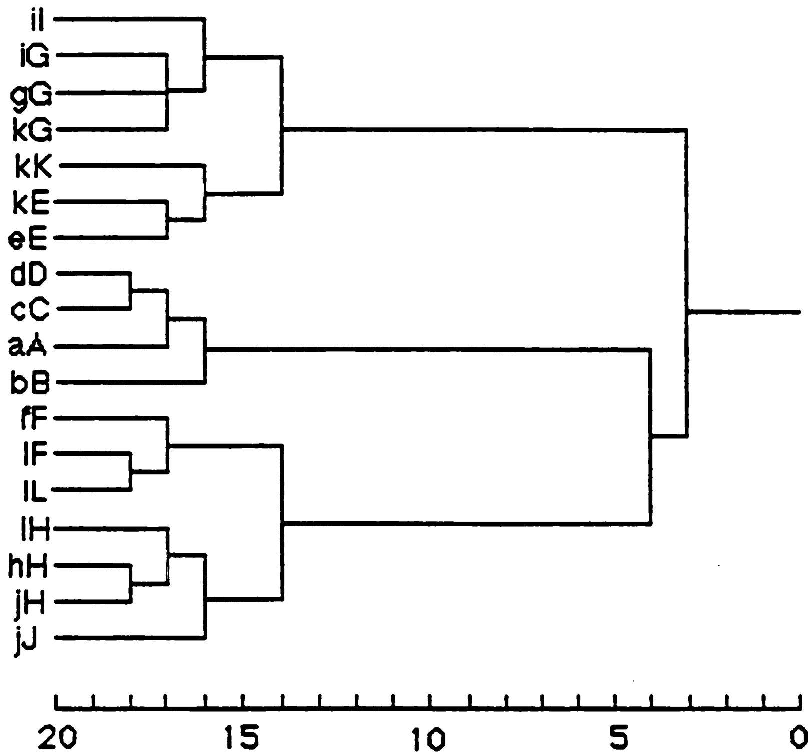
Although the card sorting data did not reveal any statistically significant organizational preferences, there were positive correlations between the card sort and recall ARCloc and ARCper scores ($r = .69$, and $r = .76$, respectively). So, although the sorting data fail to show organizational preference, for each subject they generally followed the recall organization.

Discussion

The most striking result of Experiment 2 is the dissociation between the type of organization implied by the recognition and recall measures. The person-based organization of the recognition

Figure 8

Results of a hierarchical cluster analysis for the card sorting data of Experiment 2 derived from the number of times pairs of items were grouping together.



test replicated previous person-small locations experiments, whereas the location-based organization of the recall test was consistent with Experiment 1, an organizational basis directly at odds with the recognition results. The difference in the number of recall errors for the PS and LS groups may be due in part to the location-based organization in the recall data. In particular, the LS sentences may have resulted in fewer errors because the first mentioned concept in the sentence corresponded to the recall organization basis. However, this is speculative. Finally, despite the strong organizational patterns for the recall and recognition tests, no statistically significant organizational preference was found in the card sorting task, although it was highly correlated with recall. The remaining experiments address the questions of which, if any, of these measures accurately assesses the structure of the mental model and why a dissociation occurred.

The next step is to assess the validity of the recognition results. Previous experiments (Radvansky & Zacks, 1991) have shown that the fan effect difference between the SL and ML conditions is reduced when the subjects are provided a precue of the concept around which the representation is assumed to be organized prior to the onset of the target sentences. In Experiment 3, person and location precues were provided to test the integrity of the recognition test of Experiment 2. If subjects were forming person-based mental models, then the differences between the SL and ML fan effects should be attenuated with person precues. Person precues would select the appropriate mental model from the set of related mental models that represent information related to a common location and are activated during the complex direct-access process.

CHAPTER IV

EXPERIMENT 3

Experiment 3 used a precuing procedure similar to a previous experiment using object-location materials (Radvansky & Zacks, 1991, Exp. 3). In the Radvansky and Zacks experiment, it was found that when subjects were given a location concept precue 1 s prior to the presentation of the probe sentence, the fan effect in the ML condition was attenuated. The explanation was that location concepts allowed for the selection of the appropriate location-based mental model from the set of mental models all containing a common object (see Figure 1). If person-based mental models are created for people-small location facts, then the fan effect in the SL condition should be attenuated for a person precue but not for a location precue. The notion is that, prior to the presentation of the memory probe, the person concept selects the appropriate mental model from among those containing a common location, and thereby reduces the amount of interference experienced in that condition.

Method

Subjects. Forty native English speakers were recruited from the subject pool at Michigan State University and either given partial course credit or paid $7.50 for their participation. Half of the subjects received PS sentences and the other half LS sentences.

Materials and Procedure. The materials and memorization procedure were like those for Experiment 2. Subjects were tested in a single session which lasted 1 1/2 hours. The subjects took an average of 5.1 study-test cycles to memorize the facts. There was no significant difference between the PS and LS groups, $t(38) = 1.59$, $p > .10$. The only memory test given was cued recognition. Prior to the presentation of each probe sentence, a person, location or neutral precue was presented. For the person and location precue conditions, the precue was the person or the location concept, respectively, from the probe sentence. For the neutral condition, the precue was the word "READY". The precue was presented left-justified halfway down the computer screen and remained visible for 1 s when it was replaced with the probe sentence. Subjects were instructed to try to use the cue because it would help them "to decide whether the probe sentence was studied or not". Because of the large number of conditions, subjects were presented with each precue-probe sentence combination 4 times instead of 8 times as in the previous experiments, yielding a total of 432 trials, 288 of which were included in the analysis. Because of the small number of observations per cell, no data were trimmed.

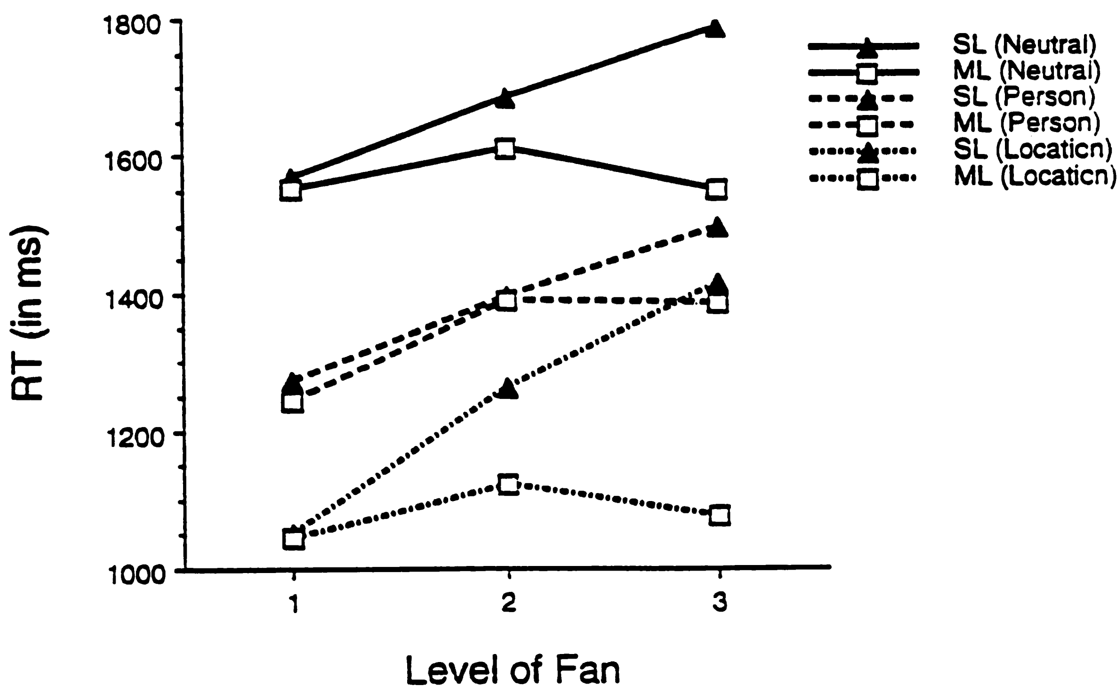
Results

RECOGNITION TEST

Reaction times. The RTs and error rates for Experiment 3 are presented in Figure 9. Overall, SL probes were responded to more slowly (1438 ms) than ML probes (1331 ms), and there was a

Figure 9

Reaction time results for the different precue conditions for Experiment 3.



general increase in retrieval time accompanying an increase in the level of fan, (fan level 1 - 1289; 2 - 1412, and 3 - 1454 ms). Furthermore, the fan effect was present for the SL condition (fan level 1 - 1298; 2 - 1449; and 3 - 1568 ms) but not the ML condition (fan level 1 - 1280; 2 - 1374; 3 - 1340 ms). In addition, subjects responded to studied probes faster (1336 ms) then nonstudied probes (1433 ms).

Insert Figure 9 about here

In terms of the precues, subjects responded fastest after location precues (1163 ms), less fast after person precues (1365 ms) and slowest after neutral precues (1627 ms). A speed-up occurred for the person and location precue conditions relative to neutral precue condition because of the early availability of information. The most probable reason that the location condition is the fastest is because more information about the probe sentence is presented in the precue. The location concepts were much longer (M - 16.0 letters for locations versus M - 7.1 letters for people).

Precue type did affect the size of the SL-ML difference. It was largest after location precues (161 ms), followed by neutral precues (110 ms), and smallest after person precues (49 ms). Furthermore, when the different cue conditions are considered separately, there is a difference in the SL and ML fan effects for location and neutral precues, but not for person precues. So, even though location precues led to the fastest RTs, probe responses in this condition led to the greatest single/multiple mental model difference. However, although the person precues did result in a elimination of a difference between the SL and ML fan effects, it did not result in an attenuation of the SL fan effect. Instead, a fan effect emerged in the ML condition. Finally, neutral precues produced a smaller fan effect (fan level 1 - 1561; 2 - 1649; 3 - 1671 ms) than location (fan level 1 - 1049; 2 - 1192; 3 - 1247 ms) and person precues (fan level 1 - 1257; 2 - 1394; 3 - 1442 ms).

The RT data were submitted to a 2 (Type of sentence) X 3 (Cue) X 2 (Studied-Nonstudied) X 2 (SL-ML) X 3 (Fan) mixed measures ANOVA. The first factor is between subjects and the rest are within subjects. There was a significant main effect of Studied-Nonstudied, $F(1,38)$ - 31.15, MSe - 107838. There were significant main effects of SL-ML, $F(1,38)$ - 22.71, MSe - 180422, and Fan, $F(2,76)$ - 16.21, MSe - 217466, as well as a significant interaction of these two variables, $F(2,76)$ - 10.43, MSe - 136280. Simple effects tests showed that the fan effect was significant for both the SL, $F(2,76)$ - 24.05, MSe- 182920, and ML conditions, $F(2,76)$ - 3.20, MSe - 170826.

There was a main effect of Cue, $F(2,76)$ - 194.67, MSe - 133824, as well as significant Cue X SL-ML, $F(2,76)$ - 5.84, MSe - 65899, Cue X Fan, $F(4, 152)$ - 2.68, MSe - 35322, and Cue X SL-ML X Fan interactions, $F(4,152)$ - 2.61, MSe - 63168. Separate analyses were performed on the RT data for the different cue conditions. The main effect of SL-ML and the SL-ML X Fan interaction were significant for neutral ($F(1,38)$ - 10.93, MSe - 132021, and

F(2,76) = 7.04, *MSe* = 73298, respectively) and location precues (*F*(1,38) = 36.88, *MSe* = 85167, and *F*(2,76) = 11.93, *MSe* = 93695, respectively), but not for person precues (*F*(1,38) = 2.99, *MSe* = 95031, *p* > .09, and *F*(2,76) = 1.24, *MSe* = 95623, *p* > .25, respectively). Simple effects tests showed significant fan effects for the SL conditions of both neutral, *F*(2,76) = 8.92, *MSe* = 108009, and location precues, *F*(2,76) = 28.30, *MSe* = 94939, but not for the ML conditions, *F*(2,76) = 1.36, *MSe* = 68125, p > .25, and *F*(2,76) = 1.14, *MSe* = 97716, *p* > .25, respectively. Significant fan effects were also found with person precues for both SL, F(2,76) = 11.56, MSe = 88851, and ML conditions, F(2,76) = 6.28, MSe = 93085.

Errors. Subjects in Experiment 3 made an average of 2.2% recognition errors. The error rate data were submitted to an analysis similar to that for the RT data. There were more errors in response to SL probes (2.6%) than ML probes (1.8%), *F*(1,38) = 9.18, *MSe* = 30. In addition there was a significant Studied-Nonstudied X Fan interaction, *F*(2,76) = 3.45, *MSe* = 19. The errors for the studied probes increased with increasing fan (fan level 1 = 2.1; 2 = 2.5; 3 = 2.7%), whereas for the nonstudied probes the pattern was more erratic (fan level 1 = 2.2; 2 = 2.3; 3 = 1.4%). However, simple effects tests did not show a significant fan effect for either probe type, both *F*s < 1. Finally, there was a significant Cue X Studied-Nonstudied X SL-ML interaction, *F*(2,76) = 4.03, *MSe* = 56. The Cue X SL-ML interaction was significant in the separate studied probe analysis, *F*(2,76) = 4.26, *MSe* = 70, but not for the nonstudied probe analysis, *F*(2,76)

= 1.78, *MSe* = 60, *p* > .10. Simple effects tests showed that for the studied probes, the SL-ML difference was significant for person precues (SL = 4.9%; ML = 0.9%), *F*(1,38) = 7.24, *MSe* = 130, but not for neutral (SL = 2.5% ; ML = 1.5%), *F*(1,38) = 2.02, *MSe* = 32, *p* > .10, or location precues (SL = 2.2%; ML = 2.6%), *F* < 1.

CARD SORTING

The results of the sorting task are presented in Figure 10. In general, the results were like the previous experiments in showing a location-based organization. The ARCloc scores were greater than the ARCper scores, *F*(1,36) = 16.02, *MSe* = .392 (see Table 3). There was no effect of sentence type, *F* < 1. The pair groupings for the two sentence groups was highly correlated and there was a strong negative correlation between the ARCloc and ARCper scores (see Table 4). Finally, there was no strong tendency to group the fan level 1 items together.

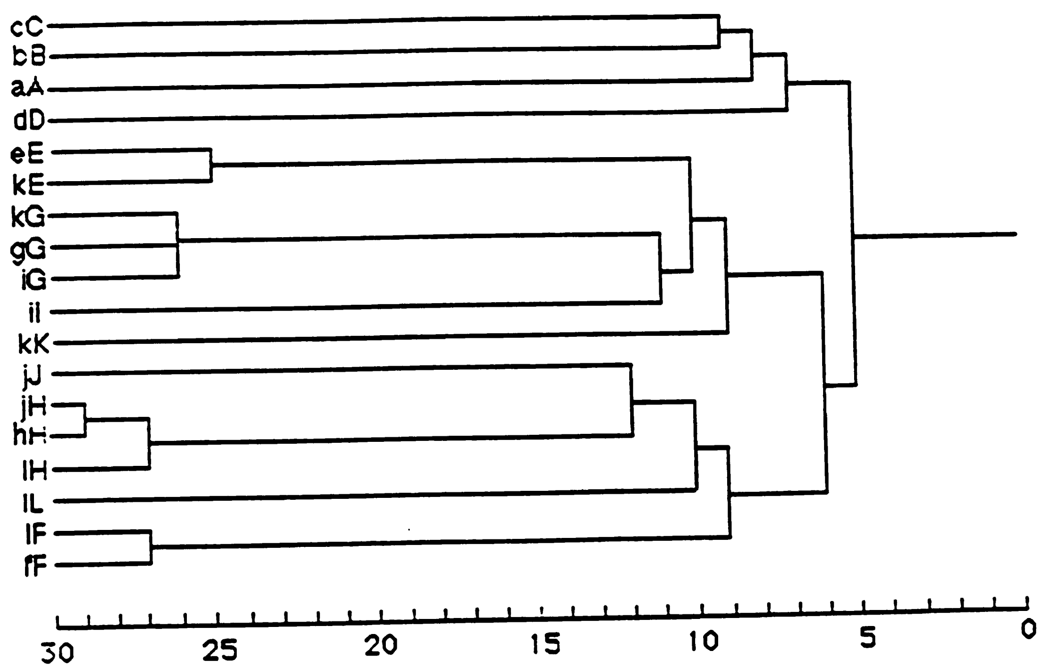
Insert Figure 10 about here

Discussion

The results of Experiment 3 supported the recognition data of the earlier experiments in suggesting a person-based organization of the facts. Furthermore, although the presentation of the person concept prior to the probe sentence did not reduce the SL fan effect, it did substantially reduce the difference between the SL and ML conditions. The difference in the retrieval rates for these two conditions is central to the argument concerning organization of information into mental models. An

Figure 10

Results of a hierarchical cluster analysis for the card sorting data of Experiment 3 derived from the number of times pairs of items were grouping together.



important point is that the person precues were able to reduce this difference, whereas the location precues were not.

The fact that a fan effect was present in both the SL and ML conditions for person precues is troublesome for the mental model explanation. The reason for this finding is uncertain at this time. It seems counter-intuitive that providing a precue to a recognition decision would actually make performance worse. This result may be an anomaly arising from unusually fast response times for the fan level 1 probes. The notion that facts with a single association are responded to faster under certain conditions is not novel (e.g., Anderson, 1983). Another possible explanation may lie in the differences with the amount of sentential information the location and person precues provided. However, this issue cannot be resolved at this time. Further testing and possibly replication of the experiment would be required. Despite the person precue SL and ML fan effect results, it could be a serious error to reject the mental model account of the effects of mental organization given the dramatic change in the SL and ML fan effect difference.

The sorting task replicated Experiment 2 in showing a preference for a location-based organization, although the recognition data suggest a person-based organization. This corroborates the notion that either overt organization or recognition does not necessarily reflect the underlying mental model organization.

Although there is no other evidence in the mental model literature for such a dissociation in organization other than this dissertation, this result does parallel some research that has been done on schemas. Basically, recall, but not recognition, is affected by the particular schema used to retrieve information from memory (Alba & Hasher, 1983). In particular, studies by Anderson and Pichert (1978) and by Hasher and Griffin (1978) have shown that the type of previously presented information retrieved by a subject varies with the schema suggested by the experimenter. Subjects tend to recall information that is consistent with the suggested schema, and avoid recalling information that is less central to the schema. If the experimenter switches the schema suggested to the subject, then the subject can recall information that was previously unrecallable.

The schemas in Hasher and Griffin (1978) and Anderson and Pichert's (1978) studies provided retrieval cue sets which aided in recall. In doing this, the schemas guided the organization of the output of information in a fashion that may be inconsistent with its organization in memory. It is possible that the retrieval plan generated in both Experiments 1 and 2 was also organized in a fashion inconsistent with the structure of the underlying representation. In particular, the output of the facts may have been regulated by a general preference for location-based organization.

CHAPTER V

EXPERIMENT 4

As mentioned in the introduction, retrieval plan organization may not correspond to mental model organizations. It could be that the facts presented in Experiments 2 and 3 conformed more closely to some explicitly preferred organization that differs from the organization of the mental model. To test whether this consideration is viable, Experiment 4 assessed the preferred explicit organization of person-small location facts in the absence of memorization and memory tests. Only the card sorting task was given. The assumption here is that the explicit organizational preference may influence the structure of a retrieval plan which is generated for recall which may demand an explicit organization. This is because both explicit organization and retrieval plan creation would rely on organizational preferences which exist in the absence of situation specific organizations.

If the result of Experiment 4 is a predominantly location-based organization, this would suggest that the recall and sorting results of Experiments 2 and 3 might be due to some organizational preference that emerges in a retrieval plan. A predominantly person-based organization would suggest that the results of Experiments 2 and 3 might be due to employing location-based organization specifically for free recall, and that location-based organization is absent when this demand is not present. Finally, differential organization preferences for PS and LS sentences would indicate an organization preference based on either the first or last mentioned concept (Wolford, 1971) and that a more consistent organization in comes about when the facts have been committed to memory.

Method

Subjects. Thirty-five native English speakers were recruited from the subject pool at Michigan State University and given partial course credit for their participation. Eighteen were in the PS group and 17 to the LS group.

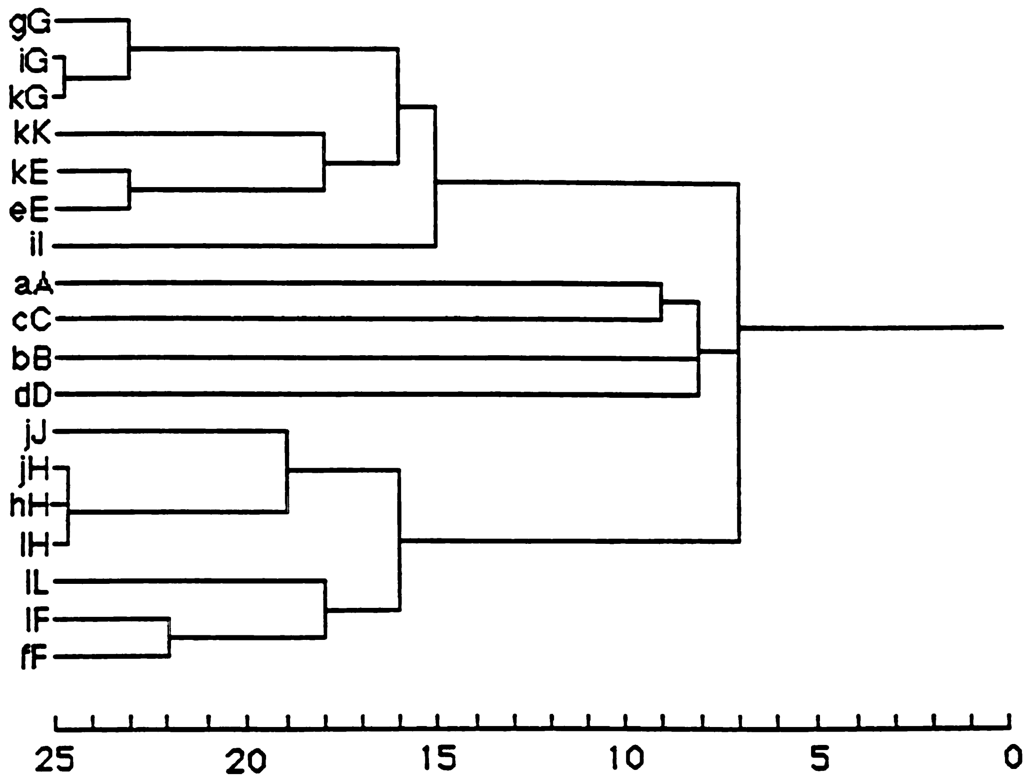
Materials and Procedure. The materials and design of each list of sentences was the same as Experiments 2 and 3. Subjects were tested in two groups in conjunction with a study on bizarre imagery. The card sorting task was presented at the end of the bizarre imagery study. The card sorting task took approximately 10 minutes and the entire experimental session lasted approximately 1/2 hour.

Results and Discussion

The results of Experiment 4 are presented in Figure 11. In general, there was less organization in Experiment 4 than Experiments 1 and 2. Though not striking, there was a general preference for location-based organization. This is consistent with Experiment 2. The ARCloc scores were higher than the ARCper scores (see Table 3), although this difference was not significant, $F(1,33) = 2.73$, $MSe = .401$, $p > .10$. Furthermore,

Figure 11

Results of a hierarchical cluster analysis for the card sorting data of Experiment 4 derived from the number of times pairs of items were grouping together.



although the scores were lower overall, the difference was in the same direction and was of the same magnitude as Experiment 2. The organizational pattern for the LS and PS groups did not differ as revealed by the ARC scores, $F < 1$.

Insert Figure 11 about here

Furthermore, the grouping pairs for these two groups were highly correlated (see Table 4). There was also a strong negative correlation between the ARCloc and ARCper scores. This indicates that subjects generally choose one organization over the other, although this organizational preference was not guided by sentence type. Finally, like Experiment 3, there was no grouping of the fan level 1 facts. This is probably because the subjects did not invest the effort required for repeated recalls and therefore did not discover the similarity among them.

Experiment 4 showed that, in the absence of memorization and memory tests, there is a tendency to explicitly organize facts about people being in small locations in a location-based fashion, although this tendency did not reach conventional levels of statistical significance. This suggests that the location-based results of free recall and sorting tasks of Experiments 2 and 3, and possibly Experiment 1 as well, were subject to this location-based preference in explicit organization.

From the theory of recall and recognition outlined above, it would seem that the recognition data of Experiments 1, 2, and 3 would be the more reliable measure of the structure of the underlying representation. This is because recognition involves the direct access to information from long-term memory, whereas recall relies on a memory search that is mediated by a retrieval plan. As seen in Experiment 4, the bias in explicit organization does not necessarily reflect the organizational structure that would be employed in long-term memory storage. The argument here is that the recall output organization happened to coincide with the mental model organization in Experiment 1 but not in Experiment 2 because of a preference for location-based organization when explicit organization is required as shown in Experiment 4.

In Experiments 1-4, the recall and card sort organizations were location-based. This type of organization was probably chosen for a number of reasons. One of the more prominent reasons might be that people are better able to recall items organized by location than people. Bellezza and Hatala (1992) found that when subjects were asked to recall a set of words, they were better able to recall words paired with rooms in their house (51%) than with aspects of themselves (40%). Not only was actual recall of words superior when related to spatial location, but so was recall of the cue (rooms in the house: 98%; aspect of the self: 90%). Bellezza and Hatala's explanation for this finding is that the self is a more abstract and dynamic entity, whereas parts of a house are more concrete and static, and therefore serve as better mental cues. This notion is supported by a survey of six sets of concreteness norms (Brown & Ure, 1969; Friendly, Franklin, Hoffman, & Rubin, 1982; Gilhooly & Hay, 1977; Gilhooly & Logie,

Table 6

Concreteness ratings from six different studies. Each value is a mean from a 7 point scale with 1 being low concreteness and 7 being high concreteness. The number of words on which the estimate is based is in parentheses.

Study	Location	Profession
Brown & Ure (1969)	6.30 (25)	6.04 (20)
Friendly, et al. (1982)	6.80 (25)	5.98 (34)
Gilhooly & Hay (1977)	5.54 (7)	5.50 (10)
Gilhooly & Logie (1980)	6.20 (73)	5.27 (90)
Pavio et al. (1968)	6.75 (57)	6.20 (56)
Spreen & Schulz (1966)	6.71 (18)	5.88 (21)

1980; Pavio, Yuille, & Madigan, 1968; Spreen & Schulz, 1966). Although both locations and profession names were rated very high in concreteness, locations were consistently rated as more concrete. The results of the survey of these studies are presented in Table 6. Subjects in the current experiments could have been relying on similar aspects of spatial locations when organizing their free recall retrieval plan.

Insert Table 6 about here

A demonstration is now needed that the person-based organization can be revealed through recall when the retrieval plan can be circumvented. Evidence in support of this notion would indicate that the recall of the facts in Experiments 1 and 2 was guided by a retrieval plan organized around location concepts. Experiment 5 used cued recall to test this possibility. Cued recall reduces the need for subjects to organize a retrieval plan. Subjects need only to retrieve those items associated with the cue (location or person).

Recall times should reflect the organization of the information into mental models. First consider conditions where only a single response is required, where there is only one concept associated with the cue. When the target concept is contained within a single mental model retrieval time should be relatively rapid because there are no other related mental models that will interfere with the retrieval of the appropriate one. When the target concept is represented in several mental models, even though the cue requires only a single concept response, these related mental models become activated as a consequence of the target mental model becoming activated, and interfere with the retrieval of the needed mental model. So, retrieval time is lengthened. The greater the number of irrelevant and related mental models, the greater the amount of interference, and the greater the retrieval slow down.

Now, consider conditions where multiple responses to a cue are required. If all of the responses are contained within a single mental model, then the retrieval of these concepts should be relatively rapid. If the target concepts are stored in several mental models, retrieval time can be expected to be quite slow. Each of these mental models must be located during the probabilistic search process and retrieved into working memory. However, while the activation of related mental models might be expected to hamper the retrieval of the first mental model, the retrieval of any subsequent representations could be expected to be aided by their activation during the previous memory sampling.

CHAPTER VI

EXPERIMENT 5

Method

Subjects. Twenty-four native English speakers were recruited from the subject pool at Michigan State University and given partial course credit for their participation. Half of the subjects received PS sentences and the other half LS sentences. An additional 3 subjects were replaced, 2 for failing to follow instructions, and 1 for failing to memorize the sentences within an hour.

Materials and Procedure. Because one of the dependent measures in Experiment 5 was the time needed to recall the appropriate associates for different cue types, it was necessary to equate the different concepts for syllable length more closely than they had been in Experiments 2 to 4. So, the production time of the actual words should be equivalent in the two cases. For Experiments 2 to 4, the location concepts were clearly longer (M = 4.6 syllables) than the person concepts (M = 2.3 syllables). In order to compensate for this, unnecessary components were dropped from most of the location concepts. (e.g., "new car's driver's seat" became "driver's seat"). Also, three person concepts were replaced with profession names that had more syllables. Thus, the two concept types were more equivalent in the number of syllables,

although the locations were still slightly longer ($\underline{M}$ = 3.0 syllables) than the person concepts ($\underline{M}$ = 2.7 syllables). The modified materials appear in Appendix B.

The procedure for memorizing the facts was like that used in Experiments 1, 2, and 4. Subjects took an average of 5.5 cycles to memorize the facts, and there was no difference between the PS and LS groups, $\underline{t}(22) = 1.32$, $\underline{p} > .20$. For the memory test, subjects were presented with the location and person concepts one at a time. The subjects' task was to name all of the people/locations associated with the cue. As soon as the cue was displayed in the center of the screen, a millisecond timer was started. When the subjects' began their responses, this triggered a voice key and that time was recorded. This voice onset time is referred to as RT1. When all of the concepts associated with a cue had been recalled the subject pressed a button on the computer mouse and the screen was erased. The time from voice onset to the button press was also recorded and is referred to as RT2. Each cue was presented once in each of a set of 8 blocks, for a total of 192 trials per subject. The order of trials within each block was randomized for each subject. The trimming procedure was applied to both the RT1 and RT2 data. A time greater than 2 standard deviations from the mean of either RT1 or RT2 resulted in the entire trial being dropped, rather than just that portion of the trial. This trimming procedure eliminated 8.9% of the RT data. Voice key responses and key presses occurring before 100 ms were also excluded from the RT analysis, as well as failures to initially trigger the voice key but were not counted as errors.

Subjects made an average of 1.7 of such voice keys errors on the cued recall test, with no difference between the PS and LS groups, $t < 1$. Errors were considered to be those responses which were wrong, or where RT1 or RT2 were greater than 10 s.

Results

CUED RECALL

Reaction Times. The RT1 data are presented followed by the RT2 data. For the RT1 data, those conditions which required only a single response are considered first. Voice onset time was greater when the response was a location concept (1559 ms) than when it was a person concept (1249 ms). There was an increase in RT1 as the number of associations with the response concept increased (1 - 1314; 2 - 1376; and 3 - 1523 ms). However, as can be seen in Figure 12, when the person and location responses are considered separately, this pattern of RT1s holds for the location responses, but not for person responses. There was no corresponding RT1 increase for the person responses. This parallels the fan effect results of the recognition test data of Experiments 2 and 3. The only other noticeable effect was a trend for the PS group to produce their responses sooner (1310 ms) than the LS group (1499 ms).

Insert Figure 12 about here

For multiple responses conditions, the data from single responses which had only 1 target concept associate was included as a baseline. Voice onsets occurred sooner for person responses

Figure 12

Reaction time and error rate results for the single response RT1 (voice onset time) conditions of Experiment 5.

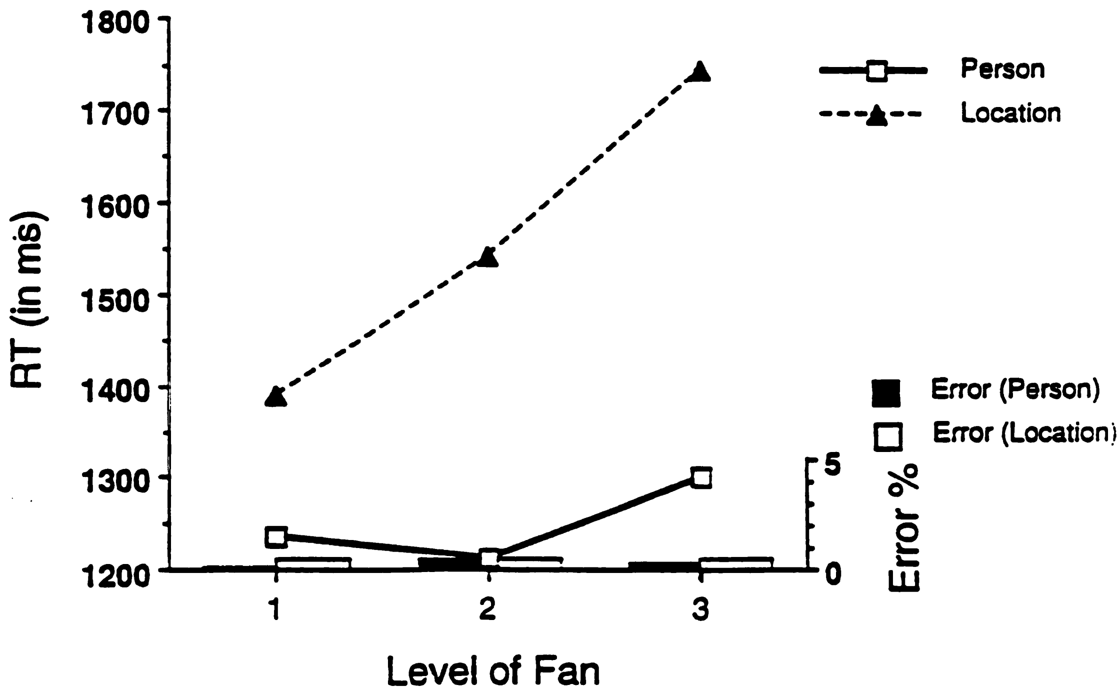
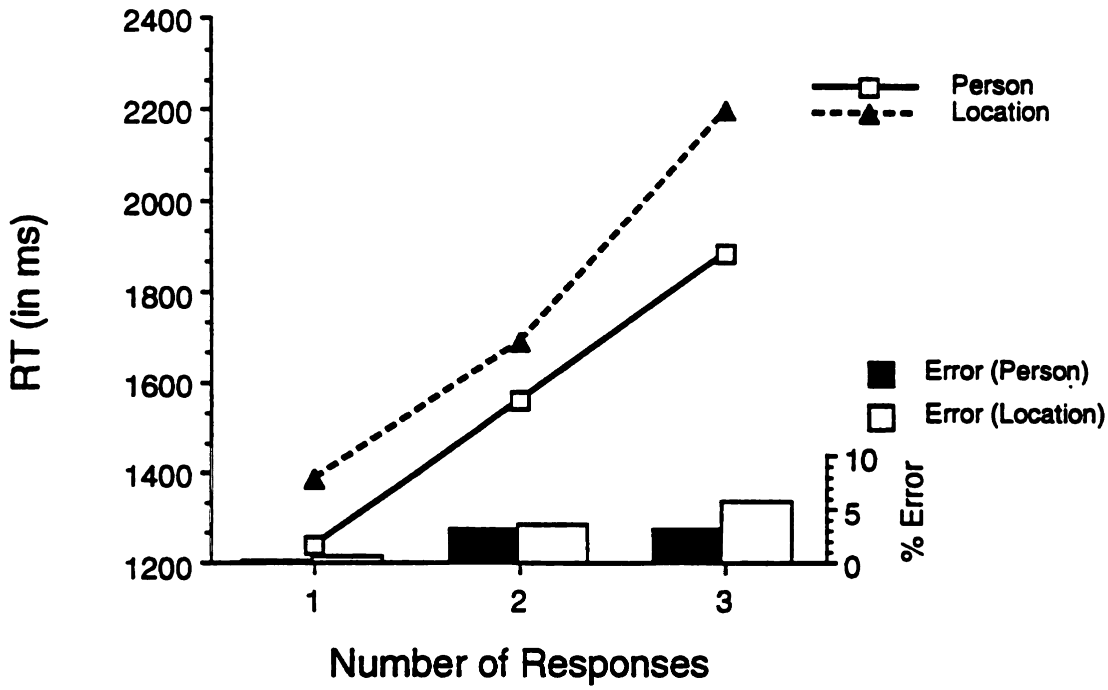


Figure 13

Reaction time and error rate results for the multiple response RT1 (voice onset time) conditions of Experiment 5.



(1560 ms) than location responses (1760 ms). Also, there was an increase in RT1 as the number of responses increased (1 = 1314; 2 = 1625; and 3 = 2041 ms). Although the difference was not significant, as can be seen in Figure 13, there is a slightly greater increase for location responses than for person responses. In addition to these findings, the PS group made their responses sooner (1531 ms) than the LS group (1789 ms).

Insert Figure 13 about here

The RT1 data for the single response conditions were submitted to a 2 (Type of sentence) X 2 (Response) X 3 (Response Fan) mixed ANOVA. The first variable was between subjects and the other two were within. There were significant main effects of Response, $F(1,22) = 70.92$, *MSe* = 48730, and Response Fan, $F(2,44) = 7.09$, *MSe* = 78615. There was also a significant interaction of these two variables, $F(2,44) = 6.93$, *MSe* = 37578. Simple effects tests showed that Response Fan was significant for the location responses, $F(2,44) = 10.45$, *MSe* = 73344, but not for the person responses, $F(2,44) = 1.19$, *MSe* = 42849, $p > .30$. Finally, the main effect of Type was marginally significant, $F(1,22) = 4.25$, *MSe* = 302585, $p < .06$.

The data for the multiple response conditions were submitted to a 2 (Type of sentence) X 2 (Response) X 3 (Number of Responses) mixed ANOVA. The first variable was between subjects and the other two were within. There were significant main effects of Response, $F(1,22) = 23.88$, *MSe* = 60607, and Number of Responses, *F*(2,44) - 66.35, *MSe* - 96377. The Response X Number of Responses interaction was not significant, *F*(2,44) - 2.08, *MSe* - 59769, *p* - .14. In addition, the main effect of Type was significant, *F*(1,22) - 6.11, *MSe* - 390760.

For the RT2 data, in single response conditions, the only noticeable difference was in the type of response. In particular, person concepts were produced faster (562 ms) than location concepts (669 ms). This may be due in part to the small advantage person concepts have in the number of syllables. As can be seen in Figure 14, there was no influence of the number of concepts associated with the target concept.

Insert Figure 14 about here

When multiple responses are considered, there is again an RT2 difference between person (1434 ms) and location concept responses (2011 ms). Not surprisingly, there was an RT2 increase accompanying an increase in the number of responses (1 - 621; 2 - 1530; and 3 - 3017 ms). As can be seen in Figure 15, the RT2 increase was greater for location than person concept responses. Finally, the RT2 increase was greater for the LS group than the PS group.

Insert Figure 15 about here

The RT2 data were submitted to similar analyses as for the RT1 data. For the single response conditions, only the main

Figure 14

Reaction time and error rate results for the single response RT2 (response production time) conditions of Experiment 5.

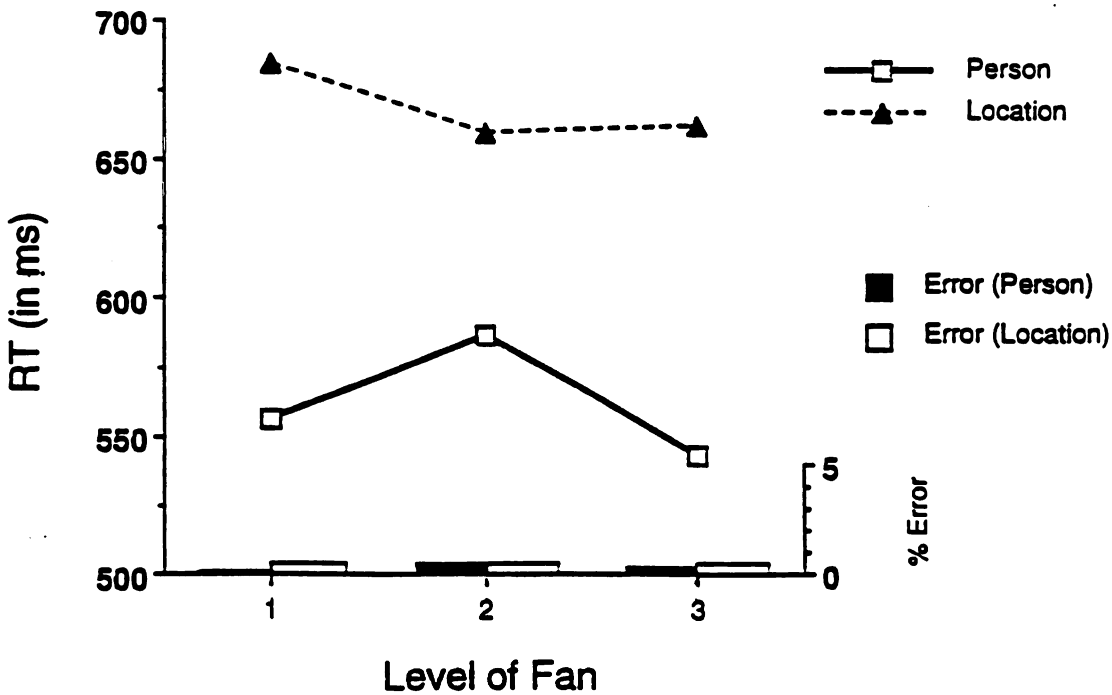
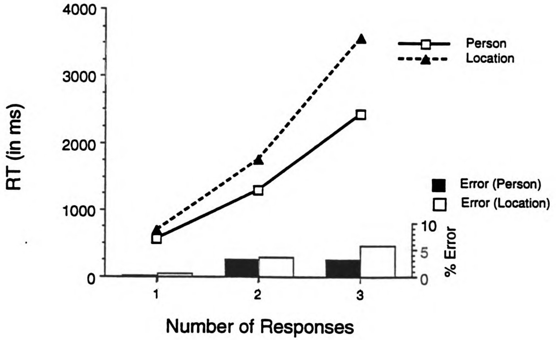


Figure 15

Reaction time and error rate results for the multiple response RT2 (response production time) conditions of Experiment 5.



effect of Response was significant, $F(1,22) = 143.23$, $MSe = 2901$. For the multiple response conditions, there were significant main effects of Response, $F(1,22) = 72.67$, $MSe = 164519$, and Number of Responses, $F(2,44) = 452.87$, $MSe = 155055$. The interaction of these two factors was also significant, $F(2,44) = 30.15$, $MSe = 105090$. Simple effects tests showed that the effect of Response was significant for 2 and 3 item responses, but not for single item responses. Finally, there was a significant Type X Number of Responses interaction, $F(2,44) = 3.73$, $MSe = 155055$.

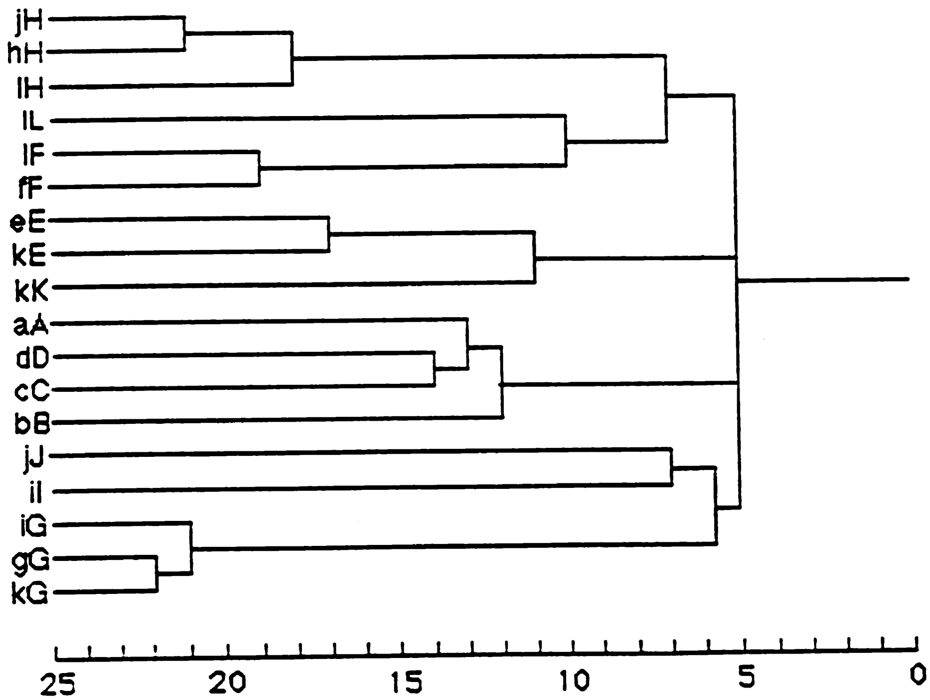
Errors. In general, there were no interesting patterns involving errors. Subjects made an average of 0.4% errors for single response conditions and there were no differences for different sentence types, response types or response fan. For the multiple response conditions, there was an increase in the number of errors as the number of responses increased (1 = 0.3; 2 = 3.4; 3 = 4.4%), $F(2,44) = 3.37$, $MSe = 65$. In addition, there was a significant Type X Response interaction, $F(1,22) = 6.29$, $MSe = 20$, with PS subjects showing no difference between response conditions (Person = 2.0%; Location = 1.3%) whereas LS subjects did show a difference (Person = 2.3%; Location = 5.3%). This difference is primarily due to 3 subjects in the LS condition who had 12%, 8.3% and 7.3% errors respectively, whereas the highest error rate for the PS group was 5.7%.

CARD SORTING

The results of the card sorting task are presented in Figure 16. The results, like Experiments 1 to 4, were generally consistent with a basic location-based organizational preference.

Figure 16

Results of a hierarchical cluster analysis for the card sorting data of Experiment 5 derived from the number of times pairs of items were grouping together.



The ARCloc scores were greater than the ARCper scores, $F(1,22) = 25.48$, $MSe = .165$ (see Table 3), and there was a negative correlation between them (see Table 4). Finally, there was no difference between the LS and PS groups, $F < 1$, and the pair groupings between them were highly correlated.

Insert Figure 16 about here

In order to further demonstrate that the organizational pattern for the card sorting task in all 5 experiments showed a similar location-based organization, all possible pair groupings from all of the experiments were correlated with each other. These correlations are presented in Table 7. As can be clearly seen, these correlations are quite high, reflecting a high degree of organizational consistency across all of the experiments.

Insert Table 7 about here

Discussion

There are two basic results for Experiment 5. The first was that, for single response conditions, as the number of associations with the target concept increased, there was an increase in voice onset time for location, but not person responses. This is consistent with the notion that person-based mental models were created. Specifically, each person response was stored in only a single mental model. Therefore, there was no retrieval interference in this condition and RT did not increase

Table 7

Correlation of card sorting pair groupings for all experiments.

	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Experiment 1	---				
Experiment 2	.85	---			
Experiment 3	.81	.70	---		
Experiment 4	.76	.73	.92	---	
Experiment 5	.89	.72	.82	.88	---

with increased response fan. However, each location response was stored in as many mental models as there were people in that location. These other mental models interfered with the appropriate one, making retrieval more difficult. This difficulty slowed down retrieval time and therefore voice onset time as a function of the number of concepts associated with the response.

The second major result was that as the number of location responses increased, the time to report them was greater than for the same number of person responses. This appears to be inconsistent with the notion that the information was integrated into person-based mental models. In recognition, there was no interference (no fan effect) for conditions in which several locations were associated with a single person. It would seem plausible that no interference should occur for cued recall under similar conditions. Specifically, all of the information in the mental model would be present in working memory where it could be checked rather rapidly.

This second result may be due to a number of possibilities. One explanation is related to the fact that the location concepts in this experiment were produced at a slower rate than the person concepts. The additive location production times could have contributed to the larger increase in RT2 for location responses. If the production times for the two response types were equated and there was no difference between the increase in production times of multiple concepts of both types, this would indicate that the RT2 results of Experiment 5 were due to a simple concept production time difference.

Another possibility is based on the notion that subjects prefer to explicitly organize by location. In this experiment, while the single response conditions did not require a retrieval plan, one may have been used for the multiple response conditions. In keeping with the previous experiments, it could be assumed that the retrieval plan would be organized in a location-based fashion. In such a case, the slow-down for the RT2 location responses would be due to a cue switching process as part of the structure of the retrieval plan. So, given equivalent production times for both response types, the recall of several locations would still take longer.

A third possibility is based on the notion that the internal structure of a memory trace may be complex and not immediately available when retrieved into working memory (Raaijamakers and Shiffrin, 1981). The SAM model does not state what sorts of internal structures are possible, and what sorts of effects of a search of the internal trace structure can be expected. However, it has been suggested by other researchers that the verification of relational information could occur after retrieval of a trace from long-term memory (Dosher, 1983). The SAM model could be modified to include an internal trace search stage that would fit with its characterization of recognition and recall and be generally consistent with the spirit of the model. If this possibility were correct, it could also lead to the multiple response results of Experiment 5. An experiment proposed at the end of the general discussion would serve to test between the second and third possibilities.

In summary, for Experiment 5 the pattern of voice onset RTs are accounted for by the organization of facts into person-based mental models, but the pattern of response production times is more problematic. At this point it cannot be determined whether this is a circumstantial result of the particular stimuli used, with locations taking longer to produce than people, or whether this pattern of data is due to some more complex processes operating in the recall of information from mental models. More research is needed.

CHAPTER VII

QUESTIONNAIRE

Finally, the results of the questionnaires from all of the experiments are provided in Table 8. The first question was open ended and concerned subjects' memorization strategies. *Location* and *Object/Person* refer to responses which indicate a location- or object-/person-based organizations. *Generic grouping* means that the subject referred to an organizational grouping, but did not specify any particular grouping orientation. *Image* responses are descriptions of memorization strategies that referred to forming an image, while *Stories* are responses that described an attempt to either create an actual story out of the memorized sentences, or to create some meaningful relation between the main concept types, such as "A carpenter drove a nail into someone's head and must go to confession." *Relate to familiar* refers to a report which conveys the attempt to relate the facts to their personal knowledge, such as imaging a lawyer friend of theirs in the situation. *Alphabetic* refers to the strategy of relying on common alphabetic characteristics of the main concept (e.g., carpenter - confessional), and *repetition* refers to trying to memorize the facts by repeating them over and over. Finally, for *No method*, subjects either reported nothing or something that gave no information (e.g., "I memorized the sentences.") The data were

Table 8
Results of Post-test Questionnaire.

	Location	Object/ person	Generic grouping	Image	Stories
Experiment 1	33.3	8.3	25.0	29.2	8.3
Experiment 2	41.7	4.2	12.5	12.5	20.8
Experiment 4	25.0	2.5	0.0	37.5	32.5
Experiment 5	37.5	0.0	25.0	20.8	37.5
Mean	34.4	3.8	15.6	25.0	24.8

	Relate to familiar	Alpha- betic	Repetition	No Method
Experiment 1	8.3	0.0	4.2	4.2
Experiment 2	4.2	12.5	0.0	8.3
Experiment 4	10.0	25.0	7.5	12.5
Experiment 5	12.5	12.5	4.2	4.2
Mean	8.8	12.5	4.0	7.3

Question #1: Briefly describe, in your own words, how you tried to memorize the sentences.

	2 Yes	3 Yes	4 Yes	5 Same	6 Same	7 Yes
Experiment 1	87.5	25.0	45.8	100.0	75.0	12.5
Experiment 2	70.8	16.7	41.7	100.0	95.8	29.2
Experiment 4	84.6	33.3	30.8	97.4	94.9	30.8
Experiment 5	91.7	20.8	25.0	100.0	95.8	12.5
Mean	83.7	24.0	35.8	99.0	90.4	21.3

Question #2: Did you try to form an image of a scene in your head?

Question #3: Did you try to fashion a story out of the sentences you were memorizing?

Question #4: Did you try to think of a sequence of events about the things described in the sentences to help you memorize?

Question #5: When a location was mentioned in more than one sentence, did you consider it to be the same location each time (e.g., the same barber shop or different barber shops)?

Question #6: When an object/person was mentioned in more than one sentence, did you consider it to be the same object/person each time (e.g., the same cola machine or different cola machines)?

Question #7: If an object/person was described as being in more than one location, did you imagine it traveling from one place to another?

sorted into these categories only by myself. The numbers in the table do not add up to 100% because some subjects reported using more than one method.

Insert Table 8 about here

In general, there were few meaningful differences among the experiments. The basic pattern reflects strategies that rely on considering the facts as descriptions of situations, such as forming images of scenes or generating stories about them. The large preference for forming images also probably lead to the high proportion of location-based overt organization output strategies. For Experiment 1 (object-location sentences) there was a greater tendency to consider each object as a different instance, whereas for the rest of the experiments (person-small location sentences) there were very few instances of considering a person in different places as being a different person. In addition, subjects in Experiment 1 were less likely to create stories or imagine an object being moved from place to place. This may be due in part to the fact that the objects are inanimate, would serve as poor story protagonists, and lack the ability to voluntarily move from place to place. Instead, the preference would be to consider each object a different instance in each location.

CHAPTER VIII

GENERAL DISCUSSION

Summary of the Current Experiments.

A set of five experiments was conducted that centered on how the retrieval of information from mental models differs during recognition and recall. This survey was guided by the SAM (Gillund & Shiffrin, 1984) model of memory. In Experiments 1 and 2, the organizational basis indicated by recognition varied with respect to the type of materials used. In particular, the recognition data suggested that facts about objects being in locations produced location-based organization, whereas facts about people being in small locations produced person-based organization. However, the organizational basis suggested by free recall was location-based in both cases. To assess the dissociation of recall and recognition, especially in Experiment 2, additional experiments were conducted.

The cued recognition test in Experiment 3 failed to fully supported the recognition test results of Experiment 2. Person precues were able to attenuate the SL and ML fan effect difference, whereas location cues did not. However, the reduction in the fan effect for the SL condition. Further research is needed to resolve whether these results actually reflect selective facilitation or interference from the person precues.

The proposed explanation for the discrepancy between the recall and recognition data of Experiment 2 is based on the notion that free recall requires a retrieval plan and that the preferred organization of the retrieval plan is one based around location (Bellezza & Hatala, 1992). Using only a sorting task, Experiment 4 demonstrated that subjects preferred a location-based organization of person-small location facts. This organizational preference, even in the absence memorization and memory tests is thought to have been carried over to the organization of the retrieval plan, which consequently leads to the retrieval of person-small location facts in a location-based fashion. In general, the card sorting task in all experiments generally paralleled the preference for location-based organization suggested by free recall.

In order to circumvent the retrieval plan, Experiment 5 employed a cued recall test in which only a small subset of items were recalled on each trial. For voice onset times, a fan effect was observed for location response preparation in accordance to the fan effect results seen in recognition tests. This was true for both single and multiple response conditions. However, the production time data showed that producing several locations took longer than producing several people. This is inconsistent with the mental model view which would seem to predict that the production of several locations would be faster because they are all contained in the same representation. Several potential explanations for the results of the multiple response conditions were presented, although none are supported at this time.

Future Directions

The next step after the experiments reported in this dissertation would be to reexamine the ambiguous results that arose at some of the more critical junctures of this investigation. This first is would be a replication of Experiment 3, the precuing experiment, to assess whether the change in the retrieval pattern following a person precue was due to an increase in the ML fan effect to more closely approximate the SL fan effect, or a decrease in the SL fan effect to more closely approximate the ML fan effect. Perhaps the most important change to make in a replication of that experiment would be to equate the concepts for length. The pattern of results observed in the person precue condition may have been due in part to the inequality in the concept lengths.

The other ambiguous result that requires closer examination is that of Experiment 5, the cued recall experiment. The first step to take would, again, be to more closely equate the two concept types for length, syllable length in this case. The difference between the two concept types in the present case, though small, may have been large enough to produce the results observed, or other factors, like frequency, or phonological structure may have been important. An absence of a difference in the multiple response production time increases for different response types, given a difference in the pattern of voice onset times that parallels that found in Experiment 5, would indicate that the RT2 results found were due to a simple difference in the production time of the different concept types.

However, if a difference between the increase in production times were to remain with an equating of general concept production time, then some means would be needed to differentiate between the retrieval plan and internal search of the memory trace possibilities outlined in the discussion section of Experiment 5. The best way to provide for such a differentiation would be to have the subjects memorize object-location facts rather than person-small location facts. Under those conditions, if the multiple response production time is greater for the location responses, this would provide support for the second possibility which is based on the notion of a location-based retrieval plan guiding response output under multiple response conditions. However, if the multiple response production time is greater for the object responses, this would provide support for the third possibility which is based on the notion of interference during an internal search of a complex memory trace.

Summary

In summary, the research presented in this dissertation demonstrate that different types memory tests are subject to different types of organizational influences on dependent measures which are used as indicators of mental organization. In particular, tasks that require the access of only small sets of information, such as recognition and cued recall, demonstrate a pattern of data that reflects the organization of information into mental models which are mediated by the information content. However, in tasks that require the access of larger sets of information, such as free recall, in which the subject needs to create some reliable output organization, the pattern of data reflect the organization of information around spatial locations, regardless of the information content. This organizational preference was also observed when subjects were explicitly asked to organize the information into groups of their own choosing.

The data reported here largely reflect the position that subjects are organizing the facts into mental models, although it is also decidedly weak at some critical points. The mental model organization that is assumed to be generated and stored does not necessarily impact on information retrieval, especially under conditions when a needed retrieval plan is organized in a fashion in contrast to the situational organization. Such a retrieval plan organization may reflect more conscious strategies, such as relying on concrete aspects of the information. like location, that could be used in generating a mental image.

Such a dissociation between the structure of the underlying representation and the overlaying retrieval scheme, even in the absence of experimenter influence (e.g., Hasher & Griffin, 1978), brings into question the purpose and utility of creating situation oriented representations when the mechanisms for retrieval may disregard this organization in the access of the desired content information. This dissociation could have impacts on processing that were not observable here. For instance, it may be that under conditions where there is a gross mismatch between the storage and retrieval structures that more retrieval errors would occur. Such effects could easily have been missed in the research reported here because of the high degree of overlearning and very low error rates. More research is needed to resolve the implications of the differences between information output organization in recall and recognition from mental models, but an important beginning has been made.

APPENDICES

APPENDIX A

Subjects and predicates used in Experiment 1.

Objects	Locations
Broken Window	Airport
Bulletin Board	Barber Shop
Ceiling Fan	Car Dealership
Cola Machine	City Hall
Digital Wall Clock	Cocktail Lounge
Fire Extinguisher	Ice Cream Parlor
Oak Counter	High School
Pay Phone	Hotel
Potted Palm	Laundromat
Revolving Door	Movie Theatre
Welcome Mat	Office Building
Wire Waste Basket	Public Library

APPENDIX B

Subjects and predicates to be used in Experiments 2-5. The people in parentheses and the portions of the location descriptions in italics were not used in Experiment 5 to make number of syllables for the people and locations equivalent.

People:	Locations:
Architect	*Old* Tire Swing
Banker	*New Car's* Driver's Seat
Carpenter	*Nearest* Voting Booth
Doctor (Janitor)	*Store's* Dressing Room
Engineer	*Back Room's* Tanning Bed
Farmer	*Greyhound* Bus's Bathroom
Grocer	*Ornate* Throne
Lawyer (Psychologist)	Witness Stand
Mechanic	*Tele*phone Booth
Salesman (Senator)	Operating Table
Teacher	*Blue and Yellow* Kayak
Writer	*Dark* Confessional

LIST OF REFERENCES

LIST OF REFERENCES

Alba, J. W., & Hasher, L. (1983). Is memory schematic? Psychological Bulletin, 93, 203-231.

Albrecht, J. E., & O'Brien, E. J. (1991). Comprehension strategies during reading. Poster presented at the meeting of the Psychonomic Society, San Francisco.

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronomial antecedents as a function of episode shifts in narrative text. Quarterly Journal of Experimental Psychology, 35A, 427-440.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. Cognitive Psychology, 6, 451-474.

Anderson, J. R. (1975). Item-specific and relation-specific interference in sentence memory. Journal of Experimental Psychology: Human Learning and Memory, 1, 249-260.

Anderson, J. R. (1983). The Architecture of Cognition. Cambridge MA.: Harvard University Press.

Anderson, R. C. , & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. Journal of Verbal Learning and Verbal Behavior, 17, 1-12.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.) Attention and Performance (Vol. 4). New York: Academic Press.

Barwise, J., & Perry, J. (1983). Situations and Attitudes. Cambridge MA.: MIT Press.

Bellezza, F. S., & Hatala, M. N. (1992). Want to remember? Don't think of yourself: Think of your house! Paper presented at the meeting of the Midwestern Psychological Association, Chicago, Illinois.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. The Journal of General Psychology, 49, 229-240.

Bousfield, W. A., Puff, C. R., & Cowan, T. M. (1964). The development of constancies in sequential organization during repeated free recall. Journal of Verbal Learning and Verbal Behavior, 3, 489-495.

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. Science, 247, 44-48.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

Brown, W. P., & Ure, D. M. J. (1969). Five rated characteristics of 650 word association stimuli. British Journal of Psychology, 60, 233-249.

Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. Journal of Memory and Language, 31, 74-98.

Byrne, R. M. J., & Johnson-Laird, P. N. (1989). Spatial reasoning. Journal of Memory and Language, 28, 564-575.

Dosher, B. A., (1983). Relational information in sentence memory. Journal of Psycholinguistic Research, 12, 377-396.

Dyne, A. M., Humphreys, M. S., Bain, J. D., & Pike, R. (1990). Associative interference effects in recognition and recall. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 813-824.

Ehrlich, K. & Johnson-Laird, P. N., (1982). Spatial descriptions and referential continuity. Journal of Verbal Learning and Verbal Behavior, 21, 296-306.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. Journal of Experimental Psychology: General, 119, 63-76.

Franklin, N., Tversky, B., & Coon, V. (in press). Switching points of view in spatial mental models. Memory and Cognition.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. Behavior Research Methods and Instrumentation, 14, 375-399.

Garnham, A. (1981). Mental models as representations of text. Memory and Cognition, 9, 560-565.

Garnham, A. (1987). Mental Models as Representations of Discourse and Text. New York: Halsted Press.

Gilhooly, K. J., & Hay, D. (1977). Imagery, concreteness, age-of-acquisition, familiarity, and meaningfulness values for 205 five-letter words having single solution anagrams. Behavior Research Methods and Instrumentation, 9, 12-17.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. Behavior Research Methods and Instrumentation, 12, 395-427.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-67.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help build mental models. Journal of Memory and Language, 31, 129-151.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute foregrounding during text comprehension. Journal of Memory and Language, 26, 69-83.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 846-858.

Hasher, L., & Griffin, M. (1978). Reconstructive and reproductive processes in memory. Journal of Experimental Psychology: Human Learning and Memory, 4, 318-330.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. Psychological Review, 93, 411-428.

Johnson-Laird, P. N. (1983). Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness. Cambridge, MA.: Harvard University Press.

Johnson-Laird, P. N. (1989). Mental Models. In M. I. Posner (Ed.) Foundations of Cognitive Science. Cambridge, MA.: MIT Press.

Johnson-Laird, P. N., Hermann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. Psychological Bulletin, 96, 292-315.

Mandler, G., & Dean, P. J. (1969). Development of serial order in free recall. Journal of Experimental Psychology, 81, 207-215.

Mani, K., & Johnson-Laird, P. N., (1982). The mental representation of spatial descriptions. Memory and Cognition, 10, 181-187.

Moeser, S. D. (1979). The role of experimental design in investigations of the fan effect. Journal of Experimental Psychology: Human, Learning, and Memory, 5, 125-134.

Morrow, D.G., Greenspan, S. L., & Bower, G. H., (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and Language, 26, 165-187.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1988). Updating situation models during narrative comprehension. Journal of Memory and Language, 28, 292-312.

Pellegrino, J. W. (1971). A general measure of the organization in free recall for variable unit size and internal sequential consistency. Behavioral Research Methods and Instruments, 3, 241-246.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. Journal of Memory and Language, 24, 503-518.

Pavio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology: Monograph, 76, 1-25.

Raaijamakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. Psychological Review, 88, 93-134.

Raaijamakers, J. G. W., & Shiffrin, R. M. (1992). Models of recall and recognition. Annual Review of Psychology, 43, 205-234.

Radvansky, G. A., Gerard, L. D., Zacks, R. T., & Hasher, L. (1990). Younger and older adults use of mental models. Psychology and Aging, 5, 209-214.

Radvansky, G. A., Spieler, D. H., & Zacks, R. T. (in press). Mental model organization. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and fact retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 940-953.

Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. Journal of Experimental Psychology, 8, 16-36.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. Cognitive Psychology, 21, 139-155.

Reder, L. M., & Anderson, J. R. (1980). A partial resolution to the paradox of interference: The role of integrating knowledge. Cognitive Psychology, 12, 447-472.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. Annual Review of Psychology, 39, 475-543.

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.

Sanford, A. J., & Garrod, S. C. (1981). Understanding Written Language: Explorations of Comprehension Beyond the Sentence. New York: John Wiley and Sons.

Sharp, D. L., & McNamara, T. P. (1990). Spatial mental models in narrative comprehension: Now you see them now you don't. Poster presented at the 1990 meeting of the Psychonomic Society in Atlanta, GA.

Shuell, T. J. (1969). Clustering and organization in free recall. Psychological Bulletin, 72, 353-374.

Smith, E. E., Adams, N., & Schorr, D. (1978). Fact retrieval and the paradox of interference. Cognitive Psychology, 10, 438-464.

Spreen, O., & Schulz, R. W. (1966). Parameters of abstraction, meaningfulness, and pronounceability for 329 nouns. Journal of Verbal Learning and Verbal Behavior, 5, 459-468.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. Journal of Memory and Language, 31, 261-292.

Tulving, E. (1962). Subjective organization in free recall of "unrelated words". Psychological Review, 69, 344-354.

White, R. T., & Kelly, M. E. (1978). Measurement of categorical clustering in free recall. British Journal of Psychology, 69, 451-465.

Wilson, S. G., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1992). Mental models of narrative comprehension. Unpublished manuscript.

Wolford, G. (1971). Function of distinct associations for paired-associate interference. Psychological Review, 78, 303-313.

Zacks, R.T., Radvansky, G. A., & Hasher, L. (in progress). Mental model retrieval and aging.