LINGLE, JOHN HUNTER
PROBE RECOGNITION SPEED AS A MEASURE OF
THOUGHT ACTIVATION DURING MEMORY-BASED
IMPRESSION JUDGMENTS.
THE OHIO STATE UNIVERSITY, PH.D., 1978

© 1978

JOHN HUNTER LINGLE

ALL RIGHTS RESERVED
PROBE RECOGNITION SPEED AS A MEASURE OF THOUGHT ACTIVATION
DURING MEMORY-BASED IMPRESSION JUDGMENTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

John H. Lingle, B.A.

The Ohio State University
1978

Reading Committee:
Timothy C. Brock
Thomas M. Ostrom
Harvey G. Shulman

Approved By

Thomas M. Ostrom
Advisor
Department of Psychology
I wish to express my appreciation to the several people who most directly contributed to the completion of this dissertation. First and foremost, my thanks go to Dr. Thomas M. Ostrom who was both a good friend and a superstar among advisors. His support, advice and good counsel served as a constant source of inspiration and encouragement.

I also thank Dr. Timothy C. Brock who provided guidance during my early graduate school years as well as Dr. Anthony G. Greenwald, Karl Rosenberg, and Lewis Hinkle who patiently guided me in the use (and occasional abuse) of the minicomputer.

My special thanks go to Janet Dukerich and Deborah Glover for their continued assistance. Extra special thanks go to Janet who literally worked night and day to help run subjects, analyze data and prepare the final report. Without her help, all would have been lost on more than one occasion.

Finally, I extend my thanks to my wife Kathie and son Michael for their uncomplaining support during both the good times and the bad.
VITA

PERSONAL DATA

Date and place of birth: February 16, 1944; Cincinnati, Ohio

Education:

B. A. Davidson College
Davidson, North Carolina
June, 1966
Major: History
Minor: Economics

Ph. D. Ohio State University
Columbus, Ohio
Summer, 1978 (expected)
Major: Social Psychology
Minors: Quantitative psychology
Experimental psychology

MEMBERSHIPS

American Psychological Association (student)
The Society for the Psychological Study of Social Issues
American Association for the Advancement of Science
The Society for the Advancement of Social Psychology

PROFESSIONAL EXPERIENCE

1978 - present Graduate Research Associate, College of Administrative Science, Ohio State University.

1976 - 1977 Graduate Research Assistant to Dr. Thomas M. Ostrom, Department of Psychology, Ohio State University.

1974 - 1976 Teaching Associate, Ohio State University.

1972 - 1975 NIMH predoctoral trainee in social psychology, Ohio State University.


1968 - 1969 Peace Corps (staff trainee) Fellow, Washington, D. C.


PUBLICATIONS

Lingle, J. H., Brock, T. C., & Cialdini, R. B. Surveillance instigates entrapment when violations are observed, task allegiance is high, and sanctions are severe. *Journal of Personality and Social Psychology*, 1977, 35, 419-429.


CONVENTION PRESENTATIONS


### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>VITA.</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES.</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. MEMORIAL PROCESSES IN PERSON PERCEPTION.</td>
<td>1</td>
</tr>
<tr>
<td>Introduction.</td>
<td>1</td>
</tr>
<tr>
<td>Three Alternative Approaches for Studying Person Judgments.</td>
<td>3</td>
</tr>
<tr>
<td>II. MEASURING SILENT THOUGHT</td>
<td>20</td>
</tr>
<tr>
<td>Limits to Current Practices</td>
<td>20</td>
</tr>
<tr>
<td>An Alternative Approach for Measuring Thought</td>
<td>23</td>
</tr>
<tr>
<td>III. THE MEASURABILITY OF INTERNALLY ACTIVATED THOUGHT:</td>
<td>34</td>
</tr>
<tr>
<td>EXPERIMENTS 1 &amp; 2</td>
<td></td>
</tr>
<tr>
<td>The Task: Hum a Happy Tune</td>
<td>34</td>
</tr>
<tr>
<td>Method.</td>
<td>36</td>
</tr>
<tr>
<td>Results</td>
<td>41</td>
</tr>
<tr>
<td>Discussion</td>
<td>44</td>
</tr>
<tr>
<td>IV. MEASURING THOUGHT ACROSS TIME AND THOUGHT:</td>
<td>49</td>
</tr>
<tr>
<td>EXPERIMENT 3</td>
<td></td>
</tr>
<tr>
<td>The Task: Think a Happy Rhyme.</td>
<td>49</td>
</tr>
<tr>
<td>Method.</td>
<td>53</td>
</tr>
<tr>
<td>Results</td>
<td>62</td>
</tr>
<tr>
<td>Discussion</td>
<td>70</td>
</tr>
<tr>
<td>V. MEASURING WHAT IS ACCESSED DURING A PERSON JUDGMENT:</td>
<td>73</td>
</tr>
<tr>
<td>EXPERIMENT 4</td>
<td></td>
</tr>
<tr>
<td>The Task: Would a Friendly Person Make a Good Nightwatchman?</td>
<td>73</td>
</tr>
</tbody>
</table>
Method .......................................................... 75
Results .......................................................... 84
Discussion ....................................................... 90

VI. SUMMARY AND CONCLUSIONS ............................... 96

Summary .......................................................... 96
Methodological Limitations ................................. 98
Applications .................................................... 102

APPENDIXES

A. Lyrics to the rhymes and songs used as stimuli in Experi­ments 1 and 2 (underlined words were used as probes). ...... 107

B. Rhymes used as practice and experimental stimuli in Ex­periment 3 with probe presentation positions indicated by single underscoring and probe words indicated by double underscoring. ............................................. 111

C. Occupations and 32 trait sets used as judgment (J), rel­evant (R), and irrelevant (I) descriptor traits in Experiment 4 ................................................................. 114

LIST OF REFERENCES ............................................. 118
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean response latencies in milliseconds to matched and unmatched probes: Experiment 1 - word naming</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>Mean response latencies in milliseconds to matched and unmatched probes: Experiment 2 - color naming</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions: Experiment 3 - Replication 1 &amp; 2 combined</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions: Experiment 3 - Replication 3 (no money)</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions: Experiment 3 - Replication 4 (money)</td>
<td>67</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Possible patterns of probe recognition speed when an early (position #3) and late (position #6) word in a sequence of eight are displayed in each position.</td>
<td>51</td>
</tr>
<tr>
<td>2</td>
<td>Mean probe recognition speed as a function of probe word type and presentation position for money and no money replications: Experiment 3</td>
<td>69</td>
</tr>
<tr>
<td>3</td>
<td>Flow chart of the experimental task used in Experiment 4</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>Mean probe recognition speed as a function of probe type and time of probe presentation in Experiment 4</td>
<td>85</td>
</tr>
</tbody>
</table>
CHAPTER I
MEMORIAL PROCESSES IN PERSON PERCEPTION

Introduction

A voter enters a polling place undecided which of two candidates to vote for. A young man contemplates whether to call and invite Sally or Sue for a date. A country's chief executive takes a lonely stroll to make a final decision on a cabinet appointment. A mother worries whether her young son is mature enough to walk to school alone.

In these and many other daily situations people draw upon their memory for people to make decisions that affect both themselves and those around them. In the process of making any single such judgment a person is unlikely to remember everything he or she knows about the other. As a consequence, a memory-based judgment that one person makes about another will often depend more upon what is selectively remembered about the person than what is actually known about the person.

In spite of the importance of memory retrieval processes to interpersonal judgments, social psychologists have spent little time investigating how people remember information about other people when making a decision. The two major areas of research that developed directly from Asch's (1946) early empirical investigations of the impression formation process have focused on (a) identifying the principal dimensions people rely on when they make similarity judgments or describe other people and (b) algebraically describing how people combine information cues into a
single, stimulus-based judgment (that is, a judgment that is made while
descriptive stimulus information is available for the subject to consider).
Neither of these two research approaches have investigated how information
about a single person is represented and organized in memory, or how
information about a person, once stored in memory, is accessed during a
decision.

More recent investigations of person perception processes that have
adopted an information processing approach have begun to research how
person information is encoded and organized in memory. This has included
such questions as (a) the degree to which personality traits serve as
impression prototypes (Cantor & Mischel, 1977), (b) how well formed self-
schemata influence memory and the interpretation of new information
(Markus, 1977), (c) the way in which facts about a person are organized
in memory (Anderson, 1977; Anderson & Hastie, 1974) and (d) how first
impression judgments affect memory for information about a person
(Lingle, Geva & Ostrom, 1975; Lingle & Ostrom, in press a). However, the
question of how people draw upon their memory of a person when making a
judgment continues to receive minimal attention. One contributing factor
to this neglect is that to date only two procedures have been developed
for identifying the thoughts a person activates in memory during a decision.
One of these (cf., Calder, Insco & Yandell, 1974; Higgins, Rholes & Jones,
1977) relies on the intercorrelation between remembered stimuli, self-
generated thought and a judgment while the other (cf., Lingle & Ostrom,
in press b) uses decision time to infer how memory is accessed. Both
methods, unfortunately, are characterized by several interpretative
difficulties.
This paper proposes an alternative method for measuring the thoughts people access in memory when making a decision. In Chapter I, three approaches that have been, and are currently being, used to study person judgments are briefly reviewed. The degree to which each has concerned itself with the role of memory in person perception is identified. In Chapter II, the two methods that have been used to infer how people access information during a judgment are critiqued and an alternative method for measuring thought activation is proposed. Chapters III and IV report three studies conducted to validate the method, while Chapter V describes a study which uses the procedure to test several hypotheses about the way in which people draw on memory when they make memory-based person judgments. Finally, in Chapter VI strengths and weaknesses of the proposed procedure as a measure of thought are considered.

Three Alternative Approaches for Studying Person Judgments

In 1946 Soloman Asch reported a series of studies designed to demonstrate that the impressions people form about one another go beyond a simple summation of known characteristics. The paradigm Asch employed to illustrate this view consisted of describing a person with a short list of personality traits and then asking a subject to (a) write a description of the person and/or (b) indicate which adjectives, in a list of bi-polar adjective pairs, would be most characteristic of the person. These early studies proved more demonstrative than explanatory. They illustrated that within some contexts certain traits appear to organize an impression. However, Asch failed to specify any memorial principles or mechanisms that might explain how people generate descriptions or make trait ratings about others. Similarly, the two groups of researchers that dominated the impression formation literature for the next two and a half decades --
the "Structuralists" and the "Integrationists" -- generally ignored ques-
tions concerning how people organize in memory and later recall information
about single individuals.

Structuralism, Implicit Personality and Memory

In a concluding comment to his research, Asch stated:

"It should now be clear that the subjects express certain
definite assumptions concerning the structure of personality. In
effect our subjects are in glaring disagreement with the elementaristic
thesis which assumes independent traits (or traits connected only
in a statistical sense) of constant content" (p. 285)

The fact that people in forming an impression rely on networks of
associations among person attributes was subsequently labeled "naive,
imPLICIT PERSONALITY THEORY" by Bruner and Tagiuri (1954) and became a
major focus of a group of impression formation researchers that might be
referred to as structuralists. Schneider (1973) has provided an excellent
review of this research.

In essence, implicit personality theory research has consisted of
obtaining and interpreting matrices of people's beliefs concerning the
association of person characteristics. Several different procedures have
been used to generate these matrices, including (a) assessing the degree
of trait co-occurrence when subjects either freely describe or assign
traits to designated stimulus persons (Rosenberg & Jones, 1972; Rosenberg,
Nelson & Vivekananthan, 1968), (b) having subjects rate the similarity
between traits or the likelihood that a person with trait X will possess
trait Y (e.g., Hays, 1958; Lay & Jackson, 1969; Messick & Kogan, 1966),
(c) asking subjects the expected covariation in one trait if some other
trait in a description were to change (e.g., Hinkle, 1965) and (d) having
subjects rate the similarity of individuals who are known to vary in a
specific set of characteristics or features (e.g., Carroll, 1972; Milord,
1978). Once a matrix of trait associations is obtained some procedure for summarizing and interpreting the matrix is required. Here again, several alternative techniques have been used including factor analysis (e.g., Norman, 1963; Peabody, 1967), multidimensional scaling (e.g., Rosenberg, 1977; Rosenberg & Sedlak, 1972) and cluster analysis (e.g., Rosenberg, 1977; Rosenberg & Jones, 1972).

It would be beyond the scope or purpose of the present discussion to consider the different advantages of these various methodologies. Of greater relevance for placing the research that is to be reported in historical perspective is the contribution implicit personality research has made (or not made) towards understanding the role of memory in the person impression process. Generally, implicit personality theorists have been most interested in identifying regularities in the structure of people's trait association networks. This has included (a) identifying cross-situational dimensions people use in comparing individuals (cf., D'Andrade, 1965; Mulaik, 1964; Passini & Norman, 1966), (b) searching for systematic ways in which people's trait networks differ (Messick & Kogan, 1966; Rosenberg & Jones, 1972; Walters & Jackson, 1966) and (c) investigating what determines trait centrality in such networks (Hays, 1958; Rosenberg et. al., 1968; Wishner, 1960). On the other hand, they have not investigated the way in which people's implicit personality theories influence the way in which new information is perceived and later remembered. Nor have they investigated how people use their networks of associated traits singularly, or in combination with remembered factual information, as the basis of memory-based judgments. In essence, structuralists and implicit personality theorists have tended to develop static maps of people's cognitive networks of person characteristics as
opposed to modeling the dynamic, fluid process by which such networks are utilized in forming an impression of, and making judgments about, a single individual.

**Integrationalism, Evaluative Judgments, and Memory**

In his early experiments, Asch (1946) had subjects make judgments about described stimulus persons on a variety of response traits (e.g., generous-ungenerorous, shrewd-wise, sociable-unsociable). Asch's purpose in doing this was to demonstrate what he believed to be the inadequacy of a simple trait summation model for describing an impression. It is somewhat ironic that a group of later researchers adopted Asch's task to demonstrate that an algebraic summation model could be used to describe the way in which people combine multiple cues when making a single evaluative judgment (cf., Anderson, 1965, 1968, 1974; Lampel & Anderson, 1968).

While both adding (cf., Fishbein & Hunter, 1964) and averaging (cf., Anderson, 1965) models have been proposed to describe judgment processes, an averaging model has generally been better supported. In such a model each stimulus cue is characterized by a scale value $s$ representing its projection on a dimension to be judged and a weight $w$ pertaining to its relevance or importance within a particular judgment context. A subject's expected response or judgment along an evaluative dimension can then be represented by:

$$R = \frac{w_0s_0 + \sum w_s}{w_0 + \sum w_s}$$

where $s_0$ and $w_0$ represent the subject's initial opinion and weight given to this opinion and $\sum w_s$ represents the sum of the products of the weights and scale values of the several stimulus cues.

This type of weighted averaging model assumes that the stimulus value of any cue is constant for a particular dimension, while the weight
given to the cue can vary. Consequently, the majority of research con­
ducted within this paradigm has tended to isolate factors which influence
the relative weighting of stimulus cues such as (a) cue valence (e.g.,
Hamilton & Zanna, 1972; Hodges, 1974; Kanouse & Hansen, 1972), (b) scale
value extremity (e.g., Warr, 1974; Warr & Jackson, 1975), (c) order of
presentation (e.g., Asch, 1946; Jones & Goethals, 1972; Luchins, 1957,
1958), (d) source credibility (Rosenbaum & Levin, 1968), and (e) cue
consistency (e.g., Bugental, Kaswan & Love, 1970).

Generally, the integrationists have employed stimulus-based judgment
tasks in which subjects are asked to make their evaluative judgments
either while the stimulus cues are all available or being presented. Con­
sequently, they have tended to disregard questions about the role of memory
in the judgment process (cf., Anderson & Hubbert, 1913, for an exception).
This has included questions concerning the network of trait associations
subjects bring with them to a judgment setting (i.e., their implicit per­
sonalities) as well as questions concerning which cues are likely to be
fixed in memory so as to be available as the basis for subsequent judgments.
Similary, the integrationists have not investigated factors that influence
which subset of cognitions, out of a person's total cognitions, are likely
to be remembered in a particular judgment context.

Information Processing, Impression Organization and Memory

Information processing as an approach to understanding and studying
human cognition is not easily defined. Certainly, one of its central
concepts is that of information transformation. As part of the infor­
mation processing approach, the human brain is viewed not as a recorder
of external events, but rather as an active transformer of input stimuli.
Such transformations can be of different varietics. They may include
transforming input stimuli into an alternative form that can be more easily stored in memory; or, they may include reorganizing input stimuli into new temporal or categorical groupings. Transformation may also include selectively recalling stimulus events when a behavioral response or evaluative judgment is required.

Viewing the brain as an active processor of information is not a new concept to either psychology or the study of person perception. Its roots can be traced at least as far back as the early 1900's and the work of the Gestalt theorists. Asch's early work could be viewed an an information processing approach towards person impression to the degree that it focused on the transformation of discrete information cues into a unitary impression. However, there are at least two features which distinguish current information processing approaches to cognition from Asch's work. The first is the breaking down of information transformation into its several component parts of encoding, storage and retrieval. The second is the development of new methodologies for identifying and measuring stimuli transformations at each of these separate stages.

Studies that have adopted an information processing approach to studying person impressions, unlike the structural and integrational approaches, have focused on the role of memory in impression formation. However, for the most part such research has investigated how information about a single individual is encoded and organized in memory, rather than how it is accessed during a judgment. Nevertheless, understanding how information is represented and organized in memory is a first step towards understanding how it is later remembered. In the remaining pages of this chapter some general findings concerning the encoding and storage of person information are discussed followed by a review of the few
studies which have examined how person information is accessed in memory during a decision.

Encoding and storage of person information. It has been demonstrated rather conclusively that when people encode stimulus information they tend to store in memory not only some type of record of what was perceived, but also the inferences, abstractions and associations they generate while encoding information (cf., Dosher & Russo, 1976; Greenwald, 1968; Posner & Keele, 1968, 1970). For purposes of the present discussion the first will be referred to as memory for episodic events while the second will be referenced as memory for inferences or implicational associates. An experiment by Dosher and Russo (1976) provides an example of how the latter type of cognitions are stored in memory. Dosher and Russo presented subjects with pairs of digit strings and asked them to first add the strings and then decide if the two were identical. Subjects were better able to recognize two strings as different when changes in digit order were relatively few, but changes in the internal partial sums that needed to be generated in solving the problems were many, as compared to pairs in which there were many digit order changes but relatively few alterations in the partial sums. Thus, memory for internally generated partial sums proved more important for deciding if two digit strings were different than memory for the episodic incidents (i.e., the actual digits displayed as stimuli).

In a recent study Anderson (1977) has demonstrated the occurrence of a similar phenomena when people form person impressions. Anderson first had his subjects learn facts about several stimulus persons identified both by a name and profession. The subjects learned these facts in predicate sentences that began either with the proper name or the
profession (e.g., "James Bartlett adopted the child"; "The lawyer caused the accident"). Later the subjects were asked to verify the truthfulness of sentences that represented either (a) false sentences, (b) true sentences that had been seen, or (c) true sentences that had been unseen but were correct by inference (i.e., "James Bartlett caused the accident" or "The lawyer adopted the child"). In addition, the subjects were asked to recall the facts they originally learned together with the particular name or profession that had accompanied the fact.

In analyzing his data, Anderson first found a high degree of confusion in the recall protocols. Subjects were unable to distinguish clearly during recall between sentences they had seen and sentences they had inferred to be true. In some conditions confusion errors (i.e., statements that subjects had inferred, but remembered as having seen) represented as high as 29% of subjects' recall protocols. Consistent with this finding, subjects' predicate verification times indicated that after repeated exposure to the various statements, all of the presented facts about a particular stimulus person became associated with either the name or the profession, rather than each statement continuing to be associated with the noun with which it had appeared. Thus, the recall data, as well as the response time data, indicated that the inferences subjects generated when encoding the stimulus cues became an integral part of their impressions of the stimulus persons.

The fact that people store in memory both episodic incidents and implicational associates when they encode person information has important implications for information retrieval during judgments. As is discussed in greater detail below, it raises the question as to which of the two types of cognitions tend to be accessed more during a memory-based
A second finding relevant to retrieval of information during a judgment is that the episodic events and implicational associates that make up people's impressions tend to be organized in memory. This fact is reflected both in the implicational associates people use to describe people as well as their memory for episodic incidents. Two general classes of variables that have been identified as affecting impression organization include (a) characteristics of the stimulus information used to describe a person and (b) situational or task demands that exist when stimulus information is encoded.

Concerning the first of these two, Asch (1946) initially suggested that "central traits" (such as warm and cold) in a stimulus set can serve to organize a person's impression. While this assertion was subsequently challenged (cf., Wishner, 1960), other more recent research has demonstrated the ability of a deviant, or otherwise unique, identifying characteristic to organize an impression (cf., Jasper, Rommetveit, Cook, Havelka, Henry, Herkner, Pechevix & Peeters, 1971; Snyder & Uranowitz, 1977).

A recent study by Cantor and Mischel (1977) demonstrates how conceptual congruity among a set of stimuli can also organize an impression, and in so doing affect memory. These investigators presented their subjects with stimulus person descriptions composed of traits that were either closely associated with the concept of extroversion or introversion. Subsequently, subjects were given a recognition test for the traits. The results indicated a bias among the subjects towards including unpresented items that were congruent, as opposed to incongruent, with the theme of the person description. This was true even when the stimulus person had not been explicitly labeled as an extrovert (or introvert). Thus, it would
appear that when a set of stimulus traits converge on a single concept, it can organize a person's impression in a manner that influences how episodic incidents are subsequently remembered.

An early demonstration of the way in which context and task demands influence impression organization was conducted by Zajonc (1960). Zajonc first developed indices of different structural properties of subjects' written descriptions of a person, including differentiation, complexity and organization. He then demonstrated that these properties of impression organization varied as a function of whether subjects thought they would have to later describe the person or would receive new information about the person.

Recently Lingle and Ostrom (in press a) report a series of studies indicating that another contextual activity which influences impression organization is an initial judgment made about a person when the person is first introduced. Lingle and Ostrom argue that such a judgment serves as an organizing "theme" around which episodic incidents and implicational associates can be ordered. To support this hypothesis, subjects were asked to make judgments concerning the suitability of described stimulus persons for different occupations. It was found that (a) subjects could better remember descriptive traits that were relevant, as opposed to irrelevant, to a judgment that they had made, (b) the implicational associates subjects used to describe the stimulus person tended to also be judgment relevant as opposed to judgment irrelevant and (c) memory for the judgment (i.e., the organizing theme of the impression) was less affected by the passage of time than memory for the traits upon which the judgment was based.

The studies that have been discussed to this point did not examine directly how people draw on memory in a decision task; rather, they
demonstrated that impression organization is an important determinant of the general availability of cognitions in memory. However, there is an important distinction between what is available and what is actually accessed during a judgment. This is because people are flexible information processors and can search their memory in many different ways. For example, a person can try to remember either negative or positive aspects of a person or remember a person in any of several different contexts. Furthermore, in a judgment situation, the judgment and its associated criteria may provide a salient set of cues that limit memory search to cognitions that form only a subset of the thoughts that appear available during a memory test.

While general availability cannot be equated with retrieval during a judgment, the studies that have been reported suggest several hypotheses concerning the way in which person information is likely to be accessed during a judgment. Since people appear to store in memory inferences, in addition to episodic incidents, the first of these is:

**Hypothesis 1:** When people make memory-based person judgments they will access in memory implicational associates they have generated about a person in addition to episodic incidents.

Secondly, to the degree that an impression is thematically organized, another hypothesis would be:

**Hypothesis 2:** In a judgment people will access in memory implicational associates and episodic incidents relevant, as opposed to irrelevant, to the central theme around which an impression is organized.
Finally, based on the findings reported by Lingle and Ostrom (in press a) concerning the relative permanence of an organizing judgment (in comparison to specific episodic incidents) a third hypothesis would be:

Hypothesis 3: In a memory-based judgment a person will access the central theme around which an impression is organized before recalling specific episodic incidents.

In the following section the relatively few studies that have investigated how people draw on memory during a person judgment are reviewed and considered in relation to each hypothesis.

Retrieval of person information during a decision. Two methods have been employed to study how people draw on memory when making person judgments. The first has been to examine how such judgments correlate with (a) people's memory for relevant stimulus information and (b) implicational associates they generate in response to stimulus information. The second method has been to use decision time to infer how information is reviewed during a judgment. Studies using both of these methods have provided evidence consistent with the notion that people access in memory and rely on implicational associates, in addition to episodic incidents, when they make memory-based person judgments.

An early study using the correlational technique was conducted by Anderson and Hubert (1963). These investigators read stimulus person descriptions to their subjects that included both positive and negative traits and then asked for evaluative impression ratings on an 8-point scale. In some instances they also asked a subject to recall as many traits as she or he could that had been used to describe a person. Based on a comparison between subjects' recall protocols for the stimulus traits
and their evaluative judgments, Anderson and Hubert concluded that they could not have been basing their evaluative judgment on a review in memory of all (or some random subset) of the verbal stimulus traits available in memory. This conclusion was based on several types of evidence. First, subjects' evaluative judgments showed a consistent primacy effect, while their memory for the stimulus traits exhibited a marked recency effect. Second, changes in subjects' ability to recall the stimulus traits which resulted from varying the number of traits in a description, were not mirrored in subjects' evaluative judgments. Thus, while Anderson and Hubert's results provide no evidence as to what subjects did remember when they made their evaluative ratings, they indicate subjects were not simply reviewing all of the traits they could remember.

A subsequent study by Calder, Insko and Yandell (1974) attempted to show that subjects rely more on memory for self-generated inferences than stimulus arguments when judging the innocence or guilt of a trial defendant. These authors first read their subjects varying numbers of defense and prosecution arguments concerning the guilt and innocence of a defendant. Immediately thereafter, one week later, or two weeks later they asked the subjects to give their impression of the defendant's guilt or innocence. In addition, based upon a methodology developed by Greenwald (1968), they had their subjects list all thoughts that occurred to them pertinent to their judgment and then classify these thoughts according to whether they favored the defense or prosecution. Calder et al. found that an index consisting of subjects' generated pro-prosecution thoughts minus their generated pro-defense thoughts closely tracked differences in believed innocence between the three groups asked to make their judgments at different times.
Several studies have attempted to show people's reliance on remembered inferences to make person judgments by demonstrating that altering subjects' categorization of stimulus information during encoding alters their subsequent judgments about a person (cf., Baumgardner, Leippe & Ostrom, 1976; Carlston, 1976; Higgins, Rholes & Jones, 1977). In the study by Higgins, et al., for example, subjects were given a person description that consisted of behaviors that could either be positively or negatively interpreted (e.g., "He felt he didn't really need to rely on anyone" could be interpreted as either "independence" or "aloofness."). Prior to describing the stimulus person, in a completely separate context, the investigators exposed the subjects to either negative or positive categories that could be used to classify the various behaviors. A manipulation check indicated that subjects tended to use the primed positive or negative trait categories to describe the stimulus person. Furthermore, when subjects were asked a week later to give their evaluative impression of the stimulus person, subjects who had been primed with negative traits evaluated the person more negatively than subjects primed with positive traits. Two types of evidence suggest that this difference resulted from subjects recalling the trait categories they had used to encode the descriptive behaviors. First, two control groups that had been exposed to positive or negative traits that were inappropriate for categorizing the behaviors did not differ in their subsequent evaluation of the stimulus person. This indicated the effect was not the consequence of demand characteristics or halo effects that resulted merely from exposing subjects to different valenced traits. Secondly, the priming manipulation did not measurably affect subjects' memory for the stimulus sentences in a subsequent recall test. This might have been expected if the difference
in impression judgments between the groups resulted from subjects remembering different stimulus behaviors as a result of the priming manipulation.

Studies that have used decision time to study memory-based judgments are consistent with the correlational studies in indicating people's reliance on memory for implicational associates. Posner and Snyder (1975), for example, conducted a study in which they first presented subjects with lists of person traits of varying length. Subsequently, they displayed a test trait and measured how long subjects took to respond whether or not the test trait was a member of the set of traits they had seen. Posner and Snyder found that in identifying a test trait as not a member of the stimulus set subjects took increasingly longer to respond as the number of traits in the sets increased. This finding suggested that subjects serially reviewed the stimulus items in memory so that it took them longer to review larger, as opposed to smaller, stimulus sets. Such a finding is intuitively reasonable and consistent with other research (cf., Sternberg, 1969) which has found a similar search process in other memory tasks. Of greater relevance to the present discussion, Posner and Snyder also found that subjects could identify an incorrect test trait significantly faster when the trait had a different valence (i.e., positive or negative in evaluative tone) than the traits in the stimulus sets. This indicated that subjects abstracted the evaluative tone of the traits and accessed this inference to help them decide if a subsequent test trait was a member of the set.

Lingle and Ostrom (in press b) used decision time in yet another way to provide evidence of people's dependence on inferences and implicational associates to make memory-based judgments. These investigators
had subjects make pairs of occupational suitability judgments about stimulus persons described by different numbers of stimulus traits. Subjects made their first judgment while examining the stimulus traits, however, they made their second judgment without the traits, thus being forced to rely on their memory-based impression of the person.

Two types of evidence suggested that subjects based their second occupational judgments on memory for their first judgment, as opposed to memory for the stimulus traits. First, the amount of time subjects took to make their second judgment did not increase as a function of the number of traits used to describe the stimulus person. This suggested that subjects were not simply serially reviewing in memory the descriptive traits in a manner analogous to the subjects in Posner and Snyder's study. While subjects' decision speed did not vary with set size, it was affected by the nature of the first occupational judgment. Subjects' mean second decision time was faster when their first occupational judgment was similar (as opposed to dissimilar) to the second. This would be expected if subjects were accessing and relying on memory for their first judgment as the basis for their second judgment.

The reported studies that have investigated judgment retrieval processes strongly support the first of the three hypotheses listed at the conclusion of the previous section: people do appear to rely on memory for generated inferences and implicational associates when making memory-based judgments. Although there are interpretative difficulties associated with each of the summarized studies, convergent validation of the hypothesis is provided by the similarity in conclusions drawn from distinct methodologies.
Less support exists for Hypotheses 2 and 3. Lingle and Ostrom's (in press b) decision time studies provide some support for Hypothesis 2 concerning the theme relevancy of accessed cognitions. The fact that subjects could make a second occupational judgment more quickly when it was similar, as opposed to dissimilar, to an initial judgment suggests that they relied on memory for the first judgment and its associated thoughts. However, it is difficult to draw definitive conclusions based on this single finding. As the authors themselves note, other interpretations of the results are possible. For example, it may have been that decision time differences resulted from subjects taking less time to recall the relevant judgment criteria when a second judgment was similar to a first, not less time in considering their first occupational judgment as a basis for their second judgment.

None of the reported studies reflect on the validity of the third hypothesis concerning the order in which people access person information during a decision. In fact, it is unclear how currently used methods might test such a hypothesis. Both the correlational and decision time approaches are characterized by several features that limit their usefulness in studying this and other questions. Chapter 2 discusses some of these limitations and proposes an alternative method for investigating how people access memory when making person judgments.
CHAPTER II
MEASURING SILENT THOUGHT

Limits to Current Practices
In the previous chapter two distinct methods for identifying how people draw on their memory to make memory-based decisions were identified. One method focuses on the correlations between an evaluative judgment, generated implicational associates and memory for stimulus information. The other examines the effect of different variables on decision time in order to make inferences about how information is accessed in memory. Each of these approaches is characterized by limitations.

Limits to the Correlational Approach
Even in the best of circumstances correlational data are difficult to interpret since the direction of causality is rarely clear. In the case of using such an approach to investigate what people remember when they make a judgment, the usual problems are compounded. This is because a covariation tends to exist between the set of inferences a subject generates when encoding a set of stimuli and subsequent memory for subsets of the stimuli (cf., Lingle & Ostrom, in press a; Carlston, 1976). Thus, a correlation between the valence of generated implicational associates and later evaluative judgments may result from subjects using memory for the former as the basis of the latter. On the other hand, the correlation may reflect the fact that subjects are accessing a small subset of stimulus cues to make a judgment whose valence correlates with previously generated implicational associates.
There are additional problems associated with the correlational approach concerned with measuring generated implicational associates and memory for stimulus information. Asking subjects to list the thoughts they have had can be reactive if done prior to a judgment. By forcing a subject to generate and publicly announce a set of implicational associates, an experimenter may cause him or her to rely on such inferences when making subsequent judgments to a much greater degree than would normally be the case. On the other hand, if an experimenter elicits a subject's thoughts following a judgment, the generated thoughts may be a consequence, not a cause, of the decision.

Measuring memory for stimulus cues can present similar problems. If a recall test is administered (or anticipated) prior to a judgment it may increase subjects' reliance on the stimulus cues as a basis for that judgment. If memory is accessed following a judgment, it may reflect memory for stimulus cues after, but not prior to, the judgment. In general, the stimulus cues a person is capable of recalling in the context of a memory test may not reflect the small subset of stimulus cues that are recalled when a judgment is made.

In summary, the correlational approach has proven provocative in raising questions about the relationship between memory for episodic events and judgments. However, it has fallen short of a precise tool for measuring specific thoughts, or even categories of thoughts, a person remembers and considers when making a judgment.

Limits to the Decision Time Approach

Decision time represents a relatively sensitive and unobtrusive measure of information processing. As noted, when used in conjunction
with other approaches it can provide convergent validation of hypothe-
sized memory processes. Nevertheless, there are problems associated
with decision time as a dependent measure. The first is that its inter-
pretation generally requires alternative judgment models that make dif-
ferent predictions about response time. Such models invariably depend
on assumptions whose number increases with the complexity of the decision
being made. For example, in the studies by Lingle and Ostrom (in press b)
the conclusion that subjects relied on memory for their first judgments,
not the stimulus traits, as the basis for their second judgments rests
on several assumptions. One of these is that the time subjects spent
making a decision after reviewing a relevant set of remembered cognitions
was the same regardless of the number of cognitions considered. If this
assumption were not made, an alternative explanation of the results would
be that subjects felt more certain (and thus were able to make their
judgments more rapidly) when they were able to recall many, as opposed
to few, stimulus cues.

Other problems associated with response time as a dependent measure
include (a) it is insensitive to parallel processing of information and
(b) different processes may take approximately equal amounts of time.
This second limitation is especially problematical in light of retrieval
models which posit parallel memory search mechanisms (cf., Quillian, 1969;
Ratcliff, 1978). For example, refering once again to the study by Lingle
and Ostrom (in press b), it may have been that subjects reviewed in memory
both implicational associates and the stimulus cues before selecting a
limited set of relevant cognitions upon which to base their decision. If
this review were done very rapidly, or in parallel, it would have been
difficult to detect increases in decision time as a function of the
number of cognitions being reviewed.

Because of the large number of assumptions required to interpret decision time, it is a much more useful measure for invalidating memory retrieval models than for providing strong support for any single model. Thus, like the correlational approach, its major contribution has been to raise questions about old assumptions concerning memory and the judgment process rather than providing a precise measure of thoughts accessed during a decision.

The several problems associated with the correlational and decision time approaches to investigate judgment retrieval processes can perhaps best be summarized by listing the qualities a measure of thought might ideally possess. First, such a measure should be capable of indicating that a particular idea or concept has been activated in memory during a decision. Second, interpreting the measure should hopefully depend on relatively few assumptions about the way in which information is processed. Third, the measure should be able to detect if a thought has been activated even when activation occurs in close proximity with other thoughts. Finally, the measure should be non-reactive and leave a subject's normal judgment processing unaltered to the greatest degree possible.

In the next section a procedure is proposed for measuring the thoughts subjects access in memory during a decision which has the potential of meeting many, if not all, of these criteria.

An Alternative Approach For Measuring Thought

Current theorizing in cognitive psychology suggests that when a concept is activated in memory, it increases the "activation potential" or the ease with which the same concept can be activated in the immediate future (cf., Collins & Loftus, 1975; Morton, 1969). Researchers interested
in how people encode verbal stimuli have developed methods for measuring such increases in activation potential. One of these, has been to measure how quickly subjects can respond to a probe word which matches a recently activated concept. Warren (1970) for example, first auditorily presented subjects with three words. After a short delay he presented a probe word on a screen which subjects were asked to read as quickly as possible. Warren found that subjects were able to read the probe significantly faster when it was one of the words that had been auditorily presented than when it was an unpresented, matched control word. Thus, activation of the probe word in memory during its initial auditory encoding made it easier for subjects to later read the word.

Warren (1970, 1972, 1974), as well as other investigators (cf., Meyer & Schvaneveldt, 1976) have used a similar procedure to demonstrate that when a word is encoded, other closely associated words are also activated in memory. For example, in one of Warren's studies (1970) subjects were auditorily presented with word triplets pertaining to a single category (e.g., pine, oak and maple). Warren (1970) found that subjects could subsequently identify the category name (e.g., tree) more quickly as a probe word than a matched control word (although the category name was not identified as fast as one of the words that had actually been presented).

Using a slightly different procedure, Meyer & Schvaneveldt (1976) also showed that the encoding of a word activates closely associated thoughts in memory. In their experiments, subjects saw strings of letters and had to respond whether or not the letters spelled a word. It was found that subjects could identify a letter string as a word more quickly when the immediately preceding string had spelled a closely associated
word (e.g., nurse was identified more quickly when it was preceded by doctor).

Warren (1970, 1972) employed an additional measure to demonstrate that increases in probe recognition speed resulted from the activation of a concept in memory, not from subjects' purposely "guessing" that the probe word would be a word they had previously seen. Warren (1970) reasoned that if the effect were due to thought activation (as opposed to guessing) it should be measurable even when performance of the experimental task could be optimized by not guessing. The measure Warren used was based on Stroop's (1935) discovery that people encounter difficulty naming the color of ink in which a word is written if the word itself spells a different color. The current explanation of this phenomenon (see Dyer, 1973 for a review) is that when two concepts (i.e., two colors) are activated simultaneously in memory they produce a response conflict which interferes with an ability to verbalize either of them. Warren (1970, 1972) used this phenomenon to design a task that would penalize subjects for guessing what a probe word would say.

Warren's procedure again consisted of first auditorily presenting subjects with a word. Immediately thereafter, a probe word appeared printed in colored ink. Subjects were instructed to disregard what the word said and simply name the color of ink in which it was printed. Warren found that color naming latencies increased when the probe words were either list words subjects had seen or superordinate categories to which the list words pertained. Thus, he was able to conceptually replicate the word identification experiment using a task which could only be optimally performed if subjects did not try and anticipate what each probe word would say.
If it were the case that internally thinking about a concept, or verbal stimulus, activated concepts in memory in an analogous manner to perceiving external verbal stimuli, procedures similar to those just described might be used to identify the thoughts a person activated in memory while making a memory-based decision. For example, an experimenter could first describe to a subject a stimulus person with a list of traits and then later ask for a judgment about the person. The experimenter might then be able to determine the degree to which the stimulus cues were activated in memory during the decision by measuring how quickly probe words identical to the stimulus cues could be recognized while the judgment was being made. Depending on the duration of increases in activation potential that accompany thought activation during a judgment, the probe word could be presented either while a subject was making a decision, or immediately thereafter.

Such a procedure for measuring thought would meet most of the previously discussed criteria for an effective measure. First, it would provide a direct measure of whether or not a particular idea had been activated in memory and its interpretation would not depend on a large number of assumptions about how information is processed. Second, up until the moment the probe was presented, the procedure would not be overtly reactive. That is, subjects would not be required to take a recall test or list their thoughts about a person prior to making a decision. Finally, the procedure should make it possible to determine if a thought has been activated in memory even when it is activated in close proximity with other thoughts.

Whether or not it would be possible to use this type of probe recognition task to measure thought during a decision would depend on the
answers to at least two questions. The first is whether internal activation of a concept is similar to external activation in that it increases activation potential to a measurable degree. The second is how long increases in activation potential of a single concept persist when other concepts are being activated in close proximity with it.

Can Internally Activated Thought Be Measured?

Studies that have employed the probe technique to measure thought have generally measured concepts that are externally activated by auditorily or visually presented stimuli. While internal and external thought activation would seem to be related phenomena, there may be important differences that make the former more difficult to measure than the latter. For example, when a concept is externally activated, it is accompanied by activation of the visual or auditory sensory system. Consequently, iconic or echoic memory for the stimulus event will be present. One possible outcome is that the effective time that the concept is activated in memory may be increased as compared to the period of activation that generally accompanies internally activated thought. At this point it is impossible to know whether any, or all, of the amplifying effects that may result from the activation of a sensory system (and its associated memory system) are necessary to increase the activation potential of an idea to a measurable degree.

To date there has been at least one study indicating that internal thought activation is measurable. This study, conducted by Geller and Shaver (1976), was designed to investigate whether a recent theory is correct in assuming that a stimulus that makes an individual’s self-image salient (such as a mirror or camera) increases the activation of self-relevant thoughts. Geller and Shaver reasoned that if such stimuli do
increase self-focused thinking they should increase interference in the Stroop color naming task when the base words are self-relevant, but not when they are non-self-relevant.

To test this hypothesis subjects were first asked to name the colors in which each of 63 rows of x's were printed. The rows were matched in length with the words that would subsequently be used to test the experimental hypothesis. Next, subjects named the colors of 63 words written in columns on a single sheet of paper. For one half of the subjects these words were evaluative person adjectives (e.g., ugly, successful). For the other half of the subjects the words were neutral words (e.g., survey, kitchen). Cross-cutting the type-of-word variable, one half of the subjects performed the task with a mirror and a TV camera facing towards them while the other half did so with the same equipment facing away from them. Analyses of the difference in color naming speed for the x's and words produced the predicted interaction. When the neutral words were used, there was no difference in subjects' color naming ability as a function of whether the camera and mirror faced towards or away from them. However, when evaluative person adjectives were used, subjects' performance in the color naming task was significantly impaired when the mirror and camera faced them, as opposed to facing away from them.

The study be Geller and Shaver thus provides some evidence that internal thought activation results in measurable increments in activation potential. A manipulation designed to increase internal activation of a category of thoughts produced increases in color naming latency. While this finding is encouraging, the question remains as to whether a single internal thought, activated within a string of thoughts, increases activation potential to a measurable degree.
How Long Do Increases in Activation Potential Persist?

A second question that must be considered in assessing the likely suitability of the probe technique for measuring thought during a judgment concerns the rate of dissipation of increases in activation potential. During a judgment a person may activate in memory several thoughts in rapid succession. The more rapidly increases in activation potential dissipate with the passage of time or the activation of other thoughts, the more difficult they will be to measure. If a probe must appear at the exact moment a thought is internally activated in order to measure its occurrence, the applicability of the procedure would be limited. In such a case, many type II errors could be expected since failure to detect increases in activation potential would often occur because the probe was not presented at the exact moment the corresponding thought was activated in memory.

Investigations of the rate at which externally induced concept activation dissipates with the passage of time and the activation of additional thoughts have produced somewhat contradictory results. There are at least three studies that provide evidence of relatively long lasting measurable increases in activation potential (i.e., up to 50 secs.) including Warren (1970, 1972), Loss and Waugh (1967) and Anisfeld and Knapp (1968). However, each of these studies employed a paradigm in which subjects had to recall the words presented to them so that they likely engaged in some type of rehearsal activity. For example, Warren (1972) used a Stroop procedure in which words printed in color ink were presented to subjects at the rate of one word every ten seconds. When each word appeared the subjects were to name the color of ink as quickly as possible and then read the word out loud. Four blocks of 20 trials were given to each subject. Each block of 20 trials consisted of two successive groups of ten trials which were
constructed as miniature control-lag lists, with one word appearing twice at each of five lag positions (once as a new item and once as an old item). The lags used were one, two, three, four and five trials representing delays of 10 through 50 seconds (and 0 through 3 intervening thoughts). Subjects were asked at the end of each block of 20 trials to recall as many of the words they had seen as possible.

Warren found that subjects exhibited significant delay in naming the colors of the words when they were presented a second time in all 5 lag positions. Furthermore, there was no differences in the amount of interference produced. Thus, in a setting in which subjects thought they would have to later remember the words (and presumably spent some time rehearsing them) there was no detectable decrement in memory activation with the passage of time.

Several studies in which subjects were not required to remember presented stimulus words have found relatively rapid decreases in measurable concept activation. In one study, Warren (1972) auditorily presented subjects with a stimulus word and then presented a probe word for which they were to name the color of ink in which it was printed. To make sure they listened to the words, after naming the color subjects were required to recall the word they had heard. Words presented as colored probes were either control words that had never been presented or test words that subjects had heard as stimuli. Probe words that matched an auditorily presented word were shown at lags of zero, one, or two trials (a lag of zero meaning the probe word matched the word auditorily presented in that trial). Thus, a probe that matched an externally activated thought was shown 0, 15, and 30 seconds (0, 1 and 2 thoughts) after the thought had initially been activated in memory. An analysis of the results indicated that only
the color naming latencies for lag position 0 were significantly greater than the corresponding control condition. Thus, any measurable effect of thought activation dissipated prior to 15 seconds after subjects had heard the stimulus word.

Meyer, Schvaneveldt and Ruddy (1972) also investigated dissipation of thought activation, although they used word identification as a dependent measure. As noted, in their procedure subjects were asked to indicate (by pressing a button) whether or not a string of letters represented a word. Meyer et al. varied the intertrial interval between presentation of letter strings, incorporating delay times of 0, 1500 and 4000 msec. Each letter string that spelled a word was either a close associate of the preceding word that had been displayed (e.g., nurse-doctor) or unrelated to the preceding word.

Meyer et al. found a significant decrease in the facilitation effect as the inter-trial interval increased. When the two words immediately followed one another the "association effect" approached 100 msec; with a 4 second lag the effect was halved. As the authors note, it is necessary to be careful in interpreting these differences since it is unclear what subjects were thinking about during the delay periods. It is conceivable that they were rehearsing the last word they had seen and that this rehearsal was responsible for the small, but measurable effect found in the four second lag conditions.

In another study, Schvaneveldt and Meyer (1973) investigated the effects of interjecting a single stimulus item between two associated words. Again subjects were shown strings of letters and asked to respond whether each of the strings spelled a word. The independent variable was the relationship between triads of sequentially presented words; the
dependent variable was subjects' latencies in identifying the third word in a triad. There was a 250 msec delay between a subject's response to one word and the presentation of the next word. Thus, there was a lag of between 1500 and 2500 msec between the presentation of the first and third words in a triad.

Schvaneveldt and Meyer (1973) found that subjects' response to a third word in a triad was faster when it was immediately preceded by a close associate (e.g., horse, doctor, nurse) than when the close associate was presented as the first word in the triad (e.g., doctor, horse, nurse). However, even in this latter condition the third word was identified significantly faster than when it was preceded by two unassociated words (e.g., horse, school, nurse). Thus, measurable residual activation was present 1.5 to 2.5 seconds after a concept had been activated in memory even when an unassociated thought was activated immediately afterwards.

To summarize, the several studies that have been reported indicate that concept activation resulting from the encoding of an external verbal stimulus persists to a measurable degree for some limited period of time. Activation of a thought has been detected up to 50 seconds following a presented stimulus when subjects believed they would later be tested on memory for the word. When subjects were not told they would have to remember the stimulus word, concept activation has been detected after a 4 second delay, although the strength of the effect was half as great as immediately following activation. Finally, external activation of a thought has been measured after a delay of 1.5 to 2.5 seconds when an unassociated word was interjected between its activation and the measure of activation. The degree to which these conclusions can be generalized to internally generated thought remains to be investigated.
This chapter began with a critique of the correlational and response time methods used to investigate how people draw on their memory when they make person judgments. Several shortcomings of each method were identified and an alternative procedure for measuring thoughts activated during a decision was proposed based upon the speed with which subjects can respond to a presented verbal probe. Such a procedure has been used by experimenters to identify concepts activated in memory when an external stimulus is encoded. Whether or not a similar procedure could be used to measure internally activated thought during a judgment was seen as depending on (a) whether a single thought, internally activated within a string of thoughts, produces measurable increases in activation potential and (b) the speed with which internal activation of a concept dissipates with the passage of time and the activation of other thoughts. Answers to these questions cannot be gleaned from existing research. However, the evidence that does exist is encouraging. The next two chapters report three experiments that were conducted to provide a more definitive answer to these questions in a situation in which concepts are internally activated in memory in rapid succession.
THE MEASURABILITY OF INTERNALLY ACTIVATED THOUGHT: EXPERIMENTS 1 & 2

The Task: Hum a Happy Tune

This chapter reports two experiments that were conducted (a) to ascertain whether a single thought, internally activated in memory within a rapid succession of thoughts, is detectable and (b) to determine which of two different probe recognition tasks is the more sensitive measure of internally activated thought.

To investigate the measurability of internal thought requires a task with unique features. First, the task must result in subjects internally activating concepts in memory in rapid succession, as might be expected if they were contemplating a decision. Second, the task needs to allow a researcher to know, at any point, what thought a subject is thinking in order to be able to measure it. Finally, to assure that activation of a thought is internal, the task should make certain that a thought is not externally activated in some way. Such assurance would not only include not presenting either a visual or auditory cue, but would also include making certain that a subject did not vocalize aloud the thought to be measured.

While there may seem to be few tasks that fill these requirements, there is, in fact, a very common day-to-day activity that satisfies each of them. The activity is humming. Although never empirically demonstrated, it seems probable that when people hum the tune to a well-known song, they activate the words of the song in memory. To be convinced of this,
one needs only to try and hum the tune to *Jingle Bells* while thinking the words to *Yankee Doodle*. The difficulty encountered in this task suggests the close association between humming the notes to a song and activation of their coinciding words in memory. The task of humming and thinking the words to a song would seem to meet the several requirements just outlined. First, it would be possible to have an idea as to what thought was being activated in memory by listening to the note being hummed; second, while humming a tune a subject would be sequentially activating thoughts in memory in rapid succession; third, if a person were humming musical notes to a song, she or he would not be able to vocalize out loud the words to the tune.

Studies 1 and 2 employed humming to investigate the measurability of internally activated thought. In a within-subjects design, subjects were asked to hum a number of well-known songs. While humming these songs they were periodically interrupted by a probe word to which they were required to respond as quickly as possible. The probe word either matched or did not match the word corresponding to the note they were humming. It was reasoned that if internal activation of a song word in memory produced measurable increases in activation potential, differences in responses to the probe would result depending on whether or not it matched the word being hummed.

The particular response subjects were required to make to the interrupting probe differed in Experiments 1 and 2. In Experiment 1 subjects were required to read the probe word out loud as quickly as possible. In Experiment 2 they were required to name the color of ink in which the word was printed. In accordance with the research summarized in Chapter II, it was expected that word recognition would be facilitated, while color naming
would be inhibited, when the probe word matched (as opposed to did not match) the word subjects were thinking. Studies employing both of these dependent measures were conducted in order to (a) provide convergent validity concerning the measurability of internally activated thought and (b) determine which of the two procedures would be the most sensitive measure. Because the design and procedure of the two studies were identical except for the dependent measure employed, the experiments are described jointly.

Method

Subjects

Subjects consisted of 361 undergraduates from the Ohio State University subject pool who participated in the experiment in partial fulfillment of an introductory psychology course requirement. Twelve subjects (4 males and 8 females) participated in Experiment 1 in which subjects were required to read the probe word. Twenty-four subjects (7 males and 17 females) participated in Experiment 2 in which subjects identified the color of ink in which the probe word was written. Subjects were randomly assigned to one of two stimulus presentation groups used to counterbalance the design.

Stimulus Materials and Design

The first two lines from 17 well-known songs and nursery rhymes were selected to be hummed (see appendix A for a complete list). Of these, five were used in a practice session while 12 were used in the experiment proper.

Each subject hummed the 12 experimental songs twice, being interrupted

1 Six additional subjects were recruited for the experiment but were unable to learn the tunes well enough to satisfactorily complete the experimental task.
each time by the appearance of a probe word. In each replication, while humming half of the tunes, a subject was interrupted by a probe word that matched the word being hummed; for the other 6 trials in the replication the probe word was from one of the other tunes and did not match the word being hummed.

For counterbalancing purposes, in each experiment subjects were randomly divided between two equal sized groups. The design was counterbalanced so that across the two subject groups and two replications each rhyme and each probe word appeared equally often in those trials in which the probe word matched and did not match the word being hummed. This was achieved as follows: Two words from each of the 12 tunes, one from the first and one from the second line were arbitrarily selected to be used as probe words. One probe from each song was then randomly selected to be displayed in the first replication while the remaining 12 words were used in the second replication (the one constraint being that each replication contained six first and six second line words). The order in which these 24 words appeared as probes was random and constant for all subjects. The difference between subject groups was in the order in which they hummed the songs.

Six of the twelve probe words to be presented in Replication 1 were randomly selected to appear as matching probes for subject group 1, with the remaining six to be used as non-matching probes. In the second replication a reversal was undertaken such that if one word from a song appeared as a matching (or non-matching) probe in Replication 1, the other selected probe from that song appeared as a non-matching (or matching) probe in Replication 2.
The songs to be hummed when a probe was to appear as non-matching were selected by randomly pairing the 12 songs. On non-matching trials, subjects hummed the selected "twin" to a song from which the probe word that appeared on that trial had been selected. The probe word was presented at one of the two points in the song that a matching probe would have appeared (the particular position being randomly determined).

For subject group 2 the "matchedness" of each probe was reversed so that probes seen as matches by subject group 1 were seen as non-matches by group 2. Thus, across replications and subject groups, (a) each probe word was seen equally often as a matched and unmatched probe, (b) each song was interrupted an equal number of times by a matched and unmatched probe and (c) matched and unmatched probes appeared in the same positions in the songs being hummed.

In Experiment 1 subjects were required to read aloud the probe word. These were displayed on a 22 cm by 14 cm rear projection screen 1.0 m in front of the subjects. The letters of the probes were approximately 1.0 square cm in size. Because Meyer and Schvaneveldt (1976) found that differences in recognition speed for words that match and do not match a concept activated in memory is enhanced by degrading the probe word, in Experiment 1 the probes were degraded by imposing a field of Xs over them.

In Experiment 2 the words were not masked, but printed in blue, red or green ink, and subjects had to identify the color in which each probe was written. The color of the probes was randomly determined with the constraint that each color appeared equally often as a matched and unmatched probe in each replication.

Experimental Procedure and Instructions

When subjects arrived they were seated in a small room approximately
1.5 square m in size facing a rear view projection screen. This screen was built into the wall and stimulus materials were projected on it by means of a slide projector located in an adjoining control room. A built-in shelf extending the width of the room was between the subject and the screen. Upon this were placed a microphone connected via an intercom to the control room and two sheets of paper on which were printed the first two lines of 17 songs and nursery rhymes (12 experimental and 5 practice tunes). Once the subject was seated, the experimenter began the training phase of the experiment in which a subject was taught any verse she or he did not already know.

It was first explained to each subject that the experiment was concerned with people's ability to perform two mental tasks at once, and that one of the tasks required that the subject know the first two lines of 17 well-known tunes. The subject was asked to study the songs printed on the sheet until she or he felt confident of knowing them. Next, the experimenter asked the subject to sing each song aloud at his or her own speed in the order they were written on the sheets. While the subject was singing, the experimenter listened via the intercom from the control room and noted any songs the subject appeared not to know.

Following this initial practice the experimenter returned to the experiment room and obligingly sang and practiced with the subject any unknown songs. Once the experimenter was satisfied that the subject knew each of the songs, he asked him (her) to sing the songs without looking at the sheet. During this session the back projection screen was used to cue the subject as to which song to sing. This was done by displaying the first two or three words from a song for 3 seconds followed by the word "sing". After finishing each song, subjects were given a 10 second rest
before the first words of the next song were displayed. Subjects sang
the songs in a different randomized order from the order in which the
songs appeared on the song sheet. As the subject sang, the experimenter
listened from the control room and noted any verses sung incorrectly.
Finally, the experimenter asked the subject to sing again any songs not
sung correctly until the subject could sing them without error.

Once a subject had learned all of the songs (a process which took
from 20 to 40 minutes), the experimenter explained the experimental task
which proceeded as follows: At the beginning of each trial the first few
words of a song were projected on the screen. Three seconds later these
words were replaced by the word "Hum". This was the subject's cue to hum,
(not sing) the indicated song. Immediately after the subject began hum­
ing, the screen went blank. However, the subject was instructed to con­
tinue humming until she or he had finished the first two lines of the tune.
At some point while the subject was humming each tune a probe word
appeared on the screen. This appearance was controlled by the experi­
menter who listened and activated a switch just as a subject hummed the
appropriate probe word in a song. This switch (a) displayed the probe
word, (b) started a Hunter Klockounter timer (model 120A) and (c) activated
an assistant experimenter's earphone. The moment the subject responded
to the probe the assistant pressed another switch which blanked the screen
and stopped the timer (because the assistant could not hear what tune was
being hummed he remained blind to the experimental conditions). The sub­
ject's probe recognition time was then recorded by hand on a scoring sheet.

After responding to a probe, subjects were instructed to continue
humming a song at the exact point they had left off. This was strongly
emphasized in order to increase the likelihood that they would concentrate
on the humming task and not simply wait for the probe to appear. It was also felt that continuing humming would increase the likelihood that subjects would think the words to the rhymes in order to monitor their position. One final action was undertaken to make sure subjects would think the words to the tunes as they hummed. This consisted of the experimenter telling subjects that it was important to hum the tunes at the same speed they had sung them and that the best way to accomplish this was to think the words to a song as they hummed it.

Before subjects began the 24 experimental trials they were allowed to practice responding to the probes. This was done in two phases. First, subjects responded to five probes sequentially displayed on the screen. In Experiment 1 this consisted of reading the words aloud; in Experiment 2 subjects named the color in which the words were written. Next, subjects practiced humming five tunes while being interrupted by a probe as would be required in the experimental trials. On two of these trials the probe matched the word being hummed. At the completion of this practice the experimenter answered any questions about the procedure and the experimental trials began. Subjects were given a five minute rest between replications of the 12 tunes. After the second replication subjects were debriefed about the purpose for the experiment.

Results

Experiment 1: Word Naming

To review the hypothesis, to the degree that internal thoughts are measurable, it was expected that in Experiment 1 subjects would be able to read a masked probe word more quickly when it matched (as opposed to did not match) a word they were thinking. To test if this were the case,
each subject's mean probe recognition time for matched and unmatched probes was computed for both replications. These mean scores were then analyzed in a 2 x 2 (matchedness by replication) within-subjects analysis of variance.

Subjects made seven errors in the 288 probe recognition trials (2.4%). Three of these occurred when the probe matched the word being hummed, while four occurred in the non-matching condition. Because of the low percentage and relative equal distribution of errors between conditions, error trials were deleted from the data analysis and subjects' scores averaged over the remaining responses in the condition.

Subjects' mean word recognition times for matched and unmatched probes for both replications are displayed in Table 1 on the following page. As can be seen, in both replications subjects took less time to recognize matched, as compared to unmatched, probe words. The analysis of variance indicated that this difference was significant ($F (1, 11) = 16.6; p < .002$) and that the difference did not vary significantly as a function of the replications factor ($F (1, 11) = 1.32; p < .30$ for the interaction). The main effect for replications was nonsignificant ($F < 1.0$).

**Experiment 2: Color Naming**

As in Experiment 1, probe recognition errors were relatively few. There were 17 errors out of a total of 576 trials (3%), nine of which occurred in the matched condition. Again, trials on which errors occurred were omitted from the data analysis.

Experiment 2 used color naming latencies to measure thought activation. If effective, the pattern of results for this measure should be opposite those of Experiment 1. In the matched condition activation in memory of
Table 1

Mean response latencies in milliseconds to matched and unmatched probes:

Experiment 1 - word naming.

<table>
<thead>
<tr>
<th>Probe Word</th>
<th>Matched</th>
<th>Unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>Song</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1529</td>
<td>1791</td>
</tr>
<tr>
<td>2</td>
<td>1380</td>
<td>1794</td>
</tr>
<tr>
<td>M</td>
<td>1455</td>
<td>1793</td>
</tr>
</tbody>
</table>

Note: Each mean represents 72 observations (12 subjects providing 6 responses per condition).
the probe word should interfere with color naming when compared to unmatched probes. To determine if this were the case, the data were analyzed similarly to Experiment 1. Subjects' mean color naming latencies are shown in Table 2 on the following page. As expected, response latencies were slower in the matched conditions. This difference was significant ($F(1, 23) = 7.19; p < .02$). Again, the matching effect did not interact with the replication factor ($F < 1.0$). There was a marginally significant main effect for replications ($F(1, 23) = 4.05; p < .06$) indicating that subjects took less time to respond to the probes in the second replication, as compared to the first.

Discussion

Measurability of Thought

The results of Experiments 1 and 2 provide evidence that internally activated verbal concepts can be measured using similar techniques to those used to measure externally activated concepts. This conclusion is strengthened by the fact that two distinct measurement instruments, word naming and color naming, were both successful in measuring thought. The effectiveness of color naming suggests the matching effect is not the result of subjects intentionally guessing the probe word will be the word they are about to hum. While a guessing strategy would be helpful for word naming, it would be detrimental to color naming and there seems little reason for subjects to use it in the latter case. This is not to say that inadvertent expectancy on the part of subjects as to what a probe might say did not play a mediating role in both tasks. However, even if this were the case, it would not preclude using the procedures to measure thought activated in memory during a judgment.
Table 2
Mean response latencies in milliseconds to matched and unmatched probes:

Experiment 2 - color naming.

<table>
<thead>
<tr>
<th>Song Replication</th>
<th>Matched</th>
<th>Unmatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1189</td>
<td>1154</td>
</tr>
<tr>
<td>2</td>
<td>1155</td>
<td>1111</td>
</tr>
<tr>
<td>( \bar{N} )</td>
<td>1172</td>
<td>1133</td>
</tr>
</tbody>
</table>

Note: Each mean represents 144 observations (24 subjects providing 6 responses per condition).
Some caution is warranted in interpreting the results of Experiments 1 and 2. There is no assurance that the probe was presented in the matching condition at the exact moment subjects activated the word in memory.

In both experiments matchedness of the probe word to the note being hummed was confounded with whether or not the word occurred somewhere in the song. It is possible, for example, that the matchedness effect resulted from subjects activating all of the words to a tune when they first began to hum it. Evidence relevant to this alternative explanation is presented in Experiment 3 where the confounding is eliminated.

Relative Power of Word vs. Color Naming

One purpose for Experiments 1 and 2 was to investigate the feasibility of measuring internal thought. A second purpose was to determine which procedure was a more sensitive measure of thought. One appropriate means for comparing the two is to compare their power (i.e., the probability that each would reject the null hypothesis assuming the differences in matched and unmatched probe response times represent true population differences). This was done by the method described by Kirk (1968, pp. 107-109) for computing the power of an analysis of variance $F$ test. The differences between the matched and unmatched means from each experiment were used as estimates of the population differences while the square root of the mean square error was employed as an estimate of the standard deviation.

Assuming an $\alpha$ of .05 and an $F$ test with 1 and 11 degrees of freedom (i.e., $N = 12$), the power of the test between the matched and unmatched groups for word naming (Experiment 1) was .94; for color naming (Experiment 2) it was .61. Thus, word naming was a more powerful measure of internal thought activation.
As is often the case, in Experiments 1 and 2 the distributions of subjects' response times were positively skewed thus violating an assumption of analysis of variance. A final analysis was therefore performed on the data. Subjects' response latencies were transformed to speed scores (i.e., 1000/response time in milliseconds or "probe words responded to per second") and reanalyzed. An advantage of this transformation was that it ameliorated the skewness problem while maintaining a dependent measure with conceptual meaning.

The use of speed scores did not alter the pattern of significant results for the two experiments. The effect for matchedness of the probe to the thought hummed was significant in both cases ($F(1, 11) = 22.5, p < .001$ for word naming); $F(1, 23) = 4.96, p < .04$ for color naming). Nor did this effect interact with the replications factor ($p > .20$ for the interaction term in both experiments). Assuming an $\alpha$ of .05, and 1 and 11 degrees of freedom, the computed powers for the speed score analyses were .99 for word naming and .46 for color naming. Thus, the use of speed scores (as compared to response latencies) increased the power of the $F$ test for word naming, but not for color naming.

Wainer (1977) argues that the best way of choosing whether to use speed scores or response latencies as a dependent measure is to have a theoretical reason for selecting one over the other. In the present context, however, no such reason exists. Rather, the best measure is the most powerful measure that violates to the least degree the assumptions upon which parametric statistic tests rest. This being the case, the most appropriate conclusions to be drawn from Experiment 1 and 2 are (a) internally activated thoughts are measurable and (b) of the several procedures examined, the most powerful method for measuring internal thought is the
speed with which subjects can read a degraded probe word. In the following chapter an experiment is described which investigated this measure further by testing its ability to discriminate between two thoughts internally activated in memory six thoughts and 3.6 seconds apart from each other.
CHAPTER IV
MEASURING THOUGHT ACROSS TIME AND THOUGHT: EXPERIMENT 3

The Task: Think a Happy Rhyme

An important consideration in assessing the applicability of the probe technique for measuring thoughts in a judgment task, is its continued sensitivity to thought activation across time and more thought. The technique would be less useful if a probe could measure the occurrence of a thought only when it appeared at the exact moment the thought was activated in memory. Experiment 3 was designed to (a) establish some idea of how quickly activation potential dissipates following the internal activation of thought and (b) determine if the probe technique is sufficiently sensitive to discriminate between two thoughts activated close together in a rapid stream of thought.

To answer these questions, a task with unique characteristics was again needed. First, as in Experiments 1 and 2, it was necessary to have subjects rapidly activate in memory a sequence of concepts and to have the experimenter know at any moment which thought was being activated. In addition, it was desirable to have subjects activate the thought sequence at a constant rate of speed in order to establish how quickly measurable activation decayed over a fixed time interval.

The task devised consisted of again having subjects think the words to well-known rhymes. However, rather than humming the rhymes, subjects were asked to think the words to the rhymes in time with a visual metronome (a flashing black bar). In this way, every subject thought the words to
the rhymes at a fixed rate, and it was possible to know what word a subject was thinking at any moment by counting the bars. This made it possible to display different rhyme words at precise distances from the point at which subjects activated them in memory.

To understand more fully the rationale and hypotheses of Experiment 3, consider the pattern of probe recognition speeds that might occur if the third word, in a sequence of eight words in a rhyme, were displayed as a probe while each word was being activated in memory. If (a) increased activation potential dissipates very quickly or (b) the facilitation effects in Experiment 1 were only the consequence of subjects guessing that a probe word was going to be the word they were thinking, a pattern of responses similar to those displayed in Figure 1, panel a would be expected (see following page). In such a case, increases in recognition speed would occur only when the probe was presented at the exact moment the thought was being activated in memory (i.e., at position 3).

On the other hand, a second possible pattern of results would be those depicted in Figure 1b. This pattern would be expected if increases in thought activation potential dissipated more slowly so that some measurable increases in activation persisted even when the third word was displayed as a probe when later words were being activated in memory.

Displaying a single word during a verse in each of eight different positions, as just described, could provide an idea of the rate of activation dissipation. However, there is a problem. It is possible that variations exist in the cognitive effort required to think the words at different points in a rhyme. For example, subjects might have to concentrate more to remember the words at the end of a rhyme as opposed to the beginning. If this were so, and probe recognition speed varied inversely
Figure 1: Possible patterns of probe recognition speed when an early (position #3) and late (position #6) word in a sequence of eight are displayed in each position.
with the difficulty of the thinking task being interrupted, the pattern depicted in Figure 1b could be approximated even though the probe did not match a thought being activated in memory. That is, if subjects found it easier to think the words early, as opposed to late, in the rhyme they might be able to recognize any type of probe faster at the beginning, rather than the end, of the verse.

One means of dealing with this problem would be to display another word in the sequence at each of the eight positions. For example, if both the 3rd and 6th words were shown, a pattern of probe recognition speeds similar to Figure 1c could be expected if activation potential dissipated gradually. Recognition speed for the early word would increase at the beginning of the sequence and begin to fall off towards the end. Recognition speed for the late word would increase at the end of the sequence as it became activated in memory. Statistically, such a pattern would result in an interaction of the quadratic component between probe word type (early vs. late in the sequence) and position in the sequence that the probe was presented (1 through 8). If probe recognition speed also varied as a function of where a subject was in thinking a rhyme, such variations would presumably affect recognition speeds for both probe words. In this case, the shape of the two curves might change absolutely, but not relative to each other. The statistical quadratic interaction would still be expected to remain.

In summary, presenting two words as probes would provide two advantages. First, it would make available two probe recognition patterns to provide an idea of variations in probe recognition speed that resulted from difficulty in thinking different parts of the rhyme. Second, and more importantly, it would provide a direct test of whether the probe
technique is sensitive enough to discriminate between two thoughts activated close together in a sequence of thoughts.

Experiment 3 used the procedure just described. Sequences of words were selected from the middle of eight well-known nursery rhymes. Two words, one occurring early and one occurring late, were chosen from each sequence to be presented as probes in eight different positions. A within-subjects design was used in which each subject saw both an early and late word appear as a probe in the eight positions while thinking the words to different rhymes in time with a visual metronome. To the degree that the probe technique was capable of differentiating between two rapidly activated thoughts, an interaction was expected to occur between word type (early vs. late) and probe position (1 through 8).

In the absence of any precedent in the area of nursery rhyme thought measurement, some uncertainty existed as to how the experimental task might best be structured to reduce extraneous variability in probe recognition speeds. All told, four replications of Experiment 3 were conducted. Replications 1 and 2 were identical to each other and conducted jointly. Replications 3 and 4 were subsequently undertaken in an attempt to replicate the results of the first two replications and to test the ability of several procedural alterations to reduce variability in subjects' probe recognition speeds.

Method

Subjects

A total of 75 subjects were recruited to participate in Experiment 3. The data from 11 of these, however, were not used either because a subject never learned the nursery rhymes well enough to perform the
experimental task, or because the subject repeatedly made mistakes in activating the voice key. The group of 64 subjects whose data were used consisted of 26 females and 38 males, 16 of whom participated in each of the four experimental replications. For Replications 1 and 2 the subjects were recruited from the Ohio State University subject pool. Participants in Replications 3 and 4 consisted of 25 individuals from the University subject pool and seven personal acquaintances of the author who were unfamiliar with the research. In each replication subjects were randomly assigned to one of the stimulus material presentation orders used to counterbalance the design.

Stimulus Materials and Design

The 10 nursery rhymes (8 experimental and 2 practice) used in the experiment were chosen because pilot testing found they were well-known by members of the subject pool (see Appendix B for a complete list).

Initial testing indicated that thinking the words to a rhyme in time with a bar which appeared every 6/10ths of a second, (remaining on for 3/10ths of a second), was a comfortable speed for subjects to think the words to the rhymes. Such a time period, however, resulted in an excessively short time period for presenting the probe words. If a probe were presented on eight successive appearances of the bar, the first and last presentation would be separated by only 6 thoughts and 3.6 seconds. In order to extend this distance, a string of 15 words was arbitrarily selected from the middle of each of the rhymes. From this string, the 5th and 11th words were selected as probes to be shown in the odd numbered position as subjects thought the rhyme (i.e., positions 1, 3, 5, 7, 9, 11, 13, and 15). In this way the distance between the first and last presentation of a probe in a rhyme was separated by 13 thoughts and 8.1
seconds. More importantly for establishing the sensitivity of the probe procedure over time and thought, word 5 (the early probe word) appeared in position 1, 4 thoughts and 2.7 seconds prior to its being activated in memory by a subject; in position 15 the same word appeared as a probe, 10 thoughts and 6.0 seconds after it had been activated in memory. Thought 11, on the other hand, when presented in position 1 appeared as a probe 10 thoughts and 6 seconds prior to its activation in memory; in position 15 it appeared 4 thoughts and 2.4 seconds after activation. Finally, while thinking the words to a rhyme, a subject would think the selected early word 6 thoughts and 3.6 seconds prior to thinking the late word.

The within-subjects design consisted of having each subject see both an early and late word from a sequence of words in a rhyme in each of the eight positions. This was accomplished by asking subjects to think through the rhymes twice. During each repetition a subject saw either the early or late words from the rhymes in one of the eight different positions. To counterbalance the order in which subjects saw the two probe words, two groups of subjects were used. One group saw the early probe words first, while the other saw the late probe words first.

For each subject, the eight rhymes, the position in which the probe appeared (1 through 8), and the order in which the rhymes were thought were counterbalanced by a Greco latin square. This required eight subjects so that 16 subjects in all (8 times 2 groups) were needed for complete counterbalancing.

It was thought that the sensitivity of the probe procedure to detect variations in thought activation over time and thought might be decreased if subjects came to expect every probe word to be a word in the rhyme they were thinking. To avoid this from occurring, subjects were made to
think the words to the rhymes twice in each replication (four times in all). In this additional repetition of the rhymes (which took place on the odd trials) subjects were interrupted by a probe word that was not in any of the rhymes.

Experimental Procedure and Instructions

**Equipment and orienting instructions.** A Data General Nova 1220 mini-computer with several subject stations was used to control the stimulus displays and record subject responses. The physical configuration of this equipment has been described in a recent publication by Ronis, Baumgardner, Leippe, Cacioppo and Greenwald (1977). Briefly, each subject station consisted of a small room whose dimensions were approximately 4 square meters. On a shelf, extending the width of the room, was a small TV monitor (with a 21 by 26 cm screen) and a response box containing 16 buttons labeled with digits and letters.

Upon arriving, a subject was seated at a station and told that the experimental instructions would be displayed on the monitor. The subject was asked to read each instructional display carefully and press a button on the response box labeled "enter" when ready for a new page of instructions. The experimenter also asked the subject to don a pair of earphones with an attached microphone that allowed the experimenter to hear the subject throughout the session.

For reasons explained in Chapter III, the probe words shown in the experiment were degraded. This was accomplished by using the graphics capability of the computer to construct degraded letters from a six by six square dot matrix approximately 2.5 square cm in size.

The visual metronome was also constructed using the computer's graphics capability. This consisted of a black bar approximately 2.5 cm
high and 7 cm wide that flashed in the center of the TV.

Rhyme training. In order to participate in the experiment, a subject first needed to know all of the rhymes and be able to think them in time with the metronome. Training progressed in several stages. The displayed instructions first explained that the experiment would test the subject's ability to perform several mental tasks simultaneously. It was further explained that one of these tasks required the subject to be able to recite the words to ten well-known nursery rhymes. To make sure the subject knew the rhymes, the subject was asked to study the rhymes as they were displayed on the TV monitor. The amount of time each rhyme was shown was controlled by the subject who pressed "enter" whenever she or he felt familiar with the currently displayed rhyme.

After studying all of the rhymes, the subject was tested to see how well she or he knew them. At the beginning of each test trial the subject was first shown the rhyme to be tested and again allowed to study it as long as desired. When the subject indicated that she or he was ready to begin the test by pressing "enter", the rhyme disappeared and the black bar began to flash on the screen every 6/10ths of a second. Beginning with the first bar, the subject was to recite the rhyme aloud, pronouncing one word each time the bar appeared. Exactly as many bars were displayed as there were words in a rhyme so that if a rhyme was recited correctly, the subject would say the last word as the final bar appeared. During this testing phase the experimenter listened to the subject via the intercom and watched the bars on his monitor in order to identify any rhymes not recited correctly.

Following this testing phase the experimenter practiced with the subject any rhymes recited incorrectly and the entire practice session was
then repeated for only the missed rhymes. This included allowing the subject to study the missed rhymes again, and then testing his (her) ability to say the rhymes in time with the bars. Again, the experimenter listened and noted any rhymes that the subject missed. The practice session continued in this manner, until the subject had recited all of the rhymes correctly in time with the metronome.

The next training phase tested subjects on the rhymes using a procedure more similar to the experimental task. Subjects were asked to think the words silently to each rhyme in time with the metronome. The moment a subject finished thinking a rhyme she or he was to press the "enter" key. The computer then subtracted the number of bars used in thinking the rhyme from the number that should have been used, and displayed the result on the subject's (and experimenter's) monitor. In this way both were able to identify any rhymes thought incorrectly. Each subject repeated this task for all of the rhymes until she or he thought the words correctly to 70% of the rhymes and could recite without error the words to any missed rhymes when quizzed by the experimenter.

Once the rhyme learning phase of training had been completed, the experimental task was explained. This task varied depending on the replication.

**Replications 1 and 2.** On each trial, in Replications 1 and 2, subjects first saw on the monitor for four seconds the first two or three words from the rhyme they would have to think. At the end of this time the screen was blanked for two seconds after which the first black bar appeared. As subjects thought the designated rhyme in time with the flashing bars they could be interrupted in one of two ways. On some trials three words from the rhyme were suddenly displayed including the word the
subject should have been thinking at that instant. On this type of trial the subject was to indicate, by pressing a numbered response button, which of the three words she or he was thinking. A message stating whether or not the subject's response was correct was then displayed, after which the trial ended. This type of trial was presented periodically (25% of the trials) to assure that the subject would concentrate on correctly thinking the words to the rhymes.

The second way in which a subject could be interrupted was by the appearance of a degraded probe word. When this occurred the subject was to press the "enter" button the moment she or he knew what the word said. When the button was pressed the computer recorded the elapsed time since the word first appeared and blanked the screen. Immediately thereafter a message asked the subject which of four words matched the probe word. This test was included to make sure the subject had read the probe correctly. After a subject indicated one of the four words by pressing the corresponding numbered button, a message appeared informing him or her whether the response had been correct and the trial ended.

Subjects were given an eight second rest between trials. During this time a message was displayed informing them (a) how many milliseconds had been used to identify the last probe word, (b) what the subject's mean probe response time was to that point in the experiment and (c) how many, if any, errors the subject had made in identifying the probe words. This summary was displayed after every trial in an attempt to reduce error variance in subjects' response speed. It was discovered in early pilot sessions that such a summary motivated most subjects to respond to the probes as quickly as possible in order to "beat their average."
To acquaint the subjects with the various components of the experimental task, they were first asked to practice two rhymes on which they were interrupted by being asked which word they were thinking. After successfully completing this task, the subjects practiced identifying ten degraded words displayed sequentially on the screen. After identifying each word subjects received feedback concerning their accuracy and speed analogous to the feedback they would receive in the experimental trials.

Upon completing these two practice activities, the training session concluded with subjects practicing four trials identical to the experimental trials. On two of these the subjects were interrupted by being asked what word they were about to think. On the two remaining trials subjects were interrupted by a degraded word. These four practice trials were repeated until all four were completed without an error.

The 32 experimental trials were presented in two 16 trial blocks separated by a five minute rest period. In each block, the eight odd trials were filler trials in which a subject was either asked what word she or he was thinking (trials 1, 7, 9, and 14), or interrupted by a degraded probe word not in any of the rhymes. The eight even trials were the experimental trials in which an early or late word from a rhyme was presented as a probe.

Replications 3 and 4. In Replications 3 and 4 several alterations in the experimental procedure were introduced in hopes of increasing the sensitivity of the probe to variations in thought activation. First, the probe words were degraded further, making them more difficult to read. It was reasoned that if a small amount of probe masking increased the matching effect (cf., Meyer & Schvaneveldt, 1976), then degrading the probe more might further strengthen the effect.
A second change introduced was to alter the procedure used to motivate subjects to concentrate on correctly thinking the rhyme words in time with the metronome. In Replications 1 and 2 subjects were tested only four times in each 16 trials as to whether they were thinking a rhyme correctly. As a result, they may have had a tendency to wait for the probe rather than to concentrate on thinking the rhymes correctly. To avoid this, in Replications 3 and 4 subjects were tested on every trial. This was accomplished by having subjects resume thinking the words to a rhyme after responding to the probe and then press "enter" when they finished the rhyme. As in the earlier practice session, the computer subtracted the number of bars a subject used from the number that should have been used and displayed the difference on the monitor. Thus, on every trial subjects were given feedback as to whether they had thought the rhyme words correctly. All of the 16 distractor trials were then used to present probe words not in any of the rhymes.

With the altered task, subjects had to remember their place in a rhyme while responding to an interrupting probe. This would have been very difficult if they had been required to first recognize a probe and then select it from a list of words as in Replications 1 and 2. Therefore, in Replications 3 and 4 subjects were only required to read the probe aloud. A voice switch was used to time their response and blank the display. The experimenter listened via the intercom to check the subject's response for accuracy.

A final procedural alteration introduced in Replication 4 was to offer subjects a few pennies for their thoughts. It was reasoned that error variance might be reduced if subjects were rewarded with money for responding quickly and accurately. A payoff schedule was therefore used.
To encourage speed, subjects were paid from one to five cents each trial depending on how quickly they responded to the probe. To encourage accuracy in recognizing the probe, subjects' total earnings were decreased by five cents for each recognition error. Finally, to encourage accuracy in thinking the rhymes, subjects received one cent for every trial in which the rhyme words were thought correctly.

Results

In accordance with the results of Experiments 1 and 2, for all replications subjects' response latencies were converted to speed scores (1000/response time in milliseconds or "probe words recognized per second").

Replications 1 and 2

None of the subjects in Replications 1 and 2 mis-identified any of the experimental probes on the multiple choice recognition tests. Consequently, all data were available to be used in the analyses.

Replications 1 and 2 of Experiment 3 were identical in every way except for the 16 subjects who participated. There was, therefore, no reason to suspect differences in their pattern of results. To confirm this, an analysis of variance of probe recognition speeds was conducted incorporating subject group as a between-subjects factor and word type (early vs. late in the rhyme) and probe presentation position (1 through 8) as within-subjects factors. This analysis indicated that the group replication factor did not interact with the word type variable, the linear or quadratic trend of response speeds across the eight probe positions, or the interaction between these two variables (all Fs < 1.0). The data from the two replications were therefore combined in examining
the experimental hypotheses.

The questions of interest in Experiment 3 were (a) whether a spiked pattern would occur indicating that probe recognition was facilitated only when the probe matched the word being thought (i.e., in positions 3 and 6) and (b) whether an interaction between word type and probe presentation position would obtain thereby demonstrating the probe's ability to discriminate between two closely occurring thoughts.

Subjects' mean probe recognition speeds for the combined replications are displayed in Table 3 on the following page. These data show that a spiked pattern did not occur. For the early probe, elevations in recognition speeds relevant to position 3 persisted in both positions 4 and 5; similarly, for the late word probe, recognition speeds did not decrease in positions 7 and 8 in comparison to position 6.

The general pattern of response speeds suggests the predicted interaction between word type and probe position. For the early probe words, recognition speed was highest at positions 3, 4 and 5. A similar elevation did not occur for the late probe words which were recognized least quickly in position 4. These divergent patterns were reflected statistically in a marginally significant quadratic interaction between word type and presentation position (F (1, 30) = 3.01; p < .10). None of the other trend analyses of the interaction approached significance. The only other marginally significant effect was a main effect for word type (F (1, 30) = 3.00; p < .10). This was of no consequence, merely reflecting that, as a group, the late words were more difficult to recognize when they appeared as probes than the early words.

Replications 3 and 4

It was hoped that Replications 3 and 4 would replicate and strengthen
Table 3

Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions:

Experiment 3 - Replications 1 & 2 combined.

<table>
<thead>
<tr>
<th>Probe Word</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (#3)</td>
<td>1.21</td>
<td>1.22</td>
<td>1.28</td>
<td>1.30</td>
<td>1.29</td>
<td>1.19</td>
<td>1.24</td>
<td>1.19</td>
</tr>
<tr>
<td>Late (#6)</td>
<td>1.17</td>
<td>1.25</td>
<td>1.16</td>
<td>1.09</td>
<td>1.20</td>
<td>1.14</td>
<td>1.16</td>
<td>1.16</td>
</tr>
</tbody>
</table>

Note: Cell N = 32.
the pattern of results found in Replications 1 and 2. To this end, several alterations in the procedure were made. To review, these consisted of (a) further degrading the probe words, (b) testing subjects on every trial as to whether they had thought the words to the rhyme correctly and having them verbally respond to the probes and (c) in Replication 4, providing subjects with a monetary incentive to respond quickly and accurately to the probes.

Subjects made four probe recognition errors in Replication 3 and two errors in Replication 4 (1.2% of the total responses in the combined replications). Because errors were few and relatively equally distributed across probe presentation positions, response scores for trials on which errors occurred were simply replaced by the corresponding cell mean.

Unlike Replications 1 and 2, Replications 3 and 4 differed from one another in an important procedural detail — the opportunity for subjects to earn money. An initial analysis of variance of subjects' probe recognition speed scores that included replications as a between-subjects factor produced only one significant result involving the group replications factor. This was a significant three-way linear interaction between Replication, probe type, and probe presentation position ($F (1, 30) = 4.47; p < .05$). Separate analyses performed on each of the two replications, and an examination of their patterns of means (displayed in Tables 4 and 5 on the following pages) disclosed the nature of this higher order interaction. The pattern of response speeds of Replication 3 (see Table 4), in which a money incentive was not used was similar to Replications 1 and 2. The early probe word was recognized relatively quickly in positions 3, 4, and 5 as compared with the other locations; for the late word probe there was no corresponding increase in recognition speed in the middle positions.
Table 4

Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions:

Experiment 3 - Replication 3 (no money)

<table>
<thead>
<tr>
<th>Probe Word</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (#3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.52</td>
<td>.57</td>
<td>.66</td>
<td>.58</td>
<td>.65</td>
<td>.57</td>
<td>.61</td>
<td>.59</td>
</tr>
<tr>
<td>Late (#6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.53</td>
<td>.52</td>
<td>.49</td>
<td>.53</td>
<td>.52</td>
<td>.52</td>
<td>.55</td>
<td>.57</td>
</tr>
</tbody>
</table>

Note: Cell N = 16
Table 5

Subjects' mean probe recognition speeds for early (position #3) and late (position #6) probe words in each of the eight positions:

Experiment 3 - Replication 4 (money).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (#3)</td>
<td>.72</td>
<td>.71</td>
<td>.68</td>
<td>.73</td>
<td>.71</td>
<td>.63</td>
<td>.62</td>
<td>.61</td>
</tr>
<tr>
<td>Late (#6)</td>
<td>.61</td>
<td>.55</td>
<td>.60</td>
<td>.60</td>
<td>.65</td>
<td>.61</td>
<td>.61</td>
<td>.62</td>
</tr>
</tbody>
</table>

Note: Cell N = 16.
Thus, once again, the general pattern was that of a quadratic interaction, although statistically this trend was not significant ($F(1, 15) = 2.10; p < .17$). As in Replications 1 and 2, in Replication 3 there was no tendency for the linear interaction term to be significant ($F < 1.0$).

The results of Replication 4, in which a money incentive was used (see Table 5), produced a pattern of results similar to the other replications for positions 3 through 8. That is, mean recognition speed for the early probes was higher for positions 3, 4 and 5 as compared with positions 6, 7 and 8, while mean recognition speeds for the late probes exhibited no such relative increase in the earlier positions.

The results of Replication 4, however, produced a pattern of results distinct from Replication 3 (and Replications 1 and 2) when the first two presentation positions are examined. A consequence of the money incentive was an elevation in recognition speed in positions 1 and 2 for the early, but not the late, probe words. This resulted in the linear component of the word type by probe position interaction being significant in Replication 4 ($F(1, 15) = 7.78; p < .02$), while the quadratic interaction component was non-significant ($F < 1.0$). The presence of this linear interaction in Replication 4, but not Replication 3, was in turn the source of the three-way interaction between replications and the other two factors.

The effects of the money incentive can best be seen in Figure 2 (see next page), where probe recognition speeds are graphically displayed for Replication 4 (panel a) together with the combined responses from the non-money replications (panel b). In the first instance the significant linear interaction is easily discernible. An analysis of the combined data from the other three non-money replications, on the other hand, produced a significant interaction for the quadratic component ($F(1, 45) =$}
Figure 2: Mean probe recognition speed as a function of probe word type and presentation position for money and no money replications.
4.96; \( p < .04 \), but not for the linear component \( (F < 1.0) \). Again, in the combined analysis of Replications 1, 2 and 3, the replications factor did not interact with word type, probe position or the interaction between them (all \( F_s < 1.0 \)).

Discussion

The results of Experiment 3 provide additional evidence of the ability of the probe technique to measure internal activation of concepts in memory. In both the replication in which subjects were offered a money incentive, as well as in the combined replications where no incentive was provided, a significant interaction occurred between probe word type and presentation position. Thus, the probe procedure was able to discriminate between two groups of thoughts internally activated in memory 6 thoughts and 3.6 seconds apart.

The results of Experiment 3 showed no tendency for probe recognition speeds to be "spiked" (i.e., show a facilitation effect only in positions 3 and 6 when a probe matched the word subjects were thinking). The early word probes were recognized rapidly when presented in positions 4 and 5, while the late word probes likewise showed no decrement in recognition speed in positions 7 and 8. Thus, some measurable activation appeared to persist for two additional probe positions (i.e., 4 thoughts and 2.4 seconds). This provides additional evidence that the probe facilitation effect was not simply the result of subjects guessing the probe would be the word they were thinking. It also suggests that in a judgment task a probe would not have to be presented at the exact moment a thought was activated to measure its occurrence.
While the outcome of Experiment 3 was generally encouraging, several unexpected results occurred. One of these involved the high variability in subjects' probe recognition scores. Although the combined analysis of the first three replications yielded a statistically significant result, none of the replications alone was significant. If the combined replications are taken as the best estimate of the expected quadratic interaction, the power of the F test for an \( \alpha \) of .05 with 1 and 15 degrees of freedom (i.e., one replication with 16 subjects) would be only .48.

The use of a monetary incentive in Replication 4 to reduce response variability also had an unexpected effect. The relative speed with which subjects identified the probe in positions 1 and 2 increased for the early words, but not for the late. Without additional data, it is only possible to speculate why this might have occurred. Since subjects were being paid bonus money for thinking the rhyme correctly, one possibility is that in the money replication subjects attempted to mentally review the rhyme during the brief interval between the time the rhyme was announced and the bars began to flash. Because this interval was short (i.e., 5 to 6 seconds) subjects may have only had time to review the first line of each rhyme. If this were so, they would have activated in memory the early, but not the late, probe words. This, in turn, could have facilitated recognition of the early probe words when they appeared in the first two positions. Whatever the correct explanation for the differences in the money and non-money replications, in both cases an interaction occurred that discriminated between words activated by subjects early and late in the rhymes.
A final possibly perplexing result of Experiment 3 was the failure
to find a marked increase in recognition speed for the late probe words
in positions 6, 7 and 8 to correspond with the increase in early word
recognition speed in positions 3 through 5. Two factors may have contrib­
uted to this failure. First, as discussed in the introduction, it may
have been more difficult for subjects to recognize probes in some posi­
tions as compared to others. If subjects were having difficulty keeping
their place late in the rhyme this could have suppressed probe recognition
speeds and masked any facilitation effect resulting from their thinking
the late probe words. A second possible cause for the failure to find a
marked increase in recognition speed for the late probe words is the like­
lihood that fewer subjects were thinking the rhyme correctly in the late,
as compared to the early, positions. If the point at which subjects were
thinking the late probe word in the rhyme tended to be dispersed it would
have diffused any increase in late word recognition speed.

In spite of the several discrepancies between the expected and ob­
tained results, Experiment 3 demonstrated (by the obtained interactions)
the ability of the probe technique to discriminate between two groups of
words internally activated in memory in rapid succession. This ability was
used in Experiment 4 to investigate how people draw on their memory in a
person judgment task.
MEASURING WHAT IS ACCESSED DURING A PERSON JUDGMENT:

EXPERIMENT 4

The Task: Would a Friendly Person Make a Good Nightwatchman?

In this chapter an experiment is reported that employed the probe procedure to test several hypotheses about the way in which people draw on memory when they make judgments about people. The hypotheses tested were those outlined at the conclusion of Chapter I. Specifically, when people make memory-based person judgments they tend to (a) access in memory judgments or inferences they have made about a person, (b) access episodic incidents that were relevant, as opposed to irrelevant, to making an earlier judgment and (c) access an organizing judgment prior to recalling episodic incidents upon which the judgment was based.

These hypotheses were tested employing the task diagramed in Figure 3 on the following page. Subjects were asked to make pairs of judgments about several stimulus persons, each described by two personality traits. The first judgment concerned whether the stimulus person was likely to be characterized by a third trait and was made while the descriptor traits were still available to the subject. To maximize the probability that this trait judgment would serve as an organizer of a subject's impression, the judgment trait was presented prior to the description of the stimulus person in a manner similar to that reported by Lingle and Ostrom (in press a and b). Of the two descriptor traits, one was relevant and one was irrelevant for deciding if the person was likely to possess the judgment
Figure 3: Flow chart of the experimental task used in Experiment 4.
trait.

The second judgment subjects made concerned the likelihood that the person would be successful at a designated occupation. This judgment was made a short time later (without the traits being available) following a brief distracter task. During the 7½ seconds subjects were given to make this memory-based occupational judgment, a probe word was presented that was either (a) the judgment trait a subject had initially made a decision about (b) the descriptor trait that had been relevant to that judgment, (c) the descriptor trait that had been irrelevant to that judgment or (d) a matched control trait that had not been associated with the person. These probes could appear either early or late in the 7½ second judgment period (i.e., approximately 2 or 6 seconds after the question had been asked).

Based upon the results of Experiments 1, 2 and 3 it was expected that the speed with which a subject could identify a probe would increase whenever the word had recently been activated in memory during the judgment. In accordance with the several hypotheses, it was expected that probe recognition speeds would show that subjects tended to (a) access the trait judgment they had made in addition to the descriptor traits, (b) activate in memory the irrelevant descriptor trait less than the judgment trait and relevant descriptor trait and (c) access the judgment trait early in the judgment period, prior to the descriptor traits upon which it had been based.

Method

Subjects

Subjects consisted of 95 students from the Ohio State University
subject pool who participated in the experiment in partial fulfillment of an introductory psychology course requirement. Of these, 24 were used to construct the stimulus materials while 71 participated in the experiment proper. Data from seven of the experimental subjects could not be used. In one case it was learned after the experiment that a subject had misunderstood the instructions and was making the judgments incorrectly. For the other six, problems in the execution of the computer program prevented a full set of data from being collected. Of the 64 subjects whose data was used, 36 were female and 28 were male. Subjects were randomly assigned to one of the 64 stimulus material presentation orders used to counterbalance the experimental design.

Stimulus Materials and Design

Counterbalancing of the design. A 4 (probe type) by 2 (time of probe presentation) within-subjects design was employed. Counterbalancing was directed toward assuring that no systematic differences existed in the words that appeared in the four probe word conditions (i.e., judgment trait, relevant descriptor, irrelevant descriptor and unassociated control trait). This was accomplished by constructing four sets of six person traits each consisting of two subgroups whose three traits were similar to each other, but not to the traits in the other subgroup (e.g., friendly, outgoing, talkative [subgroup 1]; intelligent, thoughtful, articulate [subgroup 2]). One trait was then randomly selected from each set to be used as a judgment trait along with one trait from each subgroup to serve as a relevant and irrelevant descriptor (e.g., outgoing might be the judgment trait with talkative and intelligent as the relevant and irrelevant descriptors). The three traits randomly selected from each of the four trait sets were used as stimulus traits for one group of four subjects, the particular
probe presented to a subject being latin square counterbalanced across subjects and trait triads. The unassociated control word shown to a subject was randomly selected from the three unused traits of one of the three trait sets not used as stimuli that trial.

A second group of four subjects were then included in the design who saw as stimulus traits (and probe words) the three traits in each set not seen by the first group. This meant that they saw as experimental probes the traits from which group 1's control probes had been selected. The control probes they saw, on the other hand, were randomly selected from the first group's experimental traits. As a consequence, across the two groups, two randomly selected traits from each set of six traits appeared as probes in each of the three experimental conditions, while two traits, again randomly selected from each set of six, appeared a second time as unassociated control probes.

The experimental design required that each subject see the four types of probes both early and late in the decision process. To accomplish this, each subject saw four additional six-trait sets counterbalanced in the manner just described. However, on trials when these traits were used the probes were presented late in the judgment. In order to counterbalance probe word and time of presentation, an additional eight subjects were included in the design who saw identical stimulus materials as the first eight subjects except that they saw the early probes late and the late probes early. Thus, over the 16 subjects identical probes were seen in both the early and late probe presentation conditions.

To increase power, it was decided to have each subject complete four replications of the design, thereby providing four responses in each of the eight conditions. To do this, 24(3 x 8) additional stimulus trait sets were
constructed and counterbalanced in the manner just described. The order in which subjects saw each of the four, eight trial replications was Latin square counterbalanced requiring 64 subjects in all (4 replications x 16 subjects). Within each eight-trial replication the order in which the trials were presented was randomly determined for each subject. Finally, the occupation a subject judged on each trial was determined by pairing a single occupation with each of the 32 trait sets. An effort was made to select an occupation in each case that was not particularly relevant to any of the six traits in the set.

Construction of the trait sets. To successfully implement the counterbalancing scheme, it was necessary to construct the 32 trait sets in such a way that any trait, within a three-trait subgroup, would be more relevant for deciding if a person had some other trait in that same subgroup than any trait in the other subgroup. To this end, an experimental assistant first intuitively constructed 32 such sets, drawing without replacement from the middle 3/5ths of Anderson's (1968) scaled trait list. These traits were then ordered in six different booklets with each of the six traits in a set being placed as a judgment trait at the head of the other five. Copies of these booklets were given to 12 subjects (two subjects filling out each booklet) with instructions to go through and select the two traits in each of the 32 groups that they believed would be most helpful in deciding whether or not a person would possess the trait that headed the group. These booklets were then scored to determine how often, out of the six times one of the three traits from each subgroup appeared as the judgment trait, the other two traits from the same subgroup were selected as being most relevant to it.
There are 5! or 120 ways in which five items or traits might be ranked in relevance to a 6th trait. If the five traits were randomly ranked, the probability of the two traits from the subgroup of the judgment trait being ranked in the first two positions would be 12/120 or .10. Consequently, at least three out of the six subjects who saw judgment traits from a single subgroup had to choose the other two traits from the subgroup as most relevant for the probability of occurrence to be significant (p < .02). This did not occur for at least one of the subgroups for eight of the original 32 trait sets. These sets were therefore restructured and presented to a second group of 12 subjects who rated them. Each set met the three-out-of-six criteria on the second rating for both subgroups. Of the 384 ratings of the final 32 trait sets (6 booklets x 32 judgments x 2 subjects), both traits from the judgment traits' subgroup were selected as most relevant 294 times (see Appendix C for a list of the 32 trait sets and their paired occupations).

**Experimental Procedure and Instructions**

When subjects arrived they were seated at one of the computer controlled stations described in the report of Experiment 3. The experimenter familiarized each subject with the equipment, explained that the instructions would be displayed on the monitor, and indicated a microphone for the subject to use when making verbal responses. The first page of instructions was then displayed.

The instructions began by explaining that the experiment was concerned with how job counselors and job placement officers make rapid occupation suitability judgments about many different people. Subjects were asked to role play a job placement officer and make pairs of judgments
about different stimulus persons who would be described by two trait ad-
jectives similar to those likely to be found in a letter of recommendation.
It was explained that these judgments would consist of an initial trait
judgment followed by an occupation suitability judgment. To familiarize
them with the task, subjects next made pairs of judgments about five stim­
ulus persons. The judgment procedure was similar to that later used in
the experiment trials (see Figure 3, page 74), although in this initial
practice phase subjects were not interrupted by probe words or a distrac­
ter task.

Each judgment trial began with the message "Get ready for the next
person." Three seconds thereafter a second message appeared which read,
"Would a person with the next two traits be:" followed by the judgment
trait for that trial. This remained in view for five seconds after which it
disappeared and two descriptor traits were displayed. At this point sub­
jects were to form an impression and decide whether the stimulus person
would be characterized by the earlier shown judgment trait. After 10 sec­
onds the traits disappeared and subjects were asked to respond with a 1 or
2 on the response box depending on whether they thought the person would,
or would not, possess the judgment trait. A three second pause followed
the subject's response during which the screen remained blank (in the
experimental trials this period was occupied by a 30 to 50 second distrac­
tor task). Next, the message appeared "Would this person be successful
as a:" followed by the occupation to be judged. This display remained on
the screen for 7\(\frac{1}{2}\) seconds during which subjects were to reach a decision
(in the experimental trials a probe appeared either 2 or 6 seconds into
this judgment period). At the end of this time, a 5-point numbered scale
appeared anchored by the words "very unlikely" (1) and "very likely" (5).
At this point the subjects were to enter a number between 1 and 5 indicating how successful they felt the person would be in the occupation. A six second rest period then followed before the next trial began.

It seemed unlikely that the probe technique would prove successful if subjects spent the decision period waiting for the probe to appear rather than thinking about their decision. To avoid this from happening, it was emphasized several times in the instructions that subjects should make their judgments during the allotted time and respond immediately after a response scale appeared. To further encourage subjects to do this, an error message appeared telling them to respond faster whenever an answer was not given within five seconds of the appearance of a response scale.

After a subject completed the five practice trials, the probe recognition task was introduced. It was explained to subjects that their judgments were going to be made more difficult by having them intermittently perform a second task since job placement officers encounter many distractions while making decisions. The task, subjects were told, consisted of identifying as quickly as possible degraded words that could appear on the display monitor at any time. To familiarize them with the procedure, subjects were asked to identify ten degraded practice words. The degraded alphabet and probe recognition procedure from Replications 1 and 2 of Experiment 3 were used. This consisted of having subjects push a button the minute they identified the probe and then respond to a four-word multiple choice test asking what the probe had said. This procedure was used because the results of Replication 3 (Experiment 3) did not indicate that the increased errors resulting from the use of the more difficult alphabet and voice switch could be justified by a decrease in response variability.

At the end of each trial a message reported a subject's immediately
preceding and average word recognition times as well as the number of word identification errors made.

The final practice phase consisted of having subjects complete eight judgment trials identical to the experimental trials in which they were interrupted by distracter tasks and the appearance of degraded probe words. These practice trials were structured to exemplify several dimensions of the experimental trials. Before describing them, however, it is necessary to describe in more detail several additional aspects of the experimental procedure.

It was felt that several features needed to be incorporated into the judgment task if the probe technique were to be successful in measuring how subjects draw on memory during a decision. To accomplish this, it was decided to have subjects make pairs of judgments about eight additional stimulus persons in each block of trials. These always occurred at positions 1, 2, 4, 7, 10, 11, 13 and 14 (trials 3, 5, 6, 8, 9, 12, 15, and 16 being the eight experimental trials). These trials were included only to structure the ambience of the experimental task and were not analyzed.

One use of the bogus trials was to motivate subjects to consider their judgments carefully. It was considered imperative that subjects base their responses in some way on memory for the stimulus information. Random or arbitrary responses would be of little value. To maximize the probability that subjects would concentrate in making their judgments, on three of the eight bogus trials in each block a message unexpectedly appeared asking subjects to explain aloud into the microphone the last decision they had made.

In addition to this use, the bogus trials were employed to present probes (a) other than person traits and (b) at points other than the
occupational judgment. This was done to minimize the degree to which sub-
jects could come to expect a single type of probe word to appear at a
fixed point in the judgment sequence. Points other than early and late
in the occupational judgment where probes appeared were (a) while the
two descriptor traits were being displayed and (b) immediately
following a subject's trait judgment, before the occupational judgment
appeared. The nouns and verbs used in these trials were simply arbitrarily
generated by the experimenter. The distracter task introduced between the
trait and occupational judgments during the experimental trials consisted
of having subjects identify two degraded probe words. These were pre-
sented sequentially and were always nouns or verbs rather than adjectives.
Subjects generally took from 30 to 50 seconds to recognize the probes and
respond to the multiple choice tests. Thus, on the experimental trials
three probes were presented in all (the third being presented during the
occupational judgment). Two degraded words were also presented between
the trait judgment and occupational judgments on some of the bogus trials
so that this sequence could not serve as a reliable cue for an experimental
trial.

On four of the eight practice trials that subjects completed to conclude
the practice session, probes were presented in a similar manner to the
experimental trials (i.e., two distracter probes following their trait
judgment, and a judgment trait, relevant descriptor, irrelevant descriptor
or unassociated control trait probe during their occupational judgment).
On the remaining four practice trials, traits, as well as nouns and verbs,
were presented as probes in different patterns. These were varied as
much as possible to make it difficult for subjects to predict where and
what type of probe might appear in the judgment sequence. Finally, three
times during the practice trials, once following a trait judgment and
twice following an occupational judgment, subjects were interrupted and asked to explain their last judgment.

The instructional and training phases of the experiment lasted approximately 30 minutes. Once they were finished the experimental trials were begun. Subjects completed four blocks of 16 trials, each requiring approximately 20 minutes. A five minute rest was provided between each block during which subjects were allowed to get up and move around. After completing the fourth block of trials, subjects were quizzed about what they believed the purpose of the experiment to be and debriefed.

Results

During debriefing subjects expressed no awareness that probe recognition speed was being used to measure their thoughts. Most reported they believed the probes were meant only as distracters. A few said they thought the experiment concerned whether or not the appearance of a trait as a probe would influence their occupational decision.

As a prelude to testing the several hypotheses, subjects' response speeds in the eight experimental conditions were averaged over the four replications. These scores were then used in the set of a prior orthogonal contrasts that best tested the experimental hypotheses.

The subjects made 36 probe recognition errors out of a total of 2,048 responses. Because of the low number of errors (1.6%) and their relatively even distribution across the experimental conditions, trials on which errors occurred were deleted and a subject's average response speed was computed from the remaining scores in the condition.

Subjects' mean response speeds for the eight experimental conditions are displayed in Figure 4 on the next page.
Figure 4: Mean probe recognition speed as a function of probe type and time of probe presentation in Experiment 4.
Judgment Trait and Descriptors vs. Control

The first question examined was whether subjects activated the stimulus information in memory during their occupational judgments. In Figure 3 it can be seen that recognition for the judgment trait and two descriptors was faster than recognition of the unassociated matched control words. The single degree of freedom contrast between the combined experimental groups and the control condition (collapsed over the time of presentation factor) was highly significant ($F (1, 63) = 25.4; p < .001$). Thus, there was evidence that subjects activated prior thoughts about the stimulus person (whether descriptors or inferences) when considering their occupational judgments.

Judgment Trait vs. Relevant Descriptor

The second question considered was whether subjects accessed their early trait judgment during their subsequent occupational judgment. The pattern of recognition speeds in Figure 3 indicates that this was the case. Recognition speed for the judgment trait was nearly identical to recognition speed for the relevant descriptor ($M_s = 1.084$ vs. $1.087$, collapsed over presentation time), and the contrast between them was nonsignificant ($F < 1.0$). Thus, subjects accessed their initial trait judgment as much as the best remembered descriptor.

Judgment Trait and Relevant Descriptor vs. Irrelevant Descriptor

The third hypothesis tested was that a trait would less likely be accessed during a decision if it were irrelevant (as opposed to relevant) to an initial judgment made about a person. Support for this hypothesis is provided by the fact that mean recognition speed for the irrelevant trait was slower than recognition speed for either the judgment trait or
the relevant descriptor. The single degree of freedom contrast between the combined judgment trait and relevant descriptor vs. the irrelevant descriptor yielded a value of $F(1, 63) = 3.83; p < .06$. The non-orthogonal contrast between the relevant and irrelevant descriptor traits was $F(1, 63) = 3.83; p < .06$.

**Judgment (Early-Late) vs. Descriptors (Early-Late)**

The final hypothesis examined was whether, when making a memory-based decision, subjects would access an earlier judgment they had made prior to the traits upon which it was based. If this hypothesis were true, it would be expected that the pattern of probe recognition speeds in the early and late presentation conditions would be different for the judgment trait, as opposed to the two descriptors. If a stimulus item were activated early in the judgment process, it would facilitate recognition of that word when it appeared early as a probe. This facilitation could be expected to decline across the decision period if the thought were not activated in memory again (or remain constant if the thought were recurrently considered). On the other hand, if a thought tended to be activated in memory late in a decision, recognition for a matching probe would be facilitated when the probe appeared late, but not early, in the judgment period. This would result in either a reversal or a difference (i.e., an interaction) in magnitude between the early and late probe recognition speeds when the judgment trait was compared with the two descriptors.

An examination of Figure 4 shows that this was not the case. The differences in recognition speed between the early and late probes were very small and almost identical in all four conditions. The single degree of freedom contrast of the interaction between the early and late presentation conditions and the judgment trait vs. the descriptors was nonsignificant.
Additional Analyses

Analysis of the replications factor. The tests just described were based upon subjects' probe recognition speeds averaged over the four replications. It is possible that subjects altered their decision strategy during the experiment as a result of having to make judgments about so many stimulus persons. To determine if this were the case, subjects' probe recognition scores for the first two replications were averaged together and compared with their probe recognition scores for the second two replications. The pattern of results in the two cases was almost identical. Analyses of the interaction between the two replication pairs and each of the orthogonal contrasts used to test the several hypotheses produced no significant results. The only $F$ value greater than 1.0 was the interaction between replication pairs and the contrast between the experimental and control conditions ($F(1, 63) = 3.30; p < .10$). This reflected the fact that the difference in the contrast between the combined experimental vs. the single control condition increased slightly in the second two replications when compared to the first two. If this difference were reliable, it would be consistent with subjects' reports during debriefing that they became more proficient at remembering the stimulus information as the trials progressed. The one difference that did occur between the early and late replications was a simple main effect ($F(1, 63) = 24.2; p < .001$) indicating that subjects were able to identify the probes more quickly in the last, as compared with the first, two replications.

Alternative analysis of the relevancy effect. Subgroups within the 32 trait sets were comprised of traits likely to co-occur (e.g., friendly, talkative, outgoing). This provided a check of the relevancy manipulation
since subjects should have responded "yes" whenever they based their trait judgment on the relevant descriptor. "Yes" responses occurred on 79% of the experimental trials indicating that the manipulation had generally been successful. However, this left 21% of the trials on which subjects responded "no". Since the relevant descriptor implied the presence of the judgment trait, in these instances a subject most likely based his or her decision on what was intended to be the irrelevant descriptor. As a result, on reversal (i.e., "no") trials, the relevant probe would be irrelevant and the irrelevant probe relevant. If this were so, and the relevancy effect was reliable, rescoring the relevant probes as irrelevant (and conversely, the irrelevant ones as relevant) on reversal trials should increase the relevancy effect.

To determine if this were the case, subjects' responses were rescored in the manner just described and the several orthogonal contrasts (collapsed over the replication factor) again computed. The result was that the difference in subjects' responses to the relevant and irrelevant descriptor probes was enhanced. In the initial analysis the difference between recognition speed for relevant and irrelevant descriptors was 1.087 - 1.056 or .031. After rescoring, the difference increased to 1.090 - 1.047 or .043. This increase resulted in the single degree of freedom contrast between the judgment trait plus relevant descriptor vs. the irrelevant descriptor being significant (F (1, 63) = 4.04; p < .05). The non-orthogonal contrast between the relevant and irrelevant traits was F (1, 63) = 3.90; p < .06, while the contrast between the control and three experimental groups yielded a value of F (1, 63) = 23.0; p < .001. Thus, rescoring subjects' responses in a manner that strengthened the relevancy manipulation increased the difference in probe recognition speed between the
relevant and irrelevant conditions.

**Discussion**

Three out of the four hypotheses tested in Experiment 4 received support. First, a comparison between the experimental and control conditions indicated that subjects activated prior person-associated thoughts when making memory-based occupational judgments. Second, a comparison between probe recognition speed for the judgment trait and the descriptors indicated that the judgment trait was accessed as much as the most accessed descriptor. Finally, a comparison between recognition speed for the irrelevant descriptor and the other stimulus items, indicated that a trait irrelevant to an early judgment was less activated in memory during a subsequent decision than an initial trait judgment and a relevant descriptor.

No support was found for the hypothesis that subjects would access an earlier judgment prior to the information upon which it had been based when making a subsequent memory-based decision. The time of presentation manipulation had no significant effect on probe recognition speed. Even the consistent tendency for subjects to recognize all of the probes more quickly late in the judgment period did not represent a significant difference. Unfortunately, it is impossible to discriminate between several explanations for why time of probe presentation produced no effect. One possibility is that the three stimulus items were not accessed in any consistent order. When subjects activated more than a single stimulus item the order in which this activation occurred may have depended on their relation to the judgment, not whether they were descriptors vs. an earlier judgment.
A second possibility is that subjects tended to review in memory all of the stimulus information items very rapidly. If such a review took place in the first 2 seconds of the judgment period, little difference in early and late probe recognition speeds could be expected even though subjects might have consistently accessed their trial judgment first. To detect this type of difference, the early probe would have to be presented even earlier in the judgment period.

An important question relevant to the results of Experiment 4 concerns the degree to which the differences in probe recognition speed reflect retrieval of the stimulus information during subjects' occupational judgments as opposed to residual activation of the concepts as a consequence of subjects' encoding of the stimuli during their trait judgments. Two types of evidence exist that the increments in probe recognition speed for the judgment and descriptive traits resulted from retrieval processes rather than external activation of the stimuli in memory during encoding. First, previous research (see Chapter II) indicates that in the absence of active rehearsal or retrieval, measurable increases in activation potential resulting from the encoding of external stimuli dissipates in a matter of seconds. In the present experiment, after subjects saw the stimulus materials they engaged in a distracter task that lasted over half a minute. Furthermore, during this time they activated at least eight other words in memory (i.e., the two distractor probes and the words in the multiple choice tests). To argue that increases in probe recognition speed resulted merely from subjects' encoding of the stimuli would be to ignore the evidence that such activation is short-lived in the absence of its reactivation in memory.
A second bit of evidence that increases in probe recognition speed resulted from internal retrieval, as opposed to external encoding, is that recognition speed did not track the time each stimulus trait was displayed to the subjects. Subjects saw the judgment trait for five seconds; the descriptor traits for 10 seconds thereafter. Nevertheless, the judgment trait was recognized as fast, or faster, than the descriptors while subjects were making their occupational judgments more than 30 seconds later. And finally, although the irrelevant and relevant descriptors were displayed for an equal period of time, the latter were recognized as probes more quickly than the former.

If probe recognition speed differences were the consequence of retrieval processes, a final question concerns whether this retrieval took place in the service of a person judgment. That is, it is possible that subjects spent the trials trying to remember the stimuli, not trying to make accurate occupational judgments. Again, two types of evidence suggest subjects made a concerted effort to remember the stimuli when making their person judgments. First, although not systematically recorded, subjects' explanations of their judgments on the bogus trials often included naming one or more of the stimuli. Second, during debriefing subjects would sometimes spontaneously report concern that they had occasionally forgotten one of the stimuli when making their occupational judgments. Such concern would seem more likely if subjects had been preoccupied with recalling the stimuli while making their decisions.

When testing or calibrating a new measuring device it is desirable to examine a known quantity. If this is not done, it is difficult to know whether deviations from expected outcomes result from a faulty instrument or a faulty hypothesis. It was for this reason that the probe technique
was initially used to test hypotheses that appeared to have a high probability of being confirmed. It is not startling to learn that memory biases that are found in recall and recognition tests are also found when the thoughts activated during a judgment are measured. However, such confirmation helps to validate the probe technique as an effective measuring device.

While the experimental results may not seem surprising, they do contribute to existing knowledge of person impression processes. First, for the reasons noted earlier, it is not necessary, and therefore cannot be taken for granted, that what is recalled in a memory test represents what will be accessed during a judgment. Even though the present findings parallel what has been found in recall tests, this cannot be interpreted to mean that this will always be the case. The present results only indicate that these biases do exist in some decision contexts; the degree to which this conclusion can be generalized to other judgment contexts will depend on the findings of future research. The present work does extend the findings of Lingle and Ostrom (in press b) who concluded that subjects base memory-based judgments on memory for earlier judgments. The present study differed from Lingle and Ostrom's experiments by unexpectedly asking subjects to justify their judgments. In this context, evidence emerged that subjects accessed descriptive stimulus traits in addition to their first judgment. It is interesting to note that this reliance on the descriptive information occurred in addition to, not instead of, reliance on memory for an early trait judgment.

If the present findings prove generalizable they have several theoretical, as well as practical implications. Practically, they suggest one reason why people's impressions of others are often difficult to
change. Factual information that is relevant to an initial impression judgment is later better remembered than irrelevant (and possibly less supportive) information. It seems likely that this bias not only affects judgments, but the manner in which new information about a person is interpreted (cf., Markus, 1977). This suggests that in order to change a person's impression it is first necessary to have the person reach a new organizing decision. It is interesting to speculate whether this might not most effectively be done by structuring the situation so a person is forced to make a new decision, rather than trying to change an impression by "convincing" someone with new factual information. The present research suggests it is the judgment, not necessarily the facts, that will later be remembered.

The results of Study 4 also have implications for both social and cognitive psychological theory. For example, a long standing theoretical issue in social psychology concerns whether attitude-consistent, as opposed to attitude-inconsistent, information is better remembered. While some studies have found better memory for attitude-consistent information (cf., Levine & Murphy, 1943) other more controlled studies have not found this to be true (cf., Greenwald & Sakumura, 1967). If person impressions are similar to attitudes about other objects, the present research suggests the best remembered episodic incidents will be those that are relevant to an attitude judgment. If attitude change is equated with the reaching of a new attitude judgment, it would seem that the best remembered episodic incidents will be those that are relevant to attitude change, not those that are associated with attitude constancy.

Finally, Study 4 has theoretical implications for theories of memory structure and retrieval. Most memory models make assumptions which have
implications for the availability of information in memory. For example, a recent theory by Anderson & Bower, 1973, postulates a simple last-in-first-out stack mechanism whereby recently activated memory structures are more available than less recently activated structures. The present research suggests that memory theories need to incorporate postulates that can account for internally generated structures (i.e., judgments) being more accessible than episodic incidents.
CHAPTER VI
SUMMARY AND CONCLUSIONS

Summary

The reported research grew out of two ideas: (a) memory retrieval represents a neglected, but important, dimension of the person perception process and (b) one reason for this neglect has been the absence of methodologies for investigating how people draw on memory when they make judgments about others. In Chapter I several hypotheses concerning the way in which people access person information during a judgment were proposed based on research which has investigated how people organize person information in memory. Next, in Chapter II a method was proposed for identifying thoughts accessed in memory during a judgment. This method consisted of determining how quickly, during a decision, a person could respond to a displayed probe word. A similar procedure has been used by researchers to identify concepts activated in memory when verbal stimuli are encoded. The degree to which this procedure might be effectively used to measure internally activated thought during a judgment was seen as depending on two questions: (a) does a fleeting internal thought increase activation potential to a measurable degree and (b) would the probe procedure be sensitive enough to discriminate between thoughts briefly activated in close approximation to one another.

To answer these questions, three studies were conducted in which subjects were asked either to hum or think the words to well known songs
and nursery rhymes. Words from the rhymes were then presented as probes either simultaneously with, or at some measured distance from, the moment a subject thought them in the rhyme. It was found that (a) the speed with which a subject identified a probe word, or the color of ink in which the word was written, both served to distinguish words internally activated in memory (Experiments 1 & 2); (b) of the procedures investigated, the most powerful measure of internal thought activation was the speed with which a subject could identify a degraded word (Experiments 1 & 2); (c) once a word had been activated in memory, measurable increases in activation persisted for a brief period thereafter (Experiment 3); and (d) the probe technique was sensitive enough to discriminate between two groups of words internally activated in memory 6.6 thoughts and 3.6 seconds apart from each other (Experiment 3).

In Chapter V a fourth experiment was reported that used the probe technique to test directly the several hypotheses proposed in Chapter I concerning the way in which people draw on memory in a person judgment task. Subjects were first asked to decide if a stimulus person, described by two descriptive traits, would likely possess a third trait. Thereafter, subjects were required to make a memory-based decision about the person's suitability for a designated occupation. While making this decision, probe words were presented. In support of the several hypotheses, the speed with which subjects recognized different probes indicated that in making their occupational judgment they (a) activated in memory the stimulus traits more than a group of control traits unassociated with the stimulus person, (b) accessed the trait judgment they had made about the person as much as either of
the descriptive traits, and (c) accessed the descriptive trait irrelevant to their initial trait judgment less than the judgment itself and the descriptor relevant to that judgment. A hypothesis not supported was that in making their judgments subjects would tend to access the trait judgment they had made about a person before the descriptive traits upon which the judgment had been based. Whether or not a probe was presented early or late during a judgment had no significant effect on the speed with which it was recognized.

As explained in the discussion of Experiment 4 (Chapter V), the results of Experiment 4 not only provide evidence about the manner in which people draw on memory, but help to validate the probe procedure as an effective "cognimeter" for measuring thought during a decision. In the remaining pages of this chapter several limitations associated with the procedure are considered conjointly with possible means for minimizing them. Thereafter, the chapter concludes by considering the several factors that determine what is remembered during a judgment and the unique contribution the probe technique might make towards investigating these factors.

Methodological Limitations

Need to Anticipate and Symbolically Represent Thought

The first limitation to the probe procedure is the need to anticipate, and represent symbolically, the thoughts that a person will activate during a decision. In this regard, there would seem to be no reason why degraded pictures, instead of words, might not be used. Presumably, thinking about an object would increase the speed with which
a degraded picture of it could be recognized. However, neither a verbal or pictorial symbol can be constructed without first identifying the specific concept to be represented. It may be that in many situations subjects form idiosyncratic implicational associates when encoding stimuli that are difficult to anticipate or symbolically represent (e.g., this person is just like my Aunt Tilly). If such cognitions later form the basis of decisions, the probe technique would be of little use in identifying them.

One means of dealing with this problem is to include a group of appropriately matched, unassociated control probes as was done in Study 4. In this way, at a minimum, it would be possible to identify concepts that are not accessed during the decision process. This in itself can prove informative. For example, it could be of theoretical, as well as practical interest, to identify the conditions under which an early impression judgment is not accessed during a later decision.

A second possible means for dealing with the thought anticipation problem would be to use the probe procedure in conjunction with some other method such as thought listing (cf., Calder, Insko & Yandell, 1974; Petty, Ostrom, & Brock, in press). Subjects' reported cognitive responses could be used to anticipate categories of thoughts likely to be remembered while the probe technique could then be used to verify these introspective reports under a variety of judgment contexts.

Reactivity of the Procedure

Not surprisingly, silent thought proved to be ephemeral. Many responses were required in each condition to show that small differences
in probe recognition speed were reliable. Conceivably, it may be possible to increase the sensitivity of the measure either by using monetary incentives (cf., Experiment 3, Replication 4), or by degrading the probe further and employing some type of multivariate recognition measure comprised of both response speed and recognition errors. However, even if this were to prove possible, it would likely continue to be necessary to have subjects make repeated judgments and be exposed to multiple probes. The question therefore arises as to whether administering the measure alters subjects' normal judgment processes and thereby limits the generalizability of the results.

The evidence available from Experiment 4 suggests that repeated judgments impact on the judgment process less than might be expected. First, when the first and last replication pairs were compared there were no significant differences in the pattern of subjects' probe recognition scores. The only marginally significant difference ($p < .10$) reflected an increased discrepancy between recognition speed for the stimulus traits as opposed to the matched control words. It is possible, of course, that this resulted from subjects coming to expect that 37.5% of the probes they saw while making their occupational judgments would be one of the stimulus traits they had seen. On the other hand, it may simply have reflected subjects becoming more accustomed to making decisions while being interrupted by the probes. There is some evidence that the latter explanation is more probable than the former. On 25% of the trials subjects were interrupted during their occupational judgments, the probe word was one of the descriptive traits. On the other hand, only 12.5% of the probes shown during subjects' occupational
decisions were judgment traits. If probe recognition speed tended
to track the frequency of probe occurrence, subjects should have recog­
nized the descriptors faster than the judgment trait. This was not the
case. The judgment trait, although it appeared only half as often as
a probe, was recognized as fast or faster than each of the descriptors.

Ultimately, of course, the ecological validity of conclusions based
upon the probe technique will have to be empirically validated. This
will need to include (a) identifying ways in which the probe recog­
nition task alters subjects' judgment processes and (b) cross-valid­
ating conclusions about the judgment retrieval processes with other ex­
perimental methodologies such as the correlational and decision time
strategies discussed in Chapters II and III.

**Activation of Associated Thoughts**

An additional problem associated with the probe technique is that
when a thought is internally activated in memory, other closely associat­
ed concepts are also activated. Thus, increased recognition speed of
a presented probe, in comparison to a control probe, may reflect activa­
tion of a closely associated thought, rather than a concept identical
to the probe. For example, in Experiment 4 it might be argued that
subjects accessed in memory only the relevant descriptor, and that
recognition speed for the judgment trait increased because of its close
semantic association with the relevant descriptor (or vice versa).

Studies by Warren (1970, 1972) indicate that a close associate to
a word which is externally activated in memory is not recognized as
quickly as the word itself. This suggest one means of dealing with the
association problem is to use powerful designs capable of detecting very
small differences in probe recognition speed. In the case of Experiment 4, the almost identical recognition speeds for the judgment trait and relevant descriptor make it unlikely that only one was activated in memory and facilitated recognition for the other.

An alternative means for dealing with the association problem would be to empirically establish the degree to which activation of one concept in memory facilitates recognition of an associated concept of interest. Thus, it would have been possible to present subjects with randomly selected stimulus traits and determine the degree to which their activation facilitated recognition of other words within the same subset. In this way, a base association level could have been established reflecting the degree to which activation of a relevant descriptor could have been expected to facilitate recognition for a judgment trait (and vice versa).

**Applications**

In order to identify the probe procedure's potential for contributing towards the understanding of how people make memory-based person judgments, it is first necessary to consider briefly the two groups of factors which determine what a person remembers during a decision. The first, investigated in Experiment 4, concerns the manner in which person information is encoded and organized in memory. These initial information processing stages influence the general cross-situational availability, or ease of recall, of different cognitions in two ways. First, they determine what "cognitive units" will be stored; second, they influence the relative availability of each unit. For example, if a person were given the digits 9, 3, 7, 8, and 1 to remember, she
or he might encode and organize them in memory as the numbers 13 and 789. The later availability of any single digit would be a function of both the cognitive unit formed and the relative ease with which each unit could later be remembered. Similarly, when person information is perceived and organized in memory its subsequent availability will be a function of the cognitive units into which it is transformed as well as the relative ease with which each unit can be accessed across a variety of situations.

The second important determinant of what is remembered during a decision is the judgment context. This also influences retrieval in two ways. First, the context may cue memory for a particular cognition. Second, the context will influence the way in which memory is searched. Important features of memory search strategies include both the types of cognitions that are sought and the decision rule that is used to terminate the search. For example, in one decision context someone might search for any single attribute that would indicate a person could be a success in a particular occupation. In a different decision context, the same individual might try and remember all of the negative facts she or he could about a person. In each case, different cognitive units would serve as the object of the memory search and a different decision rule would be used to terminate the search.

In summary, what someone remembers when making a person judgment is the result of a complex interaction between (a) the general cross-situational availability of different cognitions in memory resulting from encoding and memory organizational properties; (b) the contextual availability of cognitions due to cueing effects and (c) the particular
memory search strategy that is employed. The probe procedure would appear to possess a number of unique features for investigating these several factors.

**Capability to Investigate Thought Activation in Multiple Contexts**

From the above discussion, it follows that there are two important reasons for measuring memory for person information in a variety of contexts: (a) to establish the cross-situational availability of different cognitions in order to identify impression organization and (b) to identify ways in which alternative judgment contexts interact with impression organization to determine what is remembered. A problem with measuring thought availability by a recall or recognition test is that it provides information about availability within a single limited context. Consequently, such measures are limited in the answers they can provide concerning what is likely to be remembered in a multitude of judgment contexts. The probe procedure, on the other hand, makes it possible to measure thought activation within a wide variety of both judgment and non-judgment contexts. As a result, it is potentially more useful than recall and recognition measures for investigating both the cross-situational availability of thought as well as the interactions between this availability and a variety of contextual cues.

**Relative Independence of Processing Assumptions**

Like the probe procedure, decision time as a means for inferring how information is remembered during a judgment can also be used in a variety of retrieval contexts. However, unlike the probe technique, decision time measures are difficult to interpret without well defined processing models that generally rest upon multiple assumptions. For example, one such
assumption that has been mentioned is that information items are searched and activated in memory in a serial, as opposed to parallel, fashion. Interpretation of the probe procedure does not require this assumption. It is therefore well suited for use when this assumption is known, or suspected, to be untrue.

**Capability to Identify the Order of Thought Activation**

A final characteristic of the probe technique that distinguishes it from other procedures used to measure judgment retrieval processes is its potential for measuring the sequence in which several thoughts are activated in memory. Although the hypothesis in Experiment 4 concerning the order of recall was not supported, the results of Experiment 3 demonstrate the potential of the probe technique to discriminate between thoughts activated early and late during a judgment. Such order differences are important for both practical and theoretical reasons. Practically, people are often rushed in the decisions they make. Thoughts that are normally activated late during a lengthy decision may not be accessed when time is short. Furthermore, it may be the case that order effects occur in different situations such that thoughts activated either early or late during a judgment are weighted more heavily in the decision.

Theoretically, the order in which people recall information reflects the manner in which memory operates. For example, Bartlet's (1932) early reconstructive theory suggests that an organizing theme should be recalled prior to episodic incidents supportive of the theme. To date, support for this proposition has come from demonstrations that theme-consistent errors in recall increase with the passage of time (cf., Bartlet, 1932; Sulin & Dooling, 1974). However, this evidence has been gathered
within the limited contexts of recognition and recall tests. Investigation of the order in which information is recalled in a variety of judgment contexts through use of the probe technique might demonstrate that reconstructive memory search processes are much more symptomatic of memory tests than decision tasks.
APPENDIX A

Lyrics to the rhymes and songs used as stimuli in Experiments 1 and 2 (underlined words were used as probes).
Practice trials

P1 Happy birthday. . .

Happy birthday to you. Happy birthday to you.
Happy birthday dear John, Happy birthday to you.

P2 Mary had a. . .

Mary had a little lamb, a little lamb, a little lamb.
Mary had a little lamb, its fleece was white as snow.

P3 London Bridge. . .

London bridge is falling down, falling down, falling down.
London bridge is falling down, my fair lady.

P4 Daisy. . .

Daisy, Daisy give me your answer true.
I'm half crazy over the love of you.

P5 We wish you. . .

We wish you a merry Christmas. We wish you a merry Christmas.
We wish you a merry Christmas, and a happy New Year.
Experimental trials

1. Ba, ba black.

   Ba, ba black sheep have you any wool?
   Yes sir! Yes sir! Three bags full.

2. It's raining.

   It's raining, it's pouring. The old man is snoring.
   He went to bed with a cold in his head and didn't get up 'til morning.

3. Oh give me.

   Oh give me a home, where the buffalo roam,
   And the deer and the antelope play.

4. Yankee doodle.

   Yankee doodle went to town riding on a pony;
   Stuck a feather in his hat and called it macaroni.

5. Take me out.

   Take me out to the ball game. Take me out to the game.
   Buy me some peanuts and cracker jacks—I don't care If I ever get back.


   Rock-a-bye, baby, on the tree top.
   When the wind blows, the cradle will rock.

7. Oh Susanna.

   Oh Susanna, oh don't you cry for me,
   For I'm going to Alabama with a banjo on my knee.
   It rained all night, the day I left - the weather it was dry.

8. Joy to the.

   Joy to the world! The Lord is come;
   Let earth receive her king.
9. Rudolph...

Rudolph the red **nose** reindeer, had a very shiny nose;  
And if you **ever** saw it you would even say it glows.

10. Row, row, row...

Row, row, row your boat **gently** down the stream.  
Merrily, merrily, merrily, merrily **life** is but a dream.

11. Three blind...

Three blind mice. Three blind mice. See how they run.  
See how they run. They all **ran** after the farmer's wife.  
She cut off their tails with a carving knife.

12. Jingle...

Jingle bells, jingle **bells**, jingle all the way.  
Oh what fun it is to **ride** in a one **horse** open sleigh.
APPENDIX B

Rhymes used as practice and experimental stimuli in Experiment 3 with probe presentation positions indicated by single underscoring and probe words indicated by double underscoring.
Practice Rhymes

1. Jack be nimble, Jack be quick,
   Jack jump over the candle stick.

2. Hickory, dickory dock, the mouse ran up the clock.
   The clock struck one, the mouse ran down. Hickory, dickory dock.

Experimental Rhymes

1. Little Miss Muffet sat on a tuffet eating her curds and whey,
   Along came a spider and sat down beside her and frightened Miss Muffet away.

2. Mary had a little lamb, its fleece was white as snow,
   And everywhere that Mary went, the lamb was sure to go.

3. Jack and Jill went up the hill to fetch a pail of water.
   Jack fell down and broke his crown, and Jill came tumbling after.

4. Old Mother Hubbard went to the cupboard to get her poor dog a bone.
   When she got there, the cupboard was bare, so the poor dog had none.

5. Star light, star bright, first star I see this night.
   I wish I may, I wish I might have this wish I wish tonight.

6. Twinkle, twinkle little star, how I wonder what you are.
   Up above the world so high, like a diamond in the sky.

7. Humpty Dumpty sat on a wall. Humpty Dumpty had a great fall.
   All the King's horses and all the King's men could not put Humpty together again.
8. There was an old lady who lived in a shoe.

She had so many children she didn't know what to do.
APPENDIX C

Occupations and 32 trait sets used as judgment (J), relevant (R) and irrelevant (I) descriptor traits in Experiment 4.
<table>
<thead>
<tr>
<th>Trait set</th>
<th>Subgroup 1</th>
<th>Subgroup 2</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>inquisitive (J)</td>
<td>righteous (J)</td>
<td>receptionist</td>
</tr>
<tr>
<td></td>
<td>adventurous (R)</td>
<td>moral (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>eager (I)</td>
<td>religious (I)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>thorough (J)</td>
<td>high-strung (J)</td>
<td>news broadcaster</td>
</tr>
<tr>
<td></td>
<td>diligent (R)</td>
<td>jumpy (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>untiring (I)</td>
<td>nervous (I)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>uncompromising (J)</td>
<td>clownish (J)</td>
<td>security guard</td>
</tr>
<tr>
<td></td>
<td>headstrong (R)</td>
<td>entertaining (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stubborn (I)</td>
<td>light-hearted</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>calm (J)</td>
<td>indifferent (J)</td>
<td>bartender</td>
</tr>
<tr>
<td></td>
<td>objective (R)</td>
<td>listless (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rational (I)</td>
<td>unenthusiastic (I)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>extravagant (J)</td>
<td>crafty (J)</td>
<td>telephone</td>
</tr>
<tr>
<td></td>
<td>reckless (R)</td>
<td>scheming (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>impractical (I)</td>
<td>sly (I)</td>
<td>disk jockey</td>
</tr>
<tr>
<td>6</td>
<td>clumsy (J)</td>
<td>short-tempered (J)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sloppy (R)</td>
<td>oversensitive (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>inefficient (I)</td>
<td>hot-headed (I)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>easygoing (J)</td>
<td>neglectful (J)</td>
<td>hair stylist</td>
</tr>
<tr>
<td></td>
<td>relaxed (R)</td>
<td>unobservant (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>casual (I)</td>
<td>inattentive (I)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>methodical (J)</td>
<td>skeptical (J)</td>
<td>bus driver</td>
</tr>
<tr>
<td></td>
<td>systematic (R)</td>
<td>cynical (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>orderly (I)</td>
<td>pessimistic (I)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>preoccupied (J)</td>
<td>self-sufficient (J)</td>
<td>garbage collector</td>
</tr>
<tr>
<td></td>
<td>daydreamer (R)</td>
<td>proficient (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>absent-minded (I)</td>
<td>skillful (I)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>inventive (J)</td>
<td>naive (J)</td>
<td>store clerk</td>
</tr>
<tr>
<td></td>
<td>venturesome (R)</td>
<td>impressionable (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unexceptional (I)</td>
<td>gullible (I)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>unruly (J)</td>
<td>forgetful (J)</td>
<td>psychologist</td>
</tr>
<tr>
<td></td>
<td>daredevil (R)</td>
<td>unpunctual (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>boisterous (I)</td>
<td>inconsistent (I)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>persistent (J)</td>
<td>agreeable (J)</td>
<td>dentist</td>
</tr>
<tr>
<td></td>
<td>single-minded (R)</td>
<td>cordial (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>purposeful (I)</td>
<td>amiable (I)</td>
<td></td>
</tr>
<tr>
<td>Trait set</td>
<td>Subgroup 1</td>
<td>Subgroup 2</td>
<td>Occupation</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>13</td>
<td>bold (J)</td>
<td>soft-spoken (J)</td>
<td>judge</td>
</tr>
<tr>
<td></td>
<td>fearless (R)</td>
<td>composed (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>aggressive (I)</td>
<td>discreet (I)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>tender (J)</td>
<td>overconfident (J)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sensitive (R)</td>
<td>smug (J)</td>
<td>night</td>
</tr>
<tr>
<td></td>
<td>romantic (I)</td>
<td>pompous (I)</td>
<td>watchman</td>
</tr>
<tr>
<td>15</td>
<td>solemn (J)</td>
<td>suave (J)</td>
<td>house painter</td>
</tr>
<tr>
<td></td>
<td>moody (R)</td>
<td>sophisticated (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sad (I)</td>
<td>self-assured (I)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>inhibited (J)</td>
<td>scolding (J)</td>
<td>doctor</td>
</tr>
<tr>
<td></td>
<td>meek (R)</td>
<td>overcritical (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>self-conscious (I)</td>
<td>domineering (I)</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>inexperienced (J)</td>
<td>modest (J)</td>
<td>architect</td>
</tr>
<tr>
<td></td>
<td>innocent (R)</td>
<td>showy (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bashful (I)</td>
<td>theatrical (I)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>quiet (J)</td>
<td>strict (J)</td>
<td>auto mechanic</td>
</tr>
<tr>
<td></td>
<td>modest (R)</td>
<td>moralistic (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shy (I)</td>
<td>stern (I)</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>frivolous (J)</td>
<td>poised (J)</td>
<td>secretary</td>
</tr>
<tr>
<td></td>
<td>spendthrift (R)</td>
<td>cultured (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>compulsive (I)</td>
<td>gracious (I)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>fashionable (J)</td>
<td>careful (J)</td>
<td>waiter</td>
</tr>
<tr>
<td></td>
<td>popular (R)</td>
<td>cool-headed (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>modern (I)</td>
<td>prudent (I)</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>superstitious (J)</td>
<td>discontented (J)</td>
<td>carpenter</td>
</tr>
<tr>
<td></td>
<td>illogical (R)</td>
<td>melancholy (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fearful (I)</td>
<td>down-hearted (I)</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>unpredictable (J)</td>
<td>unadventurous (J)</td>
<td>assembly</td>
</tr>
<tr>
<td></td>
<td>nonconforming (R)</td>
<td>withdrawn (R)</td>
<td>line worker</td>
</tr>
<tr>
<td></td>
<td>changeable (I)</td>
<td>timid (I)</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>soft-hearted (J)</td>
<td>vivacious (J)</td>
<td>chemist</td>
</tr>
<tr>
<td></td>
<td>sentimental (R)</td>
<td>impulsive (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>emotional (I)</td>
<td>exuberant (I)</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>sarcastic (J)</td>
<td>talkative (J)</td>
<td>street</td>
</tr>
<tr>
<td></td>
<td>critical (R)</td>
<td>forward (R)</td>
<td>sweeper</td>
</tr>
<tr>
<td></td>
<td>fault-finding (I)</td>
<td>confident (I)</td>
<td></td>
</tr>
<tr>
<td>Trait set</td>
<td>Subgroup 1</td>
<td>Subgroup 2</td>
<td>Occupation</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>25</td>
<td>rebellious (J)</td>
<td>precise (J)</td>
<td>fisherman</td>
</tr>
<tr>
<td></td>
<td>hot-tempered (R)</td>
<td>meticulous (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>argumentative (I)</td>
<td>perfectionist (I)</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>daring (J)</td>
<td>authoritative (J)</td>
<td>P.E. teacher</td>
</tr>
<tr>
<td></td>
<td>adventurous (R)</td>
<td>proud (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>curious (I)</td>
<td>self-righteous (I)</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>excitable (J)</td>
<td>anxious (J)</td>
<td>stenographer</td>
</tr>
<tr>
<td></td>
<td>vigorous (R)</td>
<td>worrier (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>outgoing (I)</td>
<td>tense (I)</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>congenial (J)</td>
<td>scientific (J)</td>
<td>chef</td>
</tr>
<tr>
<td></td>
<td>sociable (R)</td>
<td>serious (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>charming (I)</td>
<td>studious (I)</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>opinionated (J)</td>
<td>neurotic (J)</td>
<td>book</td>
</tr>
<tr>
<td></td>
<td>blunt (R)</td>
<td>eccentric (R)</td>
<td>publisher</td>
</tr>
<tr>
<td></td>
<td>outspoken (I)</td>
<td>finicky (I)</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>strong-minded (J)</td>
<td>frustrated (J)</td>
<td>janitor</td>
</tr>
<tr>
<td></td>
<td>decisive (R)</td>
<td>tempermental (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>positive (I)</td>
<td>artistic (I)</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>troubled (J)</td>
<td>persuasive (J)</td>
<td>taxi driver</td>
</tr>
<tr>
<td></td>
<td>dissatisfied (R)</td>
<td>self-confident (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>worrying (I)</td>
<td>convincing (I)</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>shrewd (J)</td>
<td>submissive (J)</td>
<td>plumber</td>
</tr>
<tr>
<td></td>
<td>opportunist (R)</td>
<td>humble (R)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cunning (I)</td>
<td>obedient (I)</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


Rosenberg, S., & Jones, R. A. A method for investigating and re­
presenting a person's implicit theory of personality: Theodore 
Dreiser's view of people. Journal of Personality and Social 
Psychology, 1972, 22, 372-386.

Rosenberg, S., Nelson, C., & Vivekanthan, P. S. A multidimensional 
approach to the structure of personality impressions. Journal 

Rosenberg, S., & Sedlak, A. Structural representations of implicit 
personality theory. In L. Berkowitz (Ed.), Advances in 

Schneider, D. J. Implicit personality theory: A review. Psychological 

Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes 
in semantic memory. In S. Kornblum (Ed.) Attention and per­

Snyder, M., & Uranowitz, S. W. Reconstructing the past: Some cognitive 
consequences of person perception. Paper presented at the 
meeting of the Western Psychological Association, Seattle, 
April 1977.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction­

Stroop, J. R. Studies of interference in serial verbal reactions. 

Sulin, R. A., & Dooling, D. J. Intrusion of thematic ideas in retention 
of prose. Journal of Experimental Psychology, 1974, 103, 255- 
262.

Walters, H. A., & Jackson, D. N. Group and individual regularities in 
trait inference: A multidimensional scaling analysis. 

Wainer, H. Speed vs. reaction time as a measure of cognitive perform­

Warr, P. Inference magnitude, range, and evaluative direction as factors 
affecting relative importance of cues in impression formation. 
Journal of Personality and Social Psychology, 1974, 20, 191- 
197.

Warr, P., & Jackson, P. The importance of extremity. Journal of 

Warren, R. E. Stimulus encoding and memory. Unpublished doctoral 

